

# APLICACIÓN DE REDES NEURONALES ARTIFICIALES AL CÁLCULO DE PREVISIONES A CORTO PLAZO EN EL MERCADO ELÉCTRICO ESPAÑOL

Pino Diez, R.  
De la Fuente García, D.  
Parreño Fernández, J.  
Priore, P.  
Universidad de Oviedo

## RESUMEN

En ocasiones, la "sobrecapacidad" de información se puede convertir en un problema incluso más grave que el no disponer de series temporales suficientemente largas. Desde el punto de vista de las RNAs, es importante disponer de un conjunto de ejemplos de entrenamiento suficientemente grande, pero si es desmesuradamente grande, es más que posible que, si no se utiliza un ordenador de gran capacidad y velocidad, el tiempo necesario para que la red converja y llegue a soluciones adecuadas, sea demasiado grande en términos relativos. En este trabajo proponemos método de entrenamiento (que denominamos "selectivo y continuo"), en el que se hace una selección previa de los ejemplos de entrenamiento del Perceptrón Multicapa que se utiliza para calcular las previsiones. Hemos comprobado la efectividad del método propuesto, pronosticando una serie temporal correspondiente al Mercado de la Electricidad Español.

**PALABRAS CLAVE:** Previsión a corto plazo, Redes Neuronales Artificiales, Perceptrón Multicapa, Redes ART.

## ABSTRACT

Sometimes, having time series that are too long can be an even greater problem, even worse than having series with too few data. From a Neural Networks point of view, it is important to have a set of training samples that is big enough; however, if this set is too big the time required to reach an adequate solution may be too long. In this paper, we propose a training method we have called a selective and continuous method, in which a previous selection for the Multilayer Perceptron (MLP) training samples is made using an ART-type neural network. The MLP is then trained and finally it is used to make forecasts. We tested the effectiveness of the proposed method, making forecasts for the time series called Daily-Market Hourly Price, part of the Electricity Production Market of Spain.

**KEYWORDS:** Short-Time Forecasting, Artificial Neural Networks, Multilayer Perceptron, ART Networks.

## 1. INTRODUCCIÓN.

Hasta los primeros años 20, las previsiones se calculaban simplemente extrapolando la serie en el tiempo. El principio de lo que podemos llamar "predicción moderna" de series temporales se puede fijar en el año 1927, cuando Yule inventó la técnica autorregresiva para pronosticar el número anual de manchas solares (Yule, 1927). Su modelo pronosticaba el

siguiente valor como una suma ponderada de las observaciones previas de la serie. Para poder obtener un comportamiento interesante de estos sistemas lineales, era necesario asumir la intervención de un factor externo (ruido), que afectaba al sistema lineal. Durante el medio siglo siguiente, se asumía que las series temporales eran generadas por sistemas lineales afectados por un ruido, y todas las investigaciones culminaron en la metodología ARIMA de Box-Jenkins (Box and Jenkins, 1970).

Sin embargo, existen casos simples para los que esta última metodología es poco adecuada. El hecho de que series temporales aparentemente complicadas puedan ser generadas por ecuaciones muy simples, hace necesario un marco teórico mucho más general para el análisis y predicción de series temporales. Así fueron surgiendo estudios en los que en los que se trataban series no estacionarias y/o no lineales, con nuevos métodos: modelos bilineales, biespectrales, de umbral, etc. (Tong, 1983, 1990; Priestley, 1988; Tsay, 1991; Subba Rao, 1992).

En los años 80, ocurrieron dos acontecimientos cruciales en la evolución de los estudios sobre series temporales. Por un lado, el incremento en la potencia de los ordenadores personales, permitió el estudio de series temporales mucho más largas, la aplicación de algoritmos más complejos, y la visualización interactiva tanto de los datos como de los resultados. El segundo hecho fue el desarrollo de las técnicas de aprendizaje automático y concretamente de las Redes Neuronales Artificiales.

Las Redes Neuronales Artificiales (RNAs) son modelos matemáticos inspirados en la organización y el funcionamiento de las neuronas biológicas. Existen numerosas variantes de RNAs que están relacionadas con la naturaleza de la tarea que se ha asignado. De la misma manera, también existen distintas variaciones sobre cómo modelar la neurona; en algunos casos se asemejan mucho a las neuronas biológicas mientras que en otros, los modelos son muy diferentes.

La literatura sugiere algunas características de las RNAs que las hacen especialmente interesantes en su aplicación a la previsión de series temporales. Fundamentalmente se señalan dos: la capacidad de las RN de aproximar prácticamente cualquier función (incluso las no lineales), y la posibilidad de hacer aproximaciones “*piece-wise*” o por trozos, de las funciones.

