



**Modelación, Pronósticos y Back-testing de la Demanda Horaria MWH de la Energía Eléctrica en El Salvador, Empleando Redes Neuronales y los Software XL Miner® y Oracle Crystal Ball®**

**Escrito por:**

**Luis Francisco Zaldívar, MSE**

**Consultor en Finanzas, Minería de Datos y Riesgos**

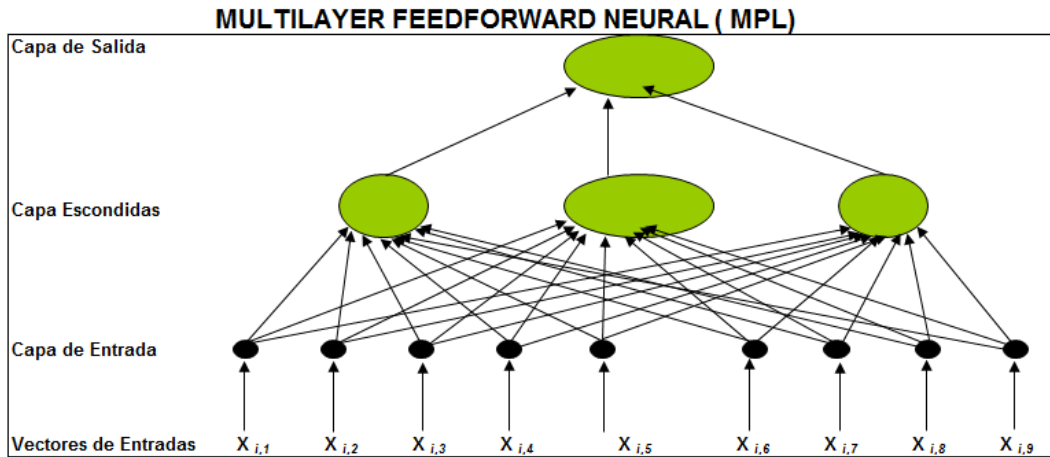
**[l.zaldivar@modelacionderiesgos.com](mailto:l.zaldivar@modelacionderiesgos.com)**

**[www.modelacionderiesgos.com](http://www.modelacionderiesgos.com)**

**Julio 2016**

# I. Introducción

Una de las técnicas más importantes en el campo de predicciones numéricas o clasificaciones (Paga/No Paga) son las Redes Neuronales. Las Redes Neuronales han tomado un crecimiento importante dado que hoy en día los software corren en ambientes de MS Excel® y computadoras personales. El tipo de Red Neuronal más popular en los negocios, es el que se conoce como MPL, Multilayer Perceptron Network. La arquitectura MPL se presenta a continuación:



La arquitectura de la Red Neuronal que se presenta en el cuadro anterior es de 3 capas escondidas. Por consiguiente, las características de las Redes Neuronales son 3: Capa de Entrada, Capas Escondidas (# capas y # Nodos por capa) y la Capa de Salida.

En la actualidad el software tiene módulos que determinan la mejor arquitectura de la Red Neuronal. En el software XL Miner®, la arquitectura de MPL que podemos emplear depende de definir:

- 1) ¿Cuántas capas escondidas requiere nuestra arquitectura? Se tiene la posibilidad de hasta 4 capas.
- 2) ¿Cuántos nodos se requieren por cada capa seleccionada en el literal #1? Se tiene la posibilidad de hasta 25 nodos por capa.

Architecture

# Hidden Layers (max 4):	4
# Nodes Layer 1:	25
# Nodes Layer 2:	25
# Nodes Layer 3:	25
# Nodes Layer 4:	25

La selección de la arquitectura MPL depende de la determinación del número de Capas y # de Nodos por capa que proporcionen el menor error de pronóstico en la validación y prueba. Es importante manifestarles que los mejores modelos no son los más complejos y de muchas variables, sino el más sencillo y que tiene las variables independientes necesarias. También es importante usar el número de # epochs o corridas necesarias para lograr que la red neuronal aprenda de la información.

Es importante saber qué es un buen modelo; un buen modelo debe tener las siguientes características:

- 1) Que sirva para tomar decisiones
- 2) Que lo entienda el personal involucrado en las decisiones
- 3) Saber qué mide
- 4) No tener ningún error de cálculo y de lógica
- 5) Que pueda medir su eficiencia en el back-testing

**Las fortalezas de las Redes Neuronales son las siguientes:**

- 1) Maneja funciones complejas tales como no lineales
- 2) Tiene alta capacidad para descifrar y aprender patrones de la información disponible
- 3) Existen softwares disponibles en el mercado que se emplean a nivel de MS Excel y computadoras personales, los cuales requieren una baja inversión
- 4) No existe una gran lista de requerimientos para aplicar la técnica, lo importante es lograr un error de pronóstico idóneo u óptimo
- 5) Puede emplear trabajos de predicción numérica y clasificaciones (Paga/No Paga)

**Las limitaciones de la Redes Neuronales son las siguientes:**

- 1) Necesita amplia base de datos
- 2) Algunas veces la solución que nos brinda es local y no global
- 3) Depende de la calidad de las entradas (GIGO)
- 4) Complejo de entender el proceso operativo
- 5) Es intensivo en cálculos y puede tomar tiempo el cálculo

Las Redes Neuronales están en la lista de las técnicas flexibles. Un buen estratega analítico, siempre debe de emplear varias técnicas para seleccionar la mejor: Error mínimo de pronóstico y alta calidad o cumplimiento de los supuestos del algoritmo. .

## II. Modelo 1

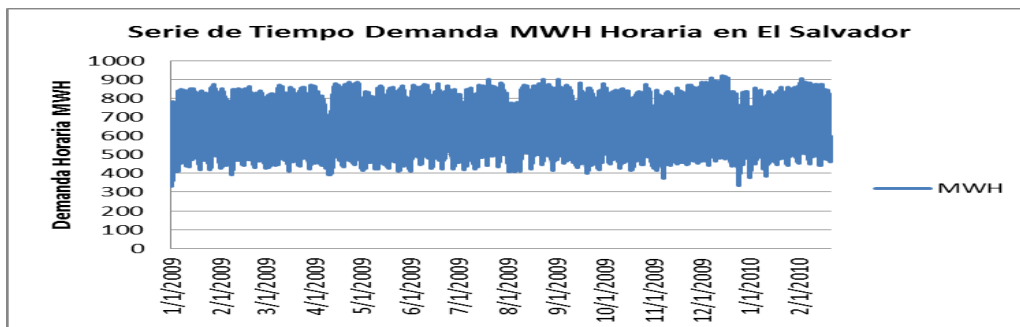
La serie de tiempo comprende la demanda de energía horaria MWH en El Salvador, iniciando desde el 1 de enero de 2009 y se mide en MWH.

