



**UNIVERSIDAD SIMÓN BOLÍVAR
DECANATO DE ESTUDIOS PROFESIONALES
COORDINACIÓN DE INGENIERÍA ELÉCTRICA**

**ESTIMACIÓN DE DEMANDA A MEDIANO PLAZO DE LA SUBESTACIÓN
LAGUNITA UTILIZANDO REDES NEURONALES**

Por:

Rogney José Sifontes Rodríguez

INFORME DE PASANTÍA
Presentado ante la Ilustre Universidad Simón Bolívar
como requisito parcial para optar al título de
Ingeniero Electricista

Sartenejas, Marzo de 2010



**UNIVERSIDAD SIMÓN BOLÍVAR
DECANATO DE ESTUDIOS PROFESIONALES
COORDINACIÓN DE INGENIERÍA ELÉCTRICA**

**ESTIMACIÓN DE DEMANDA A MEDIANO PLAZO DE LA SUBESTACIÓN
LAGUNITA UTILIZANDO REDES NEURONALES**

Por:

Rogney José Sifontes Rodríguez

Realizado con la asesoría de:

Tutor Académico: Prof. Paulo De Oliveira

Tutor Industrial: Ing. Moisés Marcano

INFORME DE PASANTÍA

Presentado ante la Ilustre Universidad Simón Bolívar
como requisito parcial para optar al título de
Ingeniero Electricista

Sartenejas, Marzo de 2010



UNIVERSIDAD SIMÓN BOLÍVAR
D E x Decanato de Extensión

Coordinación de Cooperación Técnica y Desarrollo Social

Fecha: _____

ACTA DE EVALUACIÓN DEL PERÍODO _____

TIPO DE PASANTÍA: LARGA INTERMEDIA

Título: Estimación de la Demanda a Mediano Plazo de los Circuitos de la S/E Lagunitas utilizando Redes Neuronales Artificiales

Apellidos y Nombre del Estudiante: Sifontes, Roney

Carnet N°: 04-37612 Carrera: Inq. Electrica.

CALIFICACIÓN	APROBADO <input checked="" type="checkbox"/>	REPROBADO <input type="checkbox"/>
TUTOR ACADÉMICO	Firma: Nombre: <u>PAULO H. DE OLIVEIRA</u> C.I.: <u>11024006</u>	
TUTOR INDUSTRIAL	Firma: _____ Nombre: _____ C.I.: _____	
JURADO	Firma: Nombre: <u>Alexander Buena</u> C.I.: <u>10.511.239.</u>	
<p>Artículo 15. Parágrafo Único. Reglamento de los Cursos en Cooperación: Cuando el jurado examinador considere por unanimidad que el Trabajo realizado por el pasante es EXCEPCIONALMENTE BUENO, lo hará constar en forma razonada en el Acta</p> <p>_____ _____ _____</p>		

Nota: En caso de inasistencia del Tutor Industrial, el jurado examinador quedará conformado por el Tutor Académico y Jurado, y se solicitará al estudiante la evaluación del Tutor Industrial como referencia, en este caso favor colocar No-Presente en el Acta

Puerta de Comunidad a Comunidad

ESTIMACIÓN DE DEMANDA A MEDIANO PLAZO DE LA SUBESTACIÓN LAGUNITA UTILIZANDO REDES NEURONALES

POR

Rogney José Sifontes Rodríguez

RESUMEN

Una de las etapas fundamentales en todo estudio de planificación es la estimación de demanda. La previsión acertada de la demanda de energía eléctrica es un requisito imprescindible para lograr las metas de calidad y confiabilidad del servicio. Existen diferentes métodos para estimar la demanda eléctrica a mediano plazo, uno de ellos, utilizado por la Electricidad de Caracas, es el método de Holt-Winters (triple suavizado exponencial). Por otra parte, la data histórica recopilada del sistema de adquisición de datos (SCADA) presenta una serie de distorsiones que afectan los resultados de la estimación de demanda. Para mejorarlos, es indispensable depurar las distorsiones y luego introducir los datos en el modelo de Holt-Winters. Por esta razón, se desea evaluar el potencial de las Redes Neuronales Artificiales (RNA) para la estimación de la demanda a mediano plazo y determinar la capacidad de esta herramienta para tolerar las distorsiones de la data histórica sin comprometer la precisión de sus resultados, de manera que no se requiera de un proceso previo de corrección de datos. En este estudio se utilizan los datos históricos de la demanda durante cinco años de los circuitos de la subestación Lagunita y son comparados los resultados obtenidos con las RNA y los correspondientes al modelo de Holt-Winters. Las Redes Neuronales resultaron ser una buena herramienta para el pronóstico, sin embargo es necesario contar con una data “limpia” para mejorar la precisión de los resultados.

DEDICATORIA

Dedicado con cariño a mis padres,

Zulay Rodríguez

y José Ramón Sifontes.

AGRADECIMIENTOS

Primordialmente quiero agradecer a mis padres, José Ramón y Zulay, quienes han sido mi inspiración, mi apoyo incondicional y mi fortaleza en todo momento. Gracias a ellos soy quien soy y he podido lograr todo lo que me he propuesto. Soy afortunado de tenerlos.

A mi hermano Anthony quien sin dudarle me tiende su mano para brindarme su apoyo en todos mis emprendimientos y más de una vez me sacó de apuros.

A mi novia Gabriela Rodríguez Boldú, quien a lo largo de este proceso no permitió que mis ánimos se derrumbaran y siempre estuvo a mi lado brindándome lo mejor de sí. Siempre fuiste el impulso que necesitaba en momentos de desaliento, sin ti no habría podido lograrlo. Eres una persona muy importante y especial para mí, te quiero.

A la Ing. Verónica Jiménez por brindarme sus sabios consejos y su colaboración oportuna y llena de buenas intenciones. Tu ayuda fue clave para lograr este éxito. No sé cómo podré retribuirte tanto empeño y dedicación. Fuiste quien con mucho acierto pudiste guiarme y marcarme el rumbo que debía seguir para conquistar este logro. Dios te bendiga amiga.

A las Profesoras Evelyn Abdala y Norghy Rodríguez, por no perder la fe y la entrega con la que realizan la importante labor de brindar la oportunidad de pertenecer a esta magnífica casa de estudios. Gracias por creer siempre en mí e incentivar me a alcanzar mis metas. Nunca me cansaré de reconocer que el PIO asentó las sólidas bases que me permitieron avanzar a lo largo de mis estudios universitarios y hoy los resultados hablan por sí solos.

A la Prof. Minaya Villasana por su valiosa colaboración, tiempo y esfuerzo dedicado a este trabajo. Siempre estuvo dispuesta a ayudarme y aconsejarme, muchas gracias.

A mis panas Mario, Luber, Edward (Chayanne) y Aldemaro, por todos esos excelentes momentos vividos durante todos estos años de amistad, en los que hemos superado situaciones difíciles y hemos sabido compartir los éxitos. Esto también se lo debo a ustedes.

Gracias Dios, muchas gracias.

ÍNDICE GENERAL

DEDICATORIA.....	v
AGRADECIMIENTOS.....	vi
ÍNDICE GENERAL.....	vii
ÍNDICE DE FIGURAS	xii
LISTA DE ABREVIATURAS.....	xiv
INTRODUCCIÓN.....	1
CAPÍTULO 1. DESCRIPCIÓN DE LA EMPRESA	2
1.1 La Electricidad de Caracas	2
1.2 Misión.....	2
1.3 Visión	3
1.4 Valores.....	3
1.5 Estructura del departamento de planificación	3
1.6 Objetivos de la VP de distribución	4
1.7 Actividades y objetivos del departamento de planificación	4
1.8 Organigrama de la VP de distribución.	5
CAPÍTULO 2. MARCO TEÓRICO	6
2.1 Planificación del sistema de distribución	6
2.2 Planificación a Mediano Plazo	7
2.2.1 Estimación de demanda.....	7
2.2.2 Métodos de estimación de demanda.....	8
2.2.3 Métodos analíticos.....	8
2.2.5 Métodos tendenciales	9
2.2.6 Ajuste de curvas	9
2.2.7 Descomposición de la curva de demanda.....	10

2.2.8 Medias móviles.....	11
2.2.9 Suavización exponencial simple.....	11
2.2.10 Suavización exponencial de Holt	13
2.2.11 Suavización exponencial de Winters.....	14
2.2.12 Modelo univariante de Box – Jenkins	15
2.2.13 Inteligencia artificial: Redes Neuronales.....	17
2.3 Redes Neuronales	17
2.3.1 Neurona biológica.....	18
2.3.2 Modelo de neurona artificial	19
2.3.3 Funciones de transferencia de una neurona artificial	21
2.3.4 Función escalón	21
2.3.5 Función lineal	22
2.3.7 Función sigmoïdal tangencial.....	23
2.3.8 Topología de las redes neuronales.....	24
2.3.10 Etapas en el diseño de una red neuronal.....	27
2.3.11 Procesamiento de los datos.....	27
2.3.12 Diseño de la red	27
2.3.13 Implementación	28
2.3.14 Validación.....	28
2.3.15 Perceptrón Simple	28
2.3.16 Perceptrón Multicapa.....	29
2.3.17 Descripción de una red perceptrón multicapas.....	30
2.3.18 Algoritmo de Retropropagación o “Backpropagation”	32
2.4 Descripción de las herramientas utilizadas.....	35
2.4.1 MATLAB®	35
CAPÍTULO 3. METODOLOGÍA.....	43

3.1 Etapas en el proceso de estimación de demanda	43
3.1.1 Selección de la red neuronal	44
3.1.2 Procesamiento de los datos de entrada	45
3.1.3 Selección de las variables de entrada	46
3.1.4 Filtrado	47
3.1.5 Depuración	48
3.1.6 Normalización	50
3.1.7 Entrenamiento de la red.....	50
3.1.8 Validación de la red.....	51
3.2 Realización de pronósticos	52
3.2.1 Pronóstico tipo 1	52
3.2.2 Pronóstico tipo 2.....	53
3.2.3 Pronóstico tipo 3.....	53
3.2.4 Pronóstico tipo 4.....	53
3.3 Comparación de resultados con el método de Holt-Winters	54
CAPÍTULO 4. DESCRIPCIÓN DEL CASO DE ESTUDIO	55
4.1 Subestación Lagunita.....	55
4.2 Registro histórico de la demanda de la Subestación Lagunita	56
4.3 Sistema SCADA (Supervisory Control And Data Acquisition).....	56
4.4 Definición de patrones presentes en el registro	58
4.5 Corrección de las distorsiones presentes en el registro	59
CAPÍTULO 5. RESULTADOS	62
5.1 Simulación del pronóstico tipo 1	62
5.2 Simulación del pronóstico tipo 2.....	64
5.3 Simulación del pronóstico tipo 3.....	65
5.4 Simulación del pronóstico tipo 4.....	66

5.5 Resultados obtenidos con el método de Holt-Winters	68
5.6 Comparación y análisis final	69
CONCLUSIONES Y RECOMENDACIONES	73
REFERENCIAS BIBLIOGRÁFICAS	75
ANEXO A	77
ANEXO B	91
ANEXO C	93
ANEXO D	107
APÉNDICE A	111

ÍNDICE DE TABLAS

Tabla 2.1 Planificación del sistema de distribución.....	6
Tabla 2.2 Funciones más utilizadas para el ajuste de curva de demanda.....	9
Tabla 2.3 Funciones de transferencia	23
Tabla 4.1. Características de la subestación Lagunita.....	55
Tabla 5.1. Errores porcentuales de la salida de las RNA y método de Holt-Winters respecto a la serie histórica.....	69
Tabla 5.2. Error ponderado global de los 4 tipos de pronósticos realizados con las RNA	71
Tabla 5.3. Errores porcentuales entre las estimaciones utilizando RNA y método de Holt-Winters.....	72

ÍNDICE DE FIGURAS

Figura 1.1 Organigrama de la VP de distribución.....	5
Figura 2.1 Neurona biológica.....	19
Figura 2.2 Modelo de neurona artificial.....	20
Figura 2.3 Gráfica de la función escalón.....	21
Figura 2.4 Gráfica de la función lineal.....	22
Figura 2.5 Gráfica de la función sigmoideal logarítmica.....	22
Figura 2.6 Gráfica de la función sigmoideal tangencial	23
Figuras 2.7 Estructuras de redes de capa simple y múltiple respectivamente.....	25
Figura 2.8 Perceptrón simple.....	29
Figura 2.9 Perceptrón multicapa	31
Figura 3.1 Esquema de las etapas del proceso de estimación de demanda con RNA.....	44
Figura 3.2 Filtro digital de MATLAB.....	48
Figura 4.1. Esquema General del funcionamiento del SCADA.....	56
Figura 4.2. Distorsiones típicas en el histórico de la demanda. Circuito Lagunita A4	59
Figura 4.3 Demanda del circuito Lagunita A4.....	59
Figura 4.4. Filtrado de la Data del circuito Lagunita A4	60
Figura 4.5 Depuración de la Data del circuito Lagunita A4	61
Figura 5.1 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A4	63
Figura 5.2 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita B2	63
Figura 5.3 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A4	64
Figura 5.4 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita B2	64
Figura 5.5 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A4	65
Figura 5.6 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita B2	66
Figura 5.7 Pronóstico tipo 4 de la demanda de cinco años del circuito Lagunita A4	67

Figura 5.8 Pronóstico tipo 4 de la demanda de cinco años del circuito Lagunita B2	67
Figura 5.9 Pronóstico de la demanda de cinco años del circuito Lagunita A4 según método de Holt Winters.....	68

LISTA DE ABREVIATURAS

AR	Auto Regresivo
ARIMA	Autoregresivos y Media Móvil Integrados
ARMA	Autoregresivo y Media Móvil
BP	Backpropagation (Retropropagación)
EDC	Electricidad de Caracas
MA	Media Móvil
RECM	Error Cuadrático Medio
RNA	Red Neuronal Artificial
SCADA	<i>Supervisory Control And Data Acquisition</i> (Control de Supervisión y Adquisición de Data)
S/E	Subestación
SOM	<i>Self Organizing Map</i> (Mapa de Auto-organización)

INTRODUCCIÓN

De acuerdo al desarrollo de una región, la demanda eléctrica tiende a aumentar a través del tiempo, por lo cual las empresas del sector eléctrico deben elaborar planes para suplir la magnitud de esta demanda. La planificación a mediano plazo se encarga de estimar la demanda futura para un período de cinco años y existen diferentes métodos para llevar a cabo esta labor. Algunos de los más utilizados son los métodos tendenciales, que predicen el comportamiento de la demanda en el futuro partiendo de datos históricos (series temporales).

El método adoptado por La Electricidad de Caracas para realizar estimaciones de la demanda eléctrica a mediano plazo es el de Holt-Winters, también conocido como triple suavizado exponencial, el cual logra realizar los pronósticos a partir de series históricas mensuales de al menos cinco años. Los datos históricos de los circuitos de las subestaciones de la empresa que se utilizan para realizar las estimaciones se recopilan del sistema SCADA, sin embargo los mismos presentan distorsiones debido principalmente a eventos como traspasos de carga temporales, fallas en el sistema eléctrico, errores de medición en los equipos, problemas de comunicación 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 a mediano plazo.

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

La metodología que se propone contempla la selección del tipo de red neuronal artificial a utilizar en el estudio, procesamiento de los datos histórico del sistema SCADA mediante la selección, filtrado, depuración y normalización de las variables a considerar. Además del entrenamiento de la red neuronal ajustando sus parámetros utilizando el algoritmo de retropropagación (*Backpropagation*), también la validación de la misma fortaleciendo su capacidad de generalización aplicando el principio conocido como “parada prematura” y por último realizar la comparación con el método de Holt-Winters.

CAPÍTULO 1

DESCRIPCIÓN DE LA EMPRESA

1.1 La Electricidad de Caracas

La Electricidad de Caracas, fue fundada en 1895 por el Ingeniero Ricardo Zuloaga, quién construyó la primera planta hidroeléctrica del país. Esta planta, instalada en el curso del río Güaire, comenzó sus operaciones en 1897 al poner en funcionamiento dos unidades generadoras con capacidad de 420KW.

Actualmente, La Electricidad de Caracas es una empresa perteneciente al Estado venezolano que suministra el servicio eléctrico a más de un millón de clientes y cuatro millones y medio de usuarios, cubriendo una extensión de 5200 Km^2 distribuidos entre la Gran Caracas (Vargas, Guatire, Guarenas, Los Teques) y San Felipe en el Estado Yaracuy [1].

Es una compañía que presta servicios en las áreas de Generación, Transmisión, Distribución y Comercialización, con una capacidad instalada de 2.316 Megavatios, para satisfacer la demanda de energía eléctrica de sus áreas servidas [1].

1.2 Misión

Proveer el mejor servicio eléctrico y responder las expectativas de sus clientes, trabajadores y accionistas, contribuyendo así a elevar la calidad de vida de la sociedad venezolana [1].

1.3 Visión

Ser una empresa reconocida a nivel nacional e internacional como líder innovadora, proveedora de un servicio eléctrico de alta calidad, con personal y tecnologías excelentes, financieramente sólida y factor fundamental del sector eléctrico venezolano [1].

1.4 Valores

Los trabajadores y trabajadoras de La EDC ostentan los siguientes valores [1]:

Seguridad. La seguridad siempre está en primera instancia, para su gente, los contratistas y las comunidades.

Integridad. Principalmente se basa en honestidad y responsabilidad.

Excelencia. Grandes esfuerzos para mejorar la eficiencia y desempeño de las actividades y para operar con niveles de clase mundial.

Honramiento de compromisos. Honrar los compromisos con clientes, compañeros, comunidades, accionistas, proveedores y socios, es fundamental que un negocio, en general, sea una contribución positiva a la sociedad.

1.5 Estructura del departamento de planificación

El departamento de planificación de distribución de La Electricidad de Caracas, se organiza por regiones, las cuales son: Región Este, Oeste, Centro, Sureste, Noroeste, Vargas, Los Teques, Guarenas, Guatire y Región San Felipe. Cada región debido a lo extensa que puede ser se divide en zonas, con lo cual se obtiene un mayor control de la planificación y prestación de un mejor servicio dentro de la región. Cada una de las regiones tiene la responsabilidad de realizar las siguientes labores [2, 3]:

- Operación y mantenimiento de la red de distribución eléctrica.
- Mantenimiento del alumbrado público de las vías.
- Diseño y construcción de proyectos eléctricos para los clientes.

- Ejecución de los proyectos de adecuación.
- Expansión y mejora en la red y en el servicio prestado.
- Solución de averías y reclamos de servicio.
- Servicio completo del área comercial.

1.6 Objetivos de la VP de distribución

Entre los principales objetivos están [2]:

- Auditar los procesos.
- Lograr un alto grado de satisfacción en el cliente.
- Reducir las pérdidas técnicas y no técnicas.
- Optimizar los activos y recursos financieros.
- Mejorar el desempeño mediante el uso de las mejores prácticas, la medición de resultados y la promoción de logros.
- Incrementar los ingresos.
- Garantizar una operación rentable bajo la nueva regulación y normativa interna.
- Reducir los accidentes del personal, contratistas y terceros.

1.7 Actividades y objetivos del departamento de planificación

Estos son [2]:

- Coordinar y consolidar los estudios de planificación a corto, mediano y largo plazo, realizados por cada una de las regiones y velar por su realización periódica.
- Apoyar técnicamente a cada una de las regiones de operación y mantenimiento en la realización de los estudios de planificación a corto y mediano plazo.
- Coordinar la actuación y unificación de la plataforma informática requerida en todas las regiones para realizar los estudios de planificación.
- Mantener actualizados y unificados los procedimientos, métodos y criterios técnicos usados por cada región para realizar los estudios de planificación de distribución.
- Velar por un diseño que maximice la utilización del equipamiento y establezca una expansión ordenada y oportuna.

- Asignar prioridades en los proyectos especiales que involucran a todas las regiones y velar por su ejecución.
- Elaborar planes de trabajo periódicos, especificando: alcance, recursos y costos, en conjunto con todas las regiones.
- Preparar anualmente el plan de ejecución y expansión consolidado del sistema de distribución.
- Preparar anualmente el plan de inversiones consolidado del sistema de distribución.
- Detectar requerimientos de normalización de nuevas tecnologías, criterios y procedimientos compartidos por las regiones de distribución.
- Solicitar al comité de normalización la elaboración de normas, criterios y procedimientos que satisfagan las necesidades de la planificación de distribución.
- Detectar y solucionar necesidades de entrenamiento en el área de planificación.

1.8 Organigrama de la VP de distribución.

En la Figura 1.1 se muestra el organigrama de la VP de distribución [4]:

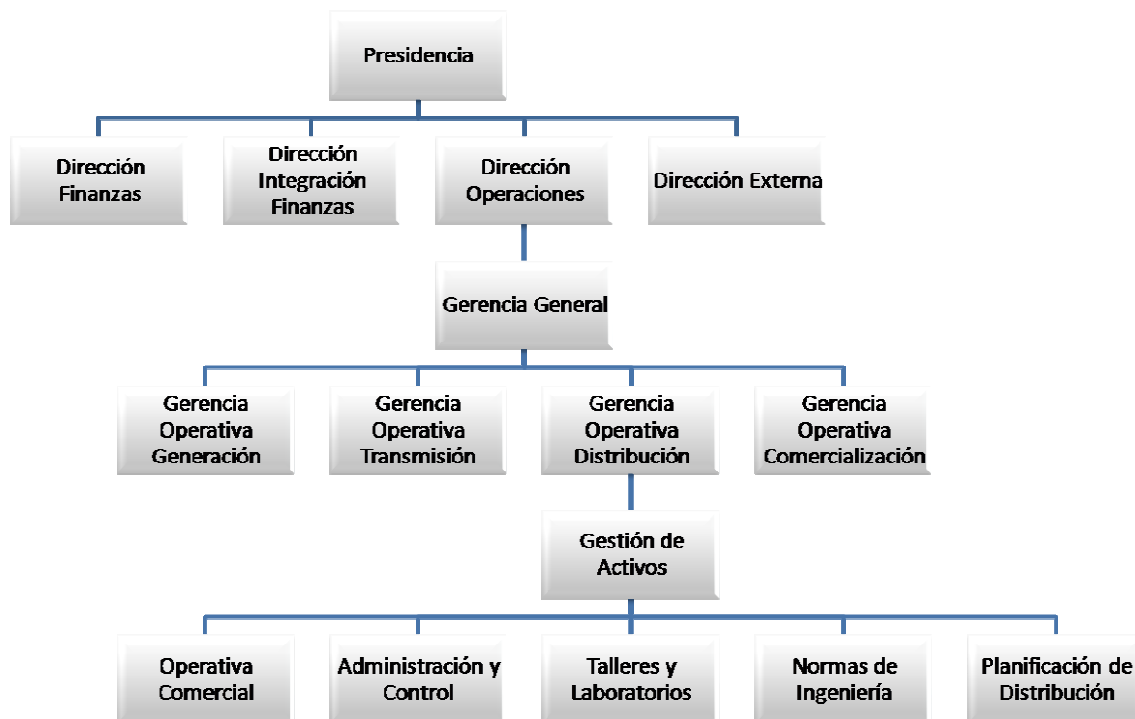


Figura 1.1 Organigrama de la VP de distribución.

CAPÍTULO 2

MARCO TEÓRICO

2.1 Planificación del sistema de distribución

La planificación de un sistema eléctrico es la que permite analizar el estado del mismo, ordenar todos sus elementos y obtener los recursos e inversiones necesarias para garantizar un suministro de energía eléctrica confiable y de calidad en el tiempo. La planificación del sistema de distribución se divide en tres etapas que se muestran a continuación en la Tabla 2.1:

Tabla 2.1 Planificación del sistema de distribución [5]

Planificación a corto plazo	Planificación a mediano plazo	Planificación a largo plazo
Mejoras y expansiones de la red durante el período de un año.	Establece estrategias de expansión de las subestaciones existentes y el desarrollo de subestaciones futuras, más los pequeños desarrollos que necesite la red, estimando la demanda para un período de 5 años.	Establece los requerimientos de futuras subestaciones, define rutas troncales de alimentadores y áreas de servicio en el futuro. Cubre períodos de 8 a 20 años. Define el terreno para la infraestructura eléctrica en general.

2.2 Planificación a Mediano Plazo

El estudio a mediano plazo define los lineamientos para el crecimiento y expansión de la red de distribución de una región, el cual es necesario para suplir la demanda eléctrica producto del desarrollo de la región. Las etapas de planificación a mediano plazo contemplan [3]:

- 1) Recopilación y actualización de información
- 2) Estimación de demanda
- 3) Creación de escenarios
- 4) Diagnóstico de las condiciones de operación
- 5) Planteamiento de estrategias y toma de decisiones

2.2.1 Estimación de demanda

La estimación de demanda es un procedimiento que permite analizar las principales características del consumo de energía eléctrica y tratar de predecir el consumo de la misma en el futuro, con el fin de ejercer acciones para garantizar su suministro en todo momento.

Por medio de la estimación de demanda, es posible determinar previsiblemente si se va a producir una insuficiencia de la capacidad generadora o por el contrario, en el futuro existirá un exceso de capacidad que sugiera desaprovechamiento de parte de la generación ya existente. Una previsión correcta de la demanda de energía eléctrica es un requisito imprescindible para lograr las metas de calidad y confiabilidad del servicio previsto, puesto que la fuerte dependencia de la electricidad aumenta los inconvenientes causados a los consumidores si se producen deficiencias en el suministro de energía eléctrica.

Una de las etapas fundamentales en todo estudio de planificación es la estimación de demanda. Sin una adecuada representación de las necesidades futuras de generación eléctrica, los problemas de exceso o deficiencia de la misma, pueden ocasionar costos realmente altos. Si las predicciones resultan ser demasiado bajas se originan carencias de energía cuyos costos son generalmente mucho mayores que el valor de la energía no suministrada. Por otro lado, si las

predicciones son demasiado altas, los costos resultan muy elevados al tener comprometidos, de manera improductiva, grandes fondos económicos por largos períodos de tiempo.

De todo esto se deriva la importancia que tiene la precisión de la estimación que se realiza. Es claro que mientras más acertada sea la estimación, menor será el riesgo de incurrir en inversiones innecesarias o insatisfacción de los usuarios.

2.2.2 Métodos de estimación de demanda

Para la estimación de la demanda, se necesitan aplicar metodologías según el tipo de estudio que se vaya a realizar (corto, mediano o largo plazo). Estas metodologías se han perfeccionado a lo largo de la historia de la evolución de los sistemas de distribución.

Existen varios métodos para realizar estimaciones de demanda y estos pueden clasificarse en [2]:

- Métodos analíticos
- Métodos no analíticos
- Métodos tendenciales

2.2.3 Métodos analíticos

Los métodos analíticos son utilizados para predecir el comportamiento de la demanda, tomando en cuenta otras variables relacionadas con ésta como la movilidad de población, empleo, zonificación y otras. Este método estudia la relación entre estas variables y la manera como éstas describen el comportamiento de la demanda, pero es un método complicado de aplicar que se utiliza usualmente para estudios de largo plazo. [2]

2.2.4 Métodos no analíticos

Los métodos no analíticos se apoyan básicamente en la experiencia que el planificador haya adquirido de la zona en estudio.

En este método se utilizan metodologías y parámetros sencillos tales como tasas interanuales de crecimiento o incrementos anuales constantes. [2]

2.2.5 Métodos tendenciales

Los métodos tendenciales son aquellos que extrapolan al futuro los patrones de la demanda en el pasado. Varios procedimientos matemáticos han sido utilizados para esta tarea de extrapolación, pero todos ellos comparten un concepto fundamental y es que sus pronósticos se basan en series de tiempo históricas de la demanda. Las series de tiempo son un registro organizado a intervalos de tiempo fijos de las características de una variable, o su observación numérica. Se usan para describir y analizar fenómenos a través del tiempo [6].

Para realizar estimaciones de demanda utilizando métodos tendenciales es necesario disponer de datos históricos de demanda y se deben tener por lo menos el mismo número de años de demanda que los que se quieren pronosticar [2]. Se conocen diversos métodos tendenciales, a continuación se da una breve descripción de los más conocidos:

2.2.6 Ajuste de curvas

Este método intenta describir mediante alguna expresión matemática el comportamiento que ha tenido la demanda en el pasado, y a partir de ésta extrapolar la curva. Para este ajuste pueden ser utilizadas varias expresiones, las más comunes se encuentran en la Tabla 2.2:

Tabla 2.2 Funciones más utilizadas para el ajuste de curva de demanda [2]

Nombre	Expresión matemática
Lineal	$f(t) = a + bt$
Exponencial	$f(t) = ae^{bt}$
Gompertz 1	$f(t) = e^{\frac{a-b}{t-D}}$
Gompertz 2	$f(t) = \frac{S}{1 + e^{a-bt}}$
Cúbica	$f(t) = a + bt + ct^2 + dt^3$

Donde:

a, b, c, d : Parámetros de ajuste de la curvas

S : Saturación

D : Desplazamiento en tiempo

t : Tiempo

También puede realizarse un ajuste de curvas utilizando polinomios de otros grados (no sólo de tercer grado como se muestra en la Tabla 2.2). Un criterio que permite saber cual polinomio utilizar es la sumatoria de errores cuadráticos, la cual debe ser mínima. A medida que se incrementa el grado del polinomio, la sumatoria del error cuadrático tiende a disminuir pero debe tenerse cuidado de no perder generalidad en el modelo ni complicarlo demasiado. Por lo general para describir el comportamiento de la demanda con el fin de obtener una extrapolación de la misma, no es necesario que el grado del polinomio sea muy elevado.

2.2.7 Descomposición de la curva de demanda

Este método parte desde el principio de que la curva de demanda está compuesta por varios elementos o patrones básicos, los cuales son:

- Componente tendencial
- Componente cíclico
- Componente estacional

El componente tendencial se refiere a la proyección lineal de la serie de temporal que marca el crecimiento o decrecimiento de la demanda a lo largo del tiempo. El componente cíclico representa grandes crecimientos o decrecimientos de los valores de demanda que pueden ocurrir en intervalos de tiempo variable. Por último, el componente estacional se refiere a determinados eventos que ocurren constantemente en intervalos de tiempo fijos, por lo que pueden ser más fáciles de reconocer.

2.2.8 Medias móviles

Es un método estadístico sencillo que promedia valores de una serie de tiempo por intervalos, es decir, se promedia un valor de la serie con valores anteriores a éste y se avanza de esta manera hasta el último valor de la misma. La elección del intervalo para la media móvil depende de la longitud de ciclos o pautas presentes en los datos originales. Por ejemplo, si se piensa que existe un ciclo en los datos que se produce cada cierto período, se elegiría una media móvil de ese mismo período para amortiguar en la medida de lo posible las fluctuaciones en el corto plazo. [7]

Entre las limitaciones más resaltantes de este método está su incapacidad para predecir picos y valles, ni máximos y mínimos. Adicionalmente, luego de implementar este método podrían identificarse falsos ciclos que los datos originales no poseen por lo que tiende a confundir al pronosticador. [7]

Las medias móviles se calculan de acuerdo a la siguiente expresión:

$$M = \frac{X_t + X_{t-1} + X_{t-2} + \dots + X_{t-n+1}}{n} \quad (2.1)$$

En la cual:

M : Medias móviles de la serie temporal X

X_t : Valor de la serie temporal en el período t

n : Número de valores que se promedian

2.2.9 Suavización exponencial simple

Es un método similar al de medias móviles que utiliza valores anteriores de la serie de tiempo para predecir valores futuros de la misma serie. Su aplicación es apropiada cuando los datos no presentan ninguna tendencia ni estacionalidad. El valor de predicción para cualquier período es la media ponderada de todos los valores previos conocidos, asignando mayor peso o importancia a

las observaciones recientes; es decir, estos pesos disminuyen geométricamente a medida que se retrocede en el tiempo. [7]

Para describir este método, se llama “ α ” a la constante de suavización, que multiplicada a la observación más reciente representa el peso asignado dicha observación. La siguiente observación más reciente se multiplica por $(1 - \alpha)\alpha$, la que sigue por $(1 - \alpha)^2\alpha$ y así sucesivamente. Todo esto puede escribirse de la siguiente manera [7]:

$$F_{t+1} = \alpha X_t + (1 - \alpha)F_t \quad (2.2)$$

Donde:

F_{t+1} : Valor de la predicción o valor suavizado para el período $(t + 1)$

α : Constante de suavización ($0 < \alpha < 1$)

X_t : Valor real de la serie en el período t

F_t : Valor de la predicción o valor suavizado para el período t

El valor de la constante de suavización debe estar comprendido entre 0 y 1. Valores cercanos a 1, proporcionan mayor ponderación a los valores más recientes de la serie de datos. De lo contrario, un valor de α cercano a 0 hace que los pesos asignados a valores lejanos en el pasado sean comparables con los pesos de los valores recientes. [7]

Un criterio útil para asignar una constante de suavización adecuada es el error cuadrático medio (*RECM*). La constante de suavización con la que se obtenga menor *RECM* sería elegida como el modelo con mayor probabilidad de producir menor error al realizar estimaciones adicionales. [7]

El error cuadrático medio (*RECM*) puede calcularse mediante la siguiente fórmula:

$$RECM = \sqrt{\frac{\sum_{t=1}^n (X_t - F_t)^2}{n}} \quad (2.3)$$

Donde:

RECM : Error cuadrático medio

X_t : Valor real de la serie en el período t

F_t : Valor de la predicción o valor suavizado para el período t

n : Número de muestras

La ventaja de este método reside en que es muy simple de implementar y necesita una cantidad limitada de datos. Por otra parte, su desventaja es que sus predicciones presentan un retardo con respecto a los datos reales y que no es capaz de ajustar la tendencia o estacionalidad de los datos.

2.2.10 Suavización exponencial de Holt

Es una ampliación de la suavización simple, que añade un factor de crecimiento (o factor de tendencia) a la ecuación de suavización simple como método de ajuste de tendencia. Este modelo consta de 3 ecuaciones y 2 constantes de suavización [7]:

$$F_{t+1} = \alpha X_t + (F_t + T_t) \quad (2.4)$$

$$T_{t+1} = \beta(F_{t+1} - F_t) + (1 - \beta)T_t \quad (2.5)$$

$$H_{t+m} = F_{t+1} + mT_{t+1} \quad (2.6)$$

Donde:

F_{t+1} : Valor de la predicción o valor suavizado para el período $(t + 1)$

α : Constante de suavización ($0 < \alpha < 1$)

X_t : Valor real de la serie en el período t

F_t : Valor de la predicción o valor suavizado para el período t

T_{t+1} : Estimación de los valores de tendencia

β : Constante de suavización para la estimación de los valores de tendencia ($0 < \beta < 1$)

m : Número de períodos futuros que se pronosticarán

H_{t+m} : Valor de estimación de Holt para el período $(t + m)$

Este método se utiliza si los valores de la serie histórica poseen estacionalidad escasa o nula, ya que no es capaz de ajustarse a valores máximos y mínimos de los datos.

2.2.11 Suavización exponencial de Winters

Es la segunda ampliación del modelo básico de suavización y es aplicable a datos que presentan tanto tendencia como estacionalidad. También puede verse como una extensión del modelo de Holt, al que se le añade una ecuación que ajusta el mismo al componente estacional. Entonces queda representado por 4 ecuaciones y 3 constantes de suavización [7]:

$$F_t = \frac{\alpha X_t}{S_{t-p} + (1-\alpha)(F_{t-1} + T_{t-1})} \quad (2.7)$$

$$S_t = \frac{\beta X_t}{F_t + (1-\beta)S_{t-p}} \quad (2.8)$$

$$T_t = \gamma(F_t - F_{t-1}) + (1-\gamma)T_{t-1} \quad (2.9)$$

$$W_{t+m} = (F_t + mT_t)S_t \quad (2.10)$$

En las cuales:

F_t : Valor de la predicción o valor suavizado para el período t

α : Constante de suavización ($0 < \alpha < 1$)

X_t : Valor real de la serie en el período t

F_{t-1} : Valor de la predicción o valor suavizado para el período $(t - 1)$

T_{t+1} : Estimación de los valores de tendencia

S_t : Estimación de estacionalidad

β : Constante de suavización para la estimación de los valores de estacionalidad ($0 < \beta < 1$)

γ : Constante de suavización para la estimación de valores de tendencia

m : Número de períodos en el avance de pronóstico

p : Número de períodos en el ciclo estacional

W_{t+m} : Valor de estimación de Winters para m períodos futuros

2.2.12 Modelo univariante de Box – Jenkins

Este modelo utiliza la técnica de combinar observaciones pasadas (esquema Autorregresivo AR) y errores pasados (esquema de Media Móvil MA) de una serie temporal para pronosticar el futuro de la variable de estudio. Se utilizan procedimientos muy particulares para identificar y determinar cuántas observaciones hay que utilizar y cual es su peso respectivo. Permite identificar, a diferencia de otras técnicas, patrones menos explícitos que los obtenidos a través de otras técnicas de serie de tiempo.

