

# Predicción de la demanda eléctrica horaria mediante redes neuronales artificiales

Carlos Mallo González.

Departamento de Economía Cuantitativa. Universidad de Oviedo.

Avda. del Cristo s/n. 33006 Oviedo

mallogc@mixmail.com

## RESUMEN

En este trabajo se realiza una introducción a los modelos neuronales y se explora su aplicabilidad en el terreno de la economía cuantitativa como instrumento de modelización y predicción no paramétrica. A tal objeto hemos desarrollado un modelo de Red Neuronal Artificial que posteriormente aplicaremos en la predicción de la demanda horaria de energía eléctrica con datos reales. Los resultados así obtenidos fueron comparados con los ofrecidos por las metodologías convencionales de ajuste y predicción.

**Palabras clave:** Red Neuronal Artificial, Perceptrón Multicapa, Regresión no paramétrica, Aproximador universal, No-Linealidad, Curva de carga eléctrica, Predicción horaria.

## 1. INTRODUCCIÓN

### 1.1. Las Redes Neuronales Artificiales en Economía

Las RNA constituyen un potente instrumento para la aproximación de funciones no lineales. Su uso resulta especialmente útil en la modelización de aquellos fenómenos complejos donde la presencia de relaciones no lineales entre las variables es habitual. La aplicación de este tipo de técnicas en el terreno de la Economía ha proporcionado interesantes resultados, sobre todo en los campos de la Teoría Económica<sup>1</sup> y la Econometría. La gran aportación de esta metodología consiste precisamente en la eliminación de la hipótesis reduccionista de linealidad que ha venido dominando el análisis económico a lo largo de su historia. Si bien dicha hipótesis simplificadora ha permitido formular modelos matemáticos con los que explicar el fenómeno económico, al no contemplar la no-linealidad inherente a los sistemas económicos, dichos modelos ofrecían una visión parcial de la realidad, Granger (1991), Shachmurove (2000). Desde el punto de vista de la econometría, principal herramienta de análisis y validación con la que cuenta el economista, no podemos olvidar que una de las hipótesis fundamentales sobre las que se construye el ya clásico *Modelo Lineal Básico* es, como su propio nombre indica, la linealidad de las ecuaciones. El empleo de técnicas econométricas sustentadas en una falsa linealidad ha conducido a la especificación de modelos incompletos, los cuales mostraban sus carencias al enfrentarse a fenómenos de dinámicas no-lineales.

Con el presente documento pretendemos explorar la capacidad de las Redes Neuronales Artificiales como instrumento de modelización y predicción económica, poniendo de relieve la superioridad mostrada por las mismas en el modelado de los fenómenos no-lineales. Para ello confeccionaremos un sencillo modelo neuronal que, implementado en *VBA*<sup>2</sup> y *Matlab*, nos permitirá predecir la evolución de la demanda horaria de energía eléctrica.

---

<sup>1</sup> Modelos de Expectativas Racionales basados en RNA, M. Heinemann (2000).

<sup>2</sup> El algoritmo utilizado para entrenar el modelo ha sido programado en *Visual Basic for Applications*, como un componente de la hoja de cálculo Excel.

Este trabajo no es más que el inicio de una línea de investigación. Actualmente tratamos de aplicar esta misma metodología en el análisis de series temporales financieras (fondos de inversión y acciones), las cuales siguen en la mayor parte de los casos complejas dinámicas no-lineales de carácter caótico (*caos determinista*).

## 1.2. La Electricidad y el Mercado Eléctrico

No cabe duda de que la electricidad constituye una de las principales fuentes energéticas con las que cuenta nuestra civilización. Su empleo abarca un amplísimo abanico de actividades que se extiende desde los usos puramente industriales hasta el consumo doméstico de las familias. Más aún, en un contexto mundial de creciente competencia, no se concibe una economía en crecimiento y competitiva en el ámbito global sin un fluido eléctrico de calidad y a precios altamente competitivos. Esta elevada dependencia ha convertido a la energía eléctrica en un input estratégico que posee unos intensos efectos de arrastre, ya no solo de carácter económico sino también social.<sup>3</sup> De modo que podemos afirmar que vivimos en un mundo que funciona con electricidad.