Les presentamos una vista parcial de la serie de tiempo que estaremos empleando en el Modelo1 a continuación:

DEMANDA HORARIO DE MWH EN EL SALVADOR, CENTRO AMERICA

FECHA	HORA	Observacion	MWH	Tiempo
1/1/2009	0	1	538.21886	1
1/1/2009	1	2	466.42092	2
1/1/2009	2	3	410.7028	3
1/1/2009	3	4	378.95449	4
1/1/2009	4	5	367.23912	5
1/1/2009	5	6	367.19522	6
1/1/2009	6	7	331.65406	7
1/1/2009	7	8	339.24121	8
1/1/2009	8	9	371.02921	9
1/1/2009	9	10	401.0895	10
1/1/2009	10	11	423.74458	11
1/1/2009	11	12	434.6461	12
1/1/2009	12	13	446.21243	13
1/1/2009	13	14	447.12965	14
1/1/2009	14	15	444.0663	15
1/1/2009	15	16	438.49791	16
1/1/2009	16	17	435.03084	17
1/1/2009	17	18	479.41424	18
1/1/2009	18	19	654.12263	19
1/1/2009	19	20	671.547	20
1/1/2009	20	21	622.86897	21
1/1/2009	21	22	536.22512	22
1/1/2009	22	23	451.34975	23
1/1/2009	23	24	403.47272	24
1/2/2009	0	25	380.13137	25

La gráfica completa de la serie de tiempo es la siguiente:



El modelo de pronósticos es el siguiente:

$$D_{ti} = F(T_{ti}, D_{Lag\ ti-1}, D_{Lag\ ti-2}, D_{Lag\ ti-3}, D_{ti-24})$$

Dónde:

D<sub>ti</sub>= Demanda de MWH en el día t y en la hora i

T<sub>ti</sub>= Tiempo de la observación en el día t en la hora i

D Lag ti-x= La demanda anterior retrasada un número de horas x respecto al día t en la hora i

De una manera más práctica el modelo es el siguiente:

$$MWH = F(T, MWH-1, MWH-2, MWH-3, \dots, MWH-24)$$

La demanda de la primera hora del día de mañana, depende de su referencia correlativa en el tiempo y de las 24 demandas horarias de energía del día anterior. Es decir, el modelo de pronósticos tiene 25 variables independientes. Lo que está haciendo la variable T es modelar la tendencia y las 24 variables independientes de demanda horaria retrasada es la estacionalidad.

En el siguiente cuadro, pueden ver la creación básica del modelo a emplear con la técnica de Redes Neuronales:

FECHA	HORA	Observacion	MWH	Tiempo	MWH-1	MWH-2	MWH-3	MWH-4	MWH-5	MWH-6	MWH-7	MWH-8	MWH-9
1/1/2009	0	1	538.218864	1									
1/1/2009	1	2	466.420924	2	538.218864								
1/1/2009	2	3	410.702797	3	466.420924	538.218864							
1/1/2009	3	4	378.954489	4	410.702797	466.420924	538.218864						
1/1/2009	4	5	367.239123	5	378.954489	410.702797	466.420924	538.218864					
1/1/2009	5	6	367.19522	6	367.239123	378.954489	410.702797	466.420924	538.218864				
1/1/2009	6	7	331.654056	7	367.19522	367.239123	378.954489	410.702797	466.420924	538.218864			
1/1/2009	7	8	339.241206	8	331.654056	367.19522	367.239123	378.954489	410.702797	466.420924	538.218864		
1/1/2009	8	9	371.029207	9	339.241206	331.654056	367.19522	367.239123	378.954489	410.702797	466.420924	538.218864	
1/1/2009	9	10	401.089498	10	371.029207	339.241206	331.654056	367.19522	367.239123	378.954489	410.702797	466.420924	538.218864
1/1/2009	10	11	423.744582	11	401.089498	371.029207	339.241206	331.654056	367.19522	367.239123	378.954489	410.702797	466.420924
1/1/2009	11	12	434.646105	12	423.744582	401.089498	371.029207	339.241206	331.654056	367.19522	367.239123	378.954489	410.702797
1/1/2009	12	13	446.212429	13	434.646105	423.744582	401.089498	371.029207	339.241206	331.654056	367.19522	367.239123	378.954489
1/1/2009	13	14	447.129652	14	446.212429	434.646105	423.744582	401.089498	371.029207	339.241206	331.654056	367.19522	367.239123
1/1/2009	14	15	444.066303	15	447.129652	446.212429	434.646105	423.744582	401.089498	371.029207	339.241206	331.654056	367.19522
1/1/2009	15	16	438.497914	16	444.066303	447.129652	446.212429	434.646105	423.744582	401.089498	371.029207	339.241206	331.654056
1/1/2009	16	17	435.030843	17	438.497914	444.066303	447.129652	446.212429	434.646105	423.744582	401.089498	371.029207	339.241206
1/1/2009	17	18	479.414244	18	435.030843	438.497914	444.066303	447.129652	446.212429	434.646105	423.744582	401.089498	371.029207
1/1/2009	18	19	654.122628	19	479.414244	435.030843	438.497914	444.066303	447.129652	446.212429	434.646105	423.744582	401.089498

En el cuadro anterior pueden observar los retrasos que experimenta la demanda de energía horaria de MWH durante el 1 de enero del 2009. Por ejemplo, si nos ubicamos en la hora 14 del día 1 de enero del 2009 y necesitamos conocer la demanda retrasada tres horas respecto a esa hora MWH-3, el resultado sería 434.646105.

Es importante manifestar que en este análisis, se perderán las primeras 24 horas de demanda. El análisis se iniciará cuando tengamos la demanda horaria con sus 24 demandas anteriores y el tiempo correlativo.

La información inicial para correr la técnica de redes neuronales se presenta a continuación:

DEMANDA HORARIO DE MWH EN EL SALVADOR, CENTRO AMERICA

FECHA	HORA	Observacion	MWH	Tiempo	MWH-1	MWH-2	MWH-3	MWH-4	MWH-5	MWH-6	MWH-7	MWH-8	MWH-9	MWH-10	MWH-11	MWH-12	MWH-13	MWH-14	MWH-15	MWH-16	MWH-17
1/2/2009	0	25	380.13137	25	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965	446.21243	434.6461	423.74458	401.0895	371.02921	339.24121
1/2/2009	1	26	367.36662	26	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965	446.21243	434.6461	423.74458	401.0895	371.02921
1/2/2009	2	27	361.42271	27	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965	446.21243	434.6461	423.74458	401.0895
1/2/2009	3	28	360.3483	28	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965	446.21243	434.6461	423.74458
1/2/2009	4	29	374.51297	29	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965	446.21243	434.6461
1/2/2009	5	30	402.07374	30	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965	446.21243
1/2/2009	6	31	406.64365	31	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663	447.12965
1/2/2009	7	32	457.29498	32	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791	444.0663
1/2/2009	8	33	539.75758	33	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084	438.49791
1/2/2009	9	34	594.26229	34	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424	435.03084
1/2/2009	10	35	624.58083	35	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263	479.41424
1/2/2009	11	36	644.52951	36	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547	654.12263
1/2/2009	12	37	638.70401	37	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897	671.547
1/2/2009	13	38	650.99363	38	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512	622.86897
1/2/2009	14	39	660.5543	39	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975	536.22512
1/2/2009	15	40	655.11301	40	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272	451.34975
1/2/2009	16	41	633.48476	41	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137	403.47272
1/2/2009	17	42	634.45299	42	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662	380.13137
1/2/2009	18	43	784.87337	43	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271	367.36662
1/2/2009	19	44	781.95195	44	784.87337	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483	361.42271
1/2/2009	20	45	727.61346	45	781.95195	784.87337	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297	360.3483
1/2/2009	21	46	633.93357	46	727.61346	781.95195	784.87337	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374	374.51297
1/2/2009	22	47	639.9357	47	633.93357	727.61346	781.95195	784.87337	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365	402.07374
1/2/2009	23	48	469.51961	48	639.9357	633.93357	727.61346	781.95195	784.87337	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498	406.64365
1/3/2009	0	49	469.51961	49	469.51961	531.16087	633.93357	727.61346	781.95195	784.87337	634.45299	633.48476	655.11301	660.5543	650.99363	638.70401	644.52951	624.58083	594.26229	539.75758	457.29498

