

Predicción de la Demanda Total de Energía Eléctrica Española Utilizando Mapas Auto-organizados como Modelo de Red Neuronal

Carolina Senabre¹, Sergio Valero¹, Antonio Gabaldón², Mario Ortiz¹ y Manuel González¹

¹ Departamento de Ingeniería de Sistemas Industriales
E.P.S.E., Universidad Miguel Hernández de Elche
Campus de Elche. Edificio Torreblanca. Avd. de la Universidad s/n, 03202 Elche (España)
Tel.:+34 96 6658969, fax:+34 96 665 89 79, e-mail: svalero@umh.es, cseembre@umh.es

² Departamento de Ingeniería Eléctrica
Escuela Técnica Superior de Ingeniería Industrial. Universidad Politécnica de Cartagena.
e-mail: antonio.gabaldon@upct.es

Resumen. La investigación está centrada en la utilización de los denominados Mapas Auto-organizados (Self-Organizing Map - SOM) como modelo de red neuronal para la predicción a corto plazo de la curva de demanda de energía eléctrica total española. Los datos utilizados han sido obtenidos del Operador del Sistema Eléctrico Español [1]. El objetivo principal es obtener la curva de demanda de algunas horas del día siguiente, día a predecir. Para validar el modelo se ha obtenido un índice de error el cual compara los resultados obtenidos con las curvas reales. El patrón de datos de entrada que se ha utilizado así como los diferentes parámetros de entrenamiento de la red han sido analizados y probados a lo largo de la investigación con el objetivo de obtener los mejores resultados. A modo resumen la investigación ha permitido establecer una herramienta de ayuda en la toma de decisiones a la hora de intentar predecir la curva de demanda de energía eléctrica global.

Palabras Clave

Predicción de la demanda eléctrica a corto plazo, Self Organizing Map (SOM), Mercado Eléctrico.

1. Introducción

Los países industrializados han experimentado un crecimiento importante de la demanda eléctrica total en la última década. A medida que crecen las necesidades de suministro eléctrico la complejidad del sistema crece y son muchos los factores que finalmente influyen en la generación y consumo de energía. En este escenario los procesos, técnicas y modelos de predicción requeridos son cada vez más complejos y necesitan de una mayor precisión. Las empresas suministradoras y las generadoras han de realizar estimaciones de la demanda prevista tanto en el corto plazo (en rangos desde unos pocos minutos u horas a varios días de antelación) como en el medio y largo plazo, para la programación de los recursos y la planificación de inversiones.

Sin embargo, la relación entre la demanda y los diferentes factores exógenos es compleja y no-lineal, lo cual hace difícil definir un modelo mediante la utilización de técnicas convencionales, como pueden ser las series temporales o análisis de regresión lineal.

Las técnicas de predicción de la demanda a corto plazo son herramientas muy útiles a la hora de tomar decisiones

con el objetivo de mantener el equilibrio del balance entre generación y consumo.

2. Caso de Estudio y Objetivos

Para llevar a cabo la investigación se han utilizado datos reales de la demanda de energía eléctrica total Española. Los mapas SOM fueron entrenados con datos históricos de curvas de demanda total con valores horarios, de los días laborables de los años 2002 y 2003, ver figura 1. Las curvas de demanda utilizadas fueron cedidas por el operador del mercado eléctrico Español, Red Eléctrica Española [1]. Para la fase de testeo y verificación de los resultados obtenidos por los mapas se utilizaron las curvas de demanda de diferentes días del 2004.

El objetivo fue aprovechar la capacidad de las redes SOM para clasificar datos históricos y posteriormente utilizar la ventaja de la memorización de dicha clasificación [2] para identificar y asociar las primeras horas de una curva de demanda de un día cualquiera del año 2004 entre las entrenadas y preclasificadas por la red de los años 2002 y 2003.