Las características más importantes para este tipo de modelación corresponden a la estacionalidad y la estacionareidad.

La metodología de Box Jenkins consiste en enfoque muy general, de alta precisión y además forma un grupo de modelos, los cuales pueden ser clasificados en tres grupos fundamentales: Autorregresivos, Media Móvil, Mixtos que usualmente son llamados ARIMA (Autorregresivos y de Media Móvil Integrado). [8]

Proceso Autorregresivo (AR)

En este tipo de modelación el valor de la serie en el instante t , $y(t)$ es expresado como una función de sus valores previos ($y(t - 1)$, $y(t - 2)$...) y de un error aleatorio $a(t)$. El orden p , de este proceso depende del valor más distante ($t - n$) con el cual $y(t)$ está más relacionado. Así, para un proceso autorregresivo AR(p) de orden p la expresión será:

$$Y(t) = \varphi_1 y(t - 1) + \varphi_2 y(t - 2) + \dots + \varphi_p y(t - p) + a(t) \quad (2.11)$$

En la literatura especializada, para que las ecuaciones sean más compactas, la ecuación anterior es expresada en términos de un operador de retardo B , el cual es definido como $y(t - 1) = B y(t)$, siendo $y(t - m) = B^m y(t)$.

Procesos de Media Móvil (MA)

Este método establece que la demanda o variable dependiente es influenciada por los errores de estimaciones anteriores. De esta manera se obtiene lo siguiente:

$$Y(t) = \theta_1 e(t - 1) + \theta_2 e(t - 2) + \dots + \theta_p e(t - p) + a(t) \quad (2.12)$$

Donde:

p : Número de términos de error

θ_p : Pesos de cada error

$e(t)$: Valor observado del error en la estimación

Proceso Autoregresivo y Media Móvil (ARMA)

Los algoritmos combinados de Autorregresión - Promedios Móviles son una mezcla de los dos anteriores, formulando la hipótesis que la variable dependiente es una función de ambos, los valores previos de la serie y los errores de estimaciones anteriores.

$$Y(t) = \varphi_1 y(t - 1) + \varphi_2 y(t - 2) + \dots + \varphi_p y(t - p) + a(t) \quad (2.11)$$

$$Y(t) = -\theta_1 e(t - 1) - \theta_2 e(t - 2) - \dots - \theta_p e(t - p) + a(t) \quad (2.13)$$

Proceso Autorregresivo y de Media Móvil Integrado (ARIMA)

Los procesos anteriormente definidos como AR, MA o ARMA están asociados a procesos estacionarios. Esto significa que la serie de tiempo tiene una media y una varianza esencialmente constantes en el tiempo.

Cuando esta condición no se cumple es necesario realizar la siguiente transformación:

$$y(t) - y(t - 1) = (1 - B) y(t) \quad (2.14)$$

Así el orden “d” (número de veces que se realizó la diferenciación), es conocido como el factor de integración y el modelo quedará:

$$\varphi(B)\nabla^d y(t) = \theta(B) a(t) \quad (2.15)$$

Donde ∇ es un operador equivalente a B

$$\nabla^d y(t) = (1 - B)^d y(t) \quad (2.16)$$

2.2.13 Inteligencia artificial: Redes Neuronales

La aplicación de redes neuronales artificiales ha tenido un auge desde los años 80 y se han publicado muchos trabajos al respecto. [9]

Las redes neuronales se utilizan para modelar sistemas no lineales, por lo tanto es un método ideal para aplicarlo en curvas de demanda eléctrica en sistemas de distribución. Esta metodología tiene una gran ventaja por su simplicidad y actualmente existen numerosas herramientas computacionales que facilitan la experiencia con las redes neuronales.

Aunque la mayoría del material disponible sobre estimación de la demanda con redes neurales está concentrado en el corto plazo, esta metodología es aplicable al mediano plazo para lo cual es necesario realizar un tratamiento previo de los datos, depurando variables de entrada de la red.

La red más utilizada para realizar la estimación de la demanda a mediano plazo es la conocida como perceptrón multicapa y así mismo el algoritmo de retropropagación es el más comúnmente utilizado en el proceso de aprendizaje o entrenamiento de la red.

2.3 Redes Neuronales

Una red neuronal artificial es un conjunto de unidades procesadoras (neuronas artificiales) distribuidas paralelamente y conectadas entre sí que intentan imitar la actividad cerebral. La red es capaz de aprender de los estímulos de su entorno (datos) y percibir sus características y propiedades.

Las neuronas presentan una tendencia natural a almacenar conocimiento por experiencia y lo ponen a disposición para su uso. Estas redes guardan ciertas similitudes con el cerebro humano, siendo principalmente estas dos:

- El conocimiento es adquirido por la red a través de un proceso de aprendizaje.
- La interconexión entre neuronas, que se conoce como pesos sinápticos o simplemente pesos, son usados para almacenar el conocimiento. [10]

Las redes neuronales artificiales han sido utilizadas para efectuar funciones complejas en varios campos de aplicación entre los que se incluyen el análisis de comportamiento, el reconocimiento de tendencias, la identificación de imágenes, la clasificación de grupos, la optimización de procesos industriales y los sistemas de control entre muchos más. [4]

2.3.1 Neurona biológica

El cerebro humano puede ser visto como una entidad de procesamiento de información. Éste recibe información de un entorno externo por medio de los sentidos y los procesa para formar modelos internos a partir de fenómenos externos. El cerebro es particularmente capaz de ajustar estos modelos, así como interpolarlos y extrapolarlos para adaptarse a nuevas situaciones con mayor agilidad que puede hacer que éste tome decisiones más confiables e incluso reconocimiento de patrones, entendimiento de conceptos y realización de predicciones a partir de una información parcial o incompleta. [11]

Las unidades de procesamiento local de información del cerebro son las neuronas, las cuales forman un gran número de redes paralelas en la corteza del cerebro y son fundamentales para esta actividad. La corteza es una delgada capa externa del cerebro que contiene un gran número de neuronas, entre 100 y 500 billones. El cerebro tiene una gran variedad de neuronas especializadas, y dependiendo del tipo, cada neurona puede enviar señales a cualquiera de cientos de miles de otras neuronas. Ahora se sabe que una repetida excitación de neuronas conduce a un crecimiento de nuevas conexiones entre ellas y entonces se crea y expande una gran red interconectada que es capaz de poseer memoria. La memoria o conocimiento adquirido es almacenado como fuertes conexiones entre neuronas. [11]

Una neurona biológica está conformada por tres componentes principales: las dendritas, que se encargan de canalizar las señales de entrada o estímulos, las cuales son ponderadas por la fuerza de las conexiones, al cuerpo celular; un cuerpo celular que acumula las señales de entrada ponderadas y además procesa estas señales; y un axón que transmite la señal de salida a otras neuronas que están conectadas a éste. [11]

La conexión entre neuronas recibe el nombre de sinapsis. Las conexiones sinápticas pueden excitar o inhibir la señal que se desea transmitir. Una sinapsis fuerte y con gran número de repeticiones tiende a reforzarse y ser más eficiente cada vez en la transmisión; mientras que una sinapsis débil o de poca utilidad tiende a disminuir su eficiencia de transmisión y por lo tanto podría desaparecer. [11]

En la Figura 2.1 se ilustra una neurona biológica.

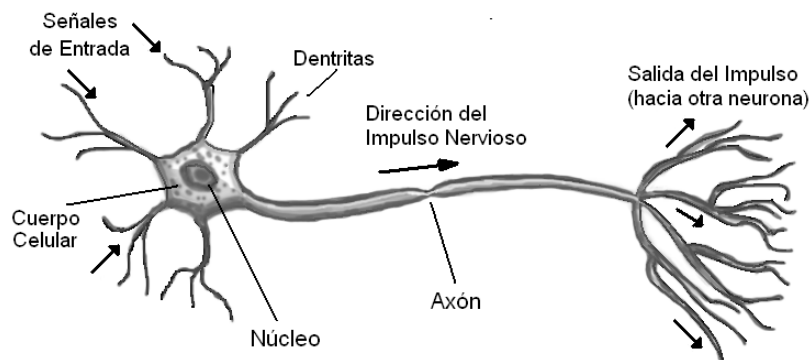


Figura 2.1 Neurona biológica

2.3.2 Modelo de neurona artificial

De la observación detallada del proceso llevado a cabo por neuronas biológicas se ha propuesto un modelo de neurona artificial que se muestra a continuación:

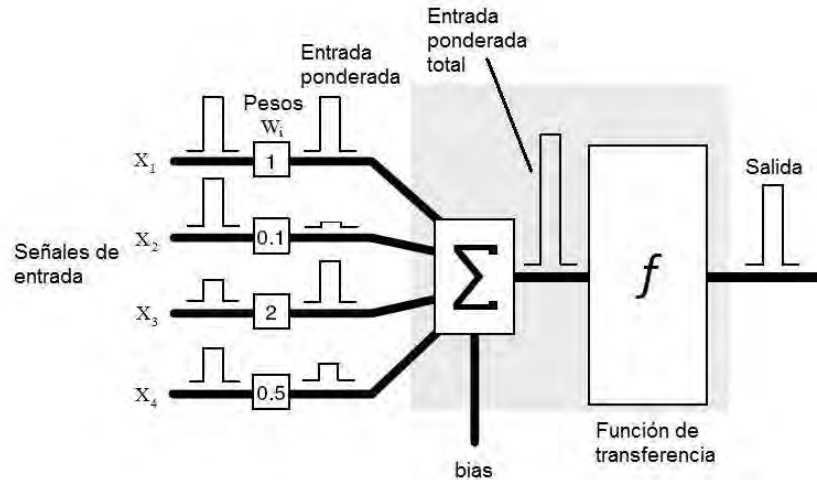


Figura 2.2 Modelo de neurona artificial [11]

De este modelo se pueden identificar tres elementos básicos [10]:

- Un grupo de sinapsis o enlaces interconexión.
- Un sumador de señales de entrada ponderadas de acuerdo a sus respectivos pesos.
- Una función de transferencia que limita la señal de salida.

En una neurona artificial las entradas X_i representan vectores o señales que provienen del exterior o son capturadas de otras neuronas. Los pesos W_i representan la intensidad de la sinapsis; tanto X_i como W_i son valores escalares. A esa señal ponderada se le adiciona el “bias”, que es un parámetro que permite polarizar convenientemente los datos de entrada y entonces se obtiene una entrada total que actúa sobre el núcleo de la neurona para activarla. La función f es la función de transferencia o función de activación que, evaluada para esa entrada total, produce la respuesta de la neurona al estímulo exterior recibido en forma similar a como ocurre en las células nerviosas biológicas.

En términos matemáticos este proceso se representa de forma sencilla como [4]:

$$Salida = f(X_{neta} + b) \quad (2.17)$$

$$X_{neta} = X_1W_1 + X_2W_2 + X_3W_3 + \dots X_nW_n \quad (2.18)$$

Tanto el vector de pesos W como el “bias” b son parámetros escalares y ajustables de la neurona.

2.3.3 Funciones de transferencia de una neurona artificial

La respuesta de salida de una neurona depende tanto de la entrada como de su función de transferencia. Existen funciones de transferencia lineal y no lineal. Se escoge una función de transferencia particular para satisfacer las especificaciones del problema que la neurona o red neuronal intenta resolver. [12]

Se conocen diversas funciones de transferencia, entre las cuales las más utilizadas son: la función escalón, lineal, sigmoideal logarítmica y la sigmoideal tangencial. A continuación se da una breve descripción de cada una de estas funciones.

2.3.4 Función escalón

Esta función hace que la salida de la red sea igual a cero si el argumento de la función es menor que cero y la lleva a uno si este argumento es mayor que cero. Esta función crea neuronas que clasifican las entradas en dos categorías diferentes. El valor umbral que separa las dos categorías mencionadas puede ser desplazado variando el valor del “bias”. Matemáticamente, esta función está definida como:

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases} \quad (2.19)$$



Figura 2.3 Gráfica de la función escalón

2.3.5 Función lineal

La salida de esta función de transferencia es igual a su entrada. Su desplazamiento se realiza mediante el “bias”. Esta función se describe matemáticamente como sigue:

$$f(x) = x \quad (2.20)$$

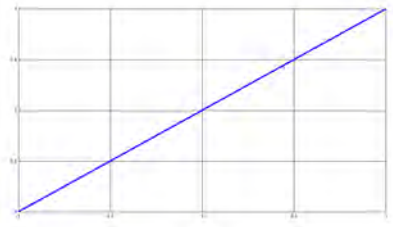


Figura 2.4 Gráfica de la función lineal

2.3.6 Función sigmoideal logarítmica

Esta función limita la salida a valores entre cero y uno, de acuerdo a la expresión:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.21)$$

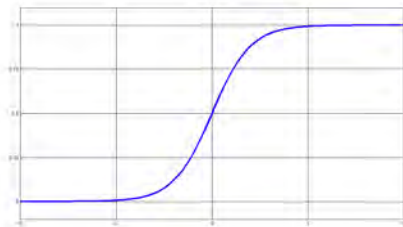


Figura 2.5 Gráfica de la función sigmoideal logarítmica

Se conoce con este nombre debido a que la gráfica correspondiente es similar a la letra “S”. Esta función es comúnmente usada en redes multicapa, entrenadas con el algoritmo de retropropagación, en parte porque se trata de una función diferenciable [12].

2.3.7 Función sigmoïdal tangencial

Es similar a la función sigmoïdal logarítmica pero difiere con ésta en limita la salida entre uno negativo y uno. La expresión que define a esta función es:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.22)$$

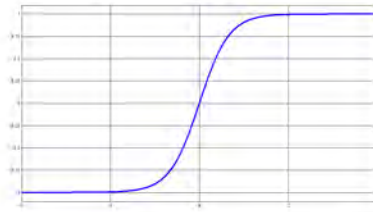
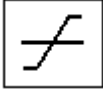

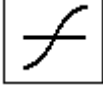


Figura 2.6 Gráfica de la función sigmoïdal tangencial

También es posible definir otras funciones de transferencia, la Tabla 2.3 resume las características de las funciones descritas anteriormente y otras que se han utilizado para caracterizar neuronas artificiales:

Tabla 2.3 Funciones de transferencia

Nombre de la función	Relación Entrada /Salida	Gráfica
Escalón o Limitador Fuerte	$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$	
Escalón Simétrico	$f(x) = \begin{cases} -1 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$	
Lineal Positiva	$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \geq 0 \end{cases}$	
Lineal	$f(x) = x$	
Lineal Saturada	$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$	

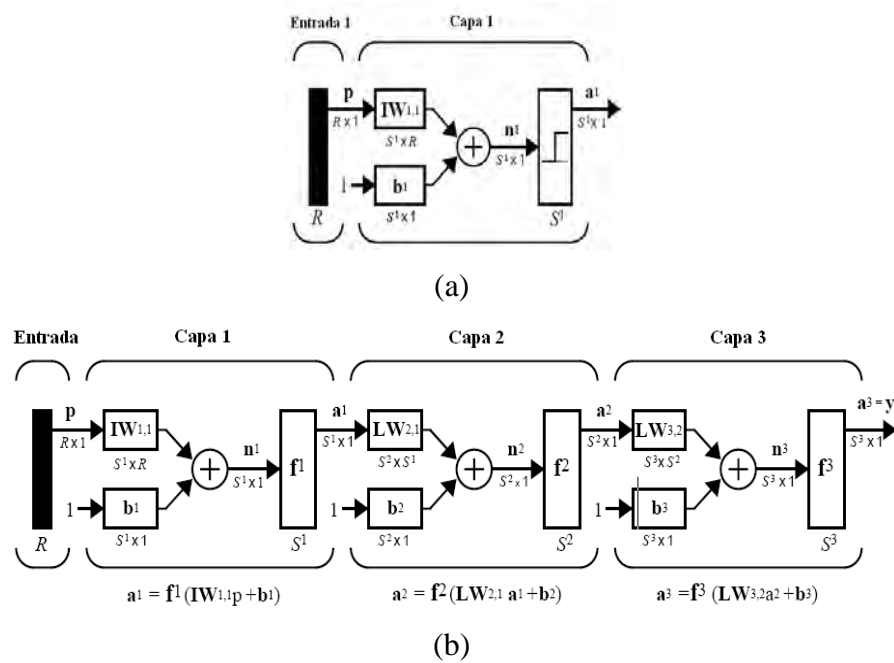
Nombre de la función	Relación Entrada /Salida	Gráfica
Lineal Saturada Simétrica	$f(x) = \begin{cases} -1 & \text{si } x < -1 \\ x & \text{si } -1 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$	
Sigmoidal Logarítmica	$f(x) = \frac{1}{1 + e^{-x}}$	
Sigmoidal Tangencial	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	

2.3.8 Topología de las redes neuronales

Una red neuronal consiste en una organización de neuronas conectadas entre sí, formando capas o agrupaciones interconectadas entre la entrada y la salida. Los componentes fundamentales de la red son: el número de capas, el número de neuronas de cada capa, el grado de conectividad y el tipo de función de transferencia de cada capa [4].

De acuerdo al número de capas, las redes pueden ser clasificadas en redes de simple o de múltiples capas. Las capas existentes entre la capa de entrada y la de salida se llaman capas ocultas. En las redes multicapa normalmente se transmiten las señales de capa a capa; las neuronas de una capa reciben las señales de la capa anterior y envían su salida a la capa contigua en búsqueda de la salida de la red. A estas conexiones se les denominan conexiones hacia delante. Sin embargo, también existe la posibilidad de conectar las salidas de las neuronas de capas posteriores a las entradas de capas anteriores, a estas conexiones se les denomina conexiones hacia atrás. También existen redes con conexiones hacia adelante y hacia atrás. [4]

Pueden observarse en la Figura 2.7 (a) y (b) las estructuras de las redes de capa simple y de múltiples capas:



Figuras 2.7 (a) y (b) Estructuras de redes de capa simple y múltiple respectivamente [15]

2.3.9 Mecanismos de aprendizaje

El aprendizaje o entrenamiento puede ser visto como el proceso mediante el cual una red neuronal modifica sus pesos y ajusta el “bias” en respuesta a una información de entrada. Los cambios que se producen durante el proceso de aprendizaje pueden interpretarse como la destrucción, modificación y creación de conexiones entre las neuronas. La creación de una nueva conexión implica que el peso de la misma pasa a tener un valor diferente de cero; una conexión se destruye cuando su peso pasa a ser cero. El proceso de aprendizaje termina, y se dice que la red ha aprendido cuando los valores de los pesos permanecen estables. [4]

El propósito del proceso de aprendizaje es entrenar la red para que ésta pueda desempeñar algunas tareas. Hay varios tipos de reglas de aprendizaje de redes neuronales que pueden ser clasificadas en tres categorías principales: supervisado, no supervisado y por refuerzo. [12]

El entrenamiento supervisado se realiza mediante un entrenamiento controlado por un agente externo (supervisor o maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida de la red y en el caso de que ésta no

coincida con la deseada, se procederá a modificar los pesos de las conexiones, con el fin de conseguir que la salida se aproxime cada vez más a la deseada. [4]

En este tipo de entrenamiento la red es ajustada mediante la comparación de su salida con un valor establecido como meta o señal de éxito obtenido de la experiencia vivida del sistema en estudio. Los valores de entrada y salida que forman los pares de entrenamiento (meta) se toman de una parte de esa experiencia que se conoce del sistema. El entrenamiento concluye cuando se ha alcanzado un error aceptable en la predicción de los parámetros del sistema por parte de la red, cuando el número de iteraciones que se han programado se excede o cuando así se decide. Una vez que el entrenamiento ha concluido se procede a validar la eficacia de ese aprendizaje al comparar valores conocidos del comportamiento del sistema con la respuesta de la simulación de la red a las entradas correspondientes. Estos pares de simulación generalmente no forman parte del segmento de entrenamiento. Finalmente, si este proceso ha ocurrido satisfactoriamente, se posee una red entrenada para responder a los intereses de simular, evaluar, clasificar, reproducir, estimar, etc., según el interés del diseñador. [4]

En el entrenamiento no supervisado, los pesos sinápticos y “biases” son modificados únicamente en respuesta a entradas presentadas a la red. Esto se debe a que no se dispone de un conjunto de salida deseado o meta. En primera instancia esto podría parecer impráctico. ¿Cómo se puede entrenar una red si no se sabe lo que se supone que ésta debe hacer? La mayoría de estos algoritmos realizan algún tipo de operación de clasificación. La red aprende a categorizar los patrones de los datos entrada dentro de un número finito de clases. La salida de la red con este tipo de entrenamiento representa el grado de familiaridad o similitud entre la información que se le está presentando en la entrada. [12]

El aprendizaje por refuerzo es similar al aprendizaje supervisado, la diferencia radica en que no es proporcionado un valor de salida que represente la respuesta correcta para cada entrada (meta). La función del supervisor consiste en indicar mediante una señal de refuerzo si la salida obtenida en la red se ajusta a la deseada (éxito o fracaso), y en función de ello se ajustan los pesos basándose en un mecanismo de probabilidades. [12]

2.3.10 Etapas en el diseño de una red neuronal

Existen algunos pasos asociados a la creación de una red neuronal que no necesariamente son rigurosos porque tienden a traslaparse en la práctica, pero resultan útiles para establecer cierto orden y tener una idea estructurada del proceso de construcción de la red. Las etapas de diseño se pueden dividir en cuatro [9]:

- Procesamiento de los datos
- Diseño de la red
- Implementación
- Validación

2.3.11 Procesamiento de los datos

Los datos iniciales pueden obtenerse de registros históricos, mediante la simulación por computadora, de tablas de resultados o de mediciones directas [4]. Antes que estos datos sean utilizados como entradas efectivas en ocasiones es necesario un procesamiento inicial con el objetivo de limitar y/o reducir la dimensión del vector de entrada y con ello el crecimiento exponencial de la complejidad del problema vinculada a la magnitud de ese vector. También el procesamiento de los datos puede ser necesario para "limpiarlos", remover errores, valores faltantes o alguna irregularidad. [9]

2.3.12 Diseño de la red

El primer paso es seleccionar una arquitectura apropiada para la red de acuerdo con la tarea que se desea realizar. Luego, se debe decidir sobre el número de capas, neuronas por capas y el tipo de función de transferencia. La entrada a la red responde al número de variables que representen al sistema y para ello se debe tener determinado conocimiento del comportamiento del mismo así como de las variables independientes que condicionan su salida. El número de neuronas a la salida es igual al número de variables de respuesta requeridas. El número de neuronas en las capas ocultas es más difícil de determinar y existen pocas reglas heurísticas para ello. Lo más común es recurrir a ensayos de prueba y error, si se seleccionan pocas, el modelo puede ser poco

flexible; y si se seleccionan muchas, tiende a conducir a problemas de memorización o sobre fijación de datos así como al incremento del tiempo de cómputo lo cual constituye una dificultad que debe conocerse. [4]

2.3.13 Implementación

Después que la red ha sido diseñada, se debe entrenar utilizando alguno de los métodos existentes. Cuando el entrenamiento es iterativo, se debe definir un criterio para detener este proceso si así se requiere. Este puede ser fijar un número máximo de iteraciones o cierta tolerancia alcanzada en la función de error utilizada.

2.3.14 Validación

Es la etapa final y debe prestarse atención al hecho de que la buena convergencia entre los datos reales y los predichos por la red durante el entrenamiento no es un indicador suficiente de que la red responda de forma adecuada a todas las entradas que puedan suceder. Una de las medidas más comunes para confirmar la calidad de la red es seccionar los datos disponibles en tres grupos, uno de ellos es el de entrenamiento, otro se dedica a la simulación de determinada zona donde algunos de sus valores pueden traslaparse con los de entrenamiento, y finalmente, uno dedicado a la validación. Los datos de validación no son presentados a la red durante su entrenamiento.

2.3.15 Perceptrón Simple

Un perceptrón simple o monocapa es una red neuronal formada por una sola capa de neuronas no conectadas entre sí, las cuales reciben las mismas entradas y cada una de ellas produce una salida individual. Fue presentado por Frank Rosenblatt en 1958, siendo objeto de gran interés dado que fue el primer modelo propuesto con aprendizaje supervisado [10]. En la Figura 2.8 puede observarse la estructura del perceptrón simple.

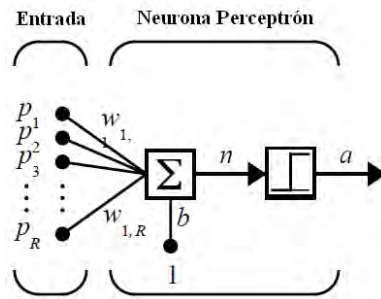


Figura 2.8 Perceptrón simple [17]

El perceptrón es la red neuronal más simple utilizada para la clasificación de patrones linealmente separables, es decir, patrones que se posicionan en lados opuestos de un hiperplano [10]. Rosenblatt demostró que bajo esta condición de separabilidad lineal, la regla de aprendizaje del perceptrón converge, es decir, se logran ajustar los parámetros de la red de manera que ésta clasifica de forma correcta todos los patrones o muestras de entrenamiento que le fueron presentadas.

El aprendizaje de tipo supervisado se refiere a un proceso iterativo donde se presenta al perceptrón un conjunto de entradas con sus correspondientes salidas que se desea que la red asocie o aprenda (meta). De esta forma, la red calcula la salida según los valores iniciales que poseen los pesos asociados a cada neurona. Luego comparando con las salidas deseadas, ajusta sus pesos proporcionalmente al error cometido. Este proceso se repetirá hasta que la diferencia entre la salida calculada por la red y la salida deseada sea nula, entonces los pesos ya no variarán sus valores entre una iteración y otra.

Las limitaciones del perceptrón simple se hallan en el hecho de que éste sólo puede clasificar conjuntos linealmente separables, lo cual implica que cualquier problema no lineal no puede resolto implementando esta estructura.

2.3.16 Perceptrón Multicapa

El perceptrón multicapa es una red neuronal alimentada hacia adelante que posee dos o más capas de neuronas. Las capas situadas entre la capa de salida y la entrada se llaman capas ocultas. Esta estructura nació con el propósito de dar solución a las limitaciones del perceptrón simple, ya

que permite aproximar con tanta precisión como se desee cualquier función continua (lineal o no lineal).

Esta tipo de red posee características propias que a continuación se enumeran [10]:

- 1) Cada neurona posee una función de transferencia no lineal que además es continua y diferenciable en cualquier punto.
- 2) La red contiene una o más capas ocultas de neuronas, las cuales hacen capaz a la red de aprender tareas complejas extrayendo progresivamente características propias de los patrones de entrada.
- 3) La red posee un alto grado de conectividad, establecido por las conexiones sinápticas entre neuronas. Un cambio en la conectividad significa cambiar la configuración o ajuste de los pesos sinápticos.

La combinación de estas características sumado a la capacidad de aprender mediante el entrenamiento, le proporcionan al perceptrón multicapa una alta capacidad computacional. Sin embargo, el hecho de que la estructura posea más capas de neuronas complica el proceso de aprendizaje y aumenta el tiempo requerido para llevar a cabo tal fin. Otra desventaja asociada a este tipo de red se encuentra a la hora de encontrar una arquitectura óptima, dado que no es posible determinar *a priori* el tamaño de la red que mejor resuelve un problema en concreto.

Existen varios algoritmos para llevar a cabo el entrenamiento del perceptrón multicapa, el más conocido y utilizado es el algoritmo de retropropagación o como comúnmente se le conoce, “*Backpropagation*” (retropropagación). Éste es un método de aprendizaje supervisado que se basa en la corrección de errores, y como su nombre lo indica consiste en ir propagando hacia atrás o hacia las capas anteriores los errores obtenidos en la capa de salida.

2.3.17 Descripción de una red perceptrón multicapas

Con el objeto de describir de forma gráfica un perceptrón multicapa, será considerada una estructura con una capa oculta y una capa de salida (Figura 2.9).

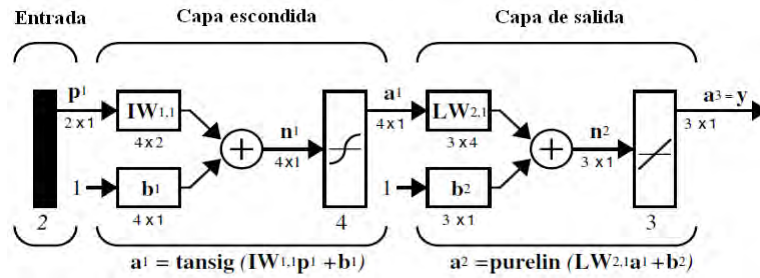


Figura 2.9 Perceptrón multicapa [11]

La entrada se representa mediante un vector $(x)(x_1, x_2, \dots, x_k, \dots, x_p)$ con p número de muestras. La capa oculta tiene q número de neuronas y la capa de salida posee r número de neuronas. Para denotar los pesos sinápticos de las neuronas de la capa oculta se utilizará el superíndice "1" y el superíndice "2" para los pesos de la capa de salida. También se asocia el subíndice "i" a las neuronas de la capa oculta y el subíndice "j" a las de la capa de salida. Entonces para cada neurona i de la capa oculta y cada neurona j de la capa de salida, la entrada está determinada por:

$$X_{net a_i} = \sum_{k=1}^p W^1_{ki} X'_k \quad i = 1, 2, \dots, q \quad (2.23)$$

$$X_{net a_j} = \sum_{i=1}^q W^2_{ij} X'_i \quad j = 1, 2, \dots, r \quad (2.24)$$

Donde:

i : Subíndice asociado a las neuronas de la capa oculta

j : Subíndice asociado a las neuronas de la capa de salida

k : Subíndice asociado a las muestras de entrada

p : Número de muestras de entrada

q : Número de neuronas de la capa oculta

r : Número de neuronas de la capa de salida

W^1_{ki} : Peso sináptico asociado a la neurona i de la capa oculta respecto a la muestra de entrada k

W^2_{ij} : Peso sináptico asociado a la neurona j de la capa de salida respecto a la salida i de la capa anterior (capa oculta).

La salida de las neuronas de cada capa se calcula tomando la función de transferencia y los biases:

$$X'_k = x_k \quad (2.25)$$

$$X'_i = f_1(Xneta_i + b_i) \quad (2.26)$$

$$X'_j = f_2(Xneta_j + b_j) \quad (2.27)$$

Donde:

X'_k : Muestras de entrada de la capa oculta

X'_i : Salidas de la capa oculta o entradas de la capa de salida

X'_j : Salidas de la red

f_1 : Función de transferencia de la capa oculta

f_2 : Función de transferencia de la capa de salida

2.3.18 Algoritmo de Retropropagación o “Backpropagation”

Es un algoritmo de aprendizaje supervisado que se fundamenta en minimizar un criterio de error que viene dado por la suma de los cuadrados de las diferencias entre los valores que proporciona la red y los valores deseados, todo ello evaluado sobre un conjunto de patrones de entrenamiento. En cada iteración los pesos sinápticos y biases son ajustados en proporción al error cometido de manera que la salida calculada por la red sea cada vez más parecida a la salida deseada.

La técnica de “*Backpropagation*” parte del cálculo de una función de error (diferencia entre la salida calculada por la red y la salida deseada). Luego que los valores de entrada se presentan, se calcula la respuesta de la RNA y se minimiza el valor del error, definido tal como se muestra a continuación [18]:

$$e = \frac{1}{2} \sum_t (y_t - x'_t)^2 \quad (2.28)$$

Donde:

y_t : Respuesta deseada

x'_t : Respuesta obtenida por la red

El concepto de la regla de gradiente descendiente consiste en hacer un cambio o ajuste en los pesos proporcional al opuesto de la derivada del error. De esta manera, se calcula la derivada de la función error con respecto a cualquier peso de la red y se actualiza el valor del peso de acuerdo a las siguientes expresiones [18]:

$$\Delta w_{ij} = -\gamma \frac{de}{dw_{ij}} \quad (2.29)$$

$$w_{ij}(n + 1) = w_{ij}(n) + \Delta w_{ij} \quad (2.30)$$

Donde:

γ : Constante de proporcionalidad

e : Función error

La constante de proporcionalidad, o también conocida como tasa de aprendizaje, es un factor importante para este método. En los procedimientos del gradiente descendente se usa una tasa de aprendizaje pequeña. Los valores grandes pueden proporcionar una rápida convergencia, pero cuando los valores son demasiado grandes podrían existir problemas de oscilaciones en la convergencia.

Escogiendo apropiadamente una función no lineal se puede realizar la diferenciación y obtener la siguiente expresión [18]:

$$\Delta w_{ij} = -\gamma \delta_j f_j \quad (2.31)$$

Así, los pesos deben ser cambiados por una cantidad proporcional al término " δ " (variación del error con respecto a la entrada de la neurona i), y el valor de la función de activación f .

Para determinar el valor de delta se usa un proceso iterativo que comienza con las neuronas de salida. Si una neurona forma parte de la capa de salida, su valor de δ está dado por la siguiente ecuación [18]:

$$\delta_j = (y_j - a_j)f'(Xneta_j) \quad (2.32)$$

$$Xneta = \sum(w_{ij}f_t + b_i) \quad (2.33)$$

Donde:

y_j : Valor deseado de la señal de activación

$Xneta$: Función de entrada a una neurona

$f'(Xneta_j)$: Derivada de la función de activación con respecto a un cambio en la entrada de la red.

La gran ventaja de la regla de BP viene de la asignación de las deltas a las unidades ocultas, que no reciben una retroalimentación directa de los patrones de entrenamiento. Estas deltas influyen en la modificación de los pesos que conectan las principales neuronas de la capa oculta. Para las neuronas de las capas ocultas donde no hay un objetivo o salida predeterminada, el valor de delta se obtiene de forma iterativa, en función de los términos deltas de las neuronas a las cuales están conectadas, y el peso de estas conexiones. Esto se representa de la siguiente manera [18]:

$$\delta_t = f'(Xneta_t) \sum_t \delta_j w_{tj} \quad (2.34)$$

La regla de aprendizaje de BP requiere que la derivada de la función de activación exista, que sea continua y no lineal. La función de transferencia que se suele usar para estos estudios de predicción y que se usará en este trabajo viene dada por la ecuación (2.21):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.21)$$

Para aplicar la regla de aprendizaje, se debe conocer la derivada de esta función con respecto a todas sus entradas, y en este caso está dada por:

$$\frac{df(x)}{dx} = f(x)(1 - f(x)) \quad (2.35)$$

Así, la señal de error δ_j para una neurona perteneciente a la capa de salida está dada por la ecuación (2.36) y para una neurona de una capa oculta está dada por (2.37)

$$\delta_j = (y_j - f_j)f_j(1 - f_j) \quad (2.36)$$

$$\delta_{jt} = f_t(1 - f_t) \sum_t \delta_j w_{tj} \quad (2.37)$$

Para mejorar la convergencia del proceso de aprendizaje, se puede agregar un término denominado momentum (α), a la regla de BP. Esto puede ser realizado de la forma como se muestra en la ecuación (2.38):

$$\Delta w_{ij}(n + 1) = \gamma(\delta_j f_j) + \alpha \Delta w_{ij}(n) \quad (2.38)$$

Donde α es una constante que determina el efecto de los cambios pasados de los pesos en la dirección correcta hacia la convergencia y el cual usualmente se toma como 0,9.