Se puede apreciar la fila que tiene el color amarillo completa. La base de datos que estamos empleando es del 2 de enero del 2009 al 21 de febrero del 2010. Después de revisar que no existen celdas vacías se debe proceder a realizar una partición de la serie de tiempo. El total de horas que hemos empleado en el análisis es de 9,892 horas de demanda de energía diaria. La partición de la información se ha hecho de la siguiente manera: 5,947 horas correlativas para construir el modelo y 3,945 horas para probar el modelo, la mezcla es de aproximadamente del 60%-40%.

Se ha empleado el software XL Miner® para realizar el pronóstico de demanda de energía horaria. Los resultados de la partición de la serie de tiempo 60%-40%, se presentan en el siguiente cuadro:

XLMiner: Time Series Data Partition Sheet

Date: 30-Sep-2015 10:34:13

Output Navigator						Elapsed Times in Milliseconds		
Summary	Time Variable	Partition Vars	Training Data	Validation Data	All Data	Read Time	Report Time	Total
						220	241	461

Data		Selected Variables																			
Workbook	Libro3	Time Variable	FECHA	HORA	Observacion	MWH	Hora0	Hora1	Hora2	Hora3	Hora4	Hora5	Hora6	Hora7	Hora8	Hora9	Hora10	Hora11	Hora12	Hora13	Hora14
Worksheet	Datos 1 (2)	Range	\$B\$3:\$AD\$9915																		
Time Variable	Tiempo	Partitioning Method	Sequential																		
Selected Variables	FECHA	HORA	Observacion	MWH	Hora0	Hora1	Hora2	Hora3	Hora4	Hora5	Hora6	Hora7	Hora8	Hora9	Hora10	Hora11	Hora12	Hora13	Hora14	Hora15	Hora16
Training Rows	5947																				
Validation Rows	3965																				

Luego, entramos al software XL Miner® y seleccionamos el menú de proyección y luego hacemos un click en Neural Network / Automatic Neural Network.

Parameters/Options	
Input variables normalized	Yes
Network Architecture	Manual
Seed: Initial Weights	12345
# Hidden Layers	2
# Nodes in Hidden Layer 1	4
# Nodes in Hidden Layer 2	25
# of Epochs	30
Step size for gradient descent	0.1
Weight change momentum	0.6
Error tolerance	0.01
Weight decay	0
Hidden layer activation function	Standard
Output layer activation function	Standard

Arriba les presentamos los parámetros del modelo de Redes Neuronales. El # de Epochs es equivalente al número de simulaciones. La arquitectura de las Redes Neuronales básica es la siguiente:

1. Dos capas
2. La primera capa # 1 tiene 4 nodos
3. La segunda capa # 2 tiene 25 nodos

Hoy es momento de analizar lo siguiente:

1. Error de Pronóstico % MAPE
2. Residuales

Los resultados son los siguientes:

#### Training Data Scoring - Summary Report

Total sum of squared errors	Average Error
6.43E+06	-2.60867

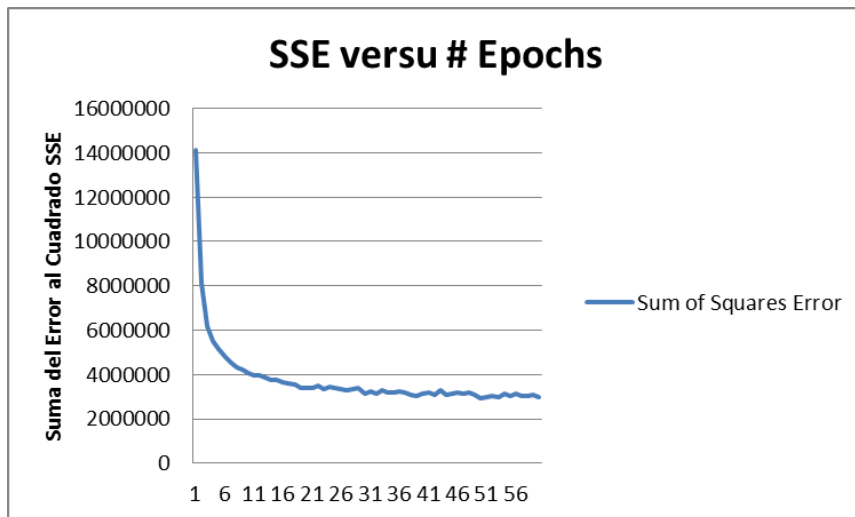
#### Validation Data Scoring - Summary Report

Total sum of squared errors	Average Error
4.97E+06	-5.517484

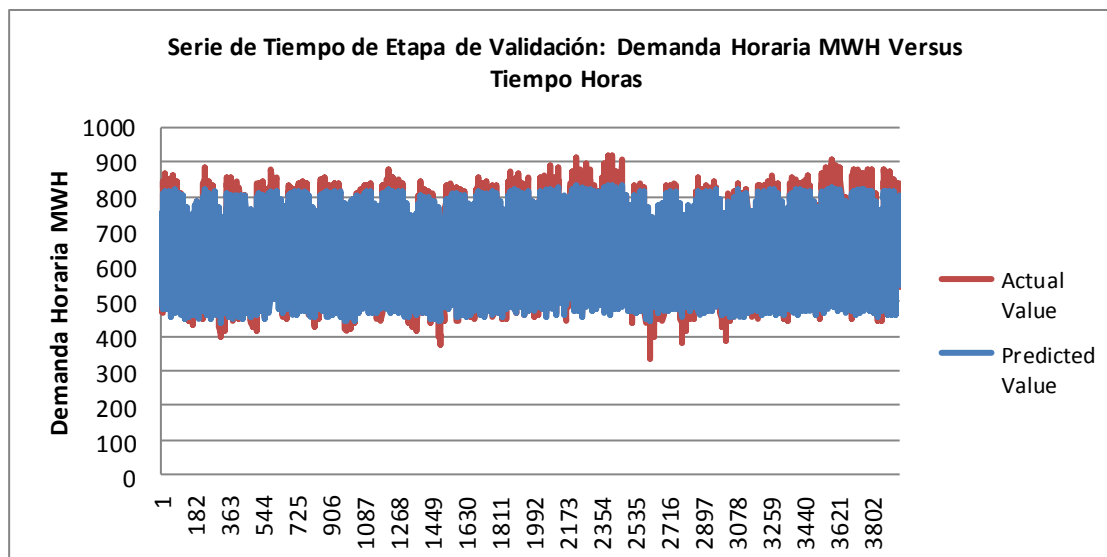
MAPE Modelo	MAPE Validación
4.40%	4.62%

Se puede apreciar que el error de pronóstico del modelo y del back-testing son bien similares. Empleando el MAPE Global, podemos ver que él % de error es aproximadamente igual.