Desde el punto de vista matemático, las RNAs se pueden considerar como aproximadores universales de funciones (Hornik et al., 1989; Cybenko, 1989). Esto significa que pueden automáticamente aproximar la función que mejor se ajuste a los datos. Mientras que esta propiedad no tiene mucho valor si la relación funcional es simple (por ejemplo lineal), permite a las RNAs extraer relaciones cuando las funciones son muy complejas. Además las RNAs son intrínsecamente no lineales (Rumelhart and McClelland, 1986), lo cual implica no sólo que pueden estimar correctamente funciones no lineales, sino que también pueden extraer elementos no lineales de los datos, una vez extraídos los términos lineales.

Por otro lado, una RNA con una o más capas ocultas, puede dividir el espacio muestral automáticamente y construir diferentes funciones en diferentes porciones del espacio. Esto significa que las redes neuronales poseen la capacidad de construir modelos no lineales “*piece-wise*”. Collopy y Armstrong (1992), revisaron las opiniones de expertos en previsión, que

coinciden en afirmar la importancia de contar con modelos con estas características, es decir, capaces de identificar y tratar cambios abruptos en los patrones de la serie temporal.

Algunos métodos estadísticos de series temporales tienen limitaciones debidas a la forma en que los modelos son estimados, por esta razón, la estimación de muchos tipos de modelos de series temporales requieren la intervención y supervisión humana. Además, la mayoría de modelos estadísticos deben ser re-estimados periódicamente cuando se dispone de nuevos datos. Por el contrario, la estimación con RNAs puede ser automatizada (Hopffroff, 1993), y no es necesario revisar los modelos puesto que las redes aprenden de forma incremental.

La falta de información, o mejor dicho, la imposibilidad de trabajar con series temporales suficientemente largas, es uno de los principales problemas de cualquier método de previsión que intenta aprender del pasado de la serie temporal para proyectarla hacia el futuro. Pero, en ocasiones, la "sobrereabundancia" de información también se puede convertir en un problema. Desde el punto de vista de las RNAs, es importante disponer de un conjunto de ejemplos de entrenamiento suficientemente grande (para una red de tamaño intermedio, 15-8-1 por ejemplo, podrían ser suficientes alrededor de 200 ejemplos, o incluso menos), pero si el conjunto de ejemplos de entrenamiento es desmesuradamente grande (5.000 ó 6.000 datos, por ejemplo), es más que posible que, si no se utiliza un ordenador de gran capacidad y velocidad, el tiempo necesario para que la red converja y llegue a soluciones adecuadas, sea demasiado grande en términos relativos. Este problema, que aumentará proporcionalmente al tamaño de la red, puede ser especialmente grave cuando lo que se pretende es obtener resultados en un breve espacio de tiempo.

Para intentar resolver este problema, se tienen varias alternativas, de las que destacaremos dos: en primer lugar se puede intentar acelerar el proceso de entrenamiento utilizando algoritmos más rápidos. El algoritmo de entrenamiento más comúnmente utilizado es el "*Backpropagation With Momentum*" (Rumelhart and McClelland, 1986), sin embargo, dadas las debilidades del algoritmo (convergencia lenta, posibilidad de mínimos locales, etc.), se han diseñado diversas variantes del algoritmo o nuevos algoritmos que intentan acelerar el proceso o buscar mínimos locales mejores en caso de no poder llegar al mínimo global. Entre estos algoritmos podemos citar el "*Backpropagation With Weight Decay*" (Werbos, 1988), en el que durante el proceso de entrenamiento, se produce una disminución progresiva de los pesos de los enlaces de forma que tienden a anularse a no ser que sean reforzados durante el entrenamiento. El algoritmo "*Quickprop*" (Fahlman, 1988), intenta acelerar el proceso usando información sobre la curvatura de la superficie de error, calcula las derivadas de segundo orden de la función del error e intenta en un único paso ir al mínimo. El algoritmo "*RPROP*" (acrónimo de **R**esilient **b**ack**P**RO**P**agation, introducido por Riedmiller y Braun en 1993), calcula el cambio para cada uno de los pesos de forma separada en función de la topología de la superficie del error. Otros algoritmos y diferentes variaciones de los anteriores, se pueden encontrar en Parker (1987), Jacobs (1988), Tolleane (1990), Pack et al. (1991a,b); Battiti (1992), Fog et al. (1995) y Cottrell et al. (1995).