Desde un punto de vista técnico suele hablarse de la energía eléctrica como una “*fente de energía secundaria*”, dado que la misma se obtiene a partir de otras fuentes denominadas primarias, básicamente: carbón, gas y petróleo (combustibles fósiles), hidráulica y nuclear, J. García et al. (1998). Al margen de estas y otras consideraciones puramente técnicas, desde un punto de vista económico debemos reseñar dos importantes características que nos ayudarán a comprender el funcionamiento del mercado eléctrico, P. Murto (1998):

1. Con independencia de la fuente primaria utilizada, ***la energía eléctrica no puede ser almacenada***<sup>4</sup>.

---

<sup>3</sup> Existe una amplia bibliografía sociológica en la que se estudia en detalle el comportamiento de determinadas poblaciones ante cortes en el suministro eléctrico. (ver R. Dynes (1970) y H. Barton (1969)).

<sup>4</sup> Dado el actual estado de la ciencia, si bien es posible almacenar pequeñas cantidades de energía eléctrica en forma de pilas químicas o dispositivos similares, no existe forma eficiente conocida para almacenar las cantidades generadas por una central estándar.

2. En la mayor parte de sus usos la electricidad **juega el papel de un input específico**, es decir, el de un factor productivo que no puede ser reemplazado en el corto plazo. (*Relación Técnica de Sustitución = 0, funciones de producción tipo Leontieff*). Pensemos que si bien la mayor parte de los instrumentos eléctricos podrían ser adaptados para funcionar con otras fuentes de energía (gasóleo, gas, etc), en el corto plazo estas modificaciones no serían factibles.

Estas características dotan al mercado eléctrico de una serie de peculiaridades sumamente interesantes.

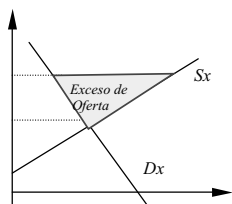


Gráfico 1.1. Exceso de Oferta

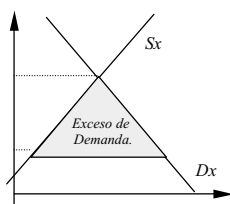


Gráfico 1.2. Exceso de Demanda.

En el gráfico 1.1. se muestra lo que en un mercado convencional sería un exceso de oferta. El área sombreada se correspondería con el valor de los stocks acumulados. Sin embargo, en el mercado eléctrico, al ser la electricidad un bien no acumulable, el exceso de oferta, es decir la electricidad producida y no consumida inmediatamente, se perdería sin más. En esta nueva situación el área sombreada representaría las pérdidas asociadas a una asignación ineficiente de las fuentes primarias, que constituyen recursos escasos y limitados.

Por otra parte, en el gráfico 1.2. se muestra una situación de exceso de demanda. En un mercado tradicional, el área sombreada representaría el valor de la demanda no atendida. Dada esta situación los demandantes tendrían la posibilidad de esperar a que se atendiera su petición (se formarían colas), ó alternativamente podrían consumir algún sustitutivo cercano del bien demandado. Sin embargo en el mercado eléctrico no existe la posibilidad de recurrir a sustitutos cercanos, al menos en el corto plazo. Respecto a la primera alternativa, (formación de colas), un retraso en el suministro eléctrico podría

tener gravísimas consecuencias económicas y sociales. En la nueva situación el área sombreada representaría un coste de eficiencia asociado a la situación de exceso de demanda.