Luego, es importante verificar que el número de corridas en la Red Neuronal sea el adecuado. A continuación, se presenta el gráfico de SSE versus Epochs que demuestra que los pasos a la red están normales.



En el siguiente gráfico, podemos comparar los valores de demanda de energía horaria reales versus los pronósticos que genera las Redes Neuronales:



En la gráfica anterior, podemos observar que el pronóstico de las Redes Neuronales subestima la demanda horaria en la zona de alta demanda y sobreestima la demanda de energía horaria en la zona baja demanda. Este factor sucede dado que las Redes Neuronales no han capturado una parte del componente estacional. Para eliminar esta situación, a la demanda horaria se le debe eliminar la estación, de esa manera el error de pronóstico MAPE, sería aún menor que el nivel determinado. Es decir, a los datos a emplear se debe aplicar la siguiente fórmula para eliminar la estacionalidad en los datos:

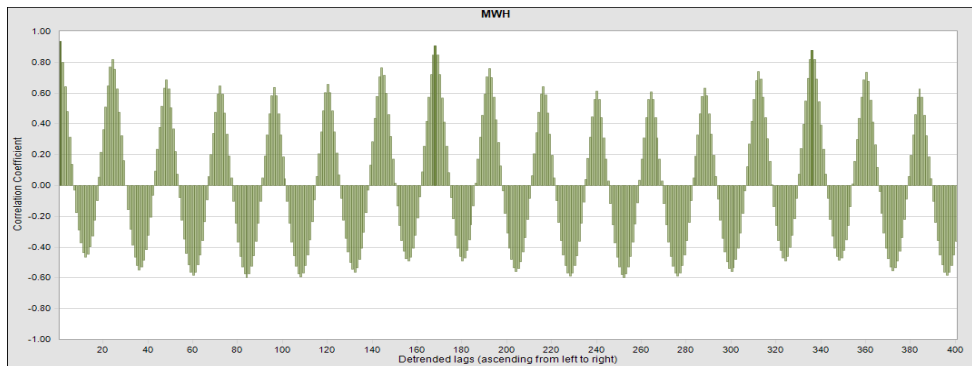


## Yi sin Estación = Yi Datos Originales / Índice estacional i

Existen varias técnicas que se pueden emplear para calcular los índices de estación. Lo importante es conocer que la estacionalidad de la energía eléctrica tiene 3 estaciones: La semana laboral, día sábado y domingo. Si empleamos los índices estacionales tradicionales, debemos de pensar en el cálculo de 168 índices.

Otra alternativa de modelar estacionalidad se presentará en la sección del Modelo 2 empleando variables ficticias, tontas o binarias (0.1)

La estacionalidad, la podemos observar en el siguiente correlograma ACF.



Una mejora al modelo sería emplear el retraso de una semana completa. Es decir, estaríamos empleando 168 variables retrasadas (MWH-1 a MWH-168) más la variable correlativo de tiempo, en total el modelo tuviera no 25 variables independientes sino que 169. Al hacer esta modelación será muy probable que el error de pronóstico MAPE sea mínimo.

En el presente trabajo, no hemos desestacionalizado la demanda de energía horaria dado que los resultados del error de pronóstico MAPE son bajos: 4.64%.

Como hemos expresado en la introducción de este trabajo, para emplear las Redes Neuronales, lo que tenemos que obtener es un back-testing aceptable para nuestras estimaciones. Esta técnica no tiene muchos requerimientos pero depende del volumen de información que se tenga y de los ciclos económicos que se tengan en ella.

A continuación, es el informe de salida de XL Miner® donde usa el Modelo encontrado y lo aplica a la base de datos de validación:

Output Navigator			
Inputs	Weights	Train. Score - Summary	Valid. Score - Summary
Training Log	Train. Score - Detailed Re	Valid. Score - Detailed Re	

Elapsed Times in Milliseconds			
Data Read Time	NNP Time	Report Time	Total
174	648	150	972

Workbook	Libro3
Worksheet	Data_PartitionTS
Range	\$B\$5968:\$AD\$9932

Predicted Value	Actual Value	Residual	Tiempo	Hora0	Hora1	Hora2	Hora3	Hora4	Hora5	Hora6	Hora7	Hora8	Hora9	Hora10
752.690726	738.233424	-14.4573016	1.73219643	0.35138471	-0.62709148	-0.64915596	-0.58351134	-0.57064334	-0.51286377	-0.495958948	-0.51713125	-0.59031792	-0.70526588	-0.83067026
742.765775	685.324725	-57.4410508	1.73277887	0.72617738	0.35146655	-0.6270656	-0.64925165	-0.58366032	-0.57075992	-0.512701197	-0.49571151	-0.5169038	-0.59011885	-0.70509197
665.305791	605.834056	-59.4717349	1.73336132	0.337921	0.72621169	0.35148007	-0.62716068	-0.64940212	-0.58377718	-0.570591845	-0.51245097	-0.49548733	-0.5167143	-0.58995747
591.76624	546.323043	-45.4431964	1.73394377	-0.24540008	0.33800454	0.72622049	0.35141182	-0.62731065	-0.64952036	-0.583607863	-0.570332	-0.51222424	-0.49530063	-0.51656091
541.64669	509.370509	-32.2761814	1.73452621	-0.68210578	-0.24524258	0.33801823	0.72616251	0.35128402	-0.62742843	-0.649344803	-0.58334585	-0.5700964	-0.51203535	-0.49514956
526.063111	486.303156	-39.7599558	1.73510866	-0.9532721	-0.6818929	-0.24522152	0.33794961	0.7260432	0.35118691	-0.627254968	-0.64907186	-0.58310826	-0.56989996	-0.51188246
495.045144	472.187499	-22.8576449	1.73569111	-1.12254572	-0.95302484	-0.68186634	-0.24530613	0.33782151	0.72595401	0.351267379	-0.6269857	-0.64882421	-0.58291013	-0.56974078

Parte del modelo de Redes Neuronales que encuentra el Software XL MIner®, lo presentamos a continuación:

**Inter-Layer Connections Weights**

Hidden Layer 1	Hora0	Hora1	Hora2	Hora3	Hora4	Hora5	Hora6	Hora7	Hora8	Hora9
Neuron 1	0.27743841	-0.49076732	0.30811341	0.31649138	0.29824626	-0.21560502	0.34088738	-0.137488443	0.43316363	-0.30200439
Neuron 2	1.37440594	0.61077423	0.18563167	-0.65260978	-0.27238046	-0.01301203	-0.13610309	-0.208485163	0.16441512	0.28495823
Neuron 3	1.27029907	-0.14019499	-0.26657471	-0.25228043	-0.15633793	0.33039619	-0.04525054	0.333459451	-0.42533375	0.0801684
Neuron 4	1.39181195	-0.71215538	-0.11498735	0.44660381	0.24263565	0.18325888	-0.53519625	-0.046413355	-0.17743782	0.45203008