2.4 Descripción de las herramientas utilizadas

Para efectuar las estimaciones de demanda es necesario realizar una serie de simulaciones utilizando herramientas computacionales que faciliten el desarrollo de los procesos requeridos. En este estudio se utiliza la plataforma MATLAB® por ser una herramienta poderosa y sencilla de manipular. A continuación se describen las principales características de esta aplicación así como el paquete contenido en éste, especializado en redes neuronales, Neural Network Toolbox.

2.4.1 MATLAB®

MATLAB® es el nombre abreviado de “MATrix LABoratory”. MATLAB® es un programa para realizar cálculos numéricos con vectores y matrices. Como caso particular puede también trabajar con números escalares (tanto reales como complejos), con cadenas de caracteres y con otras estructuras de información más complejas. MATLAB® tiene también un lenguaje de programación propio.

Es un gran programa de cálculo técnico y científico. Para ciertas operaciones es muy rápido, cuando puede ejecutar sus funciones en código nativo con los tamaños más adecuados para aprovechar sus capacidades de vectorización. En otras aplicaciones resulta bastante más lento que el código equivalente desarrollado en C/C++ o Fortran. El lenguaje de programación de MATLAB® es una magnífica herramienta de alto nivel para desarrollar aplicaciones técnicas, fácil de utilizar y que, como ya se ha dicho, aumenta significativamente la productividad de los programadores respecto a otros entornos de desarrollo. MATLAB® dispone de un código básico y de varias librerías especializadas (toolboxes).

2.4.2 Neural Network Toolbox

Neural Network Toolbox™ es un paquete que extiende la programación en MATLAB® con herramientas para el diseño, implementación, visualización y simulación de las redes neuronales.

Las redes neuronales no son recomendables para aplicaciones en las cuales el análisis formal pueda ser difícil o imposible, tales como reconocimiento de patrones e identificación y control de sistemas no lineales. El software Neural Network Toolbox™ ofrece un soporte global para muchos paradigmas de redes probadas, así como interfaces gráficas de usuario que permiten diseñar y manejar dichas redes. El diseño modular, abierto y extensible de este paquete simplifica la creación de funciones y de redes personalizadas. [13]

2.4.3 Principales funciones y comandos

Para la creación de una red tipo “Backpropagation” se utiliza la función *newff* cuya sintaxis es la siguiente:

$$\text{newff}(P, T, [S1 S2 \dots S(N - l)], \{TF1 TF2 \dots TFNl\}, TF, BLF, PF, IPF, OPF, DDF)$$

Esta función requiere que le sean especificados los siguientes parámetros:

P Matriz de dimensiones $R \times Q1$, donde $Q1$ representa el número de muestras y R la cantidad de elementos de los vectores de entrada.

- T* Matriz de dimensiones $SN \times Q2$, donde $Q2$ representa el número de muestras y SN el número de elementos del vector meta.
- Si* Tamaño de la i -ésima capa, para $(N - 1)$ capas, por defecto = $[\]$. El tamaño de la capa de salida SN es determinado por T .
- TFi* Función de transferencia de la i -ésima capa. Por defecto utiliza '*tansig*' para las capas escondidas y '*purelin*' para las capas de salida.
- BTF* Función de entrenamiento de la red de retropropagación. Por defecto utiliza '*trainlm*'.
- BLF* Función de actualización de pesos/bias de retropropagación. Por defecto utiliza '*learnngdm*'.
- PF* Función para evaluar el desempeño de la red. Por defecto utiliza '*mse*'.place
- IPF* Arreglo de filas de celdas de funciones de procesamiento de entradas. Por defecto utiliza $\{\text{'fixunknowns'}, \text{'removeconstantrows'}, \text{'mapminmax'}\}$.
- OPF* Arreglo de filas de celdas de funciones de procesamiento de salidas. Por defecto usa $\{\text{'removeconstantrows'}, \text{'mapminmax'}\}$.
- DDF* Función de división de datos, por defecto utiliza '*dividerand*'.

Las salidas son las N capas suministradas, y la red de retropropagación.

La función de transferencia *TFi* puede ser alguna función de transferencia diferenciable tales como *tansig*, *logsig* o *purelin*.

La función de entrenamiento *BTF* puede ser alguna de las funciones de entrenamiento de retropropagación tales como *trainlm*, *trainbfg*, *trainrp*, *traingd*, etc.

Para la inicialización de una red tipo “Backpropagation” puede ser utilizada alguna de las siguientes funciones:

- *Init*

- *Initlay*
- *Initnw*
- *Initwb*

La función *Init* inicia la red neuronal con valores actualizados de pesos y bias de acuerdo con la función de inicialización de la red, indicado por el comando '*net.initFcn*', y los valores de los parámetros, indicados por el comando '*net.initParam*'. Su sintaxis es la siguiente:

$$net = init(net)$$

Para inicializar la red capa por capa se utiliza la función *Initlay*, la cual inicializa cada capa de acuerdo a su propia función de inicialización '*net.layers{i}.initFcn*'.

La función *Initnw* inicializa los pesos de una capa y bias de acuerdo al algoritmo de Nguyen-Widrow. Este algoritmo elige valores de los pesos y biases con el fin de distribuirlos, aproximadamente de manera uniforme en el espacio de entrada de cada capa.

Por último, la función *Initwb* también inicializa los pesos de una capa y bias de acuerdo con sus propias funciones de inicialización y da como argumento de salida la red con los pesos de la capa y los bias actualizados.

Entrenamiento – Validación

Para el entrenamiento de una red se utilizan las siguientes funciones:

- *Train*
- *Trainlm*
- *Traingd*
- *Traingdm*
- *Traingda*
- *Traindx*

El comando *train* entrena la red neuronal de acuerdo a las funciones '*net.trainFcn*' y '*net.trainParam*'. Su sintaxis es la siguiente:

$$[net, tr, Y, E, Pf, Af] = train(net, P, T, Pi, Ai)$$

A continuación se presentan los parámetros de entrada de la función *train*:

Net Red.

P Entradas de la red.

T Salida esperada (meta).

Pi Condiciones iniciales de retardo de entrada de la capa i. Por defecto se usan ceros.

Ai Condiciones iniciales de retardo de una capa i. Por defecto se usan ceros.

Y sus salidas son las siguientes:

Net Nueva red.

TR Información sobre el entrenamiento.

Y Salidas de la red.

E Errores de la red.

Pf Condiciones finales de retardo de entrada de la capa.

Af Condiciones finales de retardo de la capa

La función de la red de entrenamiento *trainlm* actualiza los pesos y las bias de acuerdo a la optimización de Levenberg-Marquardt. Es el algoritmo más rápido para redes "*Backpropagation*" pero tiene la desventaja de requerir de un set de entrenamiento lo más estándar posible, pues de otra forma solo aproximará correctamente valores que se encuentren dentro de los patrones de aprendizaje. Si el set de entrenamiento es muy extenso, se recomienda reducir el Jacobiano.

El comando *traingd* se basa en el algoritmo de pasos descendientes, el cual actualiza pesos y ganancias variándolos en la dirección negativa del gradiente de la función del error. Tiene la desventaja que es un algoritmo de aprendizaje muy lento.

Trainngdm: equivale al algoritmo tradicional, más un nuevo coeficiente de momentum, que interviene en el proceso de actualización de los pesos.

Trainngda: se basa en el algoritmo de gradiente descendiente, y emplea una tasa de aprendizaje adaptiva durante el proceso de entrenamiento.

Trainngdx: emplea el algoritmo de gradiente descendiente con momentum y una tasa de aprendizaje adaptativa durante el proceso de entrenamiento.

Para el aprendizaje de una red se utiliza la función de aprendizaje de gradiente descendiente con *momentum* de pesos y biases (*learnngdm*) y la función de aprendizaje de gradiente descendiente de pesos y biases (*learnngd*).

Entre las funciones de error que miden el rendimiento de la red, se pueden mencionar las siguientes:

- *Mse*
- *Msereg*
- *Sse*

Mse: es una función que mide el rendimiento de la red de acuerdo a la media de los errores al cuadrado.

Msereg: calcula la media del error al cuadrado con una función de regularización del rendimiento. Esta función mide el rendimiento de la red como la suma de los pesos de dos factores: la media al cuadrado del error y la media al cuadrado de los valores de los pesos y de los biases.

Sse: función que mide el rendimiento de acuerdo a la suma de los errores al cuadrado.

Las funciones de transferencia son aquellas mediante las cuales se calcula la salida de una capa a partir de su entrada neta, entre ellas se encuentran las siguientes:

Tansig: Función de transferencia sigmooidal tangencial hiperbólica.

Logsig: Función de transferencia sigmooidal logarítmica.

Purelin: Función de transferencia lineal.

Las siguientes funciones son utilizadas para separar la entrada y el vector meta en tres conjuntos: entrenamiento, validación y prueba, y se diferencian unas de otras en la forma en que agrupan dicho vector meta.

La función *divideind* agrupa el vector meta por índices especificados, *divideint* por índices intercalados y por último *dividerand* utiliza índices aleatorios [8].

Para la simulación de una red se puede utilizar la función *sim*, su sintaxis es la siguiente:

$$[Y, Pf, Af, E, perf] = sim(net, P, Pi, Ai, T)$$

A continuación se presentan los parámetros de entrada de la función *sim*:

Net Red

P Entradas de la red

Pi Condiciones iniciales de retardo de entrada de la capa i. Por defecto se usan ceros.

Ai Condiciones iniciales de retardo de entrada de la capa i. Por defecto se usan ceros.

T Salida esperada (meta). Por defecto se usan ceros.

Y sus salidas son las siguientes:

Y Salidas de la red

Pf Condiciones finales de retardo de entrada.

Af Condiciones finales de retardo.

E Errores de la red

perf Desempeño de la red (contiene información acerca del proceso de entrenamiento).

Los argumentos de retraso se utilizan con ciertos tipos de redes, que no aplican para la red perceptrón multicapas utilizada en este estudio.

CAPÍTULO 3

METODOLOGÍA

En este trabajo se van a implementar redes neuronales artificiales para determinar su capacidad de predicción de demanda eléctrica a mediano plazo, así como evaluar su desempeño, practicidad y precisión. El período de tiempo de la estimación corresponde a 5 años debido a que se desea comparar esta metodología con la utilizada actualmente por la EDC (Método de Holt-Winters).

Son muchos los artículos donde se han reportado buenos resultados obtenidos con redes neuronales. Sin embargo, hay quienes difieren de esta opinión y no están convencidos del desenvolvimiento de las redes en la estimación de demanda, ya que afirman que no han sido probadas sistemáticamente y los resultados de las pruebas efectuadas no siempre son presentados de una manera satisfactoria [9].

Es importante aclarar que la mayoría del material que se puede encontrar está relacionado con estimación de demanda a corto plazo, dado que las redes neuronales artificiales han sido ampliamente utilizadas para realizar pronósticos en períodos de tiempo cortos. Sin embargo, es posible considerar las bases y criterios generales de estos estudios y aplicarlos en el mediano plazo.

3.1 Etapas en el proceso de estimación de demanda

A continuación se presenta la metodología propuesta en este estudio para realizar estimaciones de demanda eléctrica a mediano plazo utilizando como herramienta las redes neuronales artificiales.

Tal metodología se constituye de varias partes que abarcan la selección del tipo de red, el algoritmo utilizado para el entrenamiento, la manera en que las redes ya entrenadas son validadas y además se describen diferentes tipos de pronósticos que servirán de guía para determinar cuál es la configuración más apropiada para llevar a cabo la estimación de demanda.

A continuación se muestra un esquema que resume las diferentes etapas del proceso de estimación de demanda:

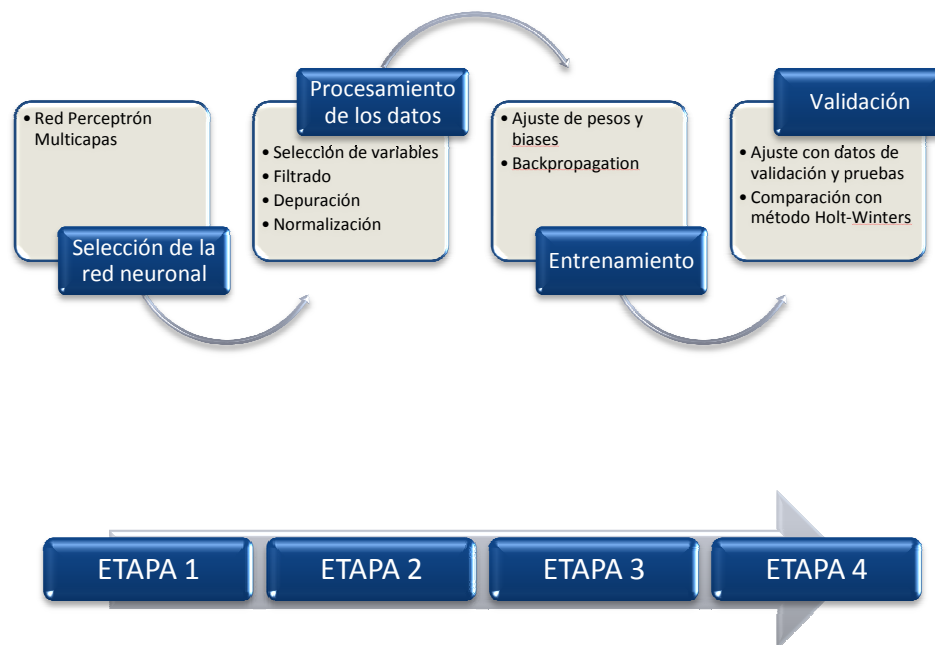


Figura 3.1 Esquema de las etapas del proceso de estimación de demanda con RNA

3.1.1 Selección de la red neuronal

Varios tipos de redes han sido utilizadas para realizar estimación de demanda, pero la que ha tenido mayor auge y ha sido implementada ampliamente es la red perceptrón multicapa [9, 14, 15, 16, 17]. Para diseñar una red multicapas alimentada hacia adelante, es necesario seleccionar el número de capas, tipo de conexión entre las capas, número de neuronas en cada capa y la función de transferencia de las neuronas de cada capa.

Una red con una capa oculta suele ser suficiente, ya que en la práctica una red completamente conectada con una capa oculta es una opción razonable [16, 19]. Todas las estimaciones que se realizarán en este trabajo estarán compuestas por una sola capa oculta.

El número de neuronas de la capa oculta determina las capacidades de aprendizaje de la red. La selección del número de neuronas de la capa oculta es la tarea fundamental en el diseño de la red. Una red con pocas neuronas ocultas no será capaz de modelar con precisión la demanda. Por otra parte, se tiene que un gran número de ellas provocarían una memorización de la red sobre los datos de entrada que se le proporcionan en el entrenamiento. Además el tamaño de la capa oculta afecta el tiempo que se requiere en el entrenamiento, por lo cual se recomienda seleccionar el mínimo número de neuronas en la capa oculta, es decir, aquellas necesarias para modelar la demanda.

Sin embargo, esto no es una tarea sencilla puesto que existen pocos métodos que ayuden a seleccionar el tamaño más apropiado, lo más común es recurrir a pruebas de ensayo y error hasta hallar una configuración satisfactoria. La capa de salida tiene sólo una neurona y esto se debe a que se tiene una sola salida (la demanda esperada en un momento dado).

En cuanto a las funciones de transferencia, se encuentran las funciones sigmoideas (en forma de “S”) como las más utilizadas. La forma exacta de la función sigmoidea tiene un ligero efecto sobre el comportamiento de la red, pero es posible que tengan una mayor influencia sobre la velocidad del entrenamiento [19]. Las dos funciones sigmoideas más comunes son la logarítmica y la tangencial hiperbólica.

3.1.2 Procesamiento de los datos de entrada

Frecuentemente los datos no son introducidos a la red neuronal directamente, sino que es necesario someterlos a una etapa de pre-procesamiento, la cual consiste en descartar, completar y “limpiar” algunas muestras de los datos de entrada. Posteriormente estos datos son sometidos a un procesamiento que implica la selección, filtrado, depuración y normalización de las variables de entrada de la red. Los datos de entrada que se utilizarán corresponden a un registro histórico

horario en KVA de cada uno de los circuitos pertenecientes a la subestación Lagunita de la Electricidad de Caracas, que se obtuvieron de la base de datos del sistema SCADA (*Supervisory Control And Data Acquisition*). Para mayor detalle ver el Capítulo 4 de este estudio.

En muchos casos existen muestras en los datos de entrada que carecen de significado físico, tales como valores negativos presentes en el registro histórico de la demanda de algunos circuitos, y es aquí donde entra en juego el proceso de descarte. También hay que considerar que esta ausencia de valores coherentes debe ser reemplazada por la existencia de nuevos valores, de manera que la información contenida en el registro histórico no sufra modificaciones que la afecten negativamente. Esto se puede realizar sustituyendo por el valor anterior o posterior al valor que se desea corregir, o bien reemplazando por un promedio de los valores vecinos que no tengan este mismo problema. Tomando todo esto en cuenta, se ha decidido reemplazar los valores negativos de la serie por el promedio de los valores vecinos, debido a que de esta forma se reúnen las características de varios valores a la vez. Los datos con los que se calcula dicho promedio también pueden ser anteriores o posteriores al que se desea corregir, por lo general con unos pocos es suficiente (cinco ó diez).

3.1.3 Selección de las variables de entrada

Las variables de entrada son seleccionadas a partir de los factores que afectan la demanda. Sólo aquellos que son realmente relevantes deben ser considerados en la estimación de demanda, ya que aunque una red neuronal puede trabajar con variables correlacionadas, un gran número de variables de entrada requiere de una red de gran tamaño y en consecuencia incrementa el tiempo de entrenamiento.

La relación entre el tamaño de una red y el tiempo de entrenamiento no es lineal. El tiempo de entrenamiento puede aumentar desproporcionadamente luego de un pequeño engrandecimiento de la red. Con muchas variables correlacionadas, la mayor parte del tiempo de entrenamiento es consumida en relaciones redundantes de aprendizaje [19].

Existen diversos tipos de variables que pueden ser consideradas para realizar la estimación de demanda, tales como: temporales, climáticas y/o económicas [20].

Se descarta el uso de las variables económicas puesto que el efecto que ejercen sobre el comportamiento de la demanda eléctrica se aprecia mejor en estudios de largo plazo y por lo general se implementan con métodos causales.

Es posible considerar variables climáticas para realizar los pronósticos, no obstante éstas son de mayor interés o brindan mayor información acerca de la demanda eléctrica en países con estaciones climáticas fuertemente diferenciadas. En un estudio de predicción de consumo de energía eléctrica realizado en Venezuela (Universidad Simón Bolívar) [21], se consideró la temperatura como variable de entrada de una red neuronal, reportando una poca influencia de los factores climáticos sobre el pronóstico. En las estimaciones que se realizarán en este trabajo no se tendrán en cuenta dichas variables, ganando de esta manera sencillez y practicidad puesto que se requerirá una red de menor tamaño y complejidad que pueda desarrollar satisfactoriamente esta labor.

Por lo tanto, sólo se van a considerar en este estudio variables temporales (aquellas que dependen del tiempo), tales como alguna que indique el mes, otra que indique el año, entre otras.

3.1.4 Filtrado

Debido a las características propias de los datos disponibles en este caso de estudio, se decide implementar un filtro que servirá para reducir el ruido y los efectos dinámicos u oscilaciones en los datos que no son de interés. Para ello se utiliza el filtro digital que ofrece el software MATLAB. El filtro consta básicamente de la combinación de dos tipos de filtros y se fundamenta en el siguiente algoritmo (ver Figura 3.2):

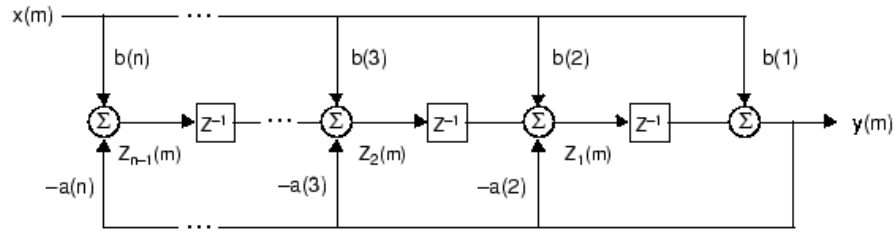


Figura 3.2 Filtro digital de MATLAB

Donde:

$n - 1$: Orden del filtro

na : Orden del filtro de retroalimentación

nb : Orden del filtro alimentado hacia adelante

La operación del filtro en una muestra m está dada por las siguientes ecuaciones en el dominio del tiempo:

$$y(m) = b(1)x(m) + z_1(m - 1) \quad (3.1)$$

$$z_1(m) = b(2)x(m) + z_2(m - 1) - a(2)y(m)$$

$$\vdots = \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$z_{n-2}(m) = b(n - 1)x(m) + z_{n-1}(m - 1) - a(n - 1)y(m)$$

$$z_{n-1}(m) = b(n)x(m) - a(n)y(m) \quad (3.2)$$

La relación entre la entrada y la salida del filtro en el nuevo dominio de la transformada z es una función de transferencia racional:

$$Y(z) = \frac{b(1) + b(2)z^{-1} + \dots + b(nb + 1)z^{-nb}}{1 + a(2)z^{-1} + \dots + a(na + 1)z^{-na}} X(z) \quad (3.3)$$

3.1.5 Depuración

Los datos de entrada representados por el registro histórico de la demanda, presentan ciertas distorsiones que podrían tener un impacto desfavorable en el pronóstico de demanda. Se observa

en la información recopilada de los circuitos que han ocurrido algunos eventos tales como traspasos de carga, fallas y otras irregularidades. Estos tipos de distorsiones aunque fueron atenuadas mediante la implementación del filtro no desaparecen.

Los traspasos de carga ocurren temporalmente e implican un crecimiento o decrecimiento repentino de la carga. En caso de que la carga tenga un decrecimiento hasta 0 KVA se considera que ha ocurrido una falla (o error de medición), o bien se ha ejecutado un “paro programado”. En muchos casos no existe información precisa acerca de los eventos que han ocurrido, o dicho de otro modo, no se conocen exactamente las causas que originaron tales eventos registrados. Por lo tanto, determinar que ha ocurrido algún evento u otro no es tarea sencilla, lo importante es que se identifiquen aquellos que realmente perjudiquen el pronóstico.

De los diferentes tipos de pronósticos que serán realizados, se encuentran algunos de ellos cuyos datos de entrada además de ser pre-procesados y filtrados, también son depurados manualmente con el objeto de determinar el efecto que tienen las distorsiones sobre éstos. Esta forma de depuración no es la más apropiada ya que debería ser un proceso automatizado que se encargue de tal labor apoyado en bases más sólidas que criterios personales o de sentido común. No obstante, resulta necesario extraer las distorsiones de los datos de entrada para poder comparar la calidad del pronóstico resultante respecto a aquellos sin depurar. Cabe destacar que existen varios métodos o procedimientos que permitirían hacer una depuración efectiva de los datos removiendo correctamente tales distorsiones, pero es un tema que va más allá de los objetivos de este trabajo y además, lo que se pretende es conocer si la red neuronal con la que se realizarán las estimaciones posee la capacidad de tolerar las distorsiones propias del registro histórico de demanda sin comprometer su precisión y acierto.

Por tal motivo se realizarán distintas pruebas que permitirán conocer y evaluar el verdadero efecto de incluir tales perturbaciones en las estimaciones.

El proceso de depuración manual sencillamente consiste en sustituir los valores de los datos de entrada correspondientes a eventos o distorsiones temporales por otros valores del mismo conjunto de entrada que no presenten este mismo inconveniente. El resultado es una data de

entrada coherente con el comportamiento propio de cada circuito, donde no aparecen cambios bruscos y aleatorios considerables.

3.1.6 Normalización

Se trata de una transformación útil en la escala de las entradas, de tal manera que estas estén dentro del mismo orden de magnitud. Todas las variables que se van a utilizar estarán escaladas entre 0 y 1. Este procesamiento hace que la elección inicial de los pesos sea más fácil, y en consecuencia el aprendizaje también será más sencillo. Si las entradas y las salidas están en un orden de magnitud unitario, los pesos también deberían estar en el mismo orden de magnitud. La normalización también hace que las variables ya dentro de un mismo orden de magnitud posean la misma importancia dentro del modelo, es decir, que no exista una mayor influencia de algunas variables respecto a otras.

3.1.7 Entrenamiento de la red

El proceso de entrenamiento en este trabajo quedará determinado por el algoritmo de retropropagación (*backpropagation*) por ser el más popular para una red de perceptrón multicapa [7, 15, 18]. A lo largo del tiempo se ha ido mejorando la eficiencia de este algoritmo, ejemplo de ello se encuentra en una de sus variantes, el método del gradiente conjugado que básicamente mejora la velocidad de convergencia.

El algoritmo de gradiente descendente generalmente es muy lento porque requiere una tasa de aprendizaje pequeño para lograr un aprendizaje estable. La incorporación de un parámetro conocido como *momentum* usualmente hace el algoritmo más rápido que el simple gradiente descendente, porque permite tasas de aprendizaje más altas manteniendo la estabilidad, pero es todavía demasiado lento para otras aplicaciones prácticas. Además se podría utilizar el entrenamiento de Levenberg-Marquardt para redes de pequeño y mediano tamaño, sólo que este método requiere de una mayor cantidad de memoria computacional.

Existen muchos algoritmos que pueden ser utilizados, pero dependiendo de los problemas de aproximación, para aquellas redes que contienen hasta unos pocos cientos de pesos, el algoritmo

de Levenberg-Marquardt tendrá la mayor convergencia. Esta ventaja es apreciable especialmente si se requiere de un entrenamiento muy preciso [13].

Debido a que el entrenamiento es iterativo, se debe definir un criterio para indicar que el proceso ha llegado a su fin, esto puede ocurrir cuando se haya alcanzado un número máximo de iteraciones, una tolerancia aceptable o bien, un límite de tiempo especificado.

No existe un rango riguroso en el que deba ser variado el número máximo de iteraciones, aún así pudieran tomarse generalmente de unas 100 a 400. En algunos casos, se obtienen errores dentro de los rangos aceptables utilizando tanto 100 iteraciones como también 400, es decir, es posible que aumentar el número máximo de iteraciones permitidas produzca una disminución del error. Sin embargo, debe tenerse en cuenta que aumentar en exceso el número de iteraciones puede producir una memorización de los datos de entrada en lugar de un aprendizaje satisfactorio del comportamiento del sistema modelado [4].

3.1.8 Validación de la red

La obtención de errores aceptables en el entrenamiento de la red no siempre asegura altos rendimientos, por lo que es necesario validar el desempeño de la red con un conjunto de datos diferentes a los que fueron utilizados en el entrenamiento [19].

Uno de los problemas que ocurren durante el entrenamiento de la red neuronal es llamado sobre fijación (memorización). El error asociado al conjunto de datos de entrenamiento tiende a ser cada vez menor, pero cuando nuevos datos son presentados a la red, el error es grande. La red ha memorizado los datos de entrenamiento, pero ésta no ha aprendido a generalizar ante nuevas situaciones.

El método para mejorar la generalización de la red es llamado *parada prematura*. En esta técnica el conjunto de datos disponibles son divididos en tres subconjuntos. El primer subconjunto es el de entrenamiento que es usado para calcular el gradiente y actualizar los pesos y biases de la red. El segundo subconjunto es el de validación. El error del conjunto de validación

normalmente disminuye durante la fase inicial del entrenamiento, tal y como lo hace el error del conjunto de entrenamiento. Sin embargo, cuando la red comienza a memorizar los datos (sobrefijación), el error del conjunto de validación típicamente comienza a crecer. Cuando esto ocurre para un número específico de iteraciones, el entrenamiento se detiene y los pesos y los biases son fijados en los valores que tenían cuando el error de validación alcanzó su punto mínimo.

El error del conjunto de prueba no se utiliza durante el entrenamiento, pero es usado para graficar el error de este conjunto durante el proceso de entrenamiento. Si el error de este conjunto alcanza un mínimo en un número de iteración significativamente diferente de donde el error del conjunto de validación lo alcanza, esto podría significar una división pobre del conjunto de datos originales [13].

En este trabajo se utilizará un 60% de los datos para entrenar, 20% para validación y 20% para pruebas, escogidos aleatoriamente [13].

3.2 Realización de pronósticos

A partir de la necesidad de conocer los efectos que puedan tener diferentes variables de entrada así como la presencia de eventos temporales o distorsiones en los mismos, se ha decidido realizar cuatro tipos de pronósticos diferentes de manera que los resultados puedan ser comparados entre ellos y también con los correspondientes a la metodología adoptada por la Electricidad de Caracas para realizar estimaciones de demanda eléctrica a mediano plazo (método de Holt-Winters).

3.2.1 Pronóstico tipo 1

- Se trabaja con valores mensuales de KVA del registro histórico.
- Se utilizan las variables de tiempo “Año” y “Mes” asociados a cada valor “KVA” para entrenar y simular la red neuronal.

3.2.2 Pronóstico tipo 2

- Se trabaja con valores diarios de KVA del registro histórico.
- Se utilizan las variables de tiempo “Año”, “Mes”, “DiaSem” y “Laborable” asociados a cada valor “KVA” para entrenar y simular la red neuronal.

3.2.3 Pronóstico tipo 3

- Similar al pronóstico 1 salvo que se implementa una depuración de los datos de entrada.
- Se trabaja con valores mensuales depurados de KVA del registro histórico.
- Se utilizan las variables de tiempo “Año” y “Mes” asociados a cada valor “KVA” para entrenar y simular la red neuronal.
- No todos los circuitos requieren que los datos del registro histórico mensual de la demanda sean depurados, por lo tanto este tipo de pronóstico sólo aplica a aquellos circuitos cuyas distorsiones en los datos de entrada serán corregidos.

3.2.4 Pronóstico tipo 4

- Similar al pronóstico 2 salvo que se implementa una depuración de los datos de entrada.
- Se trabaja con valores diarios de KVA del registro histórico.
- Se utilizan las variables de tiempo “Año”, “Mes”, “DiaSem” y “Laborable” asociados a cada valor “KVA” para entrenar y simular la red neuronal.
- No todos los circuitos requieren que los datos del registro histórico diario de la demanda sean depurados, por lo tanto este tipo de pronóstico sólo aplica a aquellos circuitos cuyas distorsiones en los datos de entrada serán corregidos.

3.3 Comparación de resultados con el método de Holt-Winters

Por medio de una función de error serán calculados los niveles de similitud entre los resultados obtenidos en cada uno de los cuatro tipos de pronósticos y aquellos obtenidos con el método de Holt-Winters. Es importante señalar que esta comparación no representa una validación de los resultados obtenidos con las redes neuronales artificiales, puesto que los resultados obtenidos con Holt-Winters o cualquier otro método también son estimaciones de la desconocida demanda eléctrica futura.

La función de error que será utilizada es la conocida como error cuadrático medio, la cual se expresa de la siguiente manera:

$$E = \sqrt{\frac{\sum_{k=1}^n (y_t - x'_t)^2}{n}} \quad (3.4)$$

Donde:

E	Error Cuadrático Medio
y_t	Respuesta deseada
x'_t	Respuesta obtenida por la red
n	Número de muestras

CAPÍTULO 4

DESCRIPCIÓN DEL CASO DE ESTUDIO: SUBESTACIÓN LAGUNITA

En este capítulo se describe el origen, características y el tratamiento de los datos utilizados para realizar el pronóstico de la demanda con redes neuronales. Se selecciona como caso de estudio la subestación Lagunita por presentar un registro completo de todos sus circuitos y reunir eventos que definen patrones especiales dentro del registro histórico de la demanda.

4.1 Subestación Lagunita

La subestación Lagunita se encuentra ubicada en el Estado Miranda y supe cargas de índole residencial y comercial. Posee tres unidades de transformación. Las características más relevantes se presentan en la tabla 4.1.

Tabla 4.1. Características de la subestación Lagunita.

Subestación	Voltaje de distribución (kV)	Capacidad Firme (MVA)	Capacidad Normal (MVA)	Número de Circuitos	Circuitos
Lagunita	8,3	7	5	9	LAG_A1 LAG_A2 LAG_A3 LAG_A4 LAG_A5 LAG_A6 LAG_B1 LAG_B2 LAG_B3

4.2 Registro histórico de la demanda de la Subestación Lagunita

Para probar el proceso de estimación de la demanda a mediano plazo con redes neuronales, se utiliza un caso de prueba con los datos del registro histórico en kVA proporcionados por el sistema SCADA de la subestación Lagunita de los nueve circuitos que la conforman. A continuación se explica brevemente el sistema SCADA.

4.3 Sistema SCADA (Supervisory Control And Data Acquisition)

La Electricidad de Caracas, como empresa de servicio eléctrico, emplea el sistema SCADA que proporciona supervisión y control sobre los dispositivos que se encuentran en la red.

El sistema SCADA se encarga de recopilar, procesar toda la información y permitir el control remoto sobre ciertos dispositivos supervisados, a través de la Unidades Terminales Remotas (*RTU- Remote Terminal Unit*), la cual es un equipo electrónico instalado en el terreno de medición, donde se encuentran las variables que se desean medir o supervisar que son enviadas como información (telemedidas). La Figura 4.2 muestra en detalle este proceso.

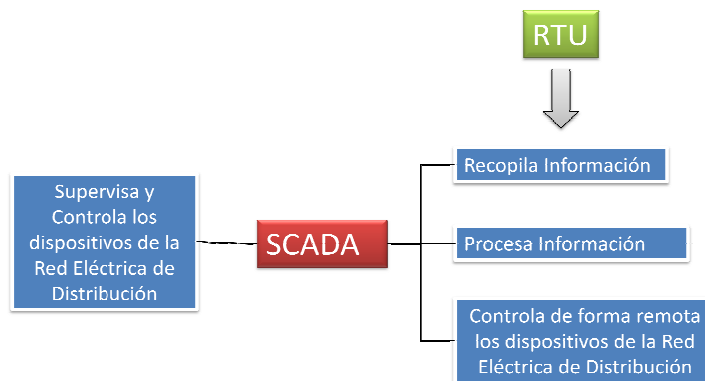


Figura 4.1. Esquema General del funcionamiento del SCADA

Este sistema tiene múltiples aplicaciones eléctricas, entre las cuales están:

- Realizar un análisis detallado de los circuitos que constituyen la red de distribución.
- Desarrollar reportes que aporten información sobre las condiciones de tensión, corriente y demanda del circuito estudiado.

Por su amplia gama de funciones es denominado DMS (*Distribution Management System* – Sistema de Gestión de Distribución).

El SCADA de Distribución, está compuesto por seis consolas de operación y una de coordinación, desde la cual los operadores de la red supervisan y realizan las maniobras necesarias todos los días del año.

El sistema operativo que lo soporta es el UNIX (Digital Unix 4.0E), en el cual todas las interfaces gráficas se basan en Xwindows (Motif), que es el manejador de ventanas de Unix.

Este sistema tiene la capacidad de alertar a aquellas personas que manejan la red, acerca de las condiciones en las que se encuentran operando los circuitos, bien sea con sobrecarga o bajo la existencia de alguna falla. Para tal fin, el sistema cuenta con varias alarmas que indican sobre los niveles de corriente y voltaje. [22]

Las telemidas históricas que proporciona el sistema SCADA se organizan en:

- Circuitos
- Principal
- Barras

- Subestación
- Reconectador

Las utilizadas en este caso de estudio es la de circuitos. Se recopiló de las bases de datos del SCADA las medidas de cada hora del día de la potencia aparente en kVA (el máximo horario) en un período de tiempo de 5 años a partir del año 2004.