Las características especiales de la energía eléctrica han motivado la creación de un organismo denominado *operador del mercado eléctrico*<sup>5</sup> que actúa como regulador, tratando de coordinar en todo momento la oferta y la demanda. Este organismo puede tener titularidad pública (Argentina, Chile el Salvador, etc.) ó privada (España, países de la UE, etc.). En el caso concreto de la UE, los mercados eléctricos de los países miembros se encuentran atravesando un proceso liberalizador con el que se pretende romper los viejos monopolios estatales y dotar así de una mayor competitividad a los mercados, siendo la última meta la creación de un mercado eléctrico europeo. Al margen de quien ostente la titularidad del operador del mercado, éste asume una serie de responsabilidades vitales:

- La sustitución de la energía que un grupo no pudo producir.
- El suministro instantáneo de la energía que los consumidores desean en cada momento.

De este modo la operación del sistema hace posible que el mercado funcione y aporta una liquidez absoluta a generadores y consumidores. El esquema de funcionamiento seguido por el operador sería el siguiente: En primer lugar debería de averiguar el valor que toma la demanda eléctrica en cada momento, tratando de llegar a la denominada *curva de carga*<sup>6</sup>. Una vez conocido este dato, el operador se dirigiría a los centros de producción, indicándoles la carga eléctrica que deben de introducir en la red en cada momento para garantizar la cobertura de la demanda y, al mismo tiempo, minimizar la sobreproducción y sus costes de eficiencia. Sin embargo la demanda eléctrica es muy variable a lo largo de los años, los meses, los días e incluso las horas. Estas características aportan una incertidumbre al sistema eléctrico que afecta de forma significativa al coste final del suministro. Para reducir este factor de coste, el gestor

---

<sup>5</sup> RED ELÉCTRICA DE ESPAÑA opera el sistema eléctrico peninsular garantizando las condiciones técnicas para que la electricidad fluya continuamente desde los centros de generación hasta los centros de consumo.

<sup>6</sup> Curva de Carga: Valor que toma la demanda eléctrica en cada momento, se correspondería con la demanda instantánea de energía eléctrica.

necesita conocer con precisión los mecanismos que hacen variar la demanda. Con la información aportada por la curva de carga se logran dos objetivos, J. Bao (2000):

1. Prever y adaptarse a la demanda en los distintos plazos. A corto, en la operación del sistema, y a largo en la creación de nuevas centrales y redes para el suministro.
2. Gestionar la demanda. Es decir, inducir a los consumidores a modificar sus pautas de consumo eléctrico para lograr una mayor eficiencia conjunta del proceso de suministro y uso.

Nuestro objetivo consiste en construir un modelo neuronal que sea capaz ajustar una curva de carga. Posteriormente trataremos de emplear este modelo en la predicción horaria de la demanda eléctrica.

## **2. FACTORES QUE AFECTAN A LA CURVA DE CARGA**

Quizás la parte más difícil en la construcción de un buen modelo de red neuronal es la relacionada con selección y recogida de la información que alimentará al mismo en la fase de entrenamiento. Una serie de investigaciones realizadas sobre la demanda eléctrica y sus determinantes (P.Murto (1998), J.Bao (1996), Proyecto Indel (1998)) en diferentes países, han permitido identificar un conjunto de variables básicas que explica la evolución de la demanda eléctrica a lo largo del tiempo, siendo este conjunto común para la mayor parte de los países desarrollados. De este modo para modelizar la curva de carga de países tan distintos como España y Suecia utilizaríamos el mismo conjunto de variables causa, dado que si bien pueden existir diferencias culturales entre ambos países, el patrón que determina la demanda eléctrica es muy similar. Comentaremos detalladamente las variables que hemos seleccionado como posibles factores explicativos de la curva de carga:

### **2.1. Condiciones meteorológicas**

Aquí englobaríamos variables como la velocidad del viento, la nubosidad, la pluviosidad, la temperatura, etc. Sin embargo nos encontramos con que no todos los factores climáticos afectan a la demanda eléctrica. Algunos de ellos son típicamente

aleatorios y otros aparecen interrelacionados. Por ejemplo, la temperatura viene explicada parcialmente por la nubosidad, la pluviometría, la humedad relativa, etc.