Otra posibilidad es la de hacer una selección previa de los ejemplos de entrenamiento, que es el método que utilizamos en este trabajo.

## 2. MÉTODO DE SELECCIÓN DE EJEMPLOS DE ENTRENAMIENTO.

Una de las características fundamentales de las RNAs es la velocidad de respuesta; así es posible obtener soluciones casi instantáneamente para ciertos tipos de redes, y en muy poco tiempo en la mayoría. Pero para que esto sea así, al menos en el perceptrón multicapa (que es la arquitectura que estamos utilizando aquí), deben darse ciertas condiciones:

1. La configuración de la red no debe ser excesivamente grande, de forma que el número de enlaces cuyo peso es necesario calcular no sea muy elevado.
2. El conjunto de ejemplos de entrenamiento no debe ser demasiado grande, cuanto más pequeño sea, mas veces por unidad de tiempo pasará cada ejemplo por la red y más rápido se llegará a la solución.
3. La composición del conjunto de ejemplos de entrenamiento deberá ser homogénea, es decir, cuanto más parecidos entre sí sean los ejemplos que la red utiliza para aprender, más rápido será el proceso de aprendizaje.

De acuerdo con esto, una forma de acelerar el proceso es hacer una selección previa de los patrones o ejemplos de entrenamiento, de forma que se escojan los estrictamente necesarios como para que todas las relaciones subyacentes en la serie temporal estén debidamente representadas (Plutowski and White, 1991; Deco et al., 1997). En este caso, el proceso de entrenamiento utilizando un conjunto de ejemplos pequeño y que además está compuesto por casos muy similares entre sí (sin llegar a ser exactamente iguales), será sensiblemente más corto, de tal manera que la RNA estará lista para calcular previsiones en un plazo de tiempo muy breve.

En los últimos años han aparecido trabajos en los que se incluyen algoritmos de selección de patrones incorporados a una metodología de previsión con redes neuronales (Peng et al, 1992; Ho et al., 1992; Vermaak, 1998). Estos algoritmos se basan en la idea del "entrenamiento selectivo y continuo", que se ha logrado implementar gracias al incremento de la potencia de las computadoras.

Si se desea calcular el valor previsto para un instante determinado, es conocido el vector de entrada a la red, que estará compuesto por los valores anteriores de todas las variables que se han incluido en el modelo. Entonces, se puede realizar una comparación entre el vector de entrada representativo de la previsión que deseamos calcular y los posibles patrones o ejemplos de entrenamiento disponibles en una base de datos que contiene todos los vectores representativos de situaciones pasadas en las que se conoce la solución (cada ejemplo está constituido por un vector de entrada y su salida correspondiente). De esta manera se puede seleccionar en cada caso los patrones con mayor parecido a la entrada (situaciones pasadas más similares a la que se desea pronosticar). Entrenando la red con los patrones elegidos, se obtiene una red neuronal lista para ser utilizada para calcular el valor previsto para ese instante. Este valor previsto puede ser guardado posteriormente como histórico virtual de cara a sucesivas previsiones. Todo el proceso aparece esquematizado en la figura 1.

Centrándonos en el proceso de selección de patrones, se puede medir la distancia entre el vector de entrada representativo de la previsión que queremos calcular, y los patrones almacenados en la base de ejemplos de entrenamiento, calculando la distancia euclídea entre ambos (Peng et al., 1992):

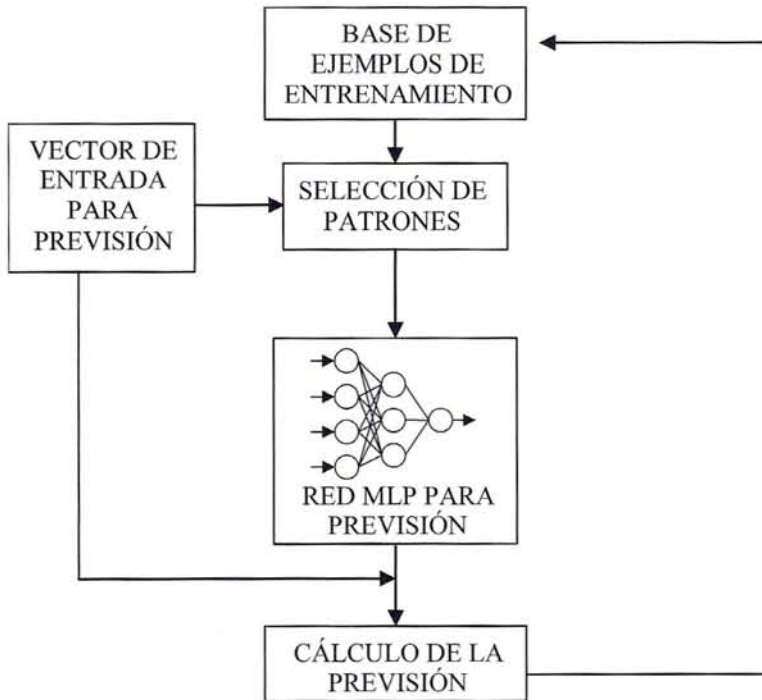
$$d(\hat{X}, X^{(k)}) = \sqrt{\sum_{i=1}^n (x_i^{(k)} - \hat{x}_i)^2}$$