Hidden Layer 2	Neuron 2	Neuron 3	Neuron 4	Bias
Neuron 1	0.2984645	-0.02016	0.1272455	0.1202138
Neuron 2	-0.001978	-0.340913	-0.332078	0.1374294
Neuron 3	0.1636278	0.2574438	0.3376981	-0.214843
Neuron 4	-0.341998	-0.522229	-0.72732	0.0039629
Neuron 5	0.0485963	-0.092886	0.1037207	0.1305004
Neuron 6	0.3025992	0.0200133	0.3677374	-0.271649
Neuron 7	0.1025605	-0.0326	-0.053378	0.008471
Neuron 8	-0.226809	-0.526979	-0.586995	0.1045311
Neuron 9	0.0343416	0.2978382	0.4060823	-0.146699
Neuron 10	-0.37165	-0.258022	-0.71811	0.196891
Neuron 11	-0.013396	-0.153982	-0.078307	0.0491379
Neuron 12	0.4437728	0.2412401	0.6906235	-0.304039
Neuron 13	-0.311347	-0.471554	-0.359268	0.2077095
Neuron 14	-0.324794	-0.361468	-0.517701	-0.073627
Neuron 15	-0.160043	-0.007104	-0.168334	-0.04171
Neuron 16	-0.470566	-0.445198	-0.683859	0.1587961
Neuron 17	-0.227758	-0.55627	-0.530831	0.0976694
Neuron 18	0.1202755	0.1437651	-0.129734	-0.15642
Neuron 19	0.0311001	-0.110856	-0.256991	0.0633033
Neuron 20	0.3720806	0.1908267	0.3186749	-0.224195
Neuron 21	0.436627	0.4448784	0.3613342	-0.301307
Neuron 22	0.057572	0.005738	0.1932752	-0.191206
Neuron 23	0.3690828	0.460925	0.7979616	-0.216703
Neuron 24	0.1552929	0.4798475	0.4659473	-0.141985
Neuron 25	-0.398325	-0.498894	-0.877193	0.0594978

<b>Output Layer</b>	<b>Neuron 2</b>	<b>Neuron 3</b>	<b>Neuron 4</b>	<b>Neuron 5</b>	<b>Neuron 6</b>	<b>Neuron 7</b>	<b>Neuron 8</b>	<b>Neuron 9</b>	<b>N</b>
<b>Response</b>	-0.392035	0.5998466	-1.221864	0.163109	0.4602494	0.0538385	-1.044296	0.6089049	.

Todos los pesos que se tienen en el modelo son los parámetros de las Redes Neuronales. Se tiene parámetros para la capa escondida1, parámetros para la capa escondida 2 y la capa de salida, la meta es poder emplear este modelo para realizar pronósticos futuros de la demanda horaria de MWH.

Supongamos que tenemos que estimar la demanda horaria MWH del siguiente cuadro:

### Cual es la demanda de energía en la HORA 13 & al 22 de febrero?

<b>Ultimo Histórico</b>	2/21/2010	12	9997	598.74	9997					
	<b>FECHA</b>	<b>HORA</b>	<b>Observación</b>	<b>MWH</b>	<b>Tiempo</b>	<b>MWH-1</b>	<b>MWH-2</b>	<b>MWH-3</b>	<b>MWH-4</b>	<b>MWH-5</b>
<b>Proyección</b>	2/22/2010	13	9998	???????	9998	598.74	592.23	581.91	559.43	533.76

Empleando el software XL Miner®, utilizamos el modelo que se ha encontrado y presentado en la página numero #7.

Los resultados son los siguientes:

#### Neural Network Prediction: Score Data

Date: 20-Jul-2016 17:12:46

<b>Model Workbook</b>	Pron D Redes Neuronales Lag R.xlsx
<b>Model Worksheet</b>	NNP_Stored
<b>Data Workbook</b>	Pron D Redes Neuronales Lag R.xlsx
<b>Data Worksheet</b>	Sheet4
<b>Data Range</b>	\$F\$5:\$AE\$6

<b>Predicted</b>	<b>Tiempo</b>	<b>Hora0</b>	<b>Hora1</b>	<b>Hora2</b>	<b>Hora3</b>	<b>Hora4</b>	<b>Hora5</b>	<b>Hora6</b>	<b>Hora7</b>	<b>Hora8</b>	<b>Hora9</b>	<b>Hora10</b>	<b>Hora11</b>	<b>Hora12</b>
<b>656.035</b>	9998	598.7397	592.2299	581.9112	559.4299	533.7638	494.9242	465.7394	524.6742	467.9429	462.1752	487.6436	484.6554	496.0763

La estimación de la demanda de energía MWH sería del 656.035, como se muestra en el cuadro anterior.

Qué pasaría si deseamos realizar estimaciones por un número de periodos más amplio? Pensemos que deseamos hacer la estimación de la hora 13 hasta la 24.

Como podemos apreciar en el siguiente cuadro, no tenemos la diagonal debajo de la demanda de MWH. Necesitamos tener completa la columna de MWH-1 para poder tener todas las casillas llenas y aplicar el modelo de Redes Neuronales.

## MATRIZ # 1 DE PRONOSTICO MWH DE LA HORA 13-23

DEMANDA HORARIO DE MWH EN EL SALVADOR, CENTRO AMERICA

FECHA	HORA	Observacion	MWH	Tiempo	MWH-1	MWH-2	MWH-3	MWH-4	MWH-5	MWH-6	MWH-7	MWH-8	MWH-9	MWH-10	MWH-11
2/22/2010	13	9998	9998	9998	598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94	462.18	487.64
2/23/2010	14	9999	9999	9999		598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94	462.18
2/24/2010	15	10000	10000	10000			598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94
2/25/2010	16	10001	10001	10001				598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67
2/26/2010	17	10002	10002	10002					598.74	592.23	581.91	559.43	533.76	494.92	465.74
2/27/2010	18	10003	10003	10003						598.74	592.23	581.91	559.43	533.76	494.92
2/28/2010	19	10004	10004	10004							598.74	592.23	581.91	559.43	533.76
3/1/2010	20	10005	10005	10005								598.74	592.23	581.91	559.43
3/2/2010	21	10006	10006	10006									598.74	592.23	581.91
3/3/2010	22	10007	10007	10007										598.74	592.23
3/4/2010	23	10008	10008	10008											598.74

Lo que haremos es emplear el módulo CB Predictor del software Oracle Crystal Ball para poder así encontrar los precios de MWH-1 de la hora 14 a la 23. Teniendo esos precios se puede completar la matriz y aplicar la técnica de Redes Neuronales empleando el software XL Miner®

Los resultados del pronóstico de CB Predictor los detallamos a continuación:

Period	Forecast	
	Lower: P20	Upper: P80
9998		600.02
9999		598.29
10000		592.84
10001		580.25
10002		581.98
10003		722.36
10004		769.77
10005		729.20
10006		639.82
10007		576.72

El proceso que sigue es llenar la diagonal de la **Matriz #1** que presenta las celdas vacías con los datos de MWH que aparecen en amarillo.