Entre todos estos factores la temperatura es el más relevante, dado influye de forma directa sobre múltiples fuentes de consumo eléctrico como sistemas calefactores, aire acondicionado, refrigeradores, etc. Esta relación entre la demanda eléctrica y la temperatura es de carácter no lineal.

En nuestro modelo, inicialmente, incluiremos dos variables climáticas: la temperatura y la velocidad del viento. La temperatura, dado a que la inmensa mayoría de los trabajos consultados señalan esta variable como principal determinante meteorológica de la demanda eléctrica; y en cuanto a la velocidad del viento, debemos señalar que si bien existen discrepancias, nosotros hemos optado por incluir inicialmente esta variable, para poder contar con un modelo de partida lo más completo posible. Posteriormente, atendiendo a los resultados obtenidos en el entrenamiento y en la predicción nos podremos plantear la eliminación de esta última.

## 2.2. Calendario

Existen diferentes efectos relacionados con el calendario que inciden sobre la curva de demanda eléctrica:

✓ **Hora del Día:** *Caeteris paribus* todas las variables que afectan a la curva de carga menos la hora, resulta evidente que la demanda eléctrica realizada a las tres de la madrugada no será igual a la realizada a las tres de la tarde.

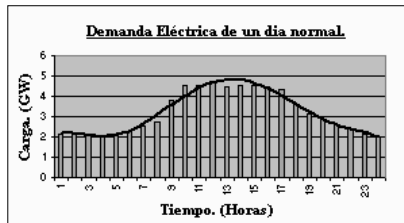


Gráfico 2.1. Curva de carga

En el gráfico 2.1. se puede apreciar como el grueso de la demanda eléctrica se acumula en el período comprendido entre las 9 de la mañana y las cinco de la tarde, algo lógico si tenemos en cuenta que dicho período que se corresponde con la jornada laboral estándar. Para que nuestro modelo sea capaz de reflejar este efecto debemos incluir una variable que represente la hora del día en la que se lleva a cabo la demanda. Para ello utilizaremos una variable numérica cuyo rango oscilará entre 0 y 23. [0,1,2...22,23].

✓ **Día de la semana:** Repitiendo el mismo planteamiento podemos intuir que la demanda de electricidad realizada a la misma hora de días diferentes será también diferente. Así la demanda llevada a cabo un miércoles a las once de la mañana probablemente no coincida con la realizada un domingo a esa misma hora.

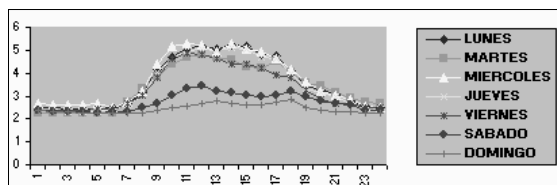


Gráfico 2.2. Curvas de cargas según los días de la semana.

En el gráfico 2.2. se puede apreciar como todos los días laborables poseen un perfil de demanda muy similar, con excepción del lunes y el miércoles en cuyas madrugadas el consumo de energía es significativamente inferior. Con respecto a los fines de semana se observa como si bien se repite el mismo patrón horario de consumo detectado en los días laborables<sup>7</sup>, en general el consumo medio es sensiblemente inferior. Para que el modelo sea capaz de identificar el perfil de carga asociado a cada día de la semana debemos de incluir una variable que recoja este efecto. Esta variable será categórica y estará codificada de la siguiente manera: Lunes=0...Domingo=6

<sup>7</sup> Máximo consumo en la franja comprendida entre las once de la mañana y las tres de la tarde. Mínimo consumo en la madrugada.



✓ **Fin de semana y fiestas**<sup>8</sup>: Teniendo en cuenta las conclusiones a las que llegamos al estudiar el gráfico 2.3., podemos observar la presencia de un factor que incide sobre la demanda eléctrica y que está asociado a los fines de semana y a los días no laborables. En estos días, tal y como comentamos anteriormente, se produce una reducción generalizada en el consumo eléctrico. Este nuevo factor puede ser incluido en nuestro modelo mediante una variable dicotómica que tome el valor 0 si la demanda se realiza en un día laborable, ó 1 en otro caso.