(1)

donde  $\hat{X}$  es el vector de entrada y  $X^{(k)}$  el k-ésimo vector de todos los almacenados en la base de ejemplos de entrenamiento.

Medir el parecido de un vector de entrada con los patrones almacenados, mediante la fórmula (1), puede llevar a que en los casos en los que una de las componentes sea muy distinta y las demás componentes muy parecidas, el resultado es que el valor de la distancia es elevado, lo que implicaría que los dos vectores son demasiado distintos como para seleccionar el patrón, y sin embargo, la realidad es que sólo se diferencian en una componente.

Figura 1. Método Selectivo y Continuo de entrenamiento de Redes Neuronales.



Fuente: Elaboración Propia.

En nuestro caso, proponemos un procedimiento de selección de patrones que se basa en los fundamentos de las redes ART (Carpenter and Grossberg, 1987), las cuales son muy utilizadas en problemas de clasificación. Estas redes son capaces de clasificar vectores que se van introduciendo como entradas a la red, dependiendo de su parecido con los ya clasificados previamente. Para ello se efectúa una "prueba de similitud" en la que se cuantifica el parecido

de un vector con los que pertenecen a un grupo determinado; después, en función de un “coeficiente de similitud” fijado de antemano, se determina si el nuevo vector pertenece o no a esa categoría. El coeficiente varía entre 0 y 100; 0 significa que no se exige ningún parecido para que ese patrón pertenezca al grupo, y 100 implica que para seleccionar un patrón, éste debe ser prácticamente igual al vector representativo de la categoría.

Si tomamos el vector de entrada (con el que intentamos pronosticar el siguiente valor de la serie temporal), como el vector representativo de una categoría, y vamos introduciendo en la red ART todos los patrones de entrenamiento de la base de ejemplos, la red seleccionará todos los ejemplos cuyas entradas son similares a las del caso que queremos pronosticar, en función del coeficiente de similitud previamente fijado. Todos estos ejemplos son almacenados en un fichero de entrenamiento a partir del cual se obtendrán (por el método tradicional), los conjuntos de entrenamiento y validación de esa RNA.

Mediante este procedimiento, es posible entrenar la RNA con conjuntos de ejemplos tan pequeños como se desee (modificando convenientemente el coeficiente de similitud). Pero sin embargo, se asegura que el conjunto de entrenamiento contiene los ejemplos más representativos para el instante que se desea pronosticar.

### **3. CÁLCULO DE PREVISIONES EN EL MERCADO ELÉCTRICO ESPAÑOL.**

El Mercado de Producción de Energía Eléctrica Español es el conjunto de transacciones derivadas de la participación de ciertos agentes en las sesiones de los mercados diario e intradiario y de la aplicación de los Procedimientos de Operación Técnica del Sistema.

Los agentes del mercado son las empresas habilitadas para actuar en el mercado de producción como vendedores y compradores de electricidad. Pueden actuar como agentes, los productores, distribuidores y comercializadores de electricidad así como los consumidores cualificados de energía eléctrica, y las empresas o consumidores, residentes en otros países, que tengan la habilitación de agentes externos.

El Mercado Diario, como parte integrante del Mercado de Producción de Energía Eléctrica, tiene por objeto llevar a cabo las transacciones de energía eléctrica para el día siguiente al que se realizan los cálculos, mediante la presentación de ofertas de venta y adquisición de energía eléctrica por parte de los agentes autorizados. El Operador del mercado procede a la casación de las mismas, con objeto de determinar el precio marginal y de asignar la producción y demanda entre los agentes que han acudido a la subasta. El resultado de este proceso es el Programa Diario Base de Funcionamiento (PDB), en el que se fijan los precios y cantidades de energía que se comprarán y venderán en cada una de las 24 horas del día siguiente.