Las curvas de demanda utilizadas en los entrenamientos de la SOM (años 2002 y 2003) necesitan un pre-tratamiento para la obtención de buenos resultados. Se consideró aplicar un incremento a los valores de demanda de energía de los años 2002/03 para compensar el crecimiento previsto o estimado para el año 2004. Este paso fue necesario después de comparar diversos años consecutivos desde el 2001 hasta el 2003 y observar que existía un incremento total de la demanda de entorno a un 4,5-6% de un año para otro. Para que este incremento previsto en la demanda del 2004 no afectara a la hora de testear sobre un mapa entrenado con datos históricos del 2002 y 2003 decidimos aumentar un promedio del 5% los datos del 2002 y 2003.

Por último los vectores fueron normalizados respecto al máximo valor de la demanda de los años 2002 y 2003, tanto para los datos entrenados como para los datos del 2004 utilizados en el testeo. Esta normalización también fue necesaria para evitar problemas con el entrenamiento de la red de haber utilizado los elevados valores de demanda.

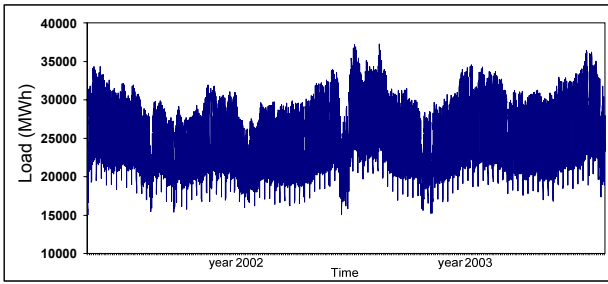


Figura 1. Curva de Demanda Eléctrica Total (años 2002 y 2003)

3. Mapas Auto-Organizados (SOM)

Un Mapa Auto-Organizado (SOM) es un tipo de red neuronal artificial cuyo entrenamiento se realiza mediante un algoritmo de aprendizaje no supervisado. Este tipo de aprendizaje se diferencia del Aprendizaje supervisado por el hecho de que no hay un conocimiento a priori de las posibles soluciones óptimas o el camino a seguir. Es decir, No existe ningún maestro externo que indique si la red neuronal está operando correcta o incorrectamente porque no se dispone de ninguna salida objetiva hacia la cual la red neuronal deba tender. La red auto-organizada debe descubrir rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e incorporarlos a su estructura interna de conexiones. Las neuronas deben auto-organizarse en función de los estímulos o datos de entrada.

El resultado es un espacio en forma de mapa bi-dimensional que contiene una representación de los datos de entrada clasificados y agrupados en las diferentes neuronas o celdas de salida del mapa mediante una etiqueta [3], ver Figura 2. El software utilizado para la realización de los entrenamientos, simulaciones y tests fue Matlab.

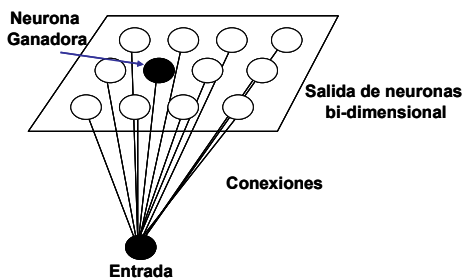


Figura 2. Estructura de un Mapa Auto-Organizado

4. Metodología

Una vez entendido el funcionamiento y las capacidades de la red neuronal SOM, los autores han seguido dos líneas de análisis diferentes y descritas en el punto A y B de este apartado.

Con el fin de expresar la exactitud de ambas metodologías, se ha definido un índice para medir el error. Este índice es el porcentaje de error medio absoluto

(Mean Absolute Percentage Error, MAPE), el cual mide la exactitud de una curva de demanda predicha comparada con la real [4].

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left[\frac{|L_{Ri} - L_{Pi}|}{L_{Ri}} \right] \times 100\%$$

Donde: N es el número de horas predichas, que variará según tomemos 8, 10 o 12 horas del día, L_{Ri} es el valor real de la demanda que se dio en la hora i y L_{Pi} es la demanda predicha para la hora i .