La matriz #1 completa y lista para emplear las Redes Neuronales se presenta a continuación:

FECHA	HORA	Observacion	MWH	Tiempo	Hora0	Hora1	Hora2	Hora3	Hora4	Hora5	Hora6	Hora7	Hora8	Hora9	Hora10	Hora11
2/22/2010	13	9998	9998	9998	598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94	462.18	487.64	484.66
2/23/2010	14	9999	9999	9999	660.02	598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94	462.18	487.64
2/24/2010	15	10000	10000	10000	598.29	660.02	598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94	462.18
2/25/2010	16	10001	10001	10001	592.84	598.29	660.02	598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67	467.94
2/26/2010	17	10002	10002	10002	580.25	592.84	598.29	660.02	598.74	592.23	581.91	559.43	533.76	494.92	465.74	524.67
2/27/2010	18	10003	10003	10003	581.98	580.25	592.84	598.29	660.02	598.74	592.23	581.91	559.43	533.76	494.92	465.74
2/28/2010	19	10004	10004	10004	722.36	581.98	580.25	592.84	598.29	660.02	598.74	592.23	581.91	559.43	533.76	494.92
3/1/2010	20	10005	10005	10005	769.77	722.36	581.98	580.25	592.84	598.29	660.02	598.74	592.23	581.91	559.43	533.76
3/2/2010	21	10006	10006	10006	729.20	769.77	722.36	581.98	580.25	592.84	598.29	660.02	598.74	592.23	581.91	559.43
3/3/2010	22	10007	10007	10007	639.82	729.20	769.77	722.36	581.98	580.25	592.84	598.29	660.02	598.74	592.23	581.91
3/4/2010	23	10008	10008	10008	576.00	639.82	729.20	769.77	722.36	581.98	580.25	592.84	598.29	660.02	598.74	592.23

Empleando la Matriz #1 completa, se puede estimar la demanda MWH de la hora 13-23 que aparece en color celeste en la matriz.

El pronóstico de MWH empleando las Redes Neuronales se presenta a continuación:

Model Workbook	Pron D Redes Neuronales Lag R.xlsx
Model Worksheet	NNP_Stored
Data Workbook	Pron D Redes Neuronales Lag R.xlsx
Data Worksheet	Pronostico Final
Data Range	\$B\$2:\$AD\$13

FECHA	HORA	Predicted	Tiempo	Hora0	Hora1	Hora2	Hora3	Hora4
2/22/2010	13	656.0354	9998	598.74	592.23	581.91	559.43	533.76
2/23/2010	14	698.2654	9999	660.02	598.74	592.23	581.91	559.43
2/24/2010	15	623.3359	10000	598.29	660.02	598.74	592.23	581.91
2/25/2010	16	604.7579	10001	592.84	598.29	660.02	598.74	592.23
2/26/2010	17	619.4894	10002	580.25	592.84	598.29	660.02	598.74
2/27/2010	18	684.716	10003	581.98	580.25	592.84	598.29	660.02
2/28/2010	19	785.1389	10004	722.36	581.98	580.25	592.84	598.29
3/1/2010	20	735.3639	10005	769.77	722.36	581.98	580.25	592.84
3/2/2010	21	696.4949	10006	729.20	769.77	722.36	581.98	580.25
3/3/2010	22	591.3144	10007	639.82	729.20	769.77	722.36	581.98
3/4/2010	23	581.2573	10008	576.00	639.82	729.20	769.77	722.36

Otra alternativa, para poder proyectar bastantes periodos es hacer uso de las variables ficticias, tontas o binarias. Esta metodología, se presentara en el Modelo 2, el cual se desarrolla a continuación.

### III. Modelo 2

En el Modelo 2 estaremos empleando la demanda de energía horaria que se demanda durante el fin de semana. Es decir, que estaríamos empleando la hora 121 a 168, que corresponde a las del fin de semana. Es importante recordar que las semana laboral es de la hora 1 a la 120.

A continuación, presentamos parte del cuadro de la serie de tiempo de la hora 121 en adelante.

El Salvador

#### DEMANDA DE ENERGIA HORARIA MWH

FECHA	HORA	Estación Ciclo	MWH
6/1/2014	22	167	563.45
6/1/2014	23	168	540.73
6/7/2014	0	121	540.14
6/7/2014	1	122	541.18
6/7/2014	2	123	545.87
6/7/2014	3	124	536.93
6/7/2014	4	125	549.65
6/7/2014	5	126	543.26
6/7/2014	6	127	572.77
6/7/2014	7	128	646.33
6/7/2014	8	129	718.02
6/7/2014	9	130	767.45
6/7/2014	10	131	789.70
6/7/2014	11	132	789.44
6/7/2014	12	133	754.42
6/7/2014	13	134	743.10
6/7/2014	14	135	737.03

En el Modelo 2 estaremos empleando variables ficticias o tontas en lugar de retrasar hasta 24 horas la demanda de energía horaria. Las variables ficticias son dicotómicas y solamente pueden tomar valores de 1 y 0. Cuando aparece el 1, indica que esa demanda pertenece a la hora que se define en la columna.

A continuación se les presenta una parte de la matriz que alimenta el modelo de pronósticos:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	El Salvador												
2	DEMANDA DE ENERGIA HORARIA MWH												
3	FECHA	HORA	Estación Ciclo	MWH	T	H121	H122	H123	H124	H125	H126	H127	
4	5/31/2014	1	122	607.46	1	0	1	0	0	0	0	0	0
5	5/31/2014	2	123	572.41	2	0	0	1	0	0	0	0	0
6	5/31/2014	3	124	572.05	3	0	0	0	1	0	0	0	0
7	5/31/2014	4	125	574.03	4	0	0	0	0	1	0	0	0
8	5/31/2014	5	126	581.23	5	0	0	0	0	0	1	0	0
9	5/31/2014	6	127	611.67	6	0	0	0	0	0	0	1	0
10	5/31/2014	7	128	685.55	7	0	0	0	0	0	0	0	0
11	5/31/2014	8	129	788.78	8	0	0	0	0	0	0	0	0
12	5/31/2014	9	130	828.01	9	0	0	0	0	0	0	0	0
13	5/31/2014	10	131	837.09	10	0	0	0	0	0	0	0	0
14	5/31/2014	11	132	829.50	11	0	0	0	0	0	0	0	0
15	5/31/2014	12	133	812.07	12	0	0	0	0	0	0	0	0
16	5/31/2014	13	134	798.09	13	0	0	0	0	0	0	0	0

El número 1 que aparece tres filas hacia abajo en la columna H124, significa que la demanda de 572.05 pertenece a la Hora 124. Las variables ficticias se hacen empleando condicionales. En este caso el condicional sería “= SI (C6 = 124,1,0)”.