Vamos a utilizar las RNAs para pronosticar este PDB, es decir, se calculará el precio en Ptas/kWh para cada una de las 24 horas de un día determinado. Contamos para ello con los datos horarios del Precio del Mercado Diario (PH) y de la Demanda de Electricidad (DEM),

desde el 1 de Enero de 1.998 (fecha de inicio del funcionamiento del Mercado), hasta el día 30 de Noviembre de 1.999. En total son más de 16.000 valores para cada una de las dos series temporales.

Entrenar una RNA con un número de ejemplos de entrenamiento tan enorme supondría que el proceso se dilataría durante varias horas o incluso días si no se cuenta con un ordenador de gran capacidad de cálculo. Si el objetivo es el cálculo de previsiones a corto plazo, es necesario acelerar el proceso utilizando un método de selección de ejemplos como el que proponemos.

Para comprobar la calidad de los pronósticos obtenidos mediante la aplicación de la técnica de selección de ejemplos propuesta, se han calculado las previsiones del Precio del Mercado Diario (PH) para las 24 horas de 4 días (2 de ellos pertenecen a fin de semana), pertenecientes a 4 meses distintos. Por supuesto, para pronosticar cada uno de los días, hemos utilizado los valores del PH y de DEM desde el día 1 de Enero de 1998 (inicio del funcionamiento del Mercado) hasta las 24 horas del día anterior al que se pronostica. Los días seleccionados y los datos disponibles se muestran en la tabla 1.

**Tabla 1. Datos disponibles para cada día a pronosticar.**

<b>Día Pronosticado</b>	<b>Desde (H y D)</b>	<b>Hasta (H y D)</b>	<b>Nº de Datos</b>
16/07/1999 (Viernes)	01:00 01/01/1998	24:00 15/07/1999	13464
14/08/1999 (Sábado)	01:00 01/01/1998	24:00 13/08/1999	14160
09/09/1999 (Jueves)	01:00 01/01/1998	24:00 08/09/1999	14784
10/10/1999 (Domingo)	01:00 01/01/1998	24:00 09/10/1999	15528

Fuente: Elaboración Propia

En cuanto a la arquitectura de la red neuronal que se ha utilizado, después de algunas pruebas, se llegó a la configuración representada esquemáticamente en la figura 2. Como se puede comprobar, se trata de un perceptrón multicapa parcialmente separado; existe una red principal, 18-10-1, que está apoyada por otra red 6-3-1 que procesa la información correspondiente a la serie DEM. Además existen otras dos redes separadas de las anteriores y entre sí, la red 5-3-1, se encarga de procesar la información correspondiente a la hora del día para la que se realiza la previsión, y la segunda, 3-2-1, procesa la información correspondiente al día de la semana que se trate. La utilización de redes separadas y/o de apoyo suele ser beneficiosa en el caso de previsión multivariante, en la que se tienen varias series temporales de distinta naturaleza (como en el caso que nos ocupa); en estas redes, se divide la capa oculta en varias secciones de forma que cada una de ellas se ocupa del procesado de un tipo de información proveniente de la capa de entrada (Chen et al., 1992; Choi et al., 1997). En la tabla 2, se muestra la composición exacta de la capa de entrada de la red.

**Tabla 2.** Composición de la Capa de Entrada de la RNA.

Nodo	Descripción
1	PH(t-1)
2	PH(t-2)
3	PH(t-3)
4	PH(t-4)
5	PH(t-5)
6	PH(t-6)
7	PH(t-24)
8	PH(t-25)
9	PH(t-26)
10	PH(t-27)
11	PH(t-28)
12	PH(t-29)
13	DEM(t-1)
14	DEM(t-2)
15	DEM(t-3)
16	DEM(t-4)
17	DEM(t-5)
18	DEM(t-6)

Nodo	Descripción					
19	0	0	0	0	.....	1
20	0	0	0	0	.....	0
21	0	0	0	0	.....	1
22	0	0	1	1	.....	1
23	0	1	0	1	.....	1

1 2 3 4 ..... 24  
Hora del Día

24	0	0	0	0	1	1	1
25	0	0	1	1	0	0	1
26	0	1	0	1	0	1	0

L M X J V S D  
Día de la Semana

Fuente: Elaboración Propia.