A. Primera Metodología

Este primer análisis se basa en el entrenamiento de la red SOM con datos de curvas de demanda eléctrica total diaria de los años 2002 y 2003. Después de los entrenamientos los mapas han sido testeados (bombardeados) con curvas de demanda de algunos días del 2004, pero únicamente las primeras horas del día a predecir. En este sentido se han realizado ensayos con las 8, 10 y 12 primeras horas del día a estimar.

El objetivo es conseguir que el mapa entrenado nos asigne que días de los clasificados y agrupados del 2002 y 2003 tienen más similitud y relación con las primeras horas de un día del 2004.

Los datos de entrada han sido etiquetados con la fecha de la curva de demanda en el formato *-mmdaa-*, mes, día y año. Tras presentar las primeras horas de un día del 2004 (8, 10 o 12 h) la red proporciona una celda ganadora que será la que más se aproxime a la curva de demanda a estimar. Estas neuronas o celdas ganadoras contienen las etiquetas de uno o de varios días del 2002 y 2003 permitiéndonos estimar la evolución de las siguientes horas hasta completar el día del 2004.

Otra de las ventajas de analizar las 10 y las 12 primeras horas del día a predecir, es que nos permite observar la posible evolución de la curva y si el mapa entrenado asigna la misma celda ganadora que nos daba desde un principio con las 8 primeras, o si asigna otras celdas con distintas fechas, que se aproximen más aún a la curva de demanda a estimar del 2004.

La Figura 3 muestra un ejemplo de mapa entrenado con las etiquetas asignadas (o agrupadas) a cada neurona.

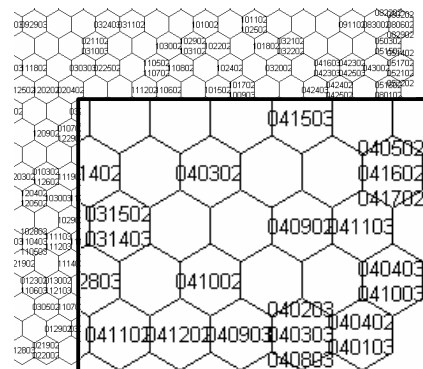


Figura 3. Mapa SOM entrenado con datos del 2002/03

La Figura 4 muestra un diagrama de la metodología seguida.

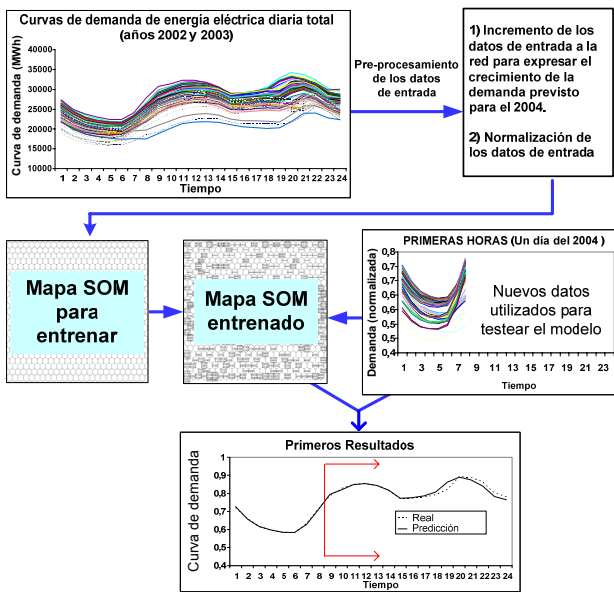


Figura 4. Diagrama de la primera metodología de la investigación

La Figura 5 muestra las 8 primeras horas del día 16 de Junio del 2004, día a testear, y las curvas de demanda de la celda (neurona) ganadora la cual contenía las etiquetas de los días 3 y 5 de Junio del 2003. La Figura 6 muestra la curva media de demanda estimada de los días de la neurona ganadora comparada con la curva real de la demanda del 16/04/2004. Para este día del 2004 los resultados obtenidos (neurona ganadora) al testear con las 8, 10 y 12 primeras horas fueron los mismos.