4.4 Definición de patrones presentes en el registro

Los datos proporcionados por el sistema SCADA presentan en muchas ocasiones distorsiones en el histórico de la medida de la demanda, entre las más importantes están:

- Traspasos de Carga
- Fallas en el Sistema Eléctrico
- Errores de Medición de la remota (RTU)
- Problemas de Comunicación entre la remota y el SCADA
- Paros Programados

Por lo tanto es necesario encontrar mecanismos para corregir estas irregularidades para trabajar luego con una data más precisa y así poder evitar errores que afecten los resultados de la estimación de la demanda. Es importante mencionar que no existe información precisa respecto a la naturaleza de las irregularidades encontradas en el registro, por lo tanto la identificación de estas perturbaciones es meramente intuitiva.

Las distorsiones típicas en la medida del histórico de la demanda se ilustran en la Figura 4.3.

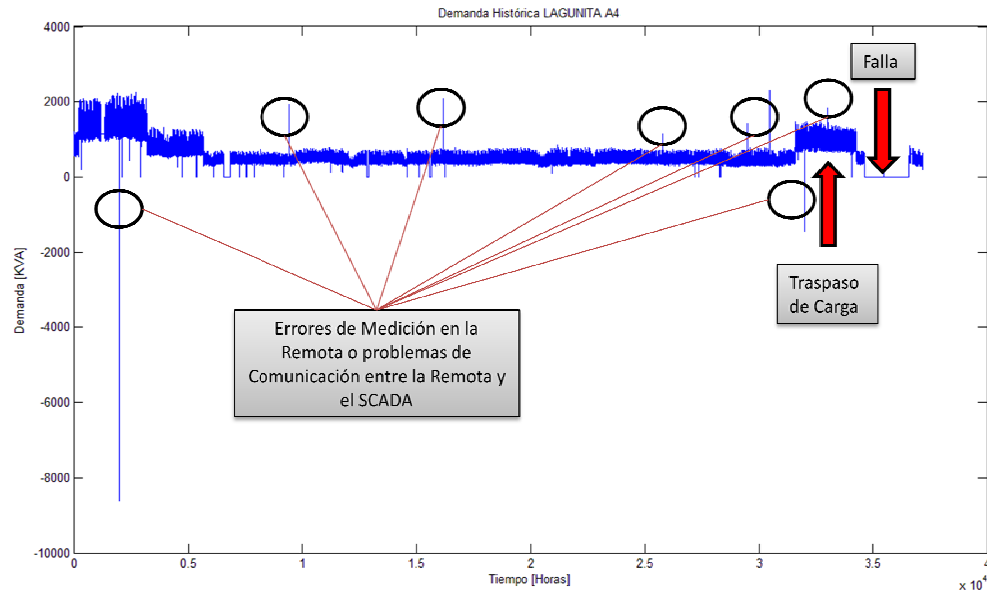


Figura 4.2. Distorsiones típicas en el historico de la demanda. Circuito Lagunita A4

4.5 Corrección de las distorsiones presentes en el registro

Mecanismos utilizados para corregir las distorsiones presentes en el historico de la demanda de los circuitos de la subestación Lagunita:

- Filtrado de la data
- Depuración de la data

En la Figura 4.4, 4.5 y 4.6 se muestra la data original, el proceso de filtrado y depuración.

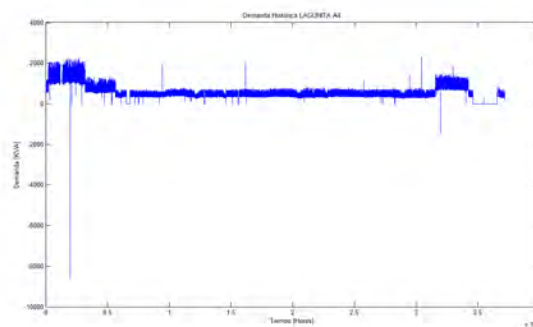


Figura 4.3 Demanda del circuito Lagunita A4

Las distorsiones en la data del circuito Lagunita A4 (Figura 4.3) se corrigen en dos pasos, primero se realiza el filtrado y luego la depuración de los datos.

Filtrado de la data del circuito Lagunita A4

El filtrado consiste en eliminar los valores negativos de la data y realizar la atenuación de las oscilaciones innecesarias. Este proceso se realiza utilizando el filtro digital contenido en MATLAB®.

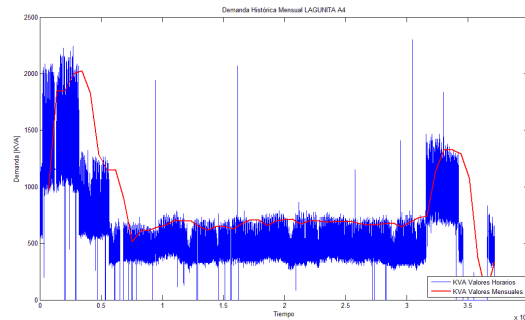


Figura 4.4. Filtrado de la Data del circuito Lagunita A4

Depuración de la data del circuito Lagunita A4

El proceso de depuración de los datos se realiza después del filtrado, debido a que aún se encuentran distorsiones que pueden afectar la estimación de la demanda. En la Figura 4.6, la gráfica azul representa la data filtrada y la roja la filtrada y depurada. Se observa en la data filtrada que existen dos perturbaciones importantes, las cuales son eliminadas en la depuración al sustituir estos valores por los datos del registro de un periodo de tiempo adyacente (ver Figura 4.6 de la derecha), con el fin de mantenerlos dentro del rango en el que se encuentran la mayoría de los valores del registro. Este proceso se explica en detalle en la sección 3.1.2 de este trabajo.

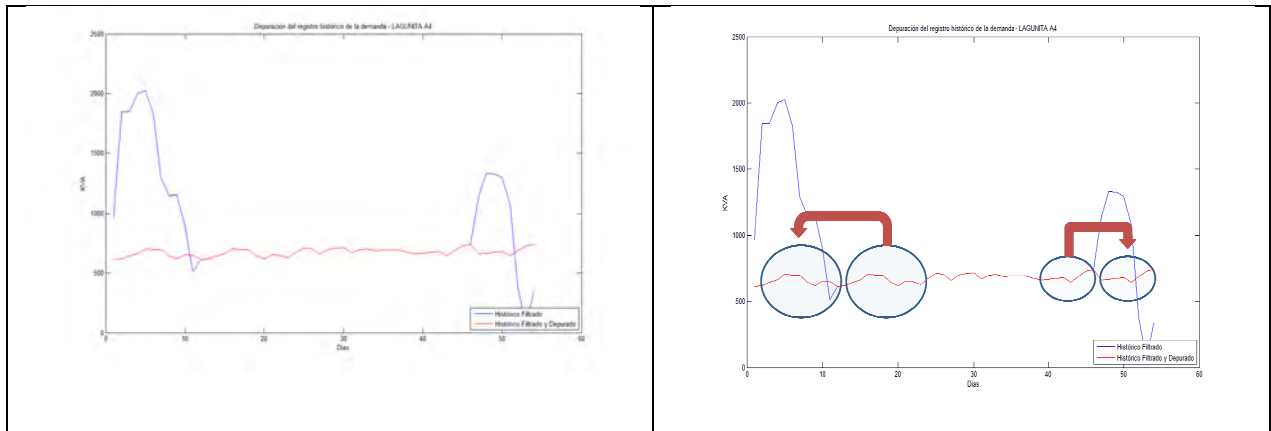


Figura 4.5 Depuración de la data del circuito Lagunita A4

Una vez filtrados y depurados los datos de los circuitos de la subestación Lagunita (ver Anexo A), entonces se puede proceder a realizar las estimaciones propuestas en este estudio.

CAPÍTULO 5

RESULTADOS

En este capítulo se presentan los resultados obtenidos por las simulaciones de la estimación de demanda con redes neuronales propuestas en la metodología. Se muestran una serie de gráficas correspondientes a cada uno de los cuatro tipos de pronósticos, donde se puede apreciar la manera en que la red neuronal reproduce las características de la serie histórica y cómo a partir de esto puede predecir los valores de la demanda futura. Finalmente, se comparan los resultados obtenidos con las estimaciones realizadas con el método de Holt-Winters. La serie histórica está representada por la línea azul y en rojo, se encuentra la respuesta de la red, que incluye la reproducción de la historia de la demanda seguido de la extrapolación de la serie en el futuro (período de cinco años). Cada RNA utilizada en cada uno de los pronósticos fue diseñada según las características propias de los datos destinados al entrenamiento de la red (como fue mencionado en la metodología). Los resultados mostrados se refieren a aquellos cuyo diseño de la red arrojó el menor error de aprendizaje final. En el Anexo B, se encuentra información acerca de la configuración de cada una de las redes, así como de su entrenamiento.

Solo serán ilustrados en este capítulo los resultados correspondientes a los circuitos Lagunita A4 y Lagunita B2 por ser aquellos cuyos datos históricos poseen distorsiones más pronunciadas. La información gráfica de las estimaciones del resto de los circuitos de la Subestación en estudio se encuentran en la sección de anexos (Ver Anexo C).

5.1 Simulación del pronóstico tipo 1

En las Figuras 5.1 y 5.2 se muestran las estimaciones de tipo 1 obtenidas con la RNA para los circuitos Lagunita A4 y B2 respectivamente.

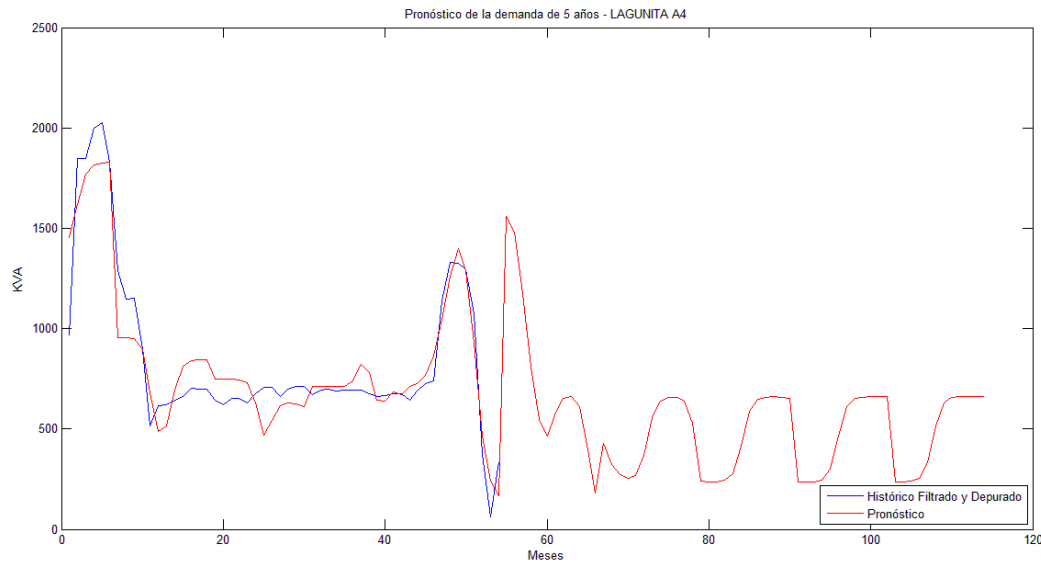


Figura 5.1 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A4

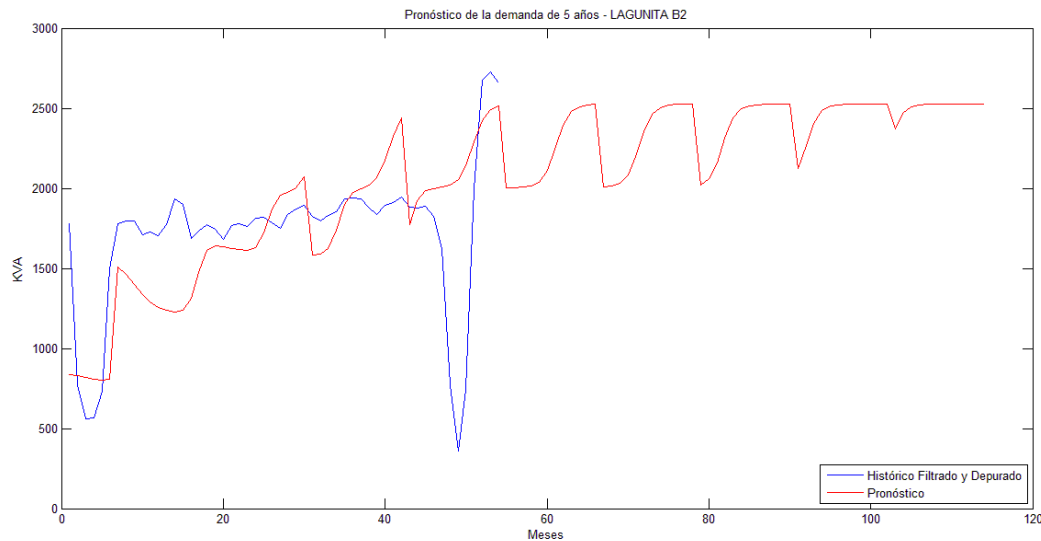


Figura 5.2 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita B2

Se puede observar que los eventos o distorsiones presentes en el registro histórico tienen una influencia directa sobre el proceso de aprendizaje, puesto que la salida proporcionada por la red no sigue fielmente la curva histórica (curva roja y azul no tienen exactamente el mismo comportamiento). Esto estaría indicando que la red no está aprendiendo apropiadamente de la demanda histórica del circuito y por lo tanto, la estimación que realiza es la correspondiente a otra serie de datos históricos (aquella que la que la red “creyó aprender”). Este caso se ve más acentuado en la figura 5.2.

5.2 Simulación del pronóstico tipo 2

En las Figuras 5.3 y 5.4 se muestran las estimaciones de tipo 2 realizadas por la RNA para los circuitos Lagunita A4 y B2 respectivamente.

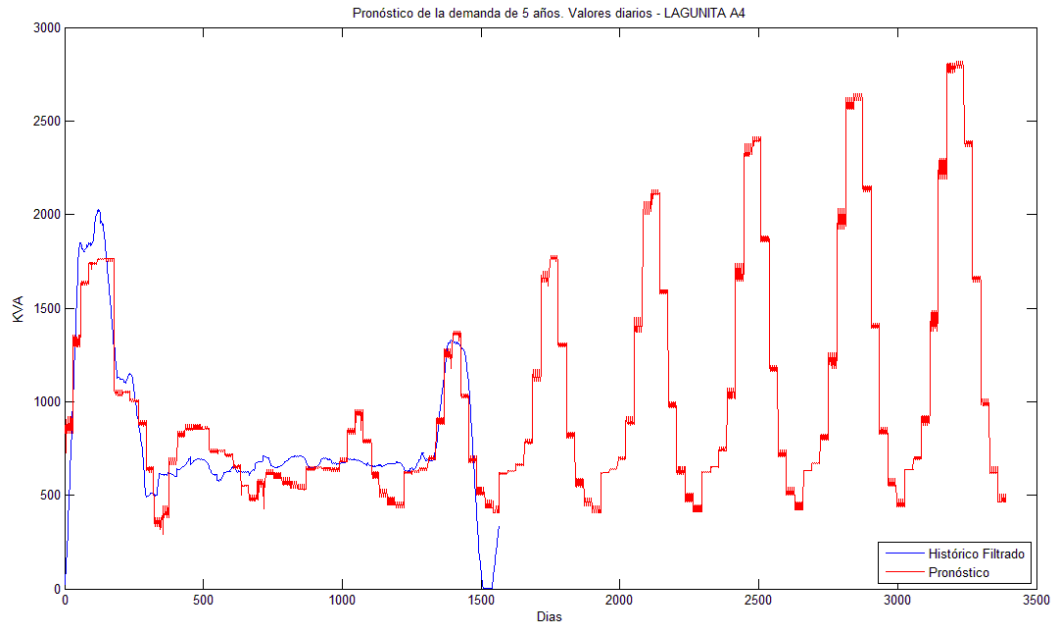


Figura 5.3 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A4

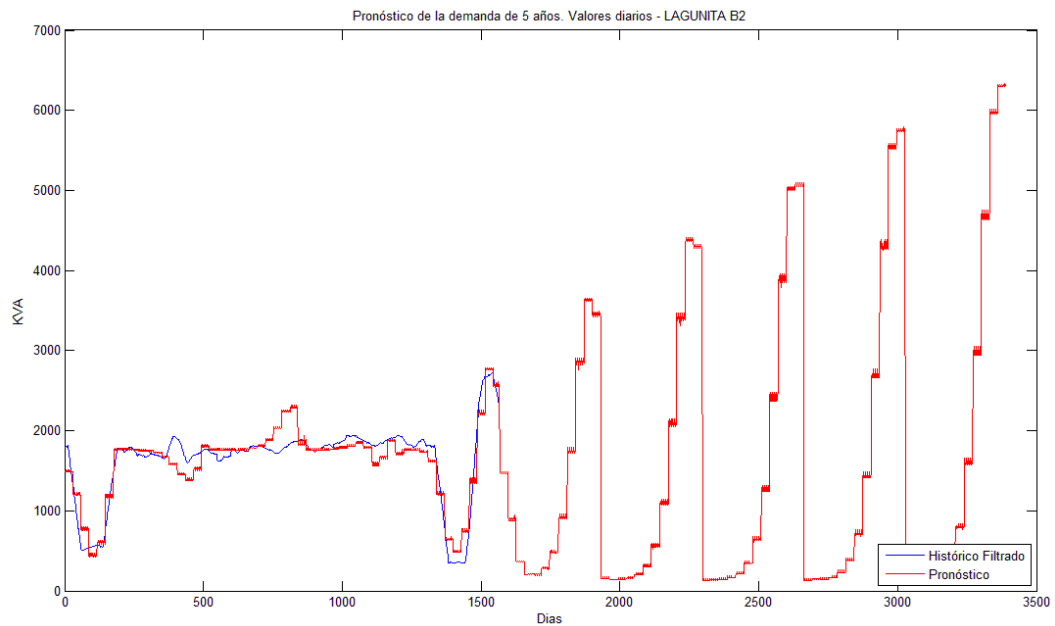


Figura 5.4 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita B2

Con las variables utilizadas en este pronóstico, la red logró realizar una mejor aproximación de los datos históricos. Sin embargo se sigue apreciando que el pronóstico está condicionado por los eventos o distorsiones presentes en los datos. Dicho de otro modo, el pronóstico está conformado por distorsiones, lo cual carece de sentido puesto que lo que se está prediciendo es la futura ocurrencia de estos eventos que son completamente aleatorios y no siguen patrones que puedan ser modelados por las variables utilizadas. Claramente puede apreciarse el impacto perjudicial que tienen estas distorsiones en la estimación de la demanda.

5.3 Simulación del pronóstico tipo 3

En las Figuras 5.5 y 5.6 se muestran las estimaciones de tipo 3 obtenidas con la RNA respectivamente para los circuitos Lagunita A4 y B2.

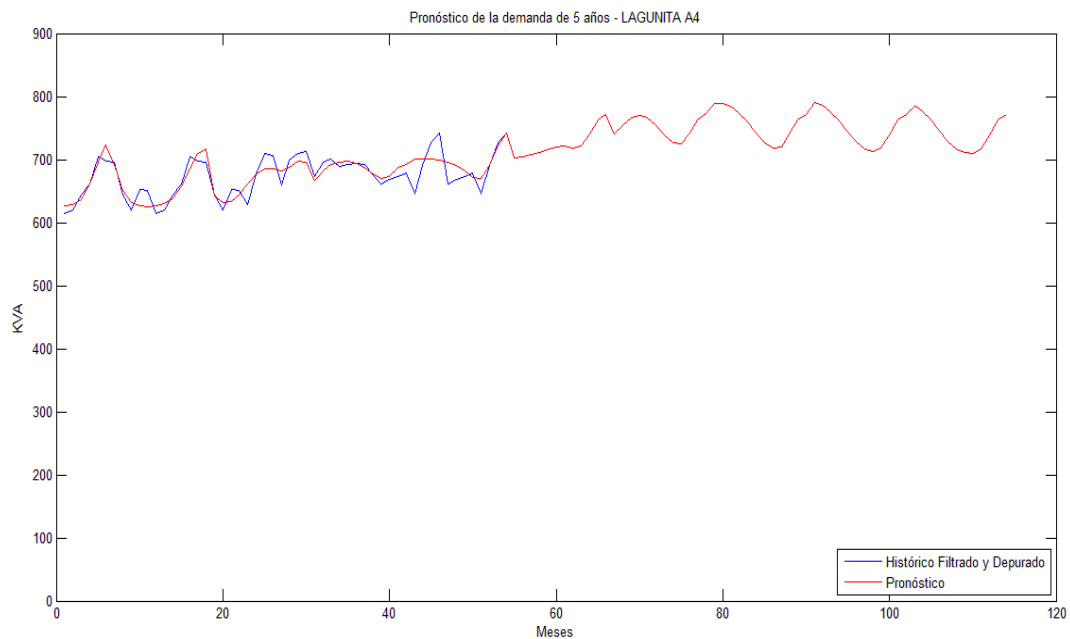


Figura 5.5 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A4

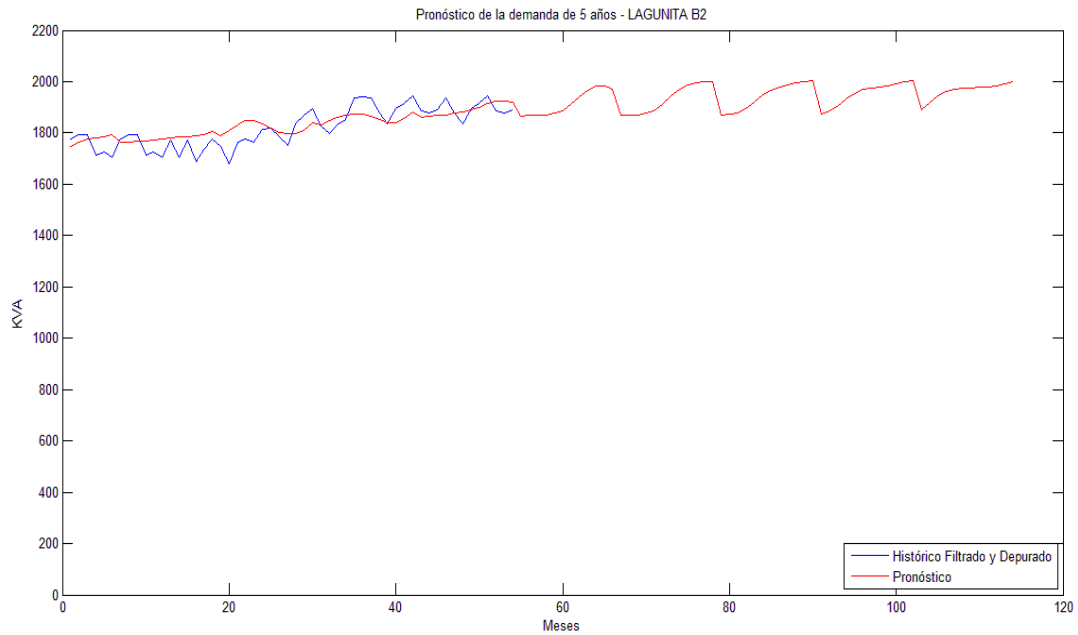


Figura 5.6 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita B2

En este tipo de pronóstico se observa claramente como en las estimaciones no aparecen distorsiones, lo que se puede atribuir al hecho de que las mismas han sido corregidas en el registro histórico. También es notable que se alcanzó un aprendizaje más apropiado (línea roja y azul muy similares en el período de tiempo correspondiente al pasado), lo que podría interpretarse como un pronóstico más acertado puesto que el aprendizaje siguió el comportamiento de la serie histórica. Esto se evidencia directamente en los errores calculados que miden la calidad del aprendizaje. (Ver sección 5.5).

5.4 Simulación del pronóstico tipo 4

En las Figuras 5.7 y 5.8 se muestran las estimaciones de tipo 4 obtenidas con la RNA para los circuitos Lagunita A4 y B2.

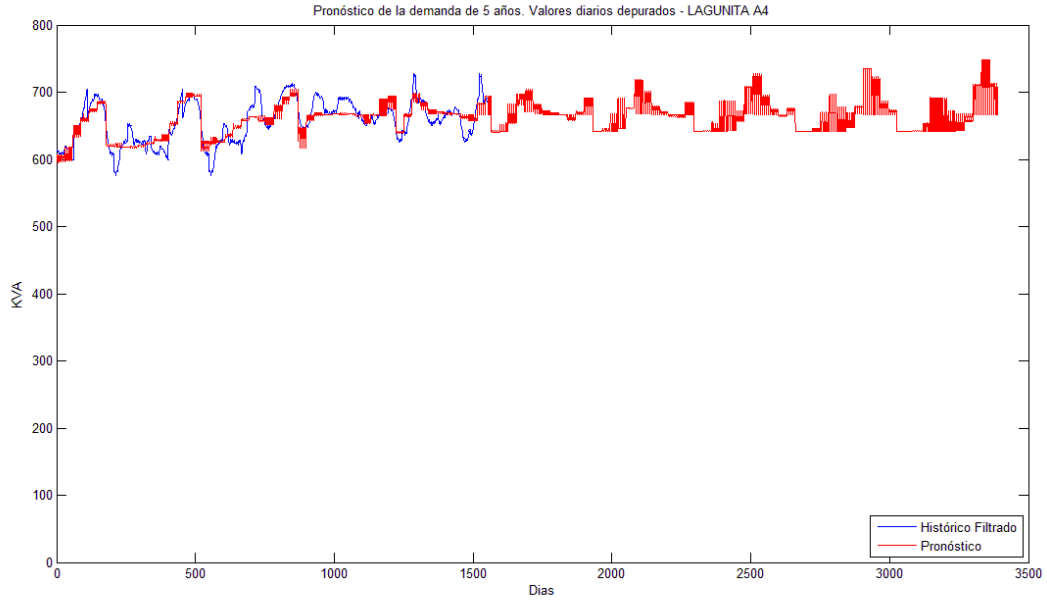


Figura 5.7 Pronóstico tipo 4 de la demanda de cinco años del circuito Lagunita A4

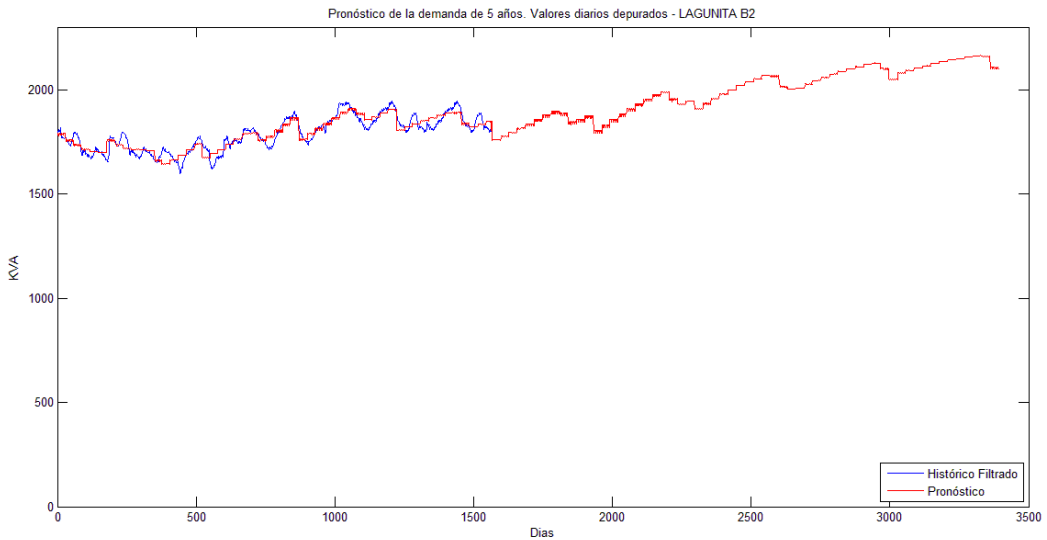


Figura 5.8 Pronóstico tipo 4 de la demanda de cinco años del circuito Lagunita B2

Similar al pronóstico tipo 3, la depuración efectuada a los datos históricos facilitan el proceso de aprendizaje de la red y se observa cómo la salida de la red y la serie histórica son bastante similares. Es importante señalar que los órdenes de magnitud del pronóstico tipo 3 y 4 realizados

por la RNA difieren considerablemente de los pronósticos cuyos datos de entrada no son depurados, observándose que los depurados constituyen estimaciones sin crecimientos, decrecimientos o variaciones bruscas, es decir, parecen reconocer la tendencia natural de la serie histórica. De nuevo se está en evidencia de las ventajas que ofrece el proceso de depuración, que se reflejan en el entrenamiento y pronóstico.

5.5 Resultados obtenidos con el método de Holt-Winters

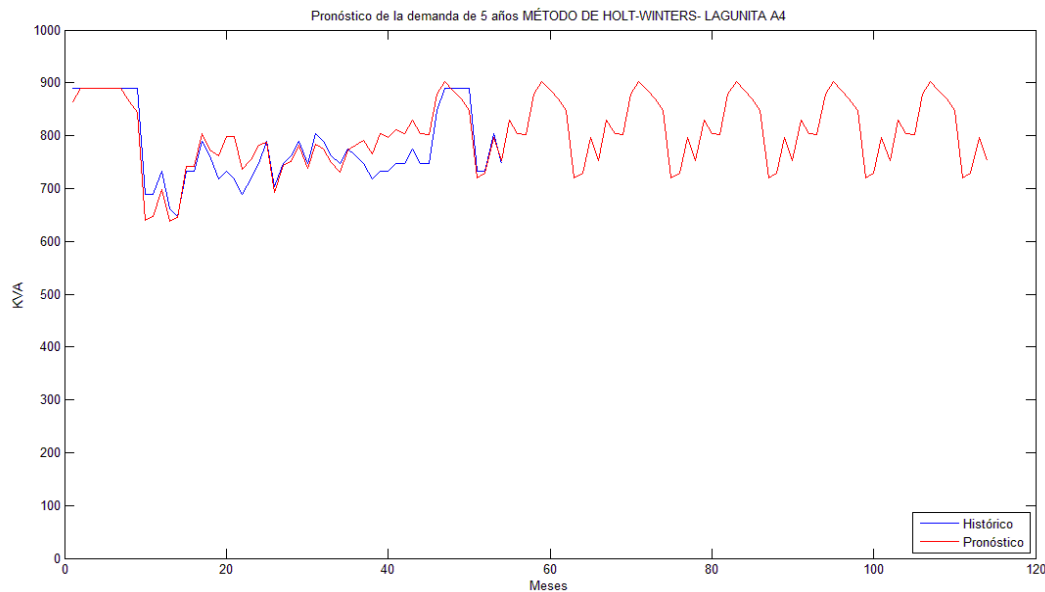


Figura 5.9 Pronóstico de la demanda de cinco años del circuito Lagunita A4 según método de Holt-Winters

Este método de estimación de la demanda utilizado por La Electricidad de Caracas incluye un proceso de depuración que no corresponde al mismo procedimiento utilizado en este trabajo. Como puede observarse, presenta una buena capacidad para imitar los datos históricos mediante el ajuste de sus parámetros. No obstante, es importante señalar que este método se ve fuertemente influenciado por el registro histórico de la demanda del último año, debido a que prácticamente se basa en éste para realizar el pronóstico. Se aprecia que los resultados obtenidos para el primer año de estimación, son los mismos valores para el resto de los años, es decir, que se repite exactamente cinco veces (ver Anexo D). Esta característica no permite visualizar crecimientos, decrecimientos o variaciones de la demanda en el futuro, sino pronósticos de aspecto llano.

5.6 Comparación y análisis final

A continuación se presenta la Tabla 5.1 que contiene información asociada a la precisión en el entrenamiento o capacidad para reproducir satisfactoriamente la serie histórica de las RNA utilizadas para cada circuito en cada tipo de pronóstico y el método de Holt-Winters. En ella se aprecian los errores porcentuales de las salidas calculadas por ambos métodos respecto al registro histórico.

Tabla 5.1. Errores porcentuales de la salida de las RNA y método de Holt-Winters respecto a la serie histórica.

Datos	No Depurados		Depurados		
	Tipo 1 (%)	Tipo2 (%)	Tipo 3 (%)	Tipo 4 (%)	Holt-Winters (%)
A1	7,39	3,49	4,09	1,94	7,06
A2	1,40	1,53	NA	NA	20,28
A3	5,21	7,04	5,17	3,20	33,04
A4	18,67	225,01	2,19	2,14	3,42
A5	6,23	2,80	1,78	1,50	2,42
A6	1,19	1,47	NA	NA	23,69
B1	1,33	1,82	NA	NA	2,41
B2	29,73	11,86	2,39	1,42	NC
B3	3,40	3,32	NA	NA	8,15

NA: No aplica

NC: No convergió método de Holt-Winters

Si se fija la atención en los pronósticos tipo 1 y 2 se puede observar que para algunos circuitos el error cometido en el pronóstico tipo 1 es mayor que el correspondiente al pronóstico tipo 2, pero para otros circuitos ocurre lo contrario, lo cual significa que utilizar un mayor número de

variables es conveniente o no, dependiendo de las características propias de los datos de entrada de cada circuito y no se puede generalizar.

Ahora bien, si se observan los errores obtenidos en los pronósticos tipo 1 y tipo 3, se pone de manifiesto que el error del pronóstico tipo 3 en todos los casos es menor a los de tipo 1. Esto se traduce en que en todos los casos donde se requiera, el proceso de depuración mejora la calidad del aprendizaje de las series históricas de los circuitos utilizando las variables con las que se realizaron estos dos tipos de pronóstico. Lo mismo ocurre cuando se utiliza un mayor número de variables como es el caso de los pronósticos tipo 2 y tipo 4, donde se puede observar que los errores del pronóstico tipo 4 (datos del registro histórico depurados) son menores que los de tipo 2 en todos los casos. Entonces se puede decir que independientemente de las variables que se utilicen para realizar las estimaciones de demanda, la depuración de las series históricas de los circuitos que lo requieran, es un proceso conveniente que mejora considerablemente la capacidad de aprendizaje de las RNA.

Por otra parte, observando los errores porcentuales de los pronósticos tipo 3 y tipo 4, ambos correspondientes a entrenamiento con series históricas depuradas pero con variables temporales diferentes para realizar las estimaciones, se aprecia que se obtienen mejores resultados con el pronóstico tipo 4. Con esto se puede decir que en el caso de que los datos de entrenamiento se encuentren depurados, resulta conveniente realizar las estimaciones de demanda modelando con un mayor número de variables puesto que se gana precisión.

En cuanto a los errores porcentuales calculados para el método de Holt-Winters, cuyos datos son depurados, se observa que éstos son mayores en todos los casos a los obtenidos en los pronósticos tipo 3 y 4 correspondientes a datos históricos depurados. Incluso los errores cometidos con este método resultan ser mayores que los obtenidos en la mayoría de los casos de los pronósticos 1 y 2 correspondientes a datos históricos sin depurar. Particularmente, en el caso del circuito Lagunita B2, el método de Holt-Winters no convergió, esto probablemente ocurre porque el procedimiento de depuración efectuado no fue suficiente para realizar la estimación. Esto demuestra la superioridad de las RNA en cuanto a la aproximación de funciones no lineales (mejor aproximación y reproducción de los datos históricos).

Resulta bastante útil calcular un error global para cada uno de los pronósticos que permita conocer el desempeño de las RNA en cada uno de ellos. Se calcula entonces un error ponderado que considera la capacidad instalada de cada circuito de la Subestación en estudio, utilizando la siguiente expresión:

$$EPG = \frac{\sum_k^n E_{p_k} * C_{inst_k}}{\sum_k^n C_{inst_k}} \quad (5.1)$$

Donde:

EPG : Error ponderado global

n : número de circuitos

E_{p_k} : Error porcentual del circuito k

C_{inst_k} : Capacidad instalada del circuito k

Luego de calcular el error ponderado global para cada tipo de pronóstico, se obtuvieron los resultados mostrados en la Tabla 5.2:

Tabla 5.2. Error ponderado global de los 4 tipos de pronósticos realizados con las RNA

Pronóstico	Tipo 1 (%)	Tipo 2 (%)	Tipo 3 (%)	Tipo 4 (%)
EPG	9,52	17,81	1,58	1,26

Puede verse en la Tabla 5.2 que se obtiene un menor error ponderado global cuando se utilizan datos depurados para realizar las estimaciones. El menor error se obtiene para el pronóstico tipo 4, por lo que puede decirse que en general es el más conveniente de los cuatro.