En algunos casos se utilizan los enlaces directos ("*shortcuts*") entre alguna neurona o neuronas de la capa de entrada y las de salida (Duliba, 1991; Chen et al., 1992), con lo que se intenta reforzar el efecto de alguna variable, o se intenta modelar alguna relación lineal entre ciertas variables de entrada. En ciertos casos se demuestra que la utilización de enlaces directos mejora la capacidad de generalización y predicción de la red, aunque en otros casos sólo lleva a que aumente el tiempo necesario para el aprendizaje de la red debido al aumento del número de enlaces. En el caso que nos ocupa, tras varias pruebas se demostró la utilidad de incluir 2 enlaces directos desde la capa de entrada hasta la neurona de la capa de salida, concretamente desde las neuronas "1" y "13" que corresponden a los valores inmediatamente anteriores del PH y DEM respectivamente.



Los resultados de las previsiones calculadas mediante el método de selección de ejemplos de entrenamiento, para cada uno de los 4 días considerados se han representado en las figuras 3, 4, 5 y 6. En cada una de ellas se muestran los valores reales de la serie PH, junto con las previsiones calculadas con RNAs y empleando la metodología de Box-Jenkins.

Figura 3. Viernes, 16 de Julio de 1999.

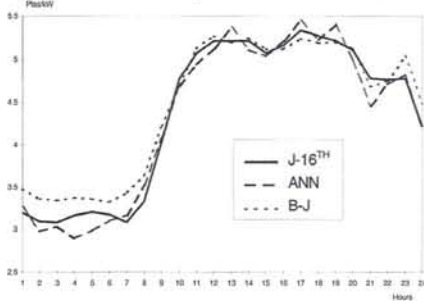


Figura 4. Sábado, 14 de Agosto de 1999.

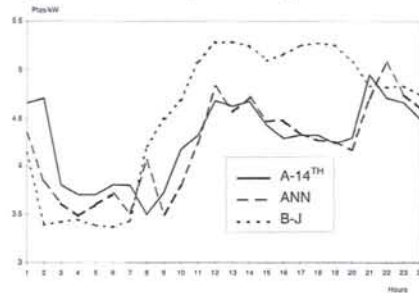


Figura 5. Jueves, 9 de Septiembre de 1999.

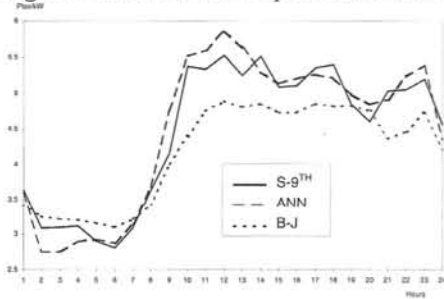
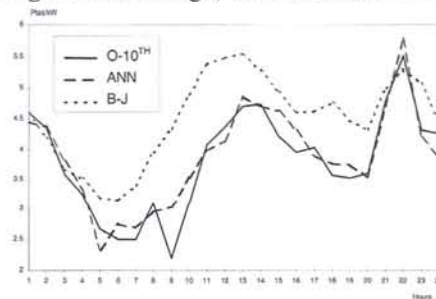


Figura 6. Domingo, 10 de Octubre de 1999.



Fuente: Elaboración Propia

Tabla 3. Medidas de la calidad de las previsiones.

	RNA	B-J		RNA	B-J
<b>MAD</b>	0.109607	0.138529		0.191094	0.374950
<b>MSE</b>	0.017548	0.030736		0.054247	0.197059
<b>MAPE</b>	2.72%	3.69%		4.42%	8.28%
<b>Viernes, 16 de Julio</b>					
<b>MAD</b>	0.202730	0.586462		0.225949	0.721833
<b>MSE</b>	0.078859	0.436600		0.079229	0.783983
<b>MAPE</b>	4.90%	13.31%		6.55%	18.35%
<b>Sábado, 14 de Agosto</b>					
<b>Domingo, 10 de Octubre</b>					

Fuente: Elaboración Propia.

En la tabla 3, se muestran los valores de varios estadísticos utilizados para comprobar la calidad de las previsiones calculadas:

- *Valor Medio del Error Absoluto* (MAD, “mean absolute deviation”):

$$MAD = \frac{\sum |e_t|}{N}$$

- *Error Cuadrático Medio* (MSE, “mean squared error”):

$$MSE = \frac{\sum (e_t)^2}{N}$$

- *Valor Medio Porcentual del Error Absoluto* (MAPE, “mean absolute percentage error”):

$$MAPE = \frac{1}{N} \sum \frac{|e_t|}{\frac{1}{2}(x_t + y_t)}$$

donde  $e_t$  es el error de previsión en el cálculo del valor correspondiente al instante  $t$ ;  $x_t$  e  $y_t$  son los valores real y pronosticado respectivamente; y  $N$  es el número de previsiones calculadas.