Los mejores resultados se obtuvieron con los parámetros de entrenamiento que muestra la Tabla I:

TABLA I.- Parámetros de Entrenamiento

Inicialización de los datos de entrada	Aleatoria
Algoritmo de entrenamiento	Secuencial
Tamaño de la red	25 X 25
Función de vecindad	Bubble
Número de iteraciones	5000

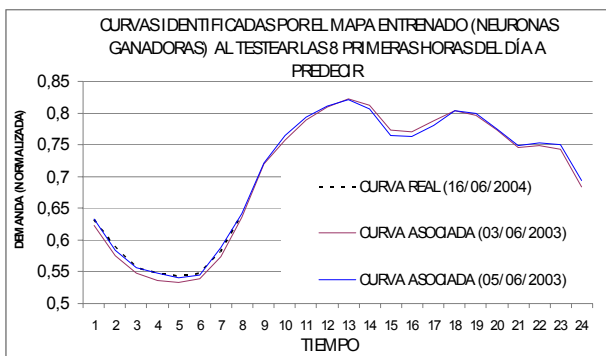


Figura 5. Resultados de las curvas de demanda identificadas en el mapa entrenado al testear las 8 primeras horas del 16 de Junio del 2004.

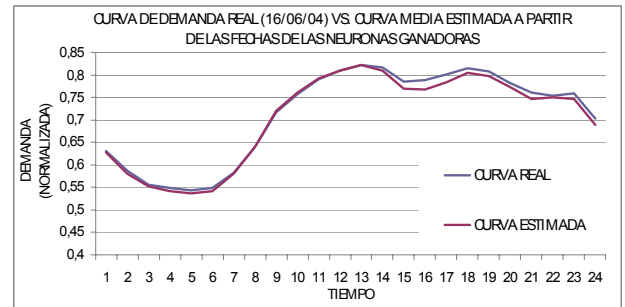


Figura 6. Curva de demanda real del 16 de Junio del 2004 vs. Curva estimada (predicha) resultado de hacer la media de las curvas de los días 3 y 5 de junio del 2003.

A continuación se muestran las Figuras 7, 8 & 9 las cuales representan la curva de demanda estimada comparada con la real del día 15 de julio del 2004, pero testeando las 8, 10 y 12 horas respectivamente. Para este día la evolución del índice del error MAPE ha ido mejorando con el aumento de las horas testeadas, obteniéndose un valor de 2,78 para 8 horas, un valor de 2,08 para 10 horas y un valor de 0,94 para 12 horas.

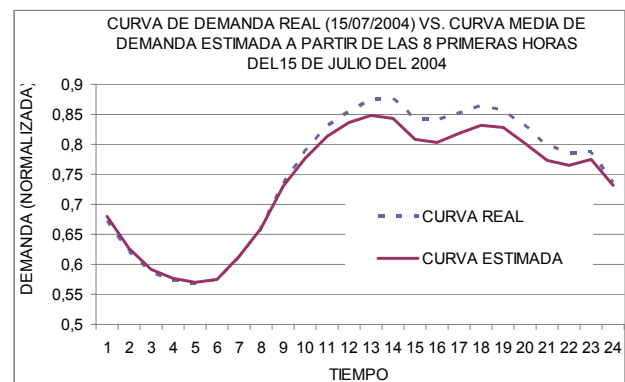


Figura 7. Curva real de demanda del 15/07/04 vs. Curva estimada testeando obtenida al testear 8 horas

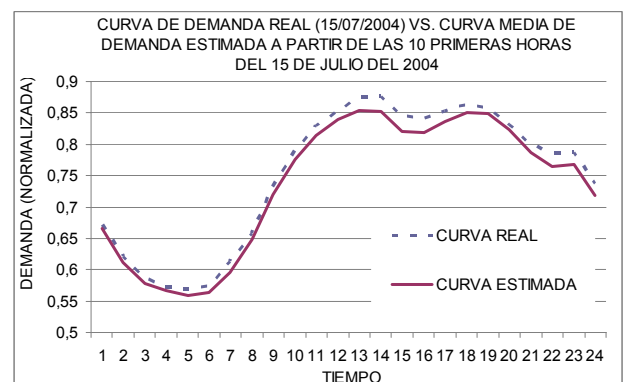


Figura 8. Curva real de demanda del 15/07/04 vs. Curva estimada testeando obtenida al testear 10 horas