El modelo tendrá un total de 49 variables independientes, las cuales serían las siguientes:

- 1) Una variable T que representa el tiempo correlativo y la tendencia.
- 2) Luego 48 variables ficticias o tontas que representan las 48 horas del fin de semana. Las horas del fin de semana inician de la 121 y termina en la hora 168. Estas variables representan la estacionalidad que se presenta en la demanda horaria de energía durante el fin de semana.

El modelo 2 no requiere que se haga estimaciones con el CB Predictor de Oracle Crystal Ball, dado que lo que hay que llenar en la matriz de pronóstico que requiere el sistema de XL Miner® son las variables tontas y el tiempo correlativo de la variable independiente T.

Si el modelo de pronóstico conocemos que depende de otras variables adicionales a la Tendencia T y Estacionalidad, se deberá de realizar estimaciones de estas variables independientes para poder incluirlas en la matriz de pronósticos. Así estarían todos los campos completos.

A continuación les presentamos la vista parcial de la matriz de variables de entrada, que son necesarias para poder hacer unos pronósticos con las Redes Neuronales:

**Pronósticos de Demanda de Energía MWH 48 Horas Fin de Semana  
En El Salvador**

			2543							
HORA	Estación Ciclo	MWH	T	H121	H122	H123	H124	H125	H126	H127
0	121		2544	1	0	0	0	0	0	0
1	122		2545	0	1	0	0	0	0	0
2	123		2546	0	0	1	0	0	0	0
3	124		2547	0	0	0	1	0	0	0
4	125		2548	0	0	0	0	1	0	0
5	126		2549	0	0	0	0	0	1	0
6	127		2550	0	0	0	0	0	0	1
7	128		2551	0	0	0	0	0	0	0
8	129		2552	0	0	0	0	0	0	0
9	130		2553	0	0	0	0	0	0	0
10	131		2554	0	0	0	0	0	0	0
11	132		2555	0	0	0	0	0	0	0
12	133		2556	0	0	0	0	0	0	0
13	134		2557	0	0	0	0	0	0	0
14	135		2558	0	0	0	0	0	0	0

Se ha empleado una serie de tiempo del 31 de mayo de 2014 al 3 de abril de 2015, se han tomado solamente las 48 horas de cada fin de semana. Es decir, se tienen 2,543 horas de fin de semana. El modelo de redes neuronales se ha construido con 2,034 horas y se ha realizado la prueba o back-testing con 509 horas. Es decir, la partición de la serie de tiempo es de 80% y 20% para construir el modelo y probarlo respectivamente.

**Training Data Scoring - Summary Report**

Total sum of squared errors	RMS Error	Average Error
3541715.223	41.72836	-3.5762

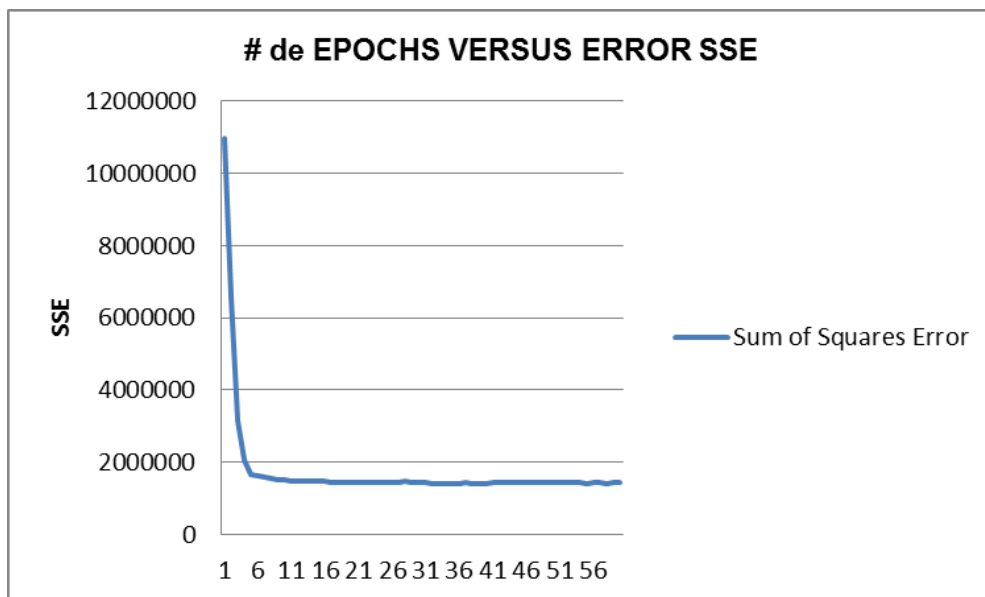
**Validation Data Scoring - Summary Report**

Total sum of squared errors	RMS Error	Average Error
923341.2479	42.59143	-2.53342

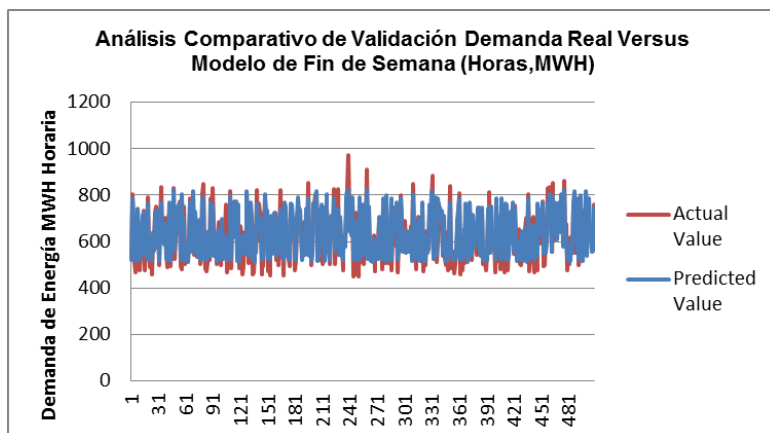
La medición con el Error del Pronósticos como porcentaje de la variación absoluta de la Demanda Estimada versus la Real, se presenta a continuación:

MAPE Modelo	MAPE Validación
5.45%	5.47%

El número de pasadas a las Redes Neuronales también ha sido adecuado, lo cual se puede apreciar en la siguiente gráfica. En esta gráfica, se ve que el error de pronóstico medido por SSE se ha disminuido.



El análisis gráfico de la Demanda Estimada del Modelo versus la Demanda Real, se presenta a continuación:



En ambos modelos que estamos presentando en este trabajo, no existe mucha exactitud en los saltos o spikes extremos que es muy típico en los precios y demanda de energía.