Se muestra a continuación una tabla que ofrece información relacionada al grado de similitud entre las estimaciones realizadas utilizando RNA y el método de Holt-Winters.

Tabla 5.3. Errores porcentuales entre las estimaciones utilizando RNA y método de Holt-Winters.

Datos	No Depurados		Depurados	
	Tipo 1 (%)	Tipo2 (%)	Tipo 3 (%)	Tipo 4 (%)
A1	52,76	29,47	12,04	6,91
A2	5,41	13,76	NA	NA
A3	39,21	15,56	11,66	19,73
A4	41,97	65,29	19,12	16,81
A5	5,85	33,46	10,60	10,31
A6	6,79	6,77	NA	NA
B1	6,65	6,26	NA	NA
B2	NC	NC	NC	NC
B3	9,05	13,96	NA	NA

NA: No aplica

NC: No convergió el método de Holt-Winters

Como puede verse, los tipos de pronósticos que utilizan series históricas sin depurar difieren considerablemente con el método de Holt-Winters, llegando hasta a 52,76% de error (Pronóstico tipo 1, caso Lagunita A1). Los errores porcentuales para los pronósticos que utilizan datos depurados o que no requieren ser depurados van desde 5,41% (Pronóstico tipo 1, caso Lagunita A2) hasta 19,73% (Pronóstico tipo 4, caso Lagunita A3).

Es importante destacar que estos errores no representan una validación de los resultados obtenidos con las RNA, simplemente establecen un modelo de comparación entre ambos métodos.

CONCLUSIONES Y RECOMENDACIONES

Las RNA representan una buena herramienta para realizar estimaciones de demanda eléctrica a mediano plazo. Es realmente necesario depurar los datos utilizados para el entrenamiento de la RNA, puesto que los pronósticos con datos distorsionados se ven seriamente afectados, es decir las RNA no son capaces de tolerar distorsiones muy acentuadas.

Los pronósticos que incluyen mayor número de variables temporales tales como los de tipo 2 y tipo 4, son realmente convenientes si los datos son depurados antes de hacer las estimaciones, según los correspondientes errores porcentuales calculados. Trabajar con valores de kVA diarios del registro histórico así como con variables temporales que indiquen el día de la semana y si el día es laborable o no, ayudan a las RNA a “aprender” mejor el comportamiento de los datos históricos y por lo tanto realizar pronósticos o estimaciones más apegados a las características propias de éstos.

La aparición de distorsiones en las estimaciones depende poco o nada de los tipos de variables de entrada de la RNA.

Aumentar el número de neuronas de la capa oculta de la red, puede hacer que el error de entrenamiento final disminuya marcadamente. Sin embargo, con esto se sacrifica calidad y precisión de los pronósticos realizados puesto que la red es incapaz de generalizar porque ha sufrido una memorización de los datos incluyendo las distorsiones presentes en el histórico de la demanda.

El tipo de pronóstico con menor error ponderado global es el de tipo 4, por lo cual se convierte en el más favorable de todos.

El método de Holt-Winters presenta limitaciones para manejar series históricas de la demanda aunque éstas sean depuradas debido a que se observó que sólo es capaz de pronosticar el primer año futuro y que repite ese mismo pronóstico para los cuatro años siguientes.

Recomendaciones

Integrar al sistema SCADA algún mecanismo que brinde información precisa acerca de los eventos que quedaron registrados en sus bases de datos. De manera que los datos que se ofrecen sean más confiables.

Se propone desarrollar una herramienta computacional basada en RNA que pueda ser capaz de realizar estimaciones de demanda a mediano plazo para cualquier circuito de la electricidad de Caracas.

Realizar una revisión exhaustiva de la aplicación de Excel basada en el método de Holt-Winters utilizado con el fin de corregir y mejorar la calidad y precisión de sus pronósticos.

Es posible utilizar la herramienta de estimación de estados contenida en el sistema EMS (SCADA de transmisión) para crear bases de datos redundantes de las variables de interés, de manera que existan patrones de comparación para determinar e identificar las distorsiones presentes en el registro del sistema DMS (SCADA de distribución).

Abarcar con mayor profundidad la aplicación de métodos de agrupamiento de datos como estrategia para categorizar los eventos o distorsiones presentes en los datos históricos, con el objeto de depurarlos. Una herramienta que podría implementarse son las redes neuronales de auto-organización (Ver Apéndice A).

REFERENCIAS BIBLIOGRÁFICAS

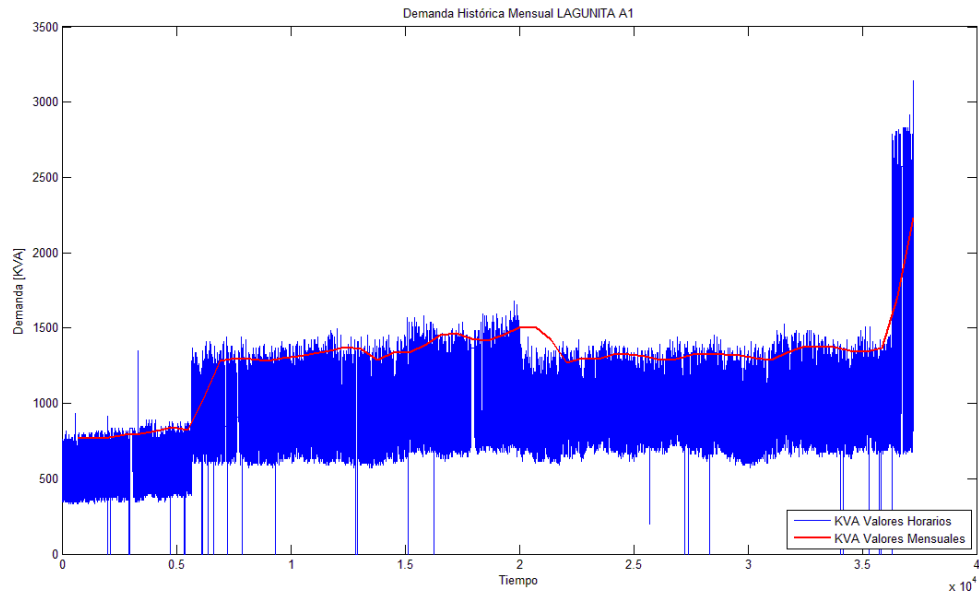
- [1] La Electricidad de Caracas. 2009. Así es la EDC. Disponible en Internet: <http://www.laedc.com.ve/ArticlesDetail.asp?CategoryId=10874&modulo=1>, consultado el 28 de Noviembre de 2009
- [2] RODRÍGUEZ, C. 2008. Estudio de planificación a mediano plazo de la S/E Esmeralda de 12.47KV de la Electricidad de Caracas. Informe de pasantía, Universidad Simón Bolívar.
- [3] GONZÁLES, R. 2004. Estudio de planificación a corto y mediano plazo: Análisis técnico-económico para el cambio de tensión de las subestaciones Caucaquita y Miranda. Informe de pasantía, Universidad Simón Bolívar.
- [4] ARMAS, M., GÓMEZ, M y PÉREZ, C. s/f. Inteligencia artificial aplicadas al análisis de sistemas energéticos con MATLAB. 1era Edición. INDUSTRIAS GRÁFICAS KORAL.
- [5] MARTÍNEZ, P. y FIGUEROA, R. 1994. Estudio de Saturación del sistema de distribución de la zona de Guarenas – Guatire – Araira. Informe de Pasantía, Universidad Simón Bolívar.
- [6] WILLIS, H. 2004. Power distribution planning reference book. 2da Edición. MARCEL DEKKER Inc, New York
- [7] WILSON, H. y BARRY, K. 1996. Previsiones en los negocios. 1era Edición. IRWIN, España.
- [8] CARRATÚ, M. 2001. Estimación de la demanda eléctrica de mediano y largo plazo para C.A. Electricidad de Valencia. Tesis de Pregrado, Universidad Simón Bolívar.
- [9] STEINHERZ, H y PEDREIRA, C. 2001. Neural Networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems 16: 44 - 55
- [10] HAYKIN, S. 2005. Neural Networks: a comprehensive foundation. 2da Edición. PEARSON, Dehli
- [11] SAMARASINGHE, S. 2007. Neural Networks for Applied Sciences and Engineering. 1era Edición. Averbach Publications, New York.
- [12] HAGAN, M., DEMUTH, H. y BEALE, M. 1996. Neural Network Design. 1era Edición. PWS Publishing Company.

- [13] DEMUTH, H., BEALE, M y HAGAN, M. 2009. Neural Network Toolbox™ 6. User's Guide. THE MATHWORKS INC., Estados Unidos.
- [14] ORTIZ, D. y SKOV, M. 1997. Accurate electricity load forecasting with artificial neural networks. IEEE Transactions on Power Systems 12:1233-1239.
- [15] KIARTZIS, S., ZOUMAS, C. 1997. Short term load forecasting in an autonomous power system using artificial neural networks. IEEE Transactions on Power Systems 12:1591-1596.
- [16] HISHAM, M. y MOUNT-CAMPBELL, C. 1997. Building a quasi optimal neural networks to solve the short term load forecasting problem. IEEE Transactions on Power Systems 12:1432-1435.
- [17] DOVEH, E., FEIGIN, P., GREIG, D. y HYAMS, L. 1999. Experience with FNN Models for medium term power demand predictions. IEEE Transactions on Power Systems 14: 538-546.
- [18] MARTÍNEZ, R. y MARTIN, E. 2001. Estimación de la demanda eléctrica asociada a la extracción de crudo mediante el uso de redes neurales. Pasantía Larga, Universidad Simón Bolívar.
- [19] CHARYTONIUK, W., CHEN, M. 2000. Neural network based demand forecasting in a deregulated environment. IEEE Transactions On Industry Applications 36: 893-898.
- [20] BAO, J. Short-term Load Forecasting Based on Neural Network and Moving Average.
- [21] GONZÁLEZ, D. y LOZA, D. 2000. Predicción del consumo de energía eléctrica a corto plazo mediante las redes neurales. Tesis de Pregrado, Universidad Simón Bolívar.
- [22] RODRÍGUEZ, G. 2008. Establecimiento de los límites de operación de los circuitos de distribución de la CA. La Electricidad de Caracas. Pasantía Larga, Universidad Simón Bolívar.
- [23] VESANTO, J., HIMBERG, J., ALHONIEMI, E. y PARHANKANGAS, J. 2000. SOM Toolbox 2.0 for Matlab 5. EXPOO, Helsinki.

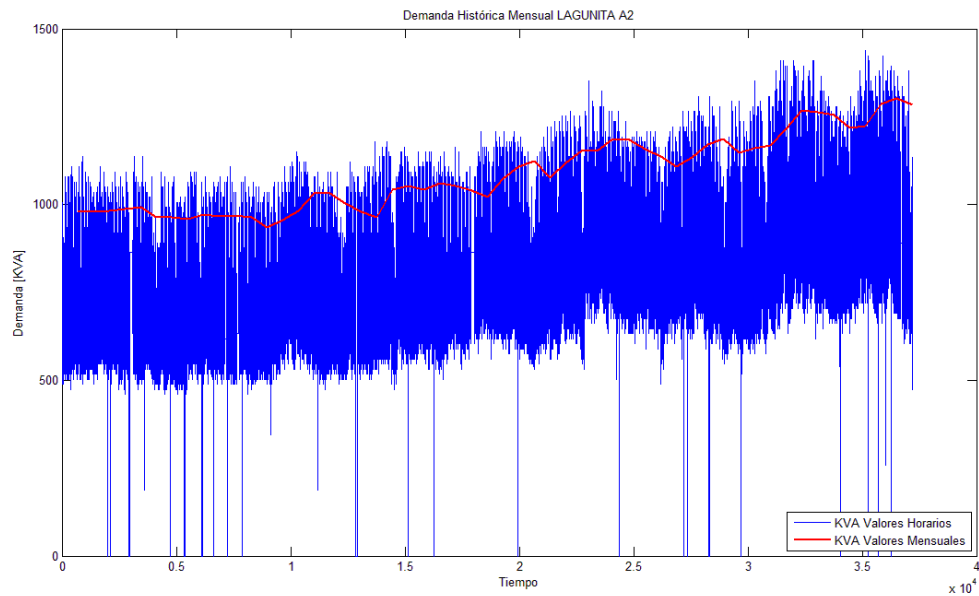
ANEXO A

DATOS FILTRADOS Y DEPURADOS DEL REGISTRO HISTÓRICO DE LA DEMANDA

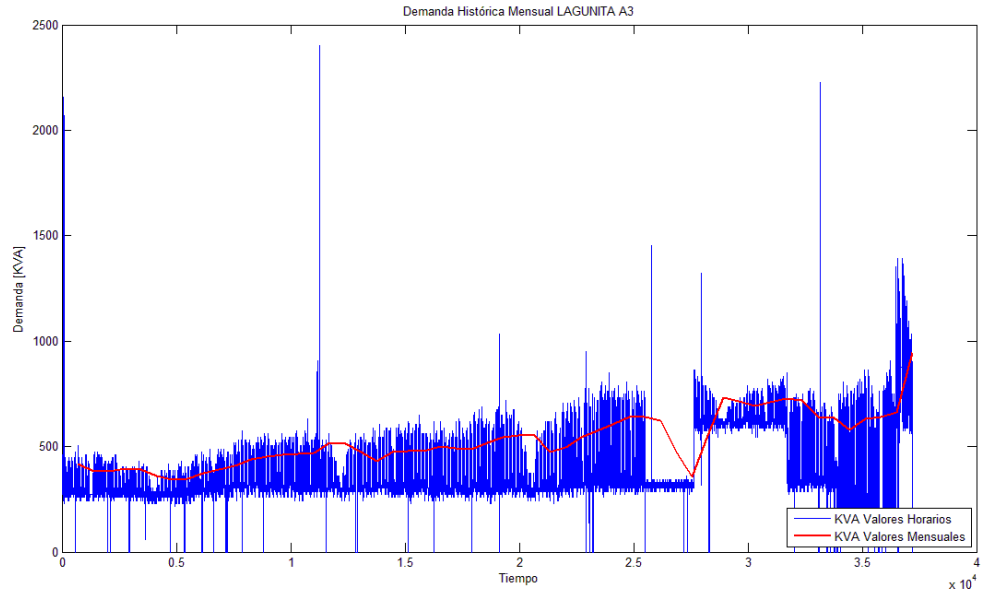
Datos en kVA utilizados para realizar los pronósticos tipo 1



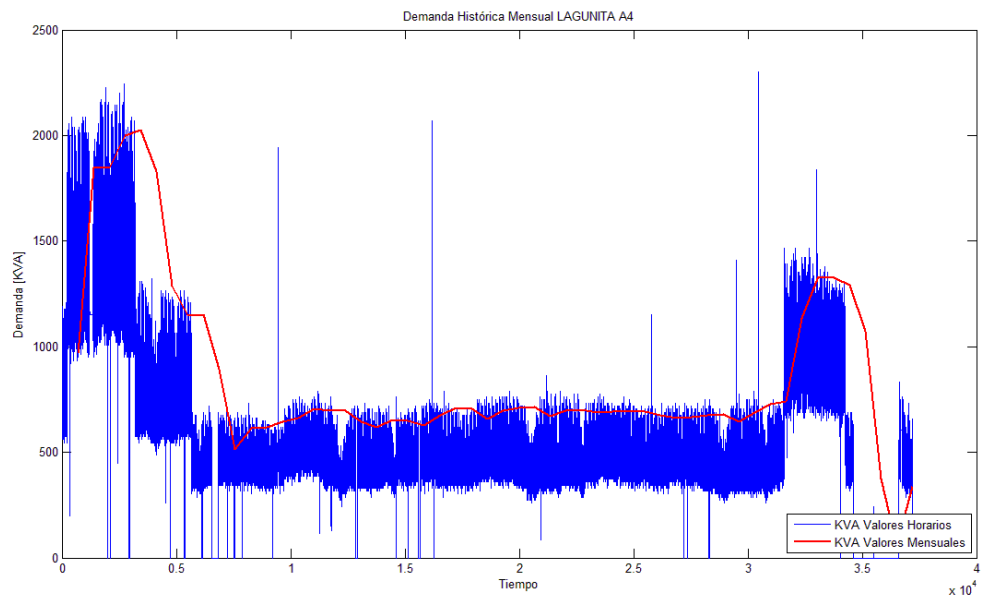
Anexo A.1. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita A1



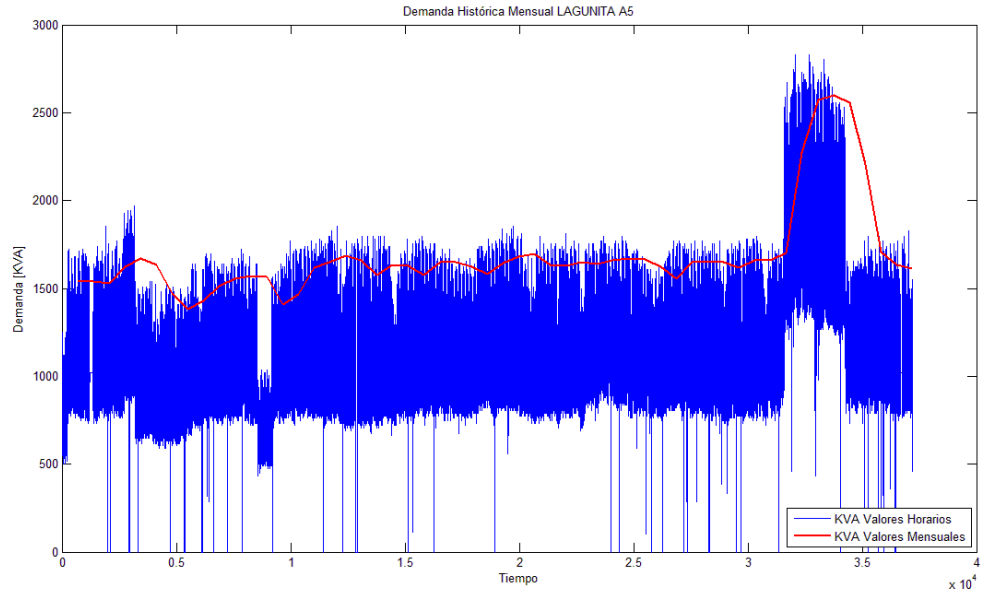
Anexo A.2. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita A2



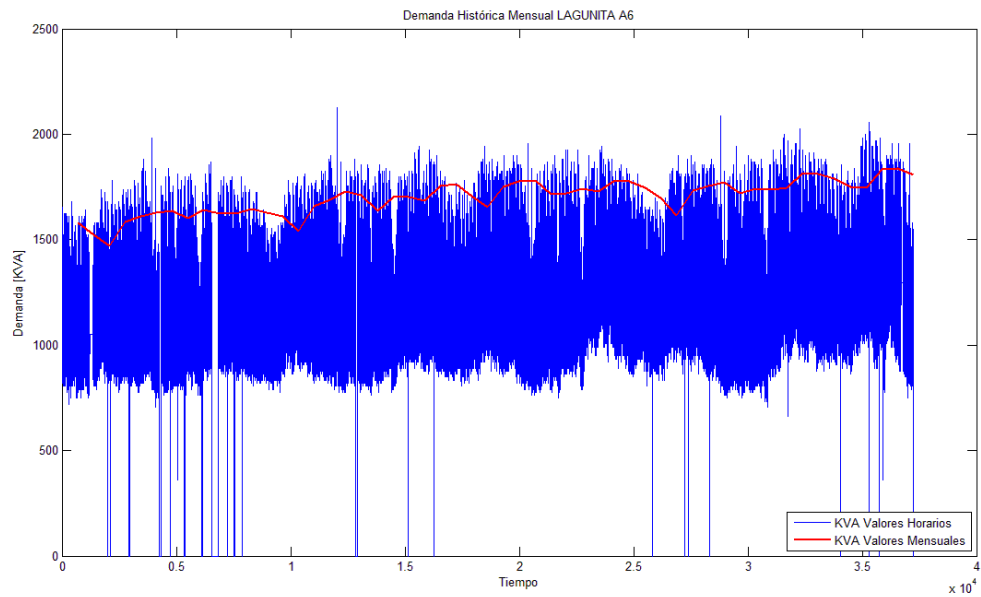
Anexo A.3. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita A3



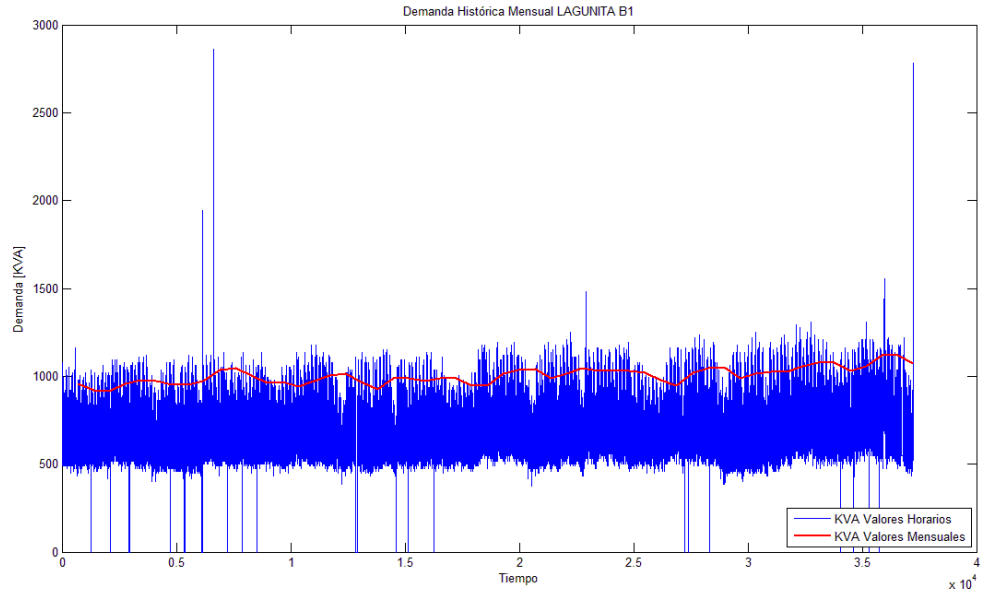
Anexo A.4. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita A4



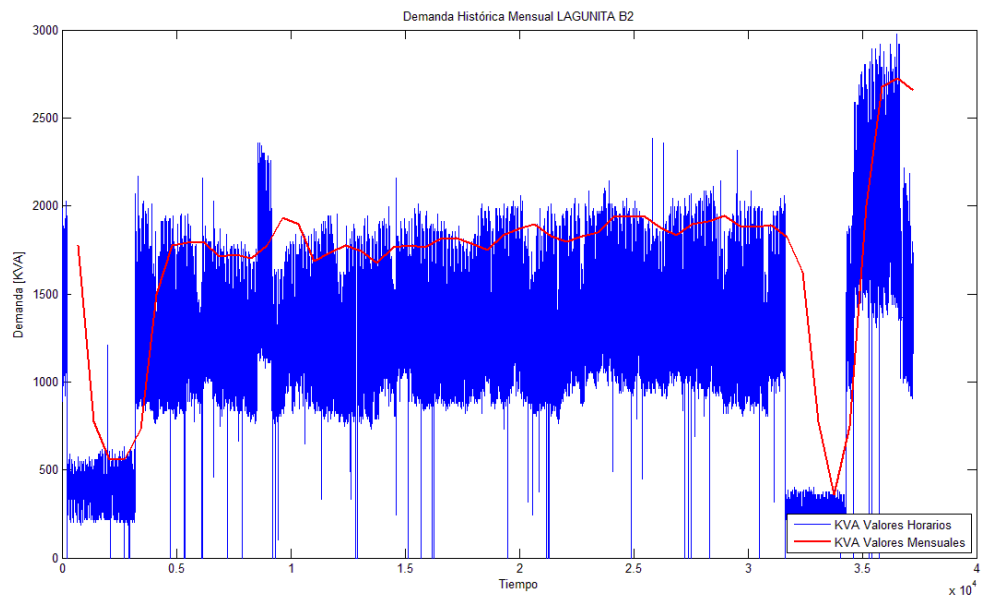
Anexo A.5. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita A5



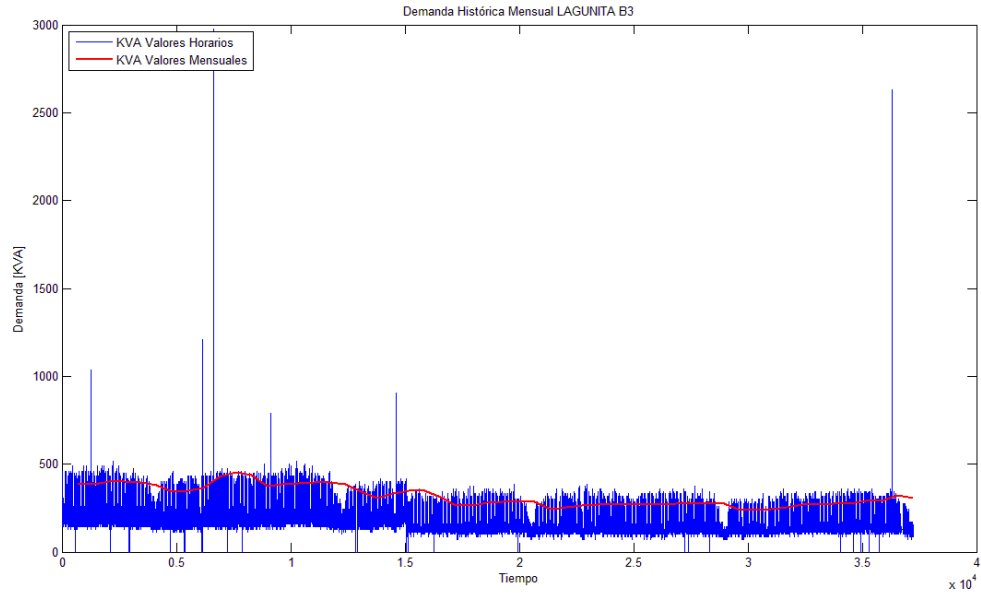
Anexo A.6. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita A6



Anexo A.7. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita B1

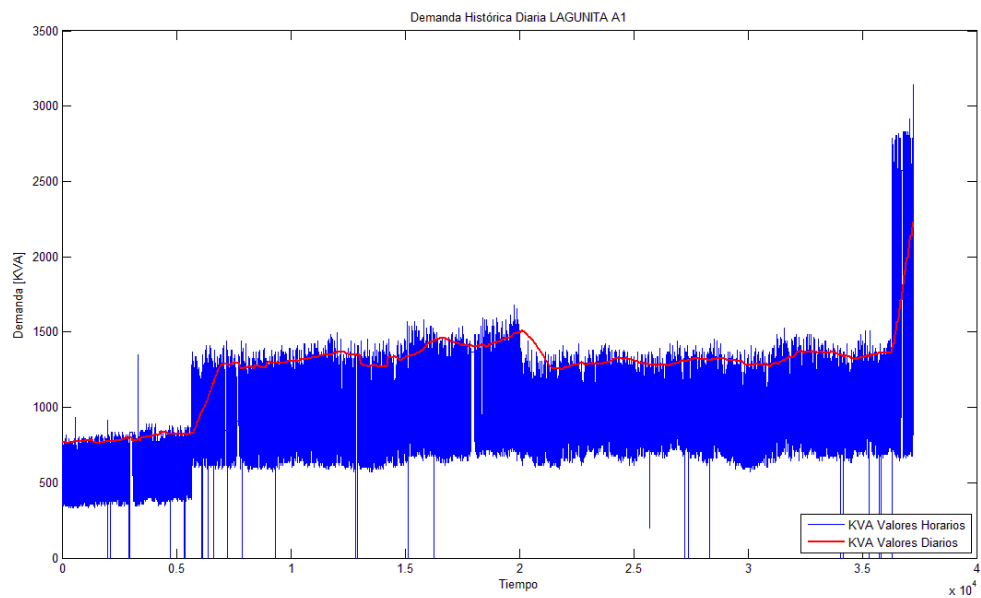


Anexo A.8. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita B2

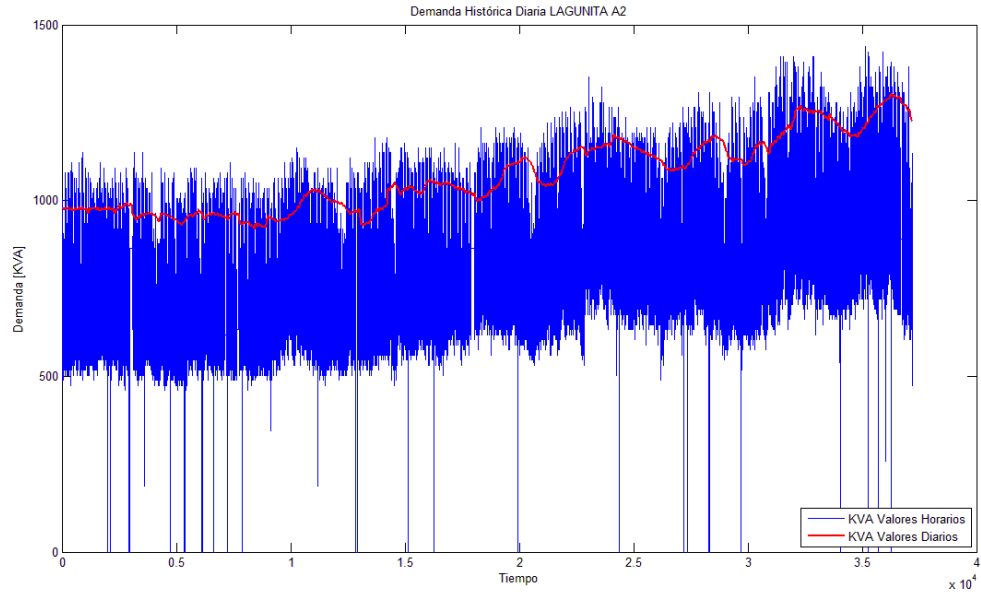


Anexo A.9. Gráfica de la demanda histórica mensual filtrada. Circuito Lagunita B3

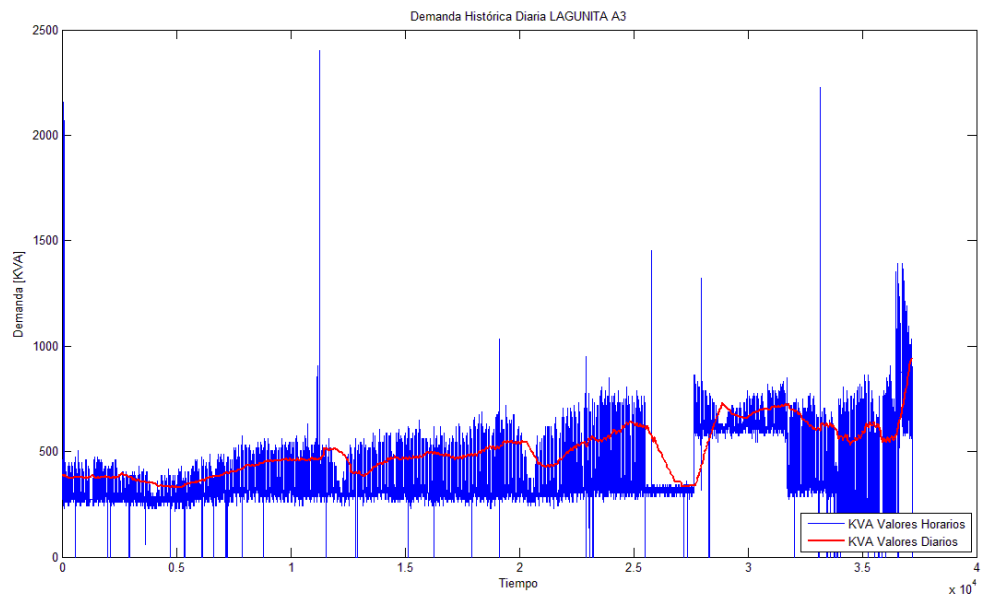
Datos en kVA utilizados para realizar los pronósticos tipo 2



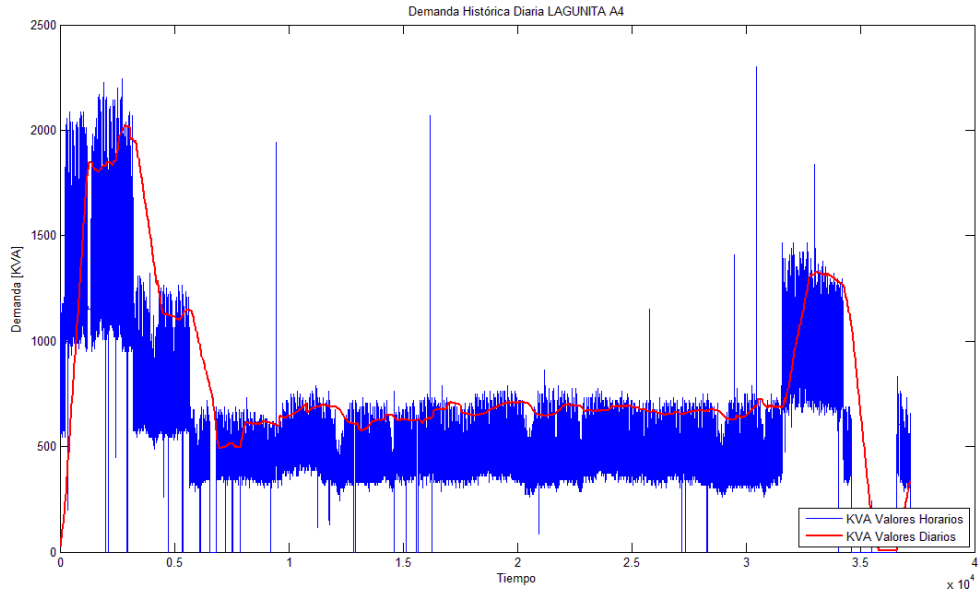
Anexo A.10. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita A1



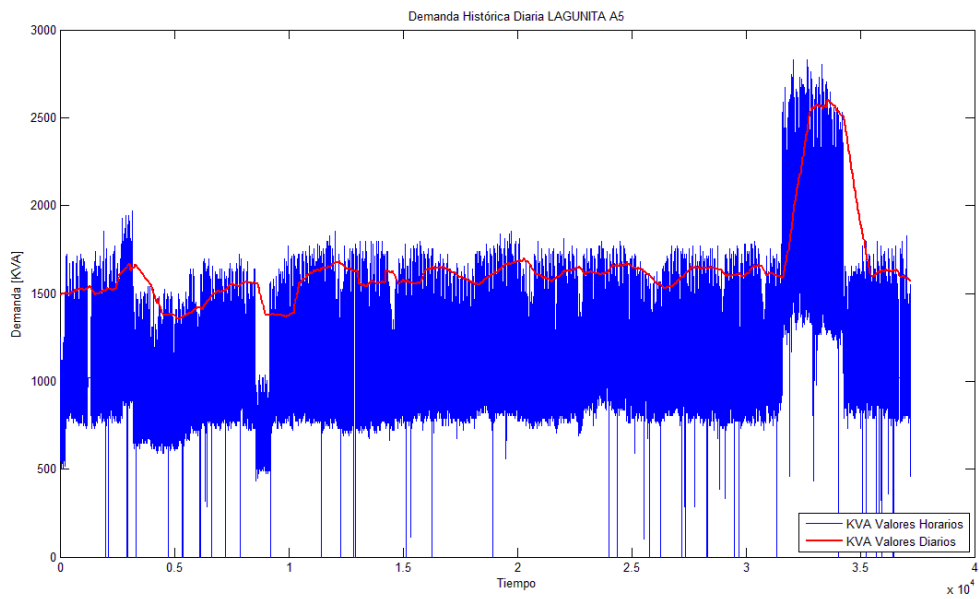
Anexo A.11. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita A2