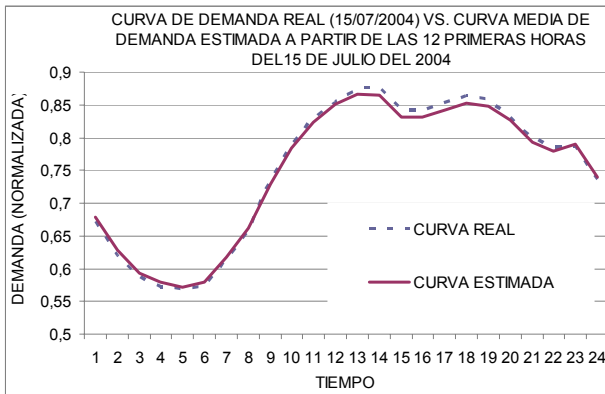


Figura 9. Curva real de demanda del 15/07/04 vs. Curva estimada testeando obtenida al testear 12 horas

La Tabla II muestra el índice de error obtenido con la primera metodología testeando las 8, 10 y 12 primeras horas de varios días del mes de Junio del 2004. El error M.A.P.E. medio mensual se sitúa en un valor de 2,27 cuando se testean solamente las 8 primeras horas del día a predecir. Sin embargo, el índice de error aumenta ligeramente a 2,34 y 2,36 cuando las horas testeadas son las 10 y las 12 primeras del día a predecir respectivamente. Existen otros meses en que los valores del índice de error mejoran al pasar de testear 8 horas a 10 y a 12 horas.

TABLA II.- Índice de Error MAPE para algunos días del mes de Junio del 2004. Primera Metodología

Días	MAPE			MAPE MEDIO DIARIO
	8 horas	10 horas	12 horas	
02/06/2004	2,44	3,88	4,26	3,52
03/06/2004	2,31	2,51	2,70	2,51
04/06/2004	0,91	1,02	1,14	1,02
07/06/2004	2,15	2,35	2,23	2,24
08/06/2004	4,09	2,44	2,61	3,05
09/06/2004	1,99	2,15	2,98	2,37
10/06/2004	0,67	0,58	0,56	0,60
11/06/2004	1,81	2,97	3,33	2,70
14/06/2004	1,25	4,39	4,98	3,54
15/06/2004	1,03	0,97	0,98	0,99
16/06/2004	1,17	1,27	1,45	1,30
17/06/2004	0,55	0,47	0,50	0,50
18/06/2004	2,84	6,44	2,92	4,06
21/06/2004	7,78	2,93	3,14	4,62
22/06/2004	1,96	2,20	2,40	2,19
23/06/2004	0,83	0,89	0,95	0,89
24/06/2004	1,20	1,30	1,29	1,26
25/06/2004	1,95	1,45	1,62	1,67
28/06/2004	4,74	3,83	4,07	4,21
29/06/2004	3,74	2,77	3,08	3,20

MAPE MEDIO	2,27	2,34	2,36	2,32
------------	------	------	------	------

La Tabla III muestra los resultados del índice de error para el mes de Julio del 2004.

TABLA III.- Índice de Error MAPE del mes de Julio del 2004. Primera Metodología.