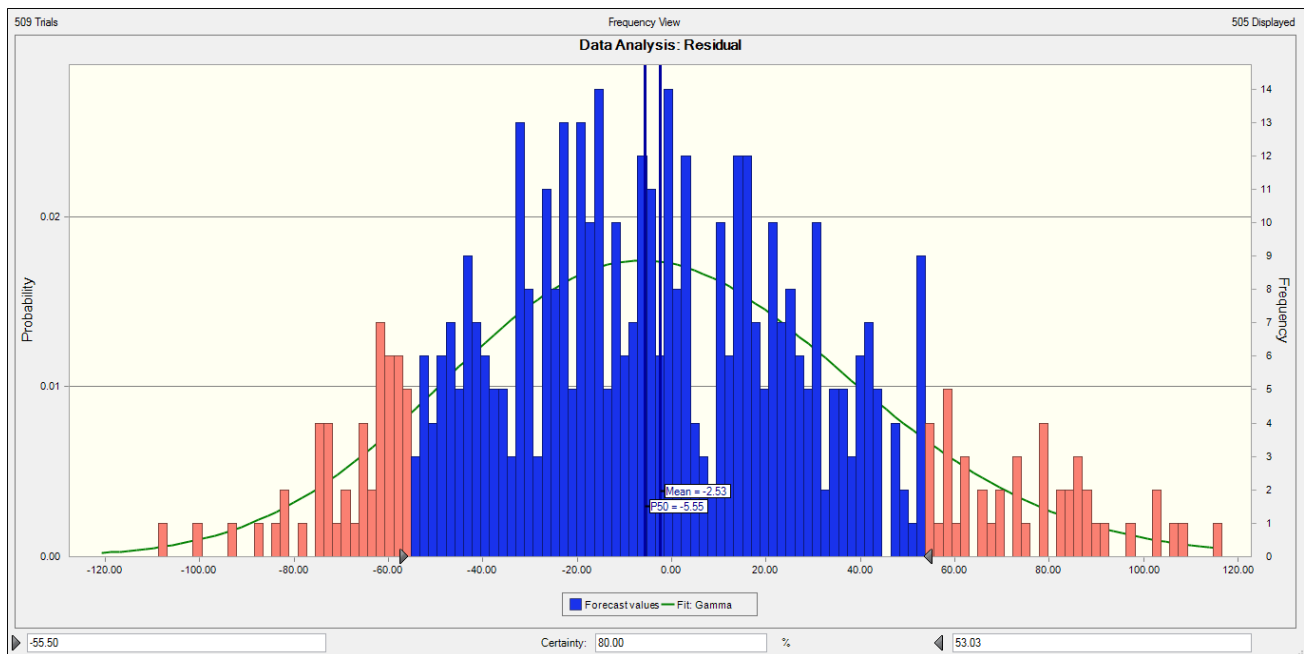
A continuación, les presentamos la estadística descriptiva de los residuales del modelo de Redes Neuronales:



Statistic	Forecast values
Trials	509
Base Case	---
Mean	-2.53
Median	-5.51
Mode	---
Standard Deviation	42.56
Variance	1,811.17
Skewness	0.3060
Kurtosis	3.75
Coeff. of Variation	-16.80
Minimum	-186.02
Maximum	154.88
Mean Std. Error	1.89

La forma de la curva de probabilidad de los residuales del modelo demuestra que sus extremos no son extremos.

Esto se puede ver en la gráfica que presentamos a continuación:



La distribución de los residuales es casi centrada, pocos valores extremos y la caracterización de la distribución es tipo Gamma. Los extremos son más ocasionados por la estacionalidad en la demanda que por otros factores.

El emplear variables ficticias para modelar pronósticos de series de tiempo con redes neuronales es mucho más fácil que la operatividad que hay que hacer empleando dos softwares como lo demostrado en el modelo1.

El error de pronóstico es bajo (5.47%) y éste puede disminuirse al emplear otros factores que afectan la demanda de la energía, tales como la temperatura, nivel de lluvias, etc. La técnica de Redes Neuronales acepta que los modelos tengan el número de variables que sea oportuno.

Lo importante es que el back-testing del modelo a emplear sea del nivel de error que es aceptable en el sector y administración, por ejemplo, en el sector bancario que se maneja una diversidad de riesgos, usualmente el error máximo que se admite es del nivel del 0.5 al 1.0%.

Finalmente, debemos hacer el pronóstico de la demanda del siguiente fin de semana y los resultados se presentan a continuación:

### Neural Network Prediction: Score Data

Model Workbook	Prueba fin de semana.xlsx
Model Worksheet	NNP_Stored
Data Workbook	Prueba fin de semana.xlsx
Data Worksheet	Pronostico
Data Range	\$D\$5:\$BC\$53

HORA	Estación Ciclo	Predicted	T	H121	H122	H123	H124	H125
0	121	557.88	2544	1	0	0	0	0
1	122	549.70	2545	0	1	0	0	0
2	123	516.73	2546	0	0	1	0	0
3	124	525.86	2547	0	0	0	1	0
4	125	554.02	2548	0	0	0	0	1
5	126	540.45	2549	0	0	0	0	0
6	127	556.19	2550	0	0	0	0	0
7	128	602.95	2551	0	0	0	0	0
8	129	715.03	2552	0	0	0	0	0

## IV. Hoja de Vida

Graduado de Licenciatura en Administración de Empresas con concentración en Gerencia Industrial de The University of Tennessee, Knoxville, Tennessee.

Posee Maestría en Ciencias Económicas con concentración en Finanzas y Estadística Aplicada de North Carolina State University, Raleigh, North Carolina.

Entrenado por Oracle Crystal Ball en Denver en Crystal Ball Introductorio y Avanzado, Opciones Reales y Seis Sigma.

Presidente de Inversiones Carrousel, S.A. de C.V. fundada en 1992.

Facilitador y Consultor Certificado en Análisis de Riesgos empleando Crystal Ball por Crystal Ball Global Unit, Denver, Colorado.

Además, es Facilitador y Consultor Certificado usando el software de riesgo más avanzado ModelRisk® de Vose Software BVBA de Bélgica.

Ha dictado cátedras en Simulación Monte Carlo, Optimización, Valorización de Empresas, Finanzas Corporativas, Inversiones en Carteras ,Minería de Datos, Pronósticos, Economía Aplicada, Riesgos, Estrategias Financieras y M&A en los programas de maestría en Finanzas y Administración de Empresas de las siguientes universidades: Universidad José Simeón Cañas of El Salvador (UCA), Universidad de El Salvador, Universidad Rafael Landívar en Guatemala (URL), Universidad Autónoma Gabriel Rene Moreno (UAGRM BUSINESS SCHOOL) en Santa Cruz, Bolivia (UARGM), Universidad Iberoamericana de Ciudad de México, Universidad de Costa Rica, Universidad Gavidia de El Salvador, Programa de Alta Dirección PAD de la Universidad de Piura en Lima, Perú, Universidad Centroamericana UCA en Nicaragua.

Dirigió Bancos Comerciales por 8 años y empresas de manufactura de exportación por 16 años en El Salvador.

Participó en Juntas Directivas en empresas manufactureras de exportación.

Conferencista en Simulación, Pronósticos y Optimización Estocástica para Oracle USA en América Latina.

Consultor especializado en Finanzas, Inversiones, Análisis de Riesgos y en Minería de Datos en América Latina.

Ha realizado entrenamientos en Análisis de Riesgo con Crystal Ball aplicado a las siguientes industrias durante los últimos nueve años: Petróleo, Gas, Minería, Banca Comercial y Central, Educación y en Empresas Corporativas durante los últimos siete años en los siguientes países: México, Guatemala, El Salvador, Nicaragua, Costa Rica, Panamá, República Dominicana, Colombia, Perú, Chile, Bolivia , Paraguay y Venezuela.