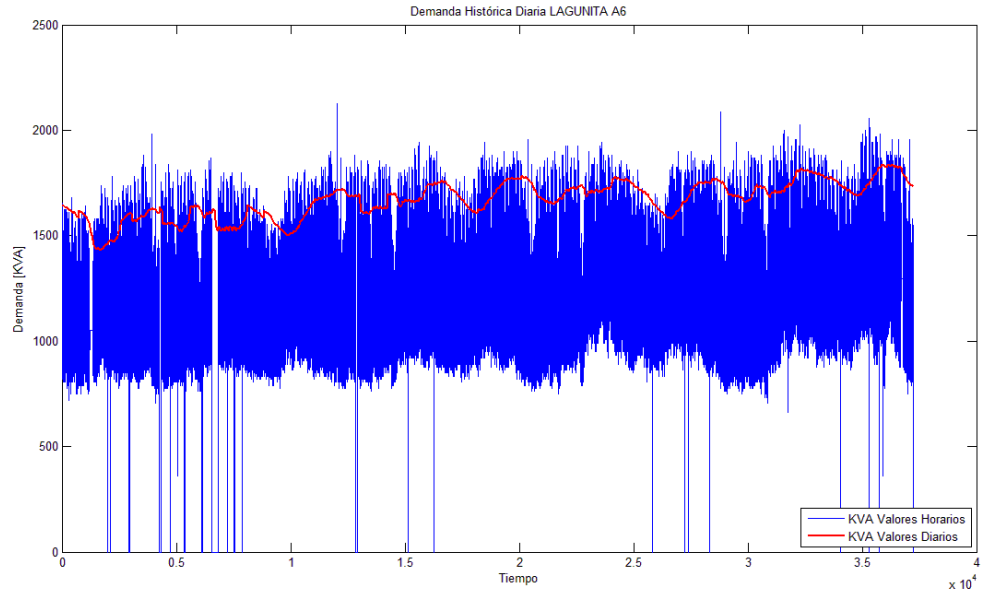
Anexo A.12. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita A3



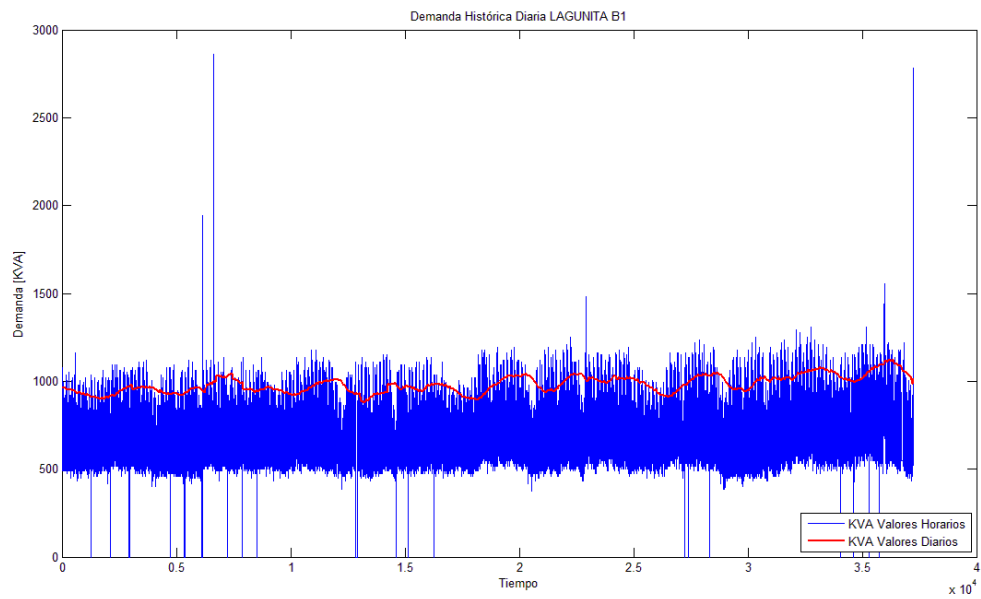
Anexo A.13. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita A4



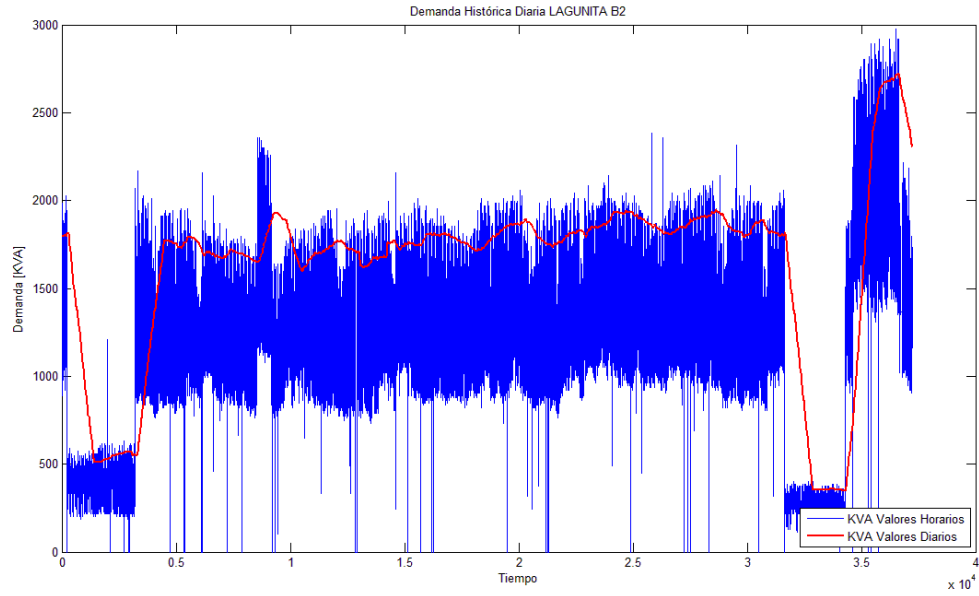
Anexo A.14. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita A5



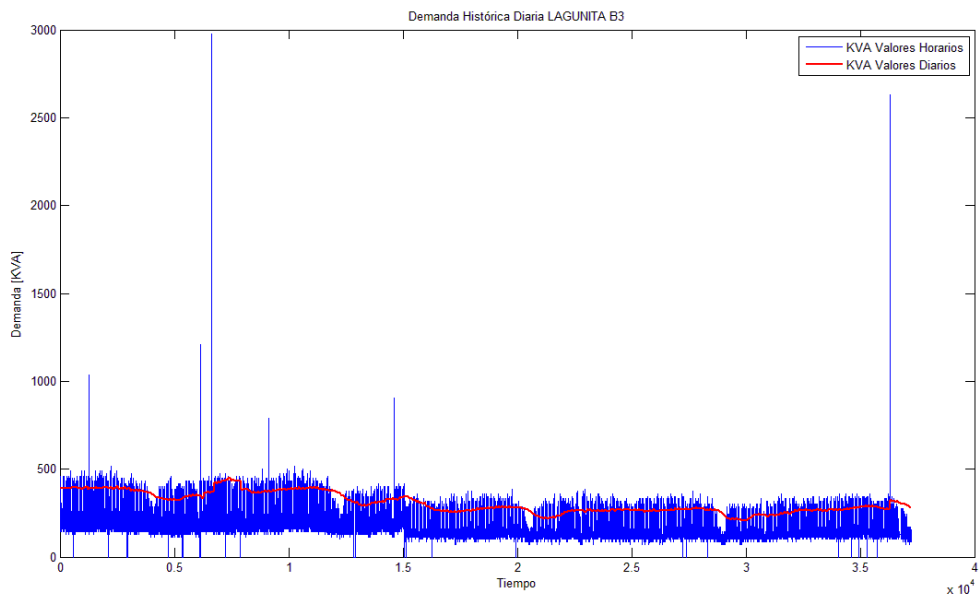
Anexo A.15. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita A6



Anexo A.16. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita B1

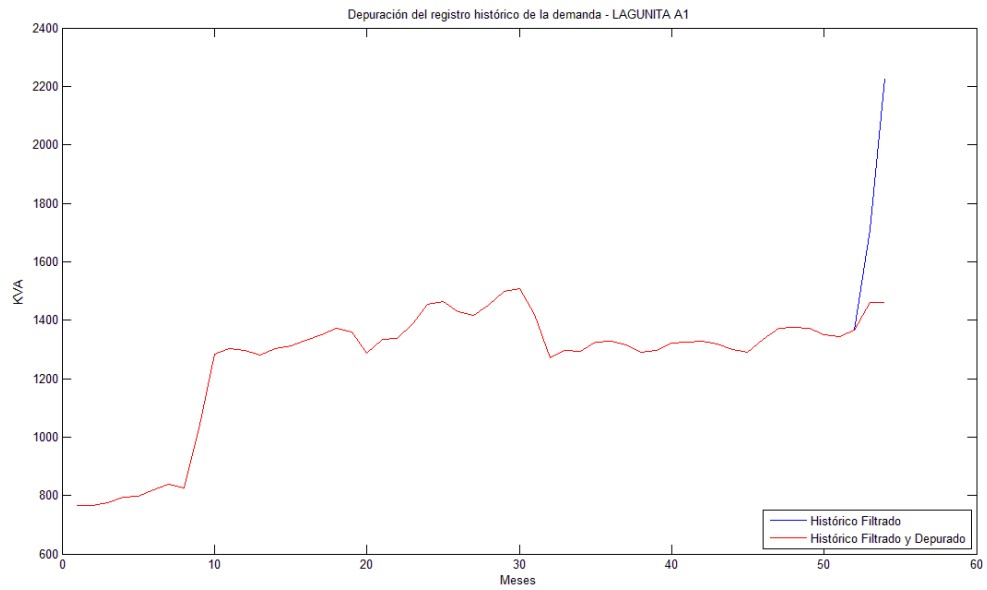


Anexo A.17. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita B2

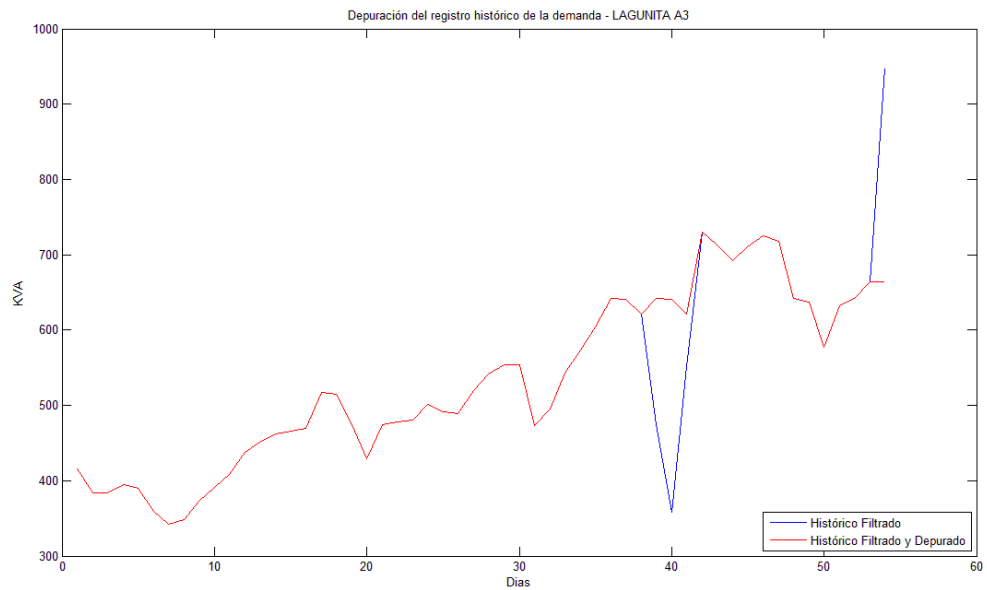


Anexo A.18. Gráfica de la demanda histórica diaria filtrada. Circuito Lagunita B3

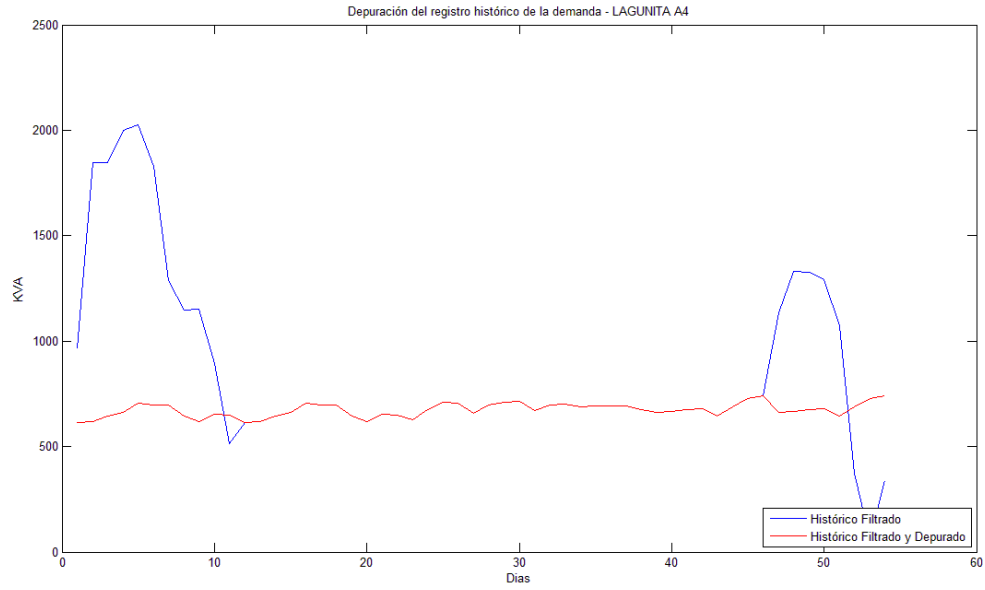
Datos en kVA utilizados para realizar los pronósticos tipo 3



Anexo A.19. Gráfica del registro histórico mensual de la demanda depurada. Circuito Lagunita A1

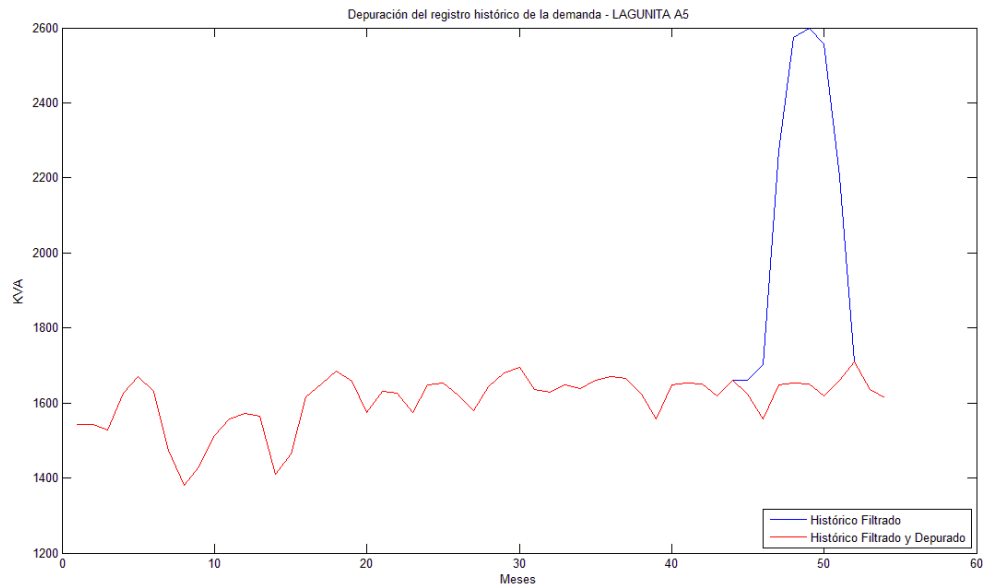


Anexo A.20. Gráfica del registro histórico mensual de la demanda depurada. Circuito Lagunita A3



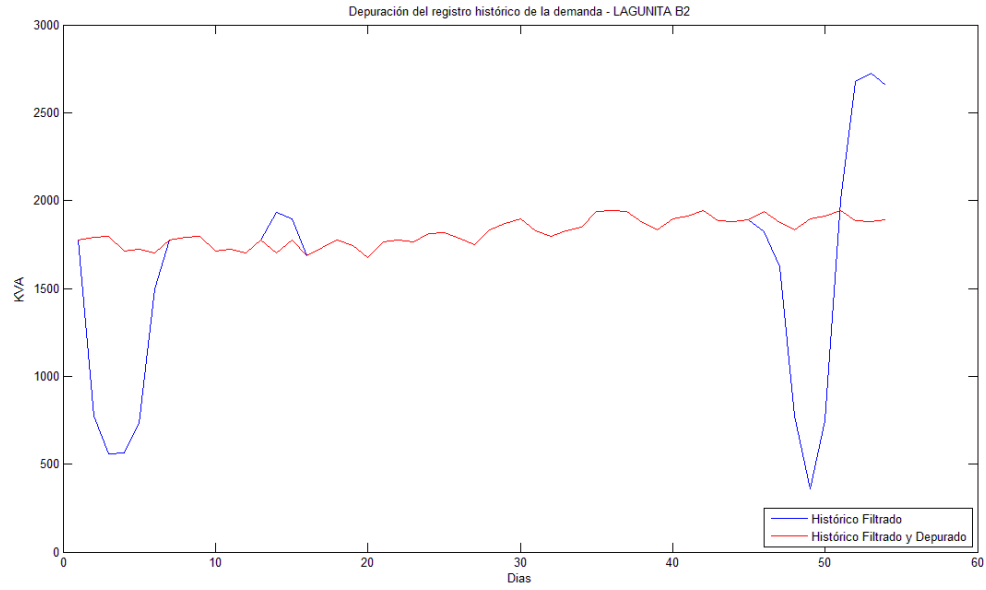
Anexo A.21. Gráfica del registro histórico mensual de la demanda depurada. Circuito Lagunita

A4



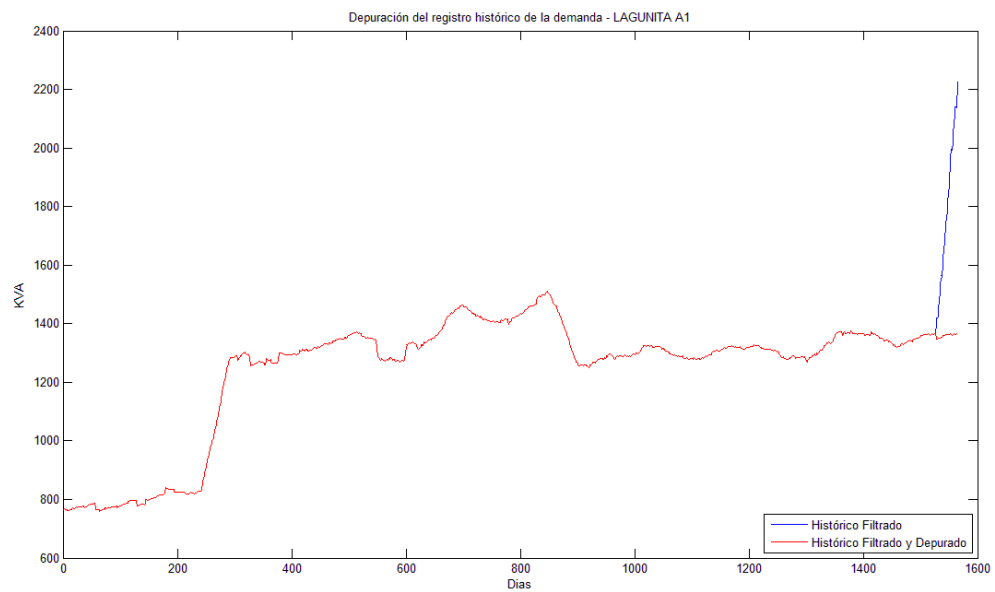
Anexo A.22. Gráfica del registro histórico mensual de la demanda depurada. Circuito Lagunita

A5

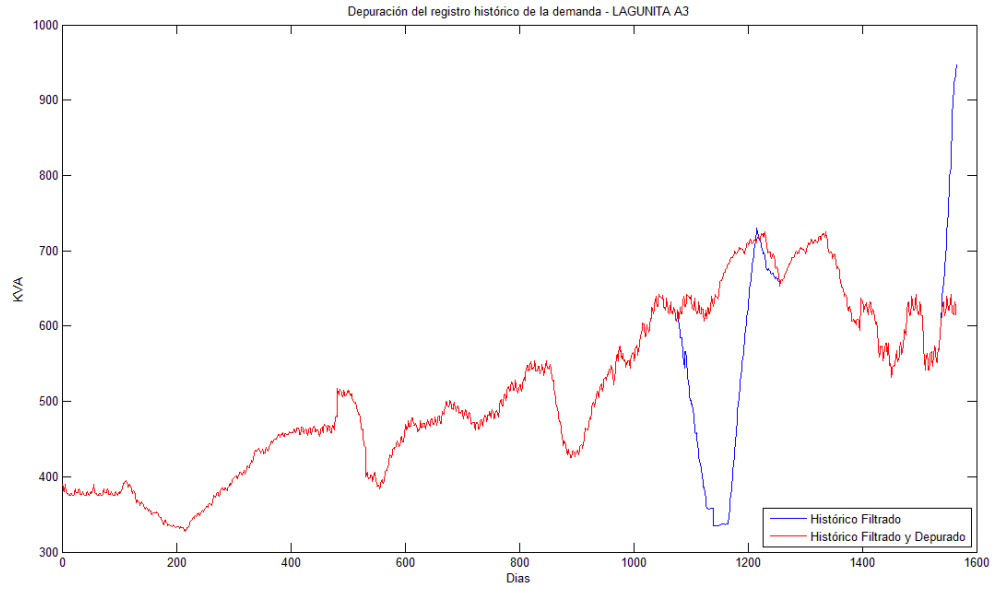


Anexo A.23. Gráfica del registro histórico mensual de la demanda depurada. Circuito Lagunita B2

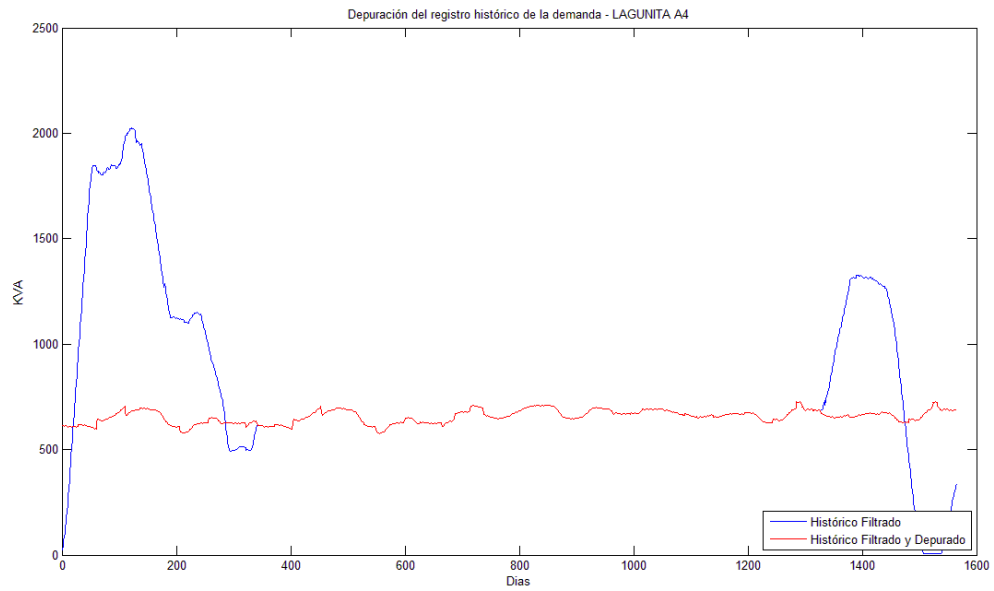
Datos en kVA utilizados para realizar los pronósticos tipo 4



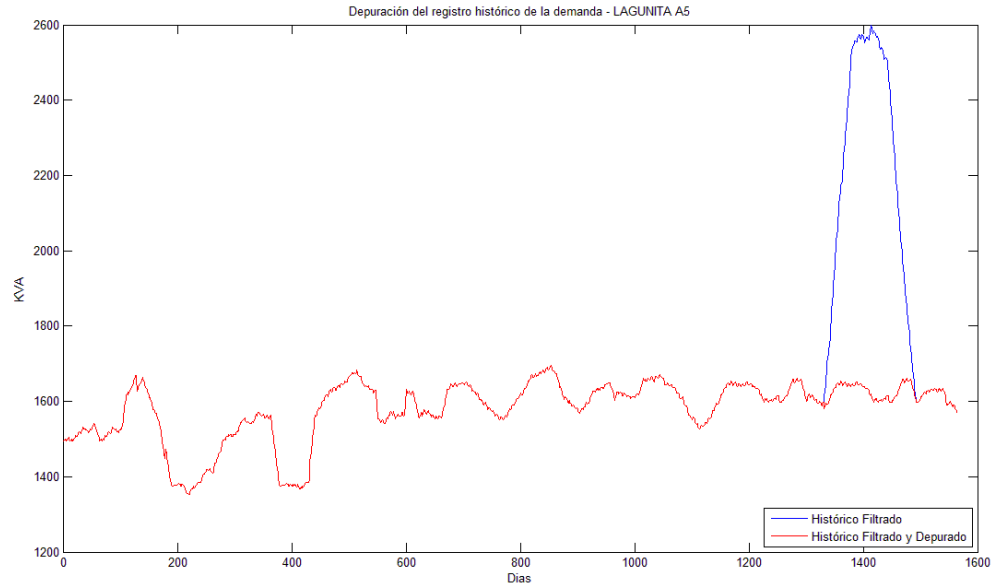
Anexo A.24. Gráfica del registro histórico diario de la demanda depurada. Circuito Lagunita A1



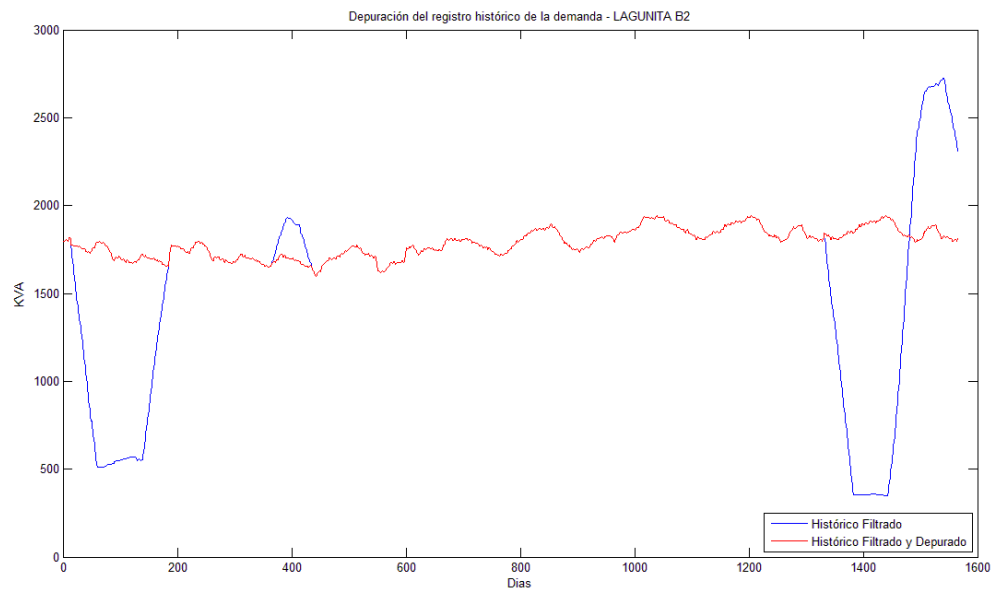
Anexo A.25. Gráfica del registro histórico diario de la demanda depurada. Circuito Lagunita A3



Anexo A.26. Gráfica del registro histórico diario de la demanda depurada. Circuito Lagunita A4



Anexo A.27. Gráfica del registro histórico diario de la demanda depurada. Circuito Lagunita A5



Anexo A.28. Gráfica del registro histórico diario de la demanda depurada. Circuito Lagunita B2

ANEXO B
CONFIGURACIÓN Y RESULTADOS DEL ENTRENAMIENTO DE LAS RNA
UTILIZADAS EN LOS PRONÓSTICOS

Tabla B.1 Configuración y resultados del entrenamiento de las RNA utilizadas en el pronóstico
 tipo 1

Circuito	No. de neuronas de la capa oculta	Error de entrenamiento	Tiempo de entrenamiento (seg)
LAG A1	6	0,0362	24,4900
LAG A2	7	0,0115	23,2160
LAG A3	8	0,0266	29,4130
LAG A4	7	0,0547	36,5560
LAG A5	6	0,0534	9,7630
LAG A6	7	0,0096	17,9470
LAG B1	8	0,0116	26,5970
LAG B2	6	0,1175	10,7800
LAG B3	7	0,0241	17,7370

Tabla B.2 Configuración y resultados del entrenamiento de las RNA utilizadas en el pronóstico
 tipo 2

Circuito	No. de neuronas de la capa oculta	Error de entrenamiento	Tiempo de entrenamiento (seg)
LAG A1	6	0,0196	42,0960
LAG A2	6	0,0126	217,5410
LAG A3	7	0,0366	26,9950
LAG A4	8	0,0679	58,1920
LAG A5	7	0,0179	135,8300
LAG A6	6	0,0131	95,5240
LAG B1	8	0,0158	166,5060
LAG B2	7	0,0498	119,3980
LAG B3	6	0,0224	69,6230

Tabla B.3 Configuración y resultados del entrenamiento de las RNA utilizadas en el pronóstico tipo 3

Circuito	No. de neuronas de la capa oculta	Error de entrenamiento	Tiempo de entrenamiento (seg)
LAG A1	6	0,0330	20,3170
LAG A2	-	-	-
LAG A3	7	0,0246	43,7990
LAG A4	6	0,0188	16,8280
LAG A5	7	0,0124	17,3000
LAG A6	-	-	-
LAG B1	-	-	-
LAG B2	8	0,0222	2,9170
LAG B3	-	-	-

Tabla B.4 Configuración y resultados del entrenamiento de las RNA utilizadas en el pronóstico tipo 4

Circuito	No. de neuronas de la capa oculta	Error de entrenamiento	Tiempo de entrenamiento (seg)
LAG A1	6	0,0155	649,2210
LAG A2	-	-	-
LAG A3	7	0,0225	617,99
LAG A4	7	0,0190	41,0550
LAG A5	8	0,0135	320,3770
LAG A6	-	-	-
LAG B1	-	-	-
LAG B2	8	0,0131	40,0510
LAG B3	-	-	-