#### 4. CONCLUSIONES.

En este trabajo hemos calculado previsiones de una serie temporal perteneciente al Mercado de Producción de Energía Eléctrica Español, un mercado que nació como resultado del proceso de desregularización del sector eléctrico en España.

Aunque se trata de un mercado con cierto parecido al mercado de valores conocido por todos, el carácter del producto que se compra y se vende (la electricidad), le confiere unas características especiales que lo hacen muy interesante para nuestros propósitos. Por ejemplo, no está influido por tantos factores exógenos y de tan distinta naturaleza como puede estarlo la Bolsa. Por estas razones, la previsión de alguna de sus series temporales debería ser más sencilla y, sobre todo, la calidad de los resultados que se calculen se podría relacionar más directamente con la técnica de previsión utilizada, con lo que será mucho más fácil evaluar la capacidad predictiva del método.

Al intentar calcular previsiones del Precio Horario del Mercado Diario de la Electricidad utilizando las RNAs, nos encontramos con el problema del exceso de información. A pesar de que el Mercado lleva tan sólo dos años de funcionamiento, la serie del Precio Horario que queríamos pronosticar, constaba de más de 15.000 valores. Esta sobreabundancia de datos, hace que el entrenamiento de las redes neuronales se pueda demorar durante varias horas o incluso días, lo cual estaba en contra del objetivo inicial que no era otro que calcular previsiones para las 24 horas siguientes a la hora en que se realizaban los cálculos.

Este problema nos llevo a diseñar un método de entrenamiento (que denominamos "selectivo y continuo"), en el que, gracias a una selección previa de los ejemplos con los que se

va entrenar la red, es posible calcular una previsión en muy pocos minutos. Dado que el entrenamiento de la RNA suele ser un proceso relativamente lento, existen numerosos intentos de acelerar el proceso para aplicarlo a la previsión a corto plazo, e incluso algunos de ellos se basan en la selección de ejemplos. La originalidad de nuestro método estriba en que se utiliza una red neuronal de tipo ART para hacer la selección. Este tipo de redes ha sido utilizado profusamente y con gran éxito en la resolución de problemas de clasificación y por ello quisimos integrarlas en nuestro sistema.

El resultado final fue que las previsiones calculadas con RNAs utilizando el método propuesto, mejoraron sustancialmente las calculadas empleando la metodología de Box-Jenkins, sobre todo cuando se trataba de pronosticar los valores correspondientes a días pertenecientes a fin de semana.

## 5. BIBLIOGRAFÍA.

- BOX, G.E. and JENKINS, G.M. (1970): *Time Series Analysis*, Holden-Day, San Francisco.
- CARPENTER, G.A. and GROSSBERG, S. (1987): "A Massively Parallel Architecture for a Selforganizing Neural Pattern Recognition Machine". *Computer Vision, Graphics and Image Processing*, 37, pp. 54-115.
- CHEN, S.T.; YU, D.C. and MOGHADDAMJO, A.R. (1992): "Weather Sensitive Short-Term Load Forecasting Using Nonfully Connected Artificial Neural Networks". *IEEE Transactions on Power Systems*, 7 (3), pp. 1098-1105.
- CHOI, H.R.; KIM, W. and AN, S.Y. (1997): "Recurrent and decomposed neural network-based hotel occupancy forecasting". *The New Review of Applied Expert Systems*, 1997, pp. 121-136.
- COLLOPY, F. and ARMSTRONG, J.S. (1992): "Expert opinions about extrapolation and the mystery of the overlooked discontinuities". *International Journal of Forecasting*, 8, pp. 575-582.
- COTTRELL, M.; GIRARD, B.; GIRARD, Y.; MANGEAS, M. and MULLER, C. (1995): "Neural modeling for time series: a statistical stepwise method for weight elimination". *IEEE Transactions on Neural Networks*, 6 (6), pp. 1355-1364.
- CYBENKO, G. (1989): "Approximation by Superpositions of a Sigmoidal Function", *Mathematics of Control, Signals and Systems*, 2, pp. 303-314.
- DECO, G.; NEUNEIER, R. and SHÜRMANN, B. (1997): "Non-parametric Data Selection for Neural Learning in Non-stationary Time Series". *Neural Networks*, 10 (3), pp. 401-407.
- DULIBA, K.A. (1991): "Contrasting neural nets with regression in predicting performance in the transportation industry". *Proceedings of the Annual IEEE International Conference on Systems Sciences*, 25, pp. 163-170.
- FAHLMAN, S.E. (1989): "Faster-learning variations of back-propagation: An empirical study". *Proceedings of the 1988 Connectionist Models Summer School*, pp. 38-51.
- FOG, T.L.; LARSEN, J. and HANSEN, L.K. (1996): "Training and Evaluation of Neural Networks for Multivariate Time Series Processing". *Proceedings of IEEE International Conference on Neural Networks*. Perth, Australia.
- HO, K.L.; HSU, Y.Y. and YANG, C.C. (1992): "Short-term load forecasting using a multilayer neural network with an adaptive learning algorithm". *IEEE Transactions on Power Systems*, 7 (1), pp. 141-149.
- HOPTRUFF, R.G. (1993): "The principles and practice of time series forecasting and business modeling using neural nets". *Neural Computing and Applications*, 1, pp. 59-66.
- HORNIK, K.; STINCHCOMBE, M. and WHITE, H. (1989): "Multilayer feedforward networks are universal approximators". *Neural Networks*, 2, pp. 359-366.