Días	MAPE			MAPE MEDIO DIARIO
	8 horas	10 horas	12 horas	
01/07/2004	7,22	8,82	1,70	5,91
02/07/2004	3,33	3,12	2,10	2,85
05/07/2004	1,85	1,02	0,99	1,28
06/07/2004	3,67	4,17	4,65	4,16
07/07/2004	1,71	0,55	0,61	0,95
08/07/2004	4,39	4,93	2,14	3,82
09/07/2004	1,85	1,83	2,00	1,89
12/07/2004	1,86	2,04	2,13	2,01
13/07/2004	1,09	1,02	1,14	1,09
14/07/2004	1,18	1,09	1,03	1,10
15/07/2004	2,79	2,08	0,94	1,94
16/07/2004	4,97	1,98	1,94	2,96
19/07/2004	8,24	1,24	1,62	3,70
20/07/2004	1,55	1,62	1,13	1,43
21/07/2004	0,63	0,25	0,25	0,38
22/07/2004	1,19	1,22	1,37	1,26
23/07/2004	7,14	7,87	2,04	5,68
26/07/2004	2,70	1,89	2,11	2,24
27/07/2004	1,44	1,51	1,65	1,53
28/07/2004	6,80	1,48	1,72	3,34
29/07/2004	1,60	1,55	1,07	1,41
30/07/2004	1,16	1,05	0,91	1,04
MAPE MEDIO	3,11	2,38	1,60	2,36

B. Segunda Metodología

En esta segunda investigación los autores han cambiado el método, ahora son dos días consecutivos los que se utilizan como datos de entrada para entrenar el mapa SOM, 48 horas en lugar de las 24 horas de la metodología anterior.

Se han tomado como datos a entrenar los lunes y martes del año 2002 y 2003, y las etiquetas tendrán el formato -*ssaa*-, *semana* y *año*. El proceso de testeo de los datos debe de cambiarse de una manera similar. Presentaremos a la red entrenada las 24 horas de un lunes y las primeras 8, 10 y 12 horas del martes del año 2004, que es el día del cual queremos saber el pronóstico.

La Figura 10 muestra un ejemplo de mapa entrenado para este segundo método.

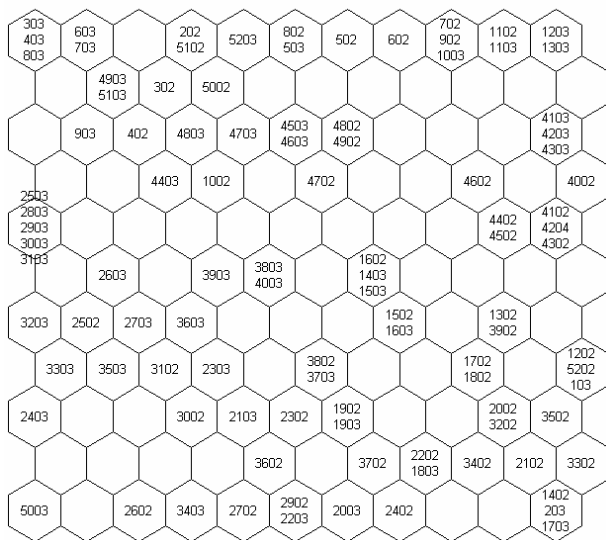


Figura 10. SOM entrenada con lunes y martes del 2002/03.

Las Figuras 11 y 12 muestran la comparación entre las curvas de demanda identificadas y la real de la semana n° 7 del año 2004.

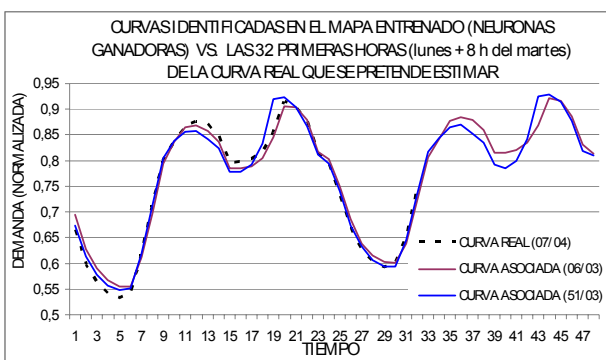


Figura 11. Curvas de demanda identificadas en el mapa entrenado vs. la real de la semana n° 7 del 2004.