ANEXO C

ESTIMACIONES DE DEMANDA DE LOS CIRCUITOS DE LA S/E LAGUNITA UTILIZANDO RNA

Pronósticos tipo 1

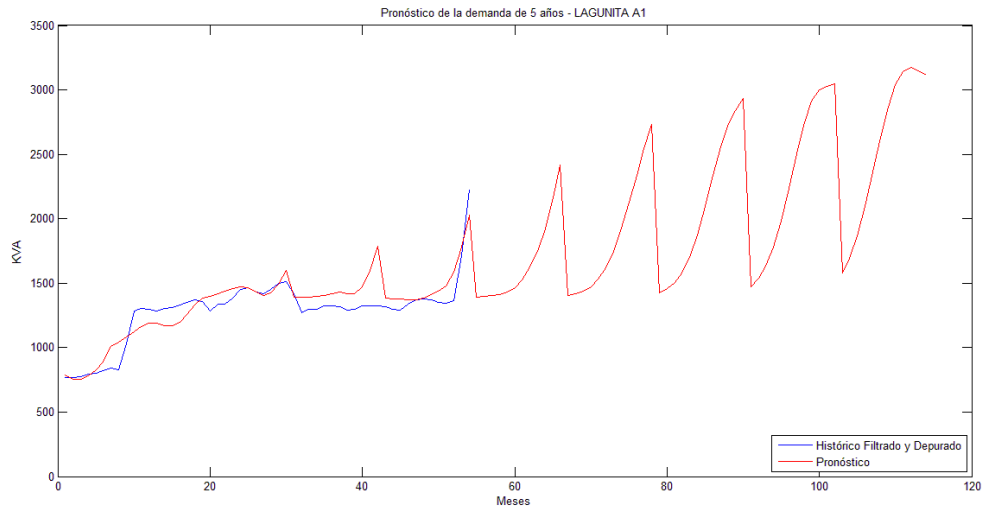


Figura C.1 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A1

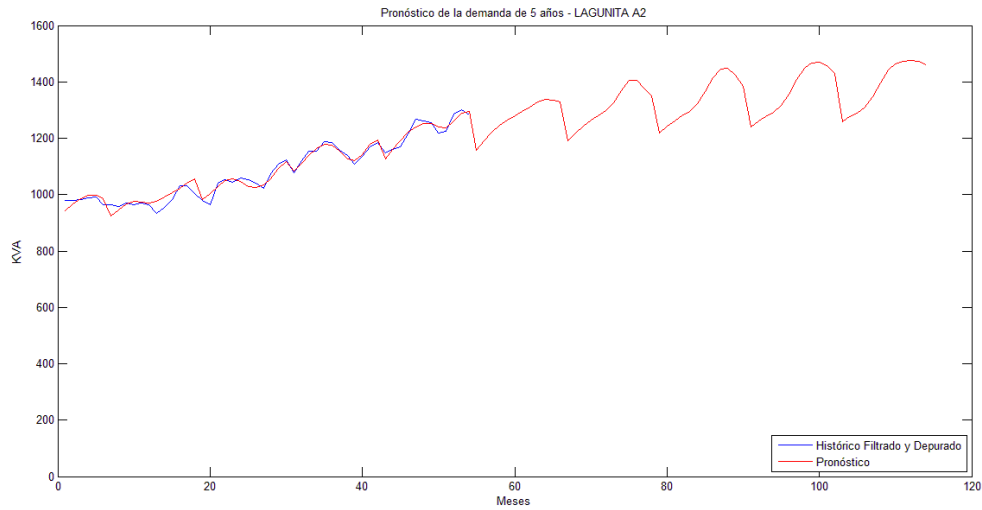


Figura C.2 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A2

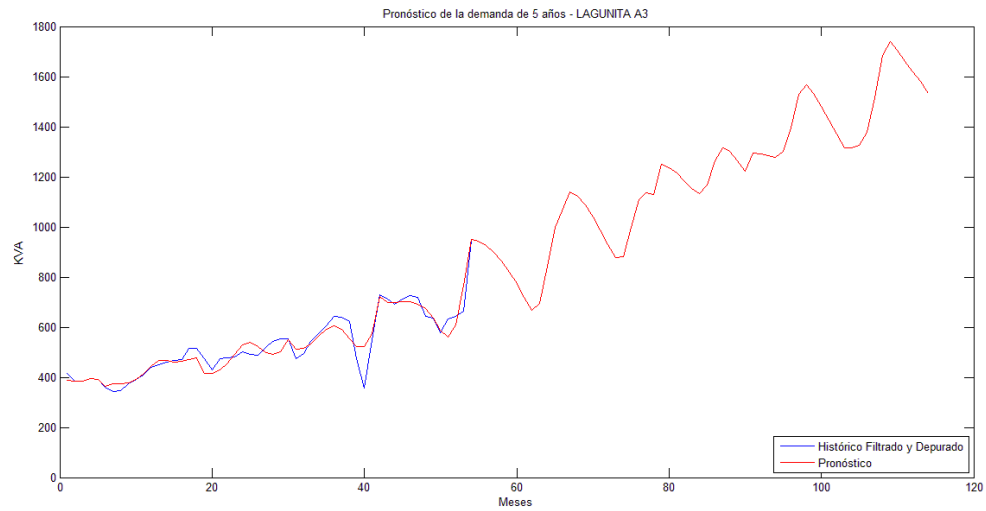


Figura C.3 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A3

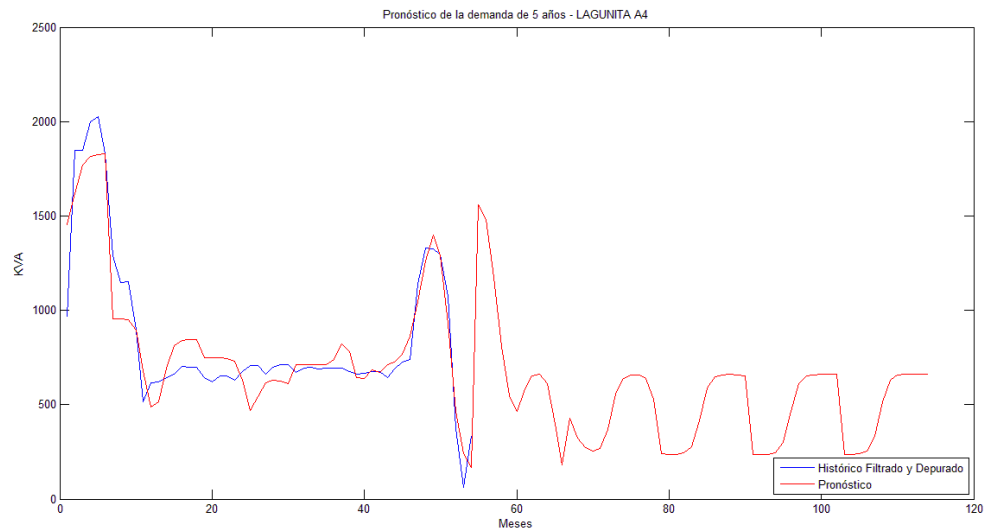


Figura C.4 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A4

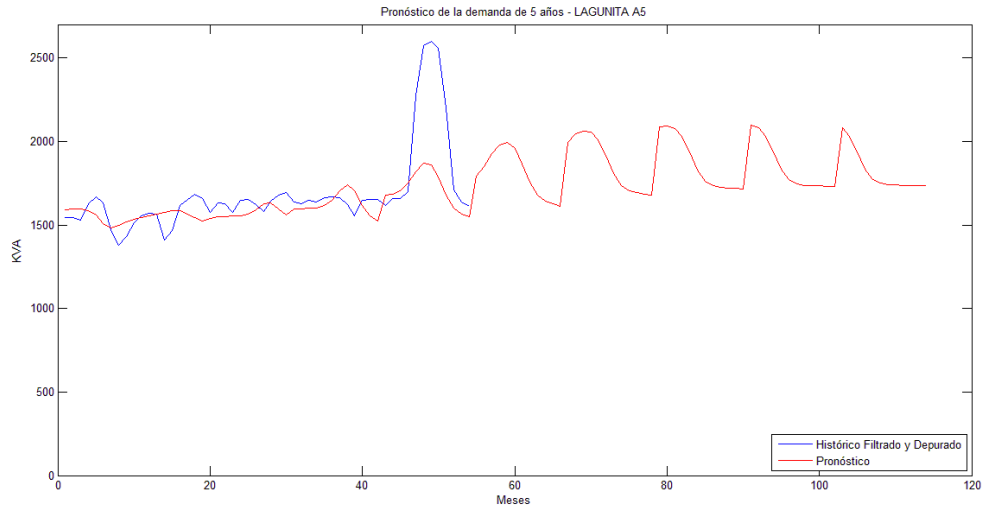
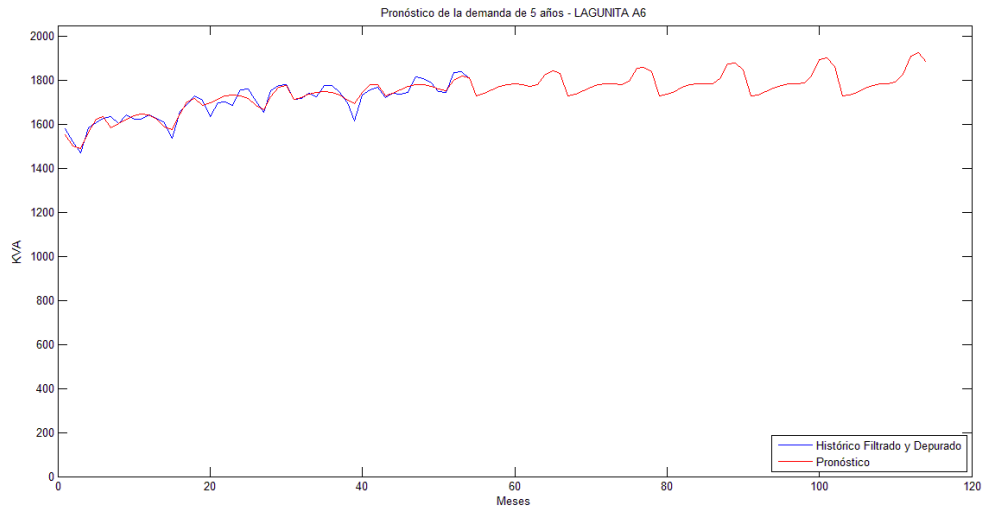
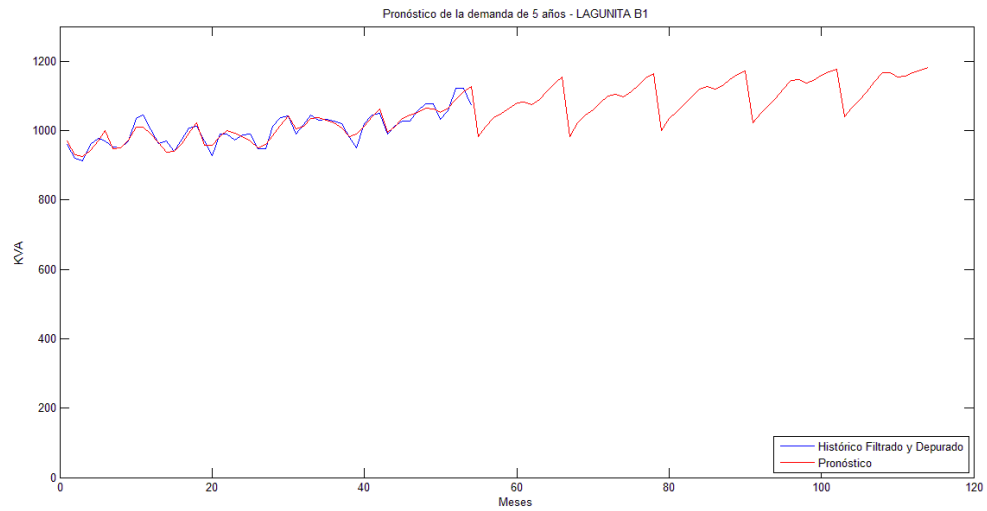


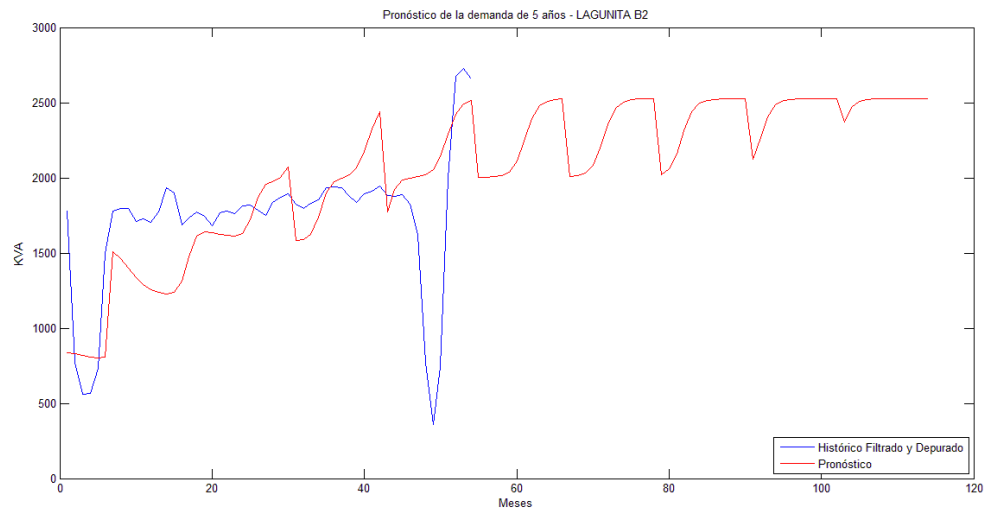
Figura C.5 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A5



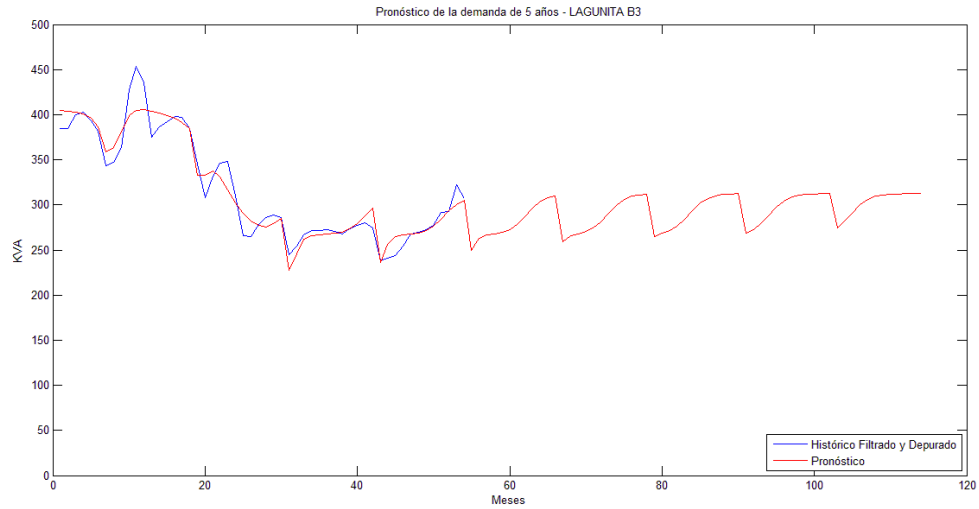
C.6 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita A6



C.7 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita B1

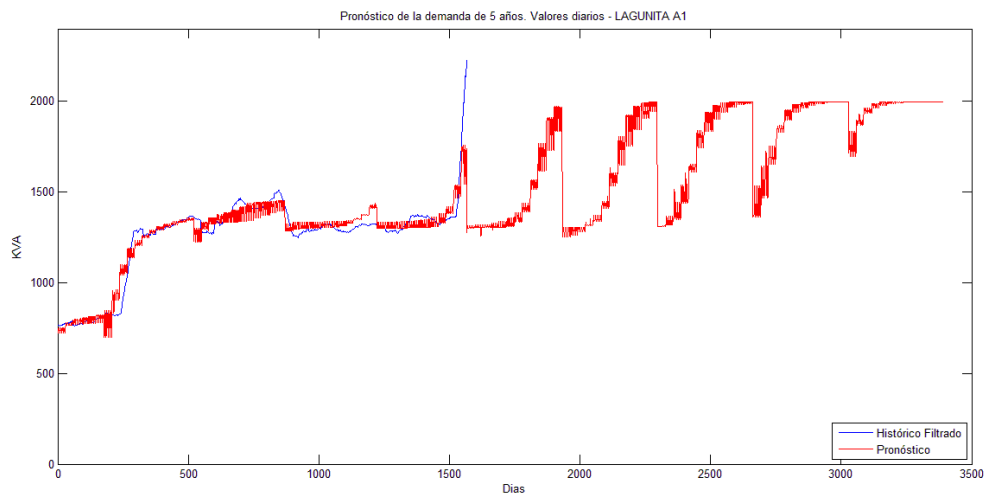


C.8 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita B2

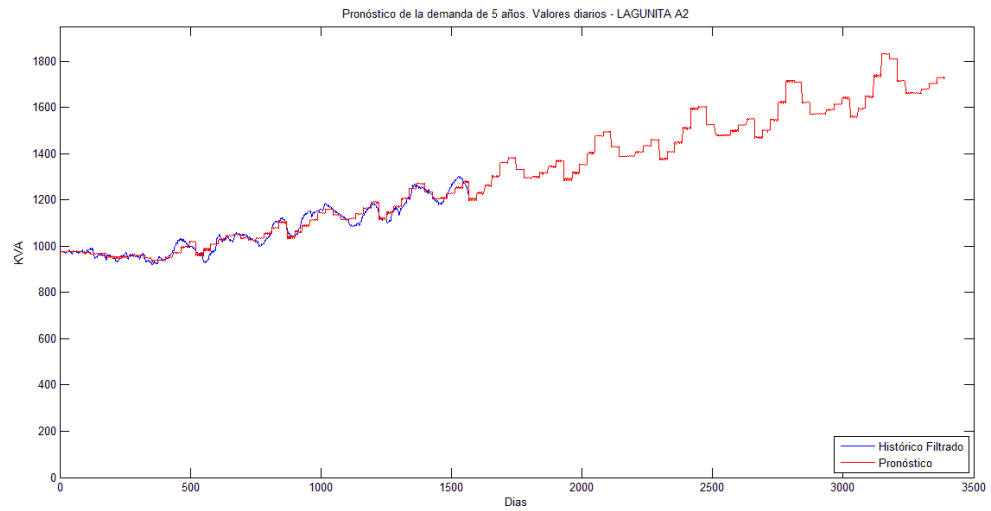


C.9 Pronóstico tipo 1 de la demanda de cinco años del circuito Lagunita B3

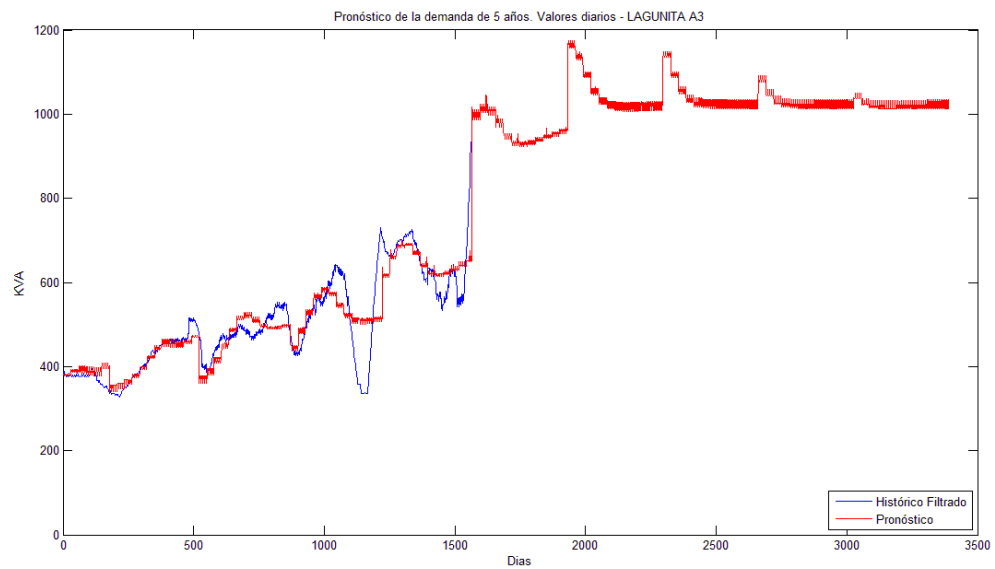
Pronósticos tipo 2



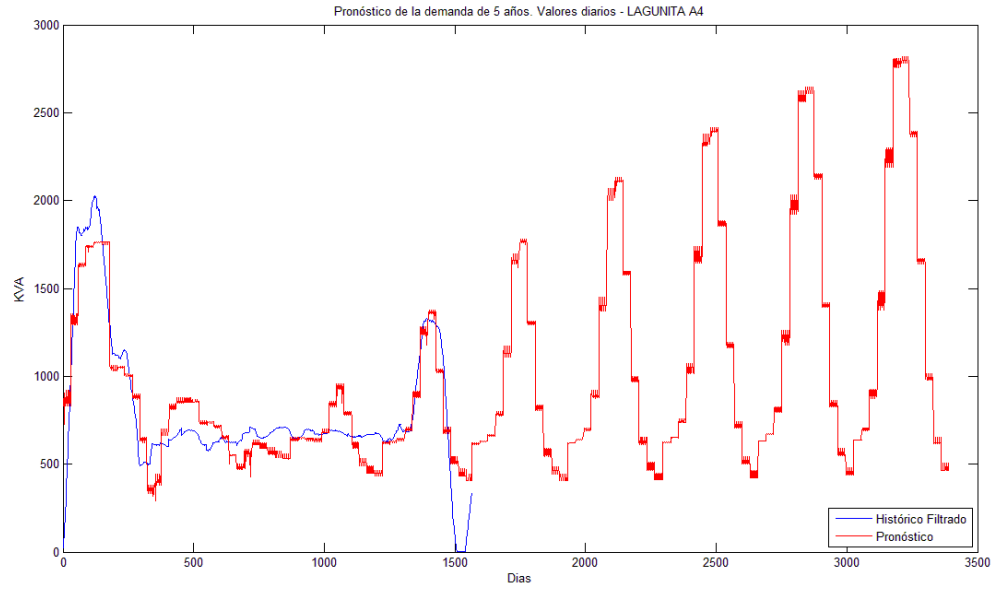
C.10 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A1



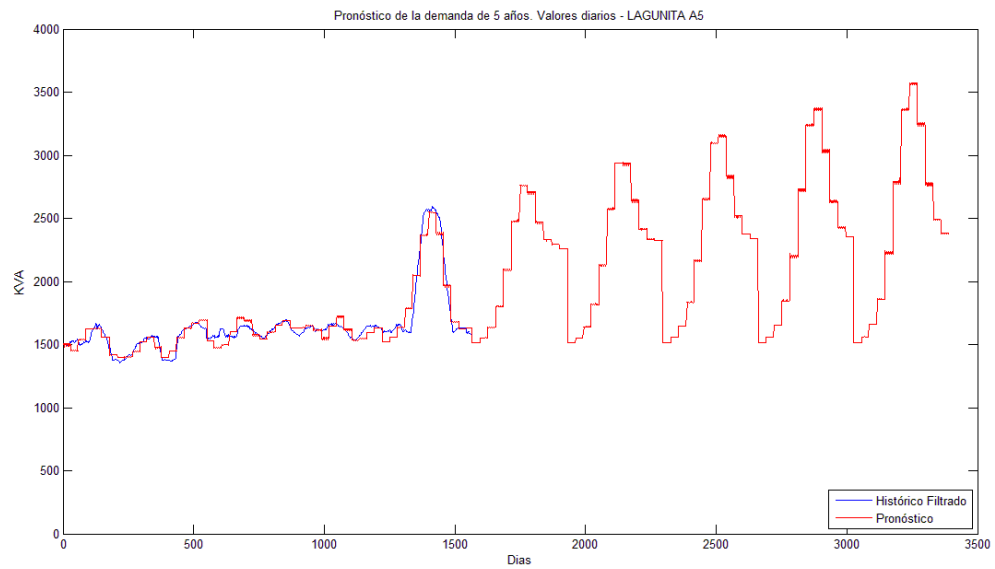
C.11 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A2



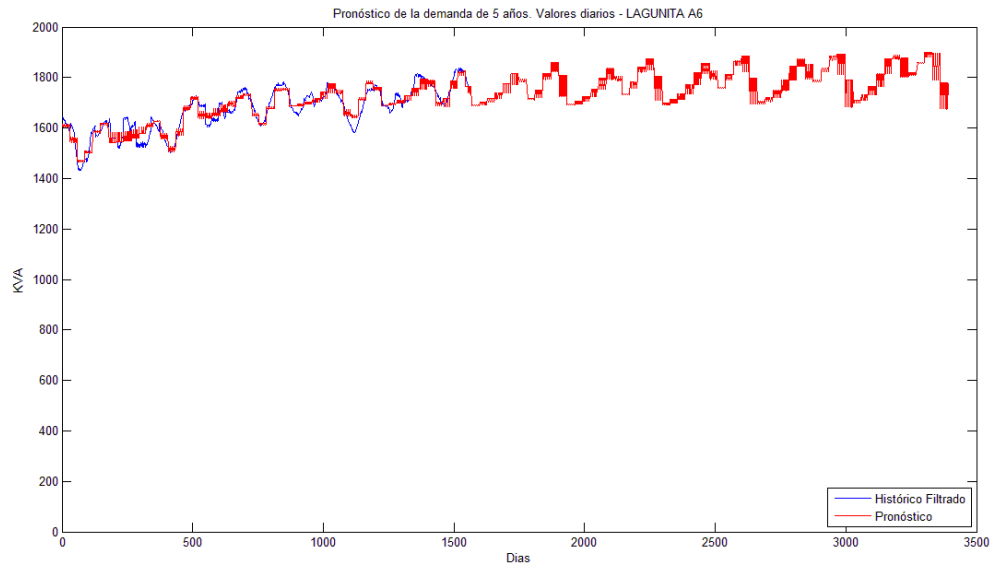
C.12 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A3



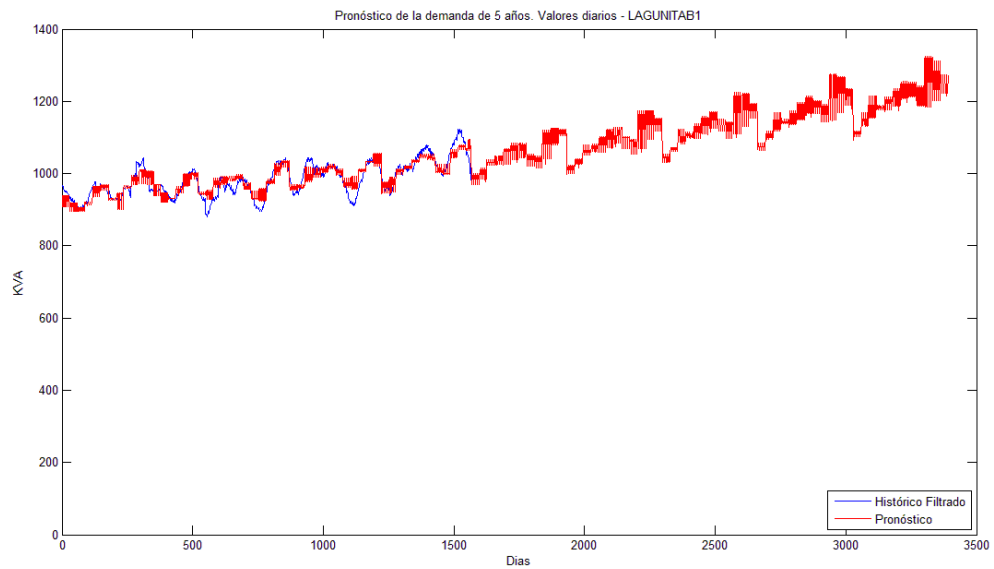
C.13 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A4



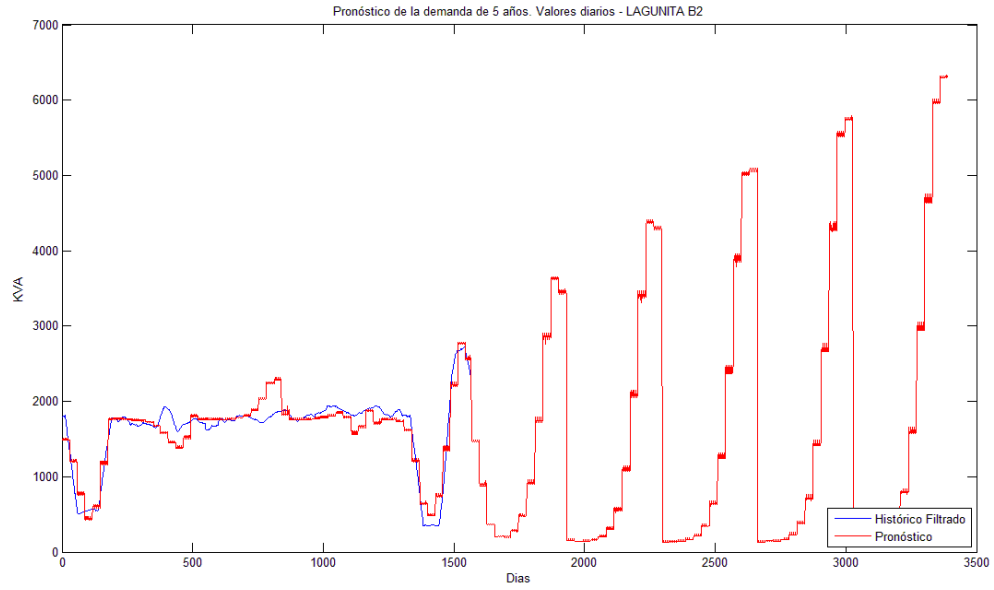
C.14 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A5



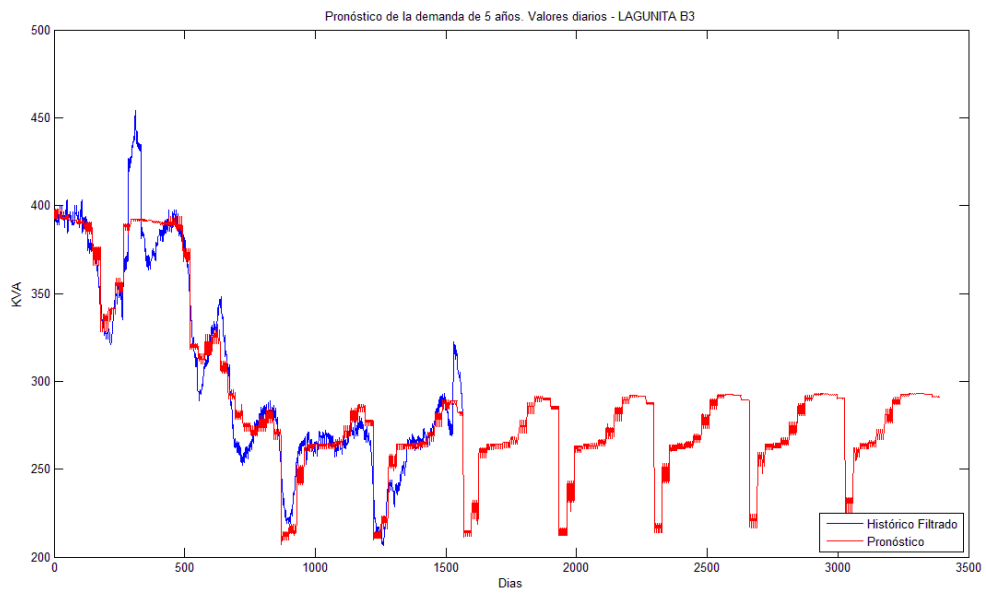
C.15 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita A6



C.16 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita B1

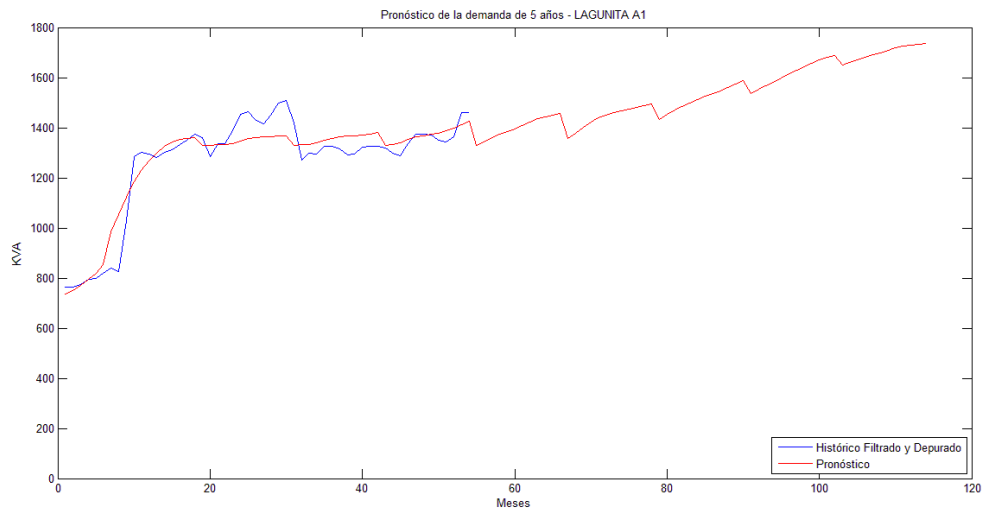


C.17 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita B2

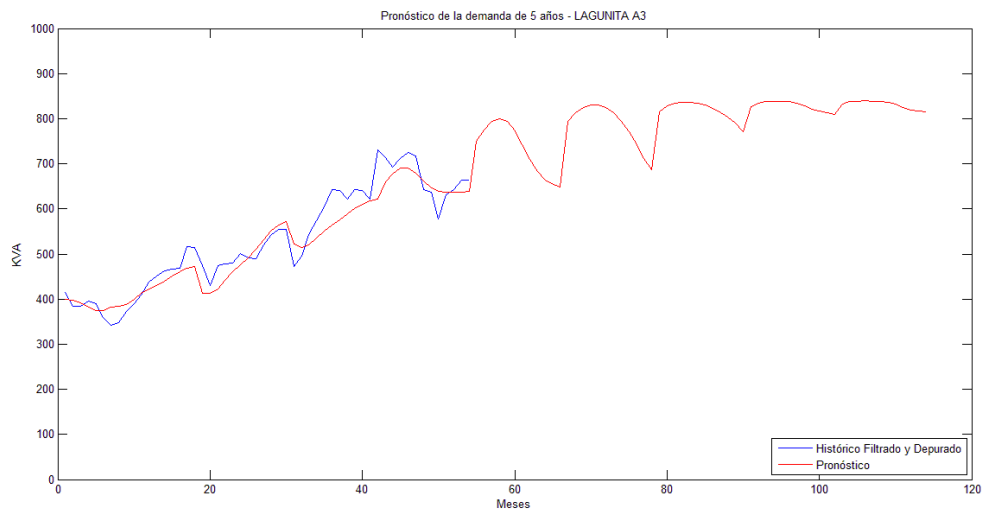


C.18 Pronóstico tipo 2 de la demanda de cinco años del circuito Lagunita B3

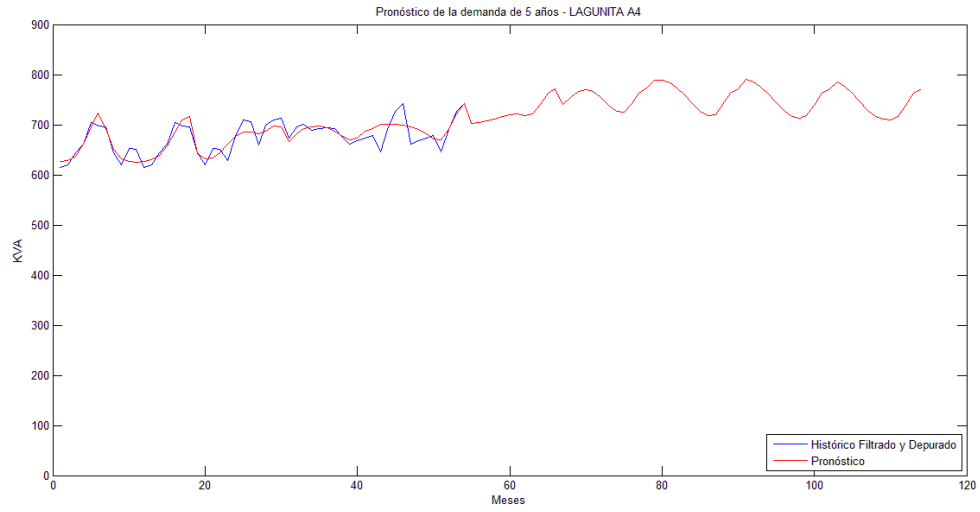
Pronósticos tipo 3



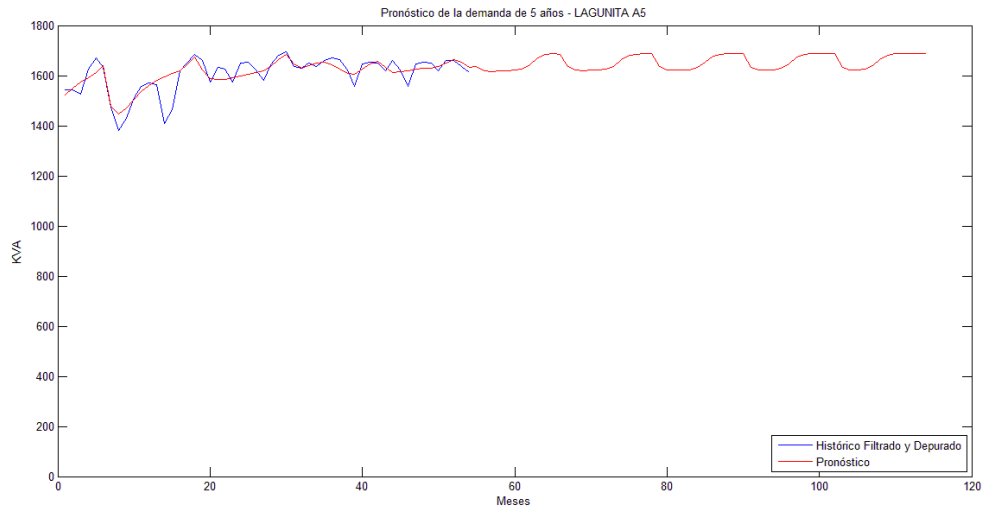
C.19 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A1



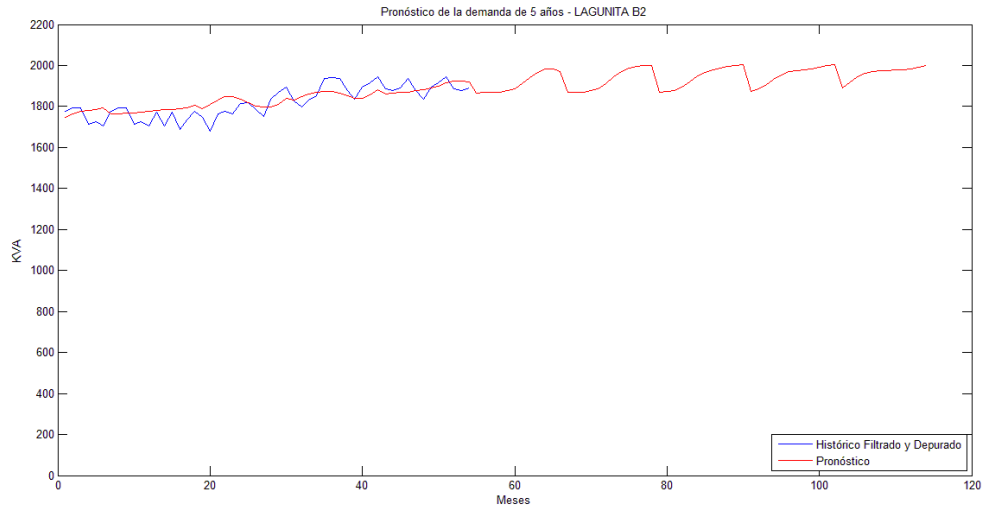
C.20 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A3