- JACOBS, R.A. (1988): "Increased rates of convergence through learning rate adaptation", *Neural Networks*, 1 (4), pp. 295-308.
- PACK, D.C.; EL-SHARKAWI, M.A. and MARKS II, R.J. (1991a): "An adaptively trained neural network". *IEEE Transactions on Neural Networks*, 2 (3), pp. 334-345.
- PACK, D.C.; EL-SHARKAWI, M.A.; MARKS II, R.J.; ATLAS, L.E. and DAMBORG, M.J. (1991b): "Electric load forecasting using an artificial neural network". *IEEE Transactions on Power Systems*, 6 (2), pp. 442-449.
- PARKER, D.B. (1987): "Optimal algorithm for adaptive network: second order back propagation, second order direct propagation, and second order Hebbian learning". *Proceedings of the IEEE International Conference on Neural Networks 2*, IEEE Press, pp. 593-600.
- PENG, T.M.; HUBELE, N.F. and KARADY, G.G. (1992): "Advancement in the application of neural networks for short-term load forecasting". *IEEE Transactions on Power Systems*, 7 (1), pp. 250-257.
- PLUTOWSKI, M. and WHITE, H. (1991): "Active selection of training examples for network learning in noiseless environments". *Technical Report No. CS91-180*. University of California, San Diego.
- PRIESTLEY, M.B. (1988): *Non-Linear and Non-Stationary Time Series Analysis*. Academic Press, London.
- RIEDMILLER, M. and BRAUN, H. (1993): "RPROP – A Fast Adaptive Learning Algorithm". *Proceeding of ISICIS VII*.
- RUMELHART, D.E. and McCLELLAND, J. (1986): *Parallel Distributed Processing*. MIT Press, Cambridge.
- SUBBA RAO, T. (1992): "Analysis of Nonlinear Time Series (and Chaos) by Bispectral Methods". In *Nonlinear Modeling and Forecasting*, (Casdagli, M. and Eubank, S., ed.). Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol XII, pp. 199-226.
- TOLLEANERE, T. (1990): "SuperSAB: Fast Adaptive Backpropagation with good Scaling Properties". *Neural Networks*, 3, pp. 561-573.
- TONG, H. (1983): *Threshold Models in Non-Linear Time Series Analysis*. Springer-Verlag, New York.
- TONG, H. (1990): *Non-Linear Time Series: A Dynamic System Approach*. Oxford University Press, Oxford.
- TSAY, R.S. (1991): "Non-Linear Time Series Analysis: Diagnostics and Modelling". *Stat. Sinica*, 1, pp. 432-451.
- VERMAAK, J. and BOTHA, E.C. (1998): "Recurrent Neural Networks for Short-Term Load Forecasting". *IEEE Transactions on Power Systems*, 13 (1), pp. 126-132.
- WERBOS, P. (1988): "Backpropagation: Past and Future". *Proceedings of the IEEE International Conference on Neural Networks*, IEEE Press, pp. 343-353.
- YULE, G. (1927): "On a Method of Investigating Periodicity in Disturbed Series with special reference to Wolfer's Sunspot Numbers". *Phil. Trans. Royal Society London*, 226, pp. 267-298.

La Revista *Investigaciones Europeas de Dirección y Economía de la Empresa* recibió este artículo el 15 de octubre de 2001 y fue aceptado para su publicación el 1 de diciembre de 2003.