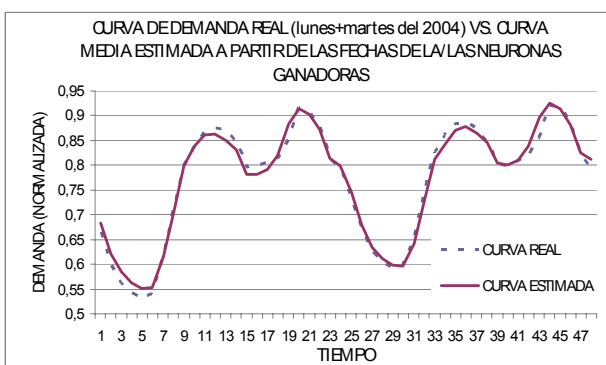


Figura 12. Curva de demanda real de la semana n° 7 del 2004 (lunes+martes) vs. curva media estimada.

La Tabla IV muestra el índice de error obtenido con la segunda metodología, es decir, testeando las 24 horas del lunes y las primeras 8, 10 y 12 horas del martes. Los índices MAPE de la tabla corresponden a las semanas 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12 del año 2004. El error M.A.P.E. medio de las semanas consideradas se sitúa en

un valor de 2,06 cuando se testean solamente las 24h. (lunes) + 8h. del martes a predecir. El índice de error alcanza un valor de 2,10 y 1,83 cuando las horas testeadas son las 10 y las 12 primeras del día a predecir respectivamente.

TABLA IV.- Índice de Error MAPE. Segunda Metodología

Semana	MAPE			MAPE MEDIO DIARIO
	8 horas	10 horas	12 horas	
3 del 2004	0,59	0,56	0,52	0,55
4 del 2004	1,09	0,92	0,87	0,96
5 del 2004	1,26	1,32	1,22	1,27
6 del 2004	1,97	2,21	2,51	2,23
7 del 2004	1,21	1,10	1,08	1,13
8 del 2004	1,24	1,24	1,23	1,24
9 del 2004	4,07	4,46	1,66	3,40
10 del 2004	3,59	3,62	3,54	3,58
11 del 2004	2,70	2,69	2,65	2,68
12 del 2004	2,93	2,96	3,08	2,99
MAPE MEDIO	2,07	2,11	1,84	2,00

5. Conclusiones

Los resultados obtenidos con esta metodología muestran un índice de error bastante bajo al comparar la curva de demanda real con la estimada a partir del histórico de datos de entrada de los años 2002 y 2003.

Han sido estudiadas muchas configuraciones distintas de parámetros de la red hasta obtener un modelo con errores entorno al 1,8-2 % para la mayoría de los días testeados del año 2004. No obstante, dentro de esta línea de investigación el objetivo es realizar más ensayos con distintas configuraciones de entrenamiento de datos históricos así como la incorporación a los mismos de factores externos climáticos como la temperatura para mejorar los índices de error.

Los mapas Auto-organizados (SOM) y las metodologías presentadas en este artículo pueden ser una herramienta muy útil tanto para compañías suministradoras como para el operador del sistema eléctrico, con el objetivo de intentar estimar a corto plazo, un día o unas horas determinadas, la demanda de energía eléctrica.

Agradecimientos

El trabajo de investigación descrito en este artículo ha sido financiado con la ayuda económica del Proyecto de Investigación con referencia UMH-BANCAJA-IPOG06 (2008-2009) dirigido por el Dr. Sergio Valero Verdú y

por el Proyecto de Investigación con referencia ENE2007-67771-C02-01&02/CON (Ministerio de Educación y Ciencia) dirigido por el Dr. Antonio Gabaldón Marín.

Referencias

- [1] REE, Red Eléctrica de España, www.ree.es
- [2] Fan S., Chen L., “Short-term load forecasting based on an adaptive hybrid method”, Power Systems, IEEE Transactions on Volume 21, Issue 1, Feb. 2006 Page(s):392 – 401
- [3] Kohonen, T.: ‘Self-organisation and associative memory’ (Springer-Verlag, Berlin, 1989, 3rd edn.)
- [4] Tafreshi S.M.M., Farhadi M., “Improved SOM based method for short term load forecast of Iran power network”, Power Engineering Conference, 2007. IPEC 2007. International 3-6 Dec. 2007 Page(s):1377 – 1384