C.21 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A4

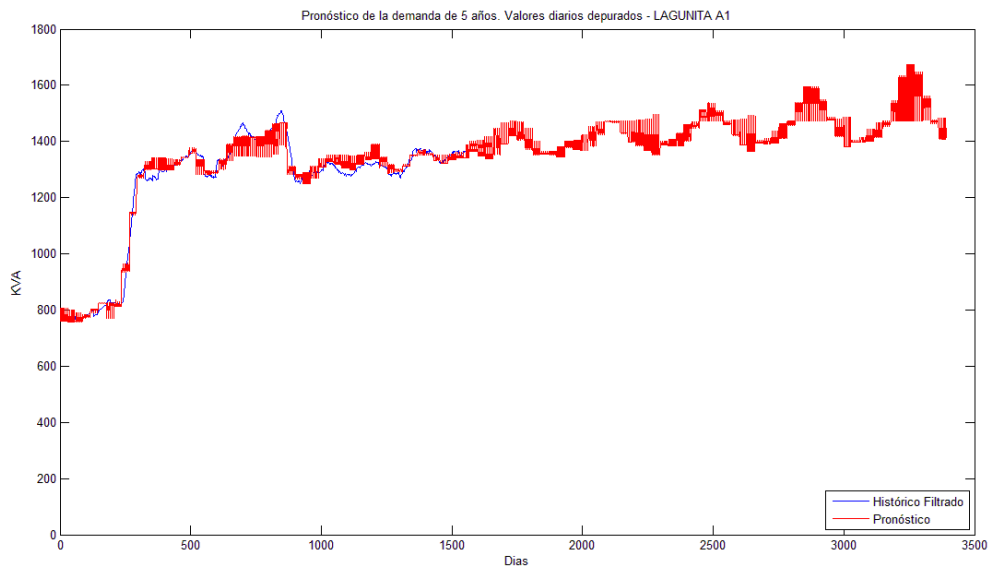


C.22 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A5

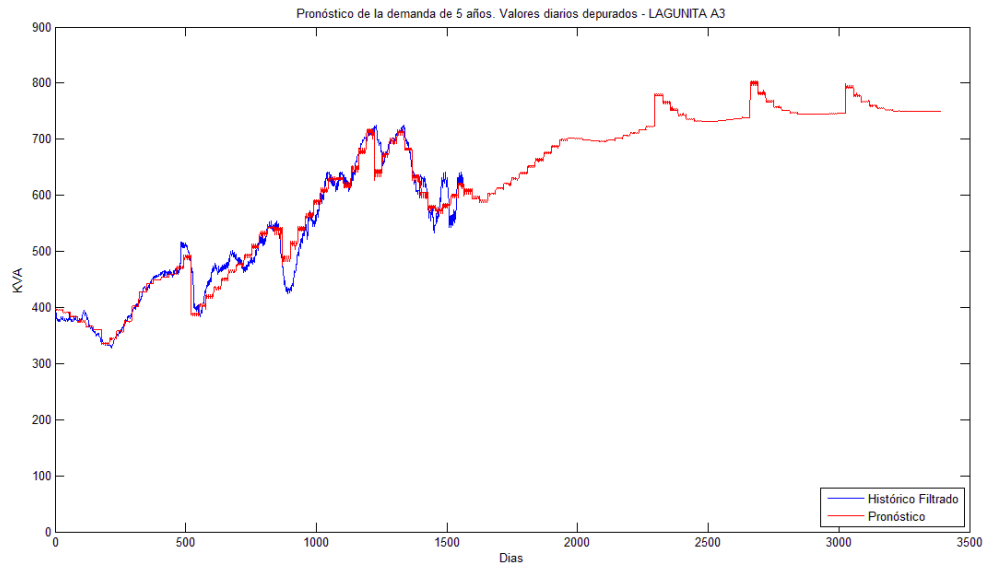


C.23 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita B2

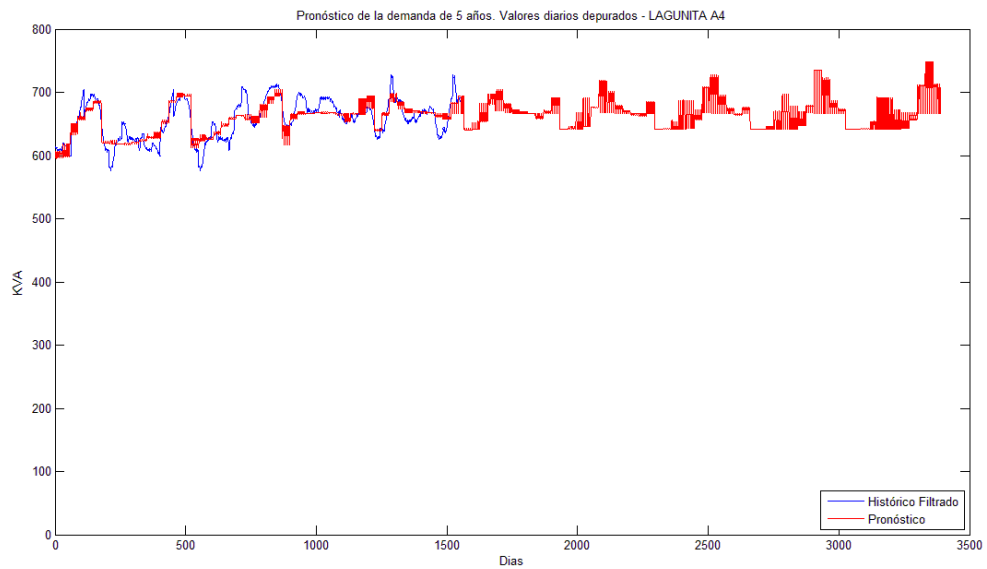
Pronósticos tipo 4



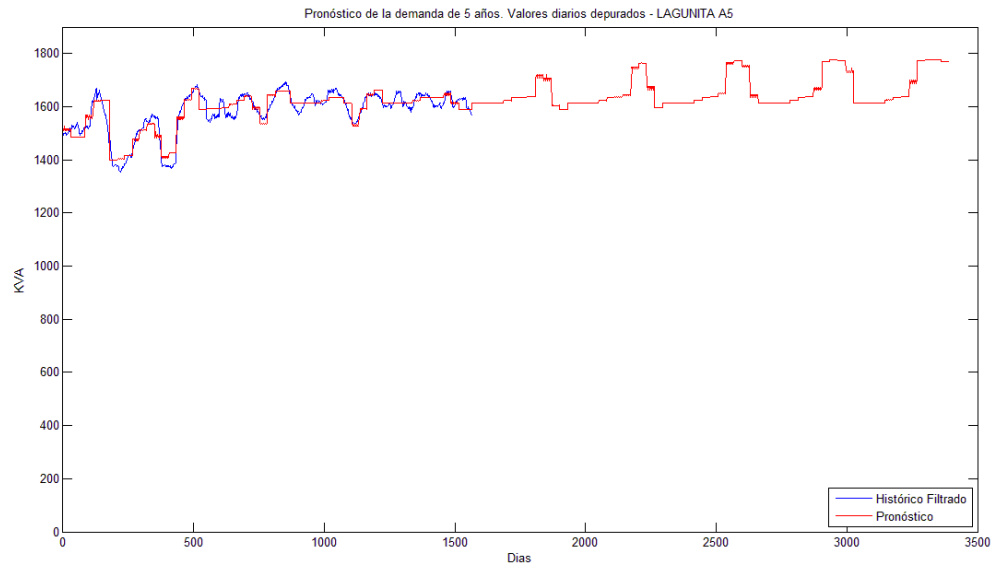
C.24 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A1



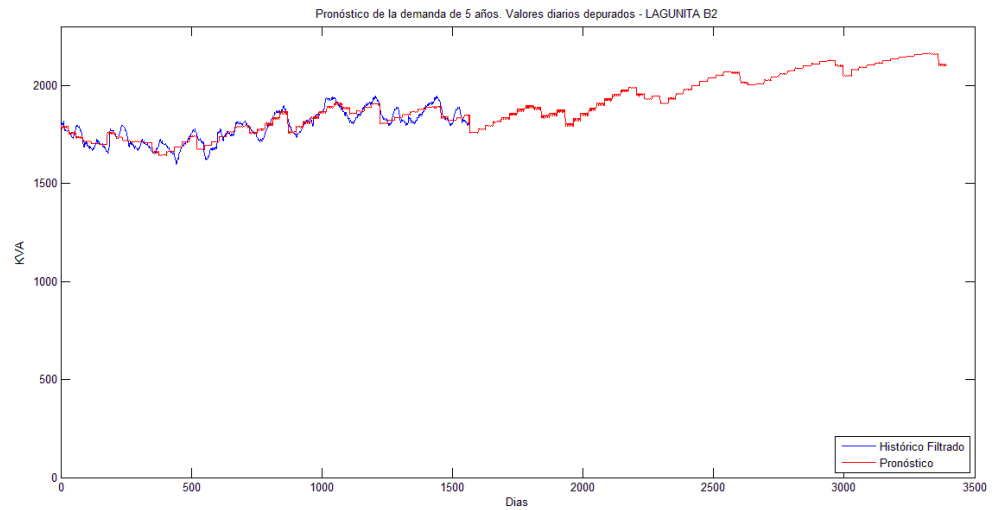
C.25 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A3



C.26 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A4



C.27 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita A5



C.28 Pronóstico tipo 3 de la demanda de cinco años del circuito Lagunita B2

ANEXO D

ESTIMACIONES DE DEMANDA SEGÚN EL MÉTODO DE HOLT-WINTERS DE LOS CIRCUITOS DE LA S/E LAGUNITA

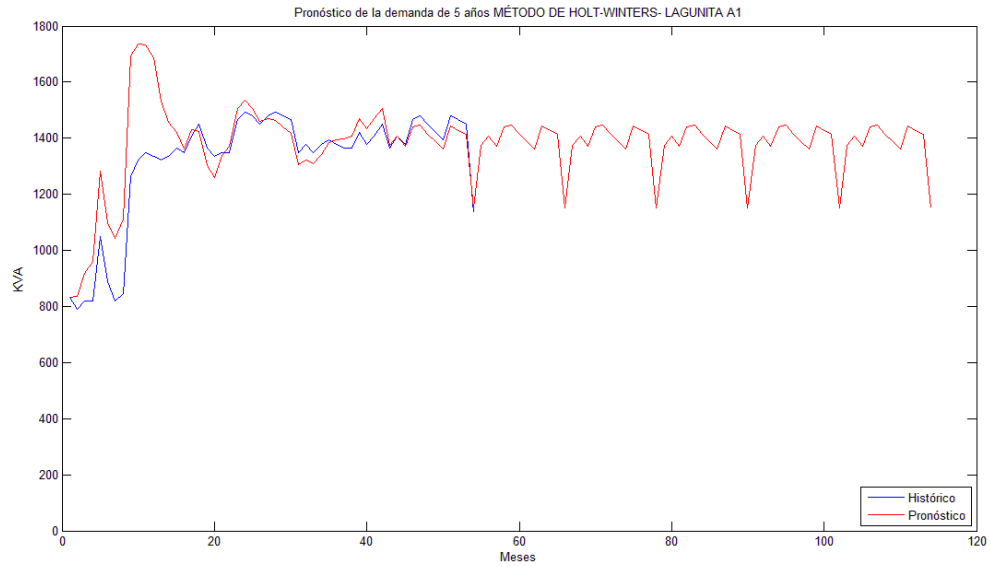


Figura D.1 Pronóstico de la demanda de cinco años. Circuito Lagunita A1

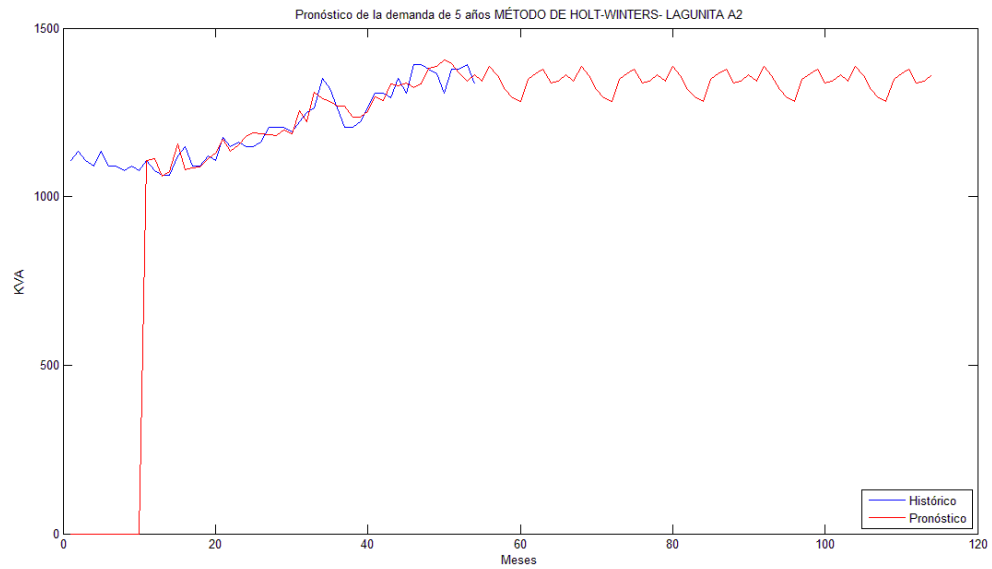


Figura D.2 Pronóstico de la demanda de cinco años. Circuito Lagunita A2

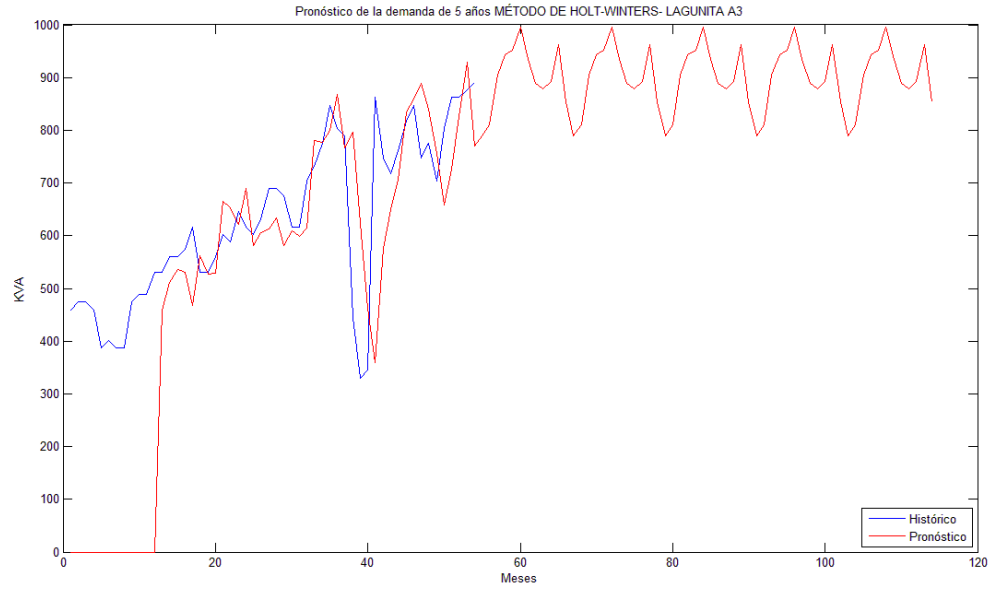


Figura D.3 Pronóstico de la demanda de cinco Años. Circuito Lagunita A3

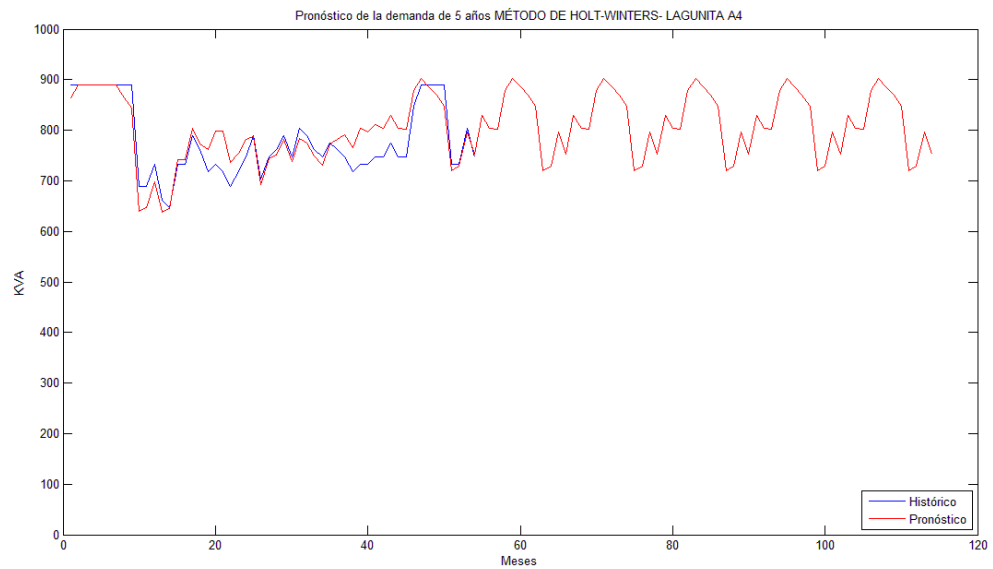


Figura D.4 Pronóstico de la demanda de cinco años. Circuito Lagunita A4

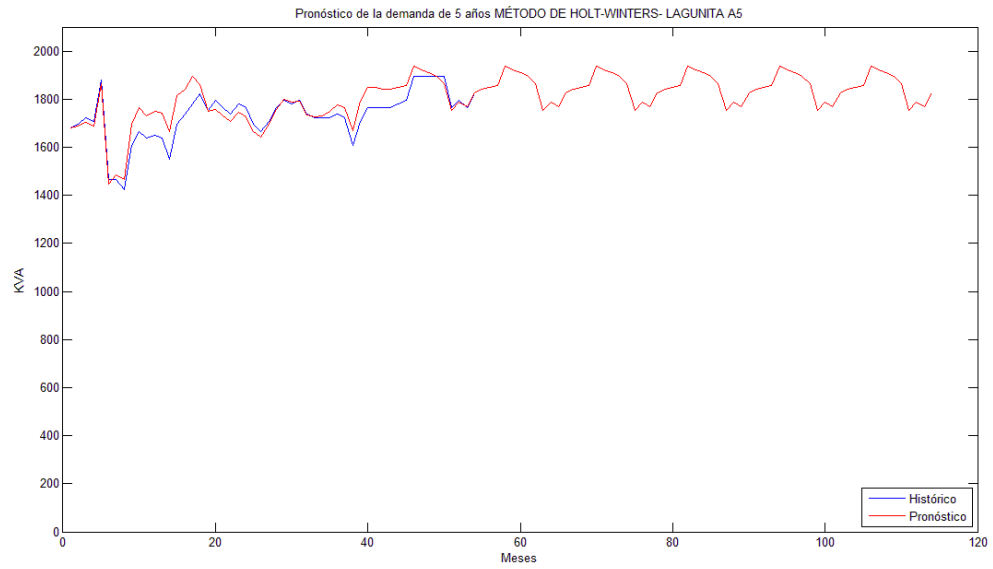


Figura D.5 Pronóstico de la demanda de cinco años. Circuito Lagunita A5

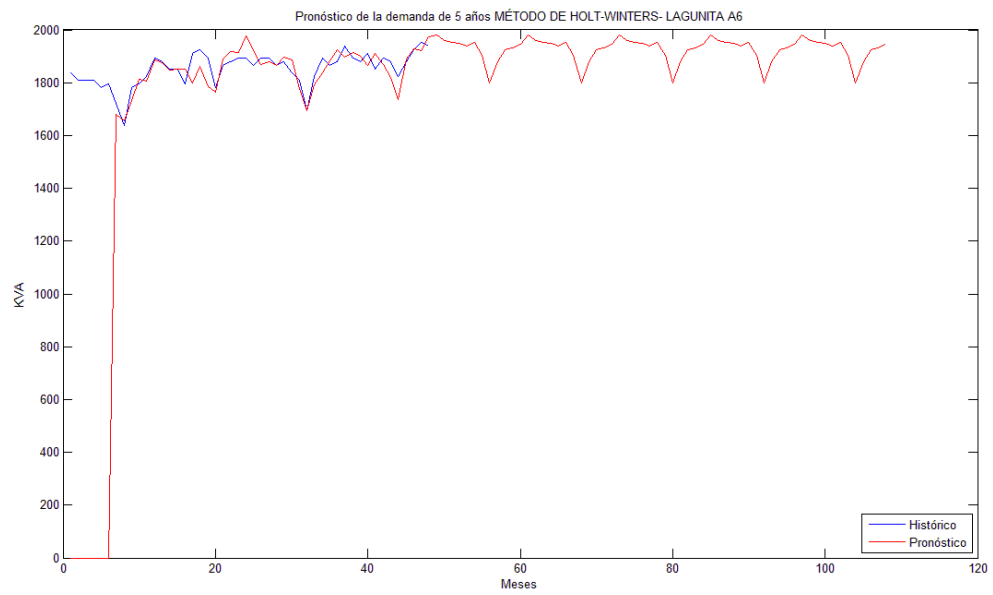


Figura D.6 Pronóstico de la demanda de cinco años. Circuito Lagunita A6

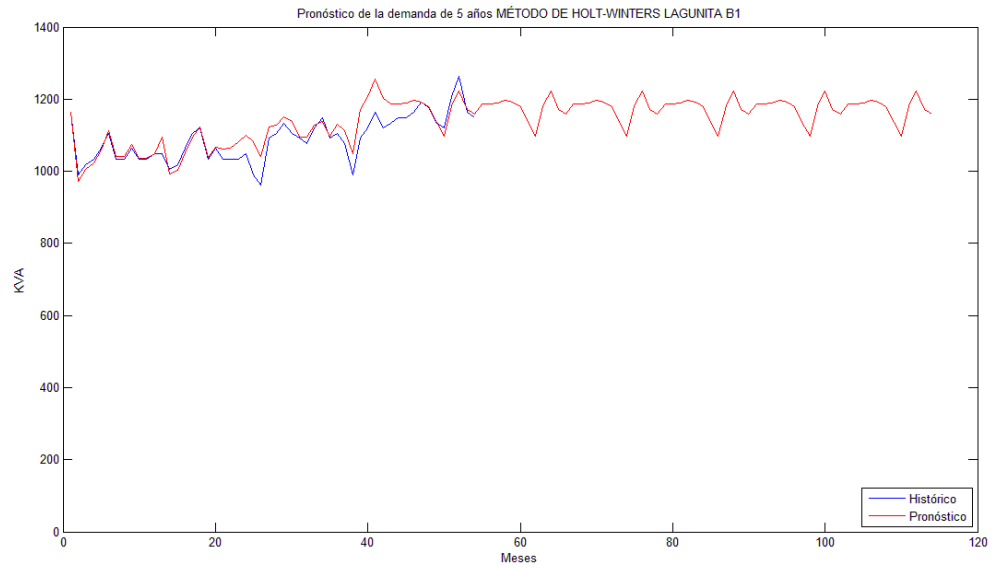


Figura D.7 Pronóstico de la demanda de cinco años. Circuito Lagunita B1

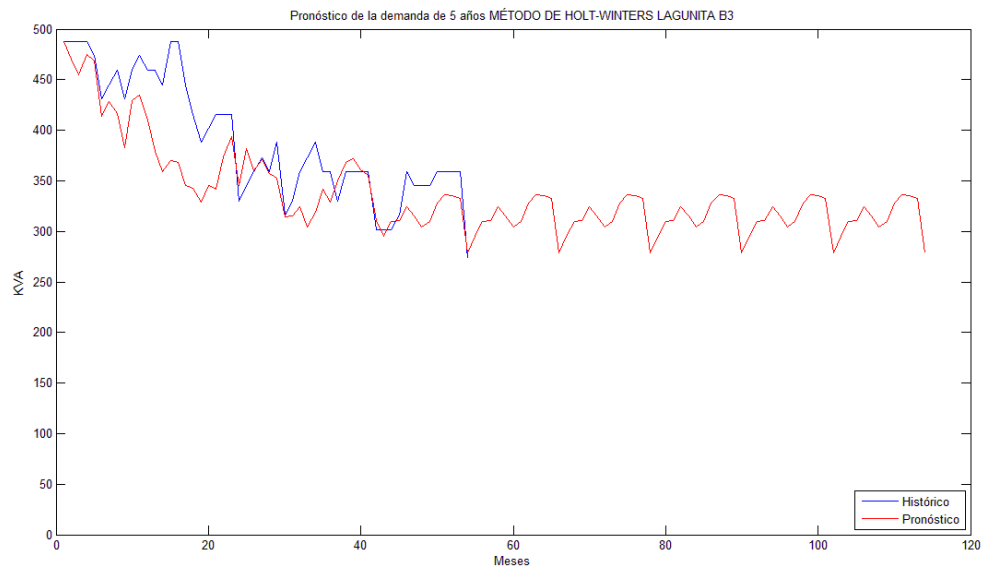


Figura D.8 Pronóstico de la demanda de cinco años. Circuito Lagunita B3

APÉNDICE A

Los mapas de auto-organización son tipos de redes neuronales no supervisadas, competitivas, distribuidas generalmente en una rejilla de dos dimensiones. Cada celda del mapa representa una neurona. Durante el entrenamiento se calcula la similitud entre cada neurona y las entradas que se le presentan a ésta. La neurona más parecida es la ganadora. Este sistema está basado en la capacidad que posee el cerebro de autoorganizarse al recibir diferentes estímulos que son procesados en regiones específicas a partir de la creación de patrones o características importantes. Esto se logra al observar como los estímulos son distribuidos (mapeados) en diversas regiones mostrando así las diferentes características o patrones presentes. Para ello se desarrollan algoritmos de entrenamiento no supervisados que están divididos en dos grandes categorías:

Hard Competitive learning:

- Las neuronas de la red compiten entre sí y la neurona ganadora produce la respuesta al estímulo.
- En el entrenamiento sólo los pesos sinápticos de la neurona ganadora se actualizan.
- No existe interacción entre neuronas vecinas.

Soft Competitive Learning:

- Las neuronas en la red compiten entre sí y solo una de ellas (neurona ganadora) produce la respuesta al estímulo.
- En el entrenamiento los pesos sinápticos de todas las neuronas se actualizan
- Existe interacción entre las neuronas vecinas.

Procedimiento del mecanismo de competencia

Consiste en considerar una medida de similitud y buscar la neurona cuyo vector de pesos sinápticos se asemeja más al estímulo presentado a la red. La medida de la similitud usada varía dependiendo de las aplicaciones.

Para medir la similitud entre neuronas y entradas se utiliza usualmente la distancia euclídea. En este caso la neurona ganadora es aquella para la cual la distancia es la más pequeña. En la Figura A.1 se ilustra un mapa de auto-organización.

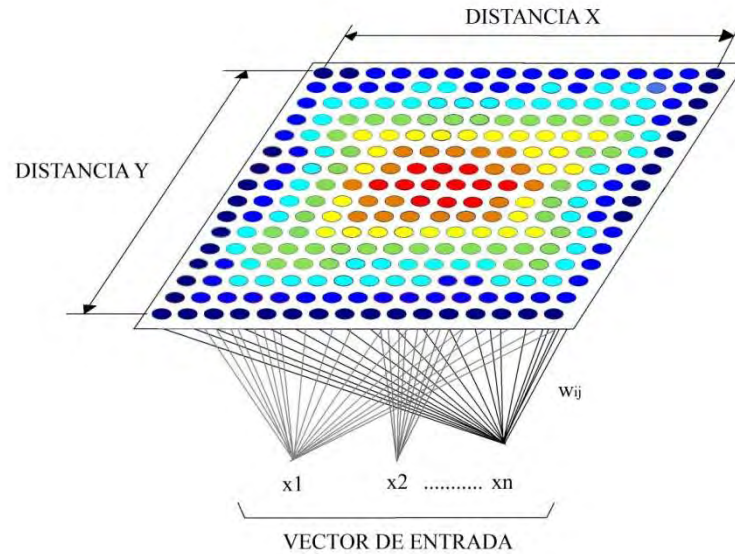


Figura A.1. Mapa de auto-organización

SOM Toolbox

Existen diversas herramientas computacionales que permiten simular los mapas de auto-organización, una de ellas es el SOM Toolbox. SOM se refiere a mapas auto-organizados en sus siglas en inglés (*Self Organizing Map*).

Es un paquete computacional que nació de la necesidad de implementar de una forma fácil y práctica los mapas auto-organizados en MATLAB® con propósitos de investigación. En particular, los investigadores responsables del paquete trabajan en el campo de la minería de datos, y por lo tanto dicho paquete está orientado hacia la formación de funciones de visualización de gran alcance.

Las principales funciones del paquete son las siguientes [23]:

- Entrenar el SOM con diferentes topologías de la red y aprender diferentes parámetros.
- Calcular diferentes errores, calidad y medidas para el SOM.

- Visualizar el SOM usando matrices de distancia, planos de componentes, codificación de colores del agrupamiento y conexión de colores entre el SOM y otros métodos de visualización.
- Hacer correlaciones y análisis de agrupaciones con SOM.

Implementación de los mapas de auto-organización al caso de estudio

Se desea crear un mapa de auto-organización con la ayuda del paquete de MATLAB®, SOM Toolbox, para visualizar de qué manera puede esto ser útil como proceso de identificación y extracción de eventos o distorsiones presentes en la serie de datos que se analizan. Para ello se utiliza la curva de demanda de un circuito de La Electricidad de Caracas en un período de 5 años, como se muestra a continuación en la Figura A.2:

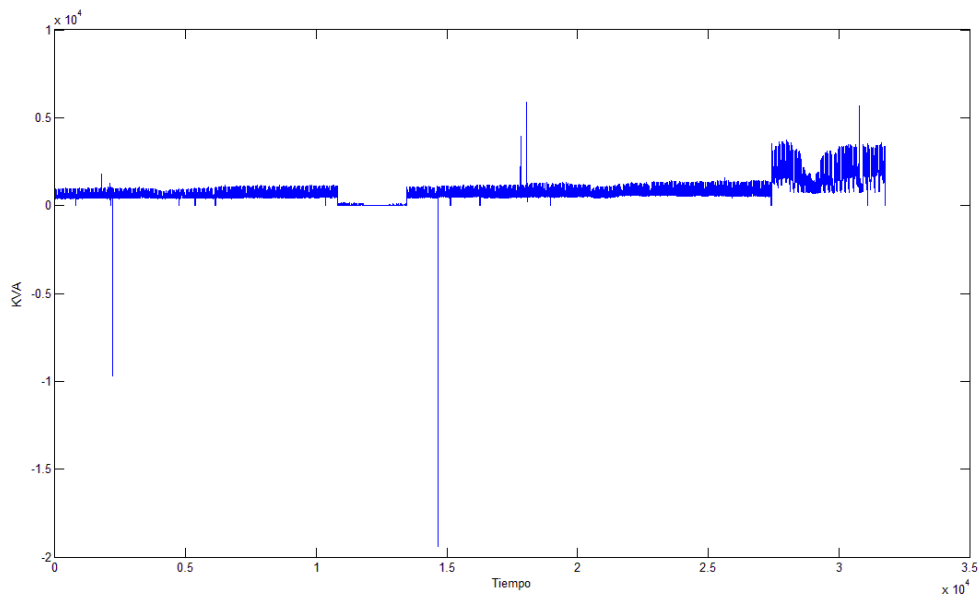


Figura A.2. Curva de demanda de un período de 5 años de un circuito de La EDC

Luego de crear y entrenar el mapa de auto-organización se obtuvo el resultado mostrado en la Figura A.3:

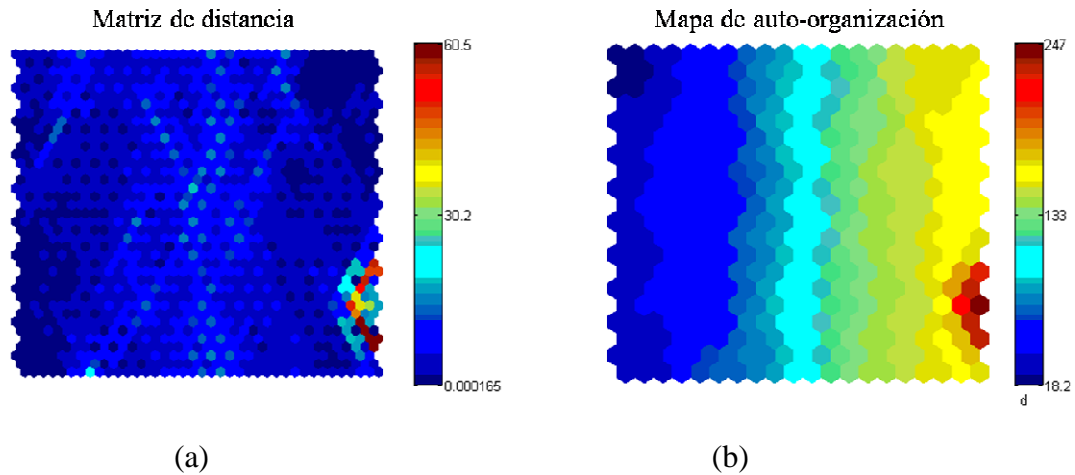


Figura A.3 (a) y (b). Matriz de distancia y mapa de auto-organización respectivamente, obtenidos a partir de la curva de demanda de un circuito de La EDC

La Figura A.3 (a) corresponde a una matriz de distancia, donde se visualizan, como su nombre lo indica, las distancias entre neuronas vecinas del mapa de auto-organización de la Figura A.3 (b). En la parte derecha de cada figura se observa una barra que representa una escala que va desde el menor hasta el mayor de los valores que se exhiben tanto en la matriz de distancia como en el SOM. Asociados a los valores de la escala se encuentra un espectro de colores que van en forma creciente desde el azul oscuro hasta el rojo oscuro. De esta manera, los colores ayudan a visualizar el valor correspondiente a cada casilla o neurona del mapa. En este orden de ideas, las casillas del mismo color pertenecen a un mismo grupo. Entonces puede observarse que el mapa de auto-organización ha formado varios grupos de datos según su valor numérico.

En especial llaman la atención dos grupos de menor tamaño respecto a los demás grupos. Estos son: el de color azul oscuro (referido a los valores más pequeños) y el de color rojo oscuro (referido a los valores más grandes). Si se observa la curva de carga mostrada en la Figura A.2, pueden diferenciarse dos intervalos en donde los valores son de menor y mayor valor numérico respecto al resto de los datos. Este hecho le da paso a la hipótesis de que los grupos de menor tamaño observados en el mapa de auto-organización corresponden a los intervalos de menor tamaño y con valores diferentes a la mayoría de los datos de la curva de carga. La Figura A.4 ilustra esta hipótesis.

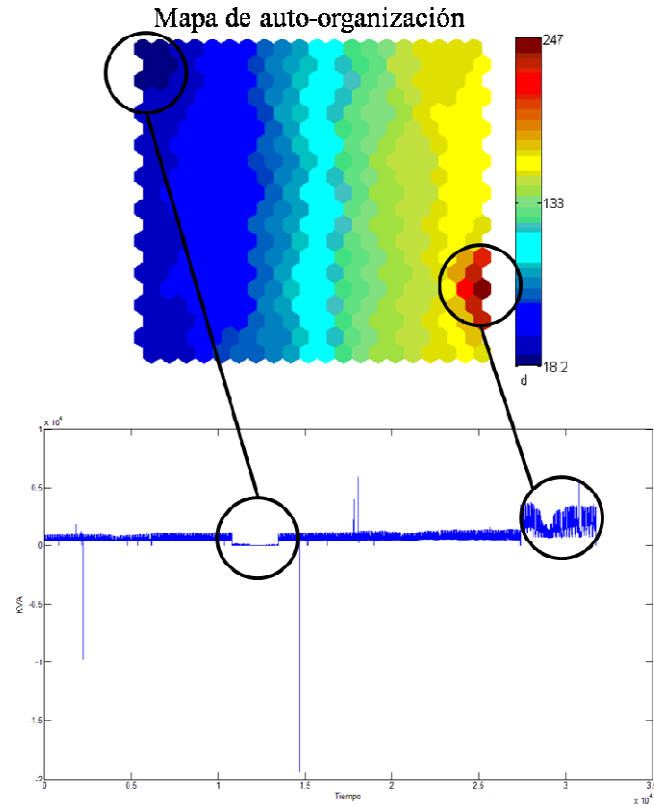


Figura A.4. Identificación de patrones con mapas de auto-organización

Esta hipótesis que se ha formulado aún no ha sido comprobada y podría realizarse profundizando en el tema de la extracción adecuada de la categorización de la información plasmada en el mapa. Tal extracción de información referida al agrupamiento realizado por este tipo de red neuronal de auto-organización aún no es tarea sencilla y actualmente se encuentra en desarrollo.