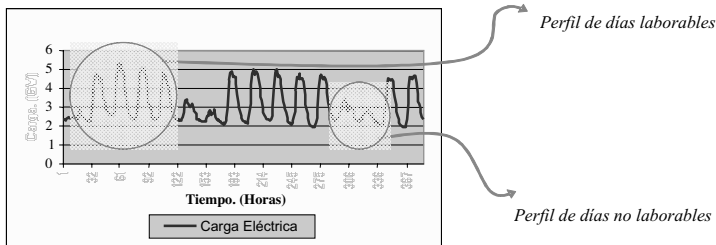


Gráfico 2.3. Curva de carga a lo largo de la semana

✓ **Mes del año**: En función del mes en el que nos encontremos la demanda diaria media de energía varía significativamente. Esto se debe principalmente al efecto de las estaciones. Pensemos que en los meses de invierno, al margen de los efectos climáticos ya considerados (básicamente temperatura), se dispone de un menor número de horas solares, la gente tiende a pasar más tiempo en casa, etc. En lo referente al verano el consumo de energía también puede experimentar un considerable aumento como resultado del uso de sistemas de aire acondicionado. Atendiendo al gráfico 2.4. podemos comprobar como efectivamente la curva de carga anual alcanza sus valores máximos en los meses de verano e invierno. Por tanto debemos de incluir en nuestro modelo una variable que recoja esta *componente estacional* de la demanda eléctrica. Para ello utilizaremos una variable categórica cuyos valores representarán los distintos meses del año. Esta variable será codificada, asignándole valores que irán desde el Enero = 0, hasta el Diciembre = 11.

<sup>8</sup> Con esta variable se pretende capturar la estacionalidad asociada a los fines de semana así como el efecto de los días no laborables, complementado de este modo a la variable “día de la semana”, que como ya comentamos recoge el perfil propio de cada día sin tener en cuenta posibles efectos estacionales.

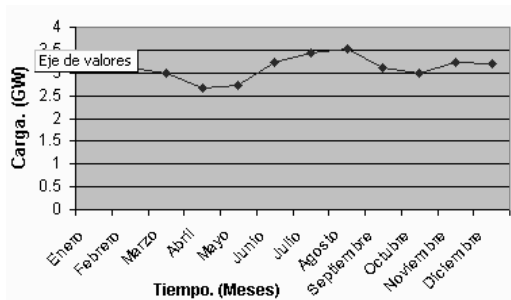


Gráfico 2.4. Curva de carga anual

### 2.3. Información económica

Dada la extensa penetración de la electricidad en la actividad económica del país, existen multitud de variables económicas que influyen en su demanda. La mayoría de ellas covarían entre sí en mayor o menor medida. Introducir varias por separado en un modelo explicativo daría lugar a problemas de multicolinealidad que harían imposible lograr unos resultados aceptables. Dada esta situación, una posible estrategia a seguir consistiría en sintetizar la información disponible mediante la aplicación del método de componentes principales a un conjunto de indicadores económicos. Algunos autores consultados argumentan que esta estrategia conduce a componentes difíciles de explicar y que no aportan mejoras significativas al modelo. Como alternativa proponen la construcción de un indicador de actividad que tome índices de producción industrial de 9 sectores económicos, índices de activos líquidos en manos del público y consumo telefónico.

Si bien todas estas variables pueden incidir significativamente en la demanda eléctrica, hemos optado por no incluirlas en nuestro modelo. La mayor parte de ellas efectivamente afectan a la demanda, pero sus efectos son apreciables en el medio y en largo plazo, siendo irrelevantes en el muy corto plazo, horizonte sobre el que se construye nuestro modelo. Para modelos que pretendan captar la evolución de la demanda a lo largo de los meses o los años dichas variables y otras como la demografía, la evolución de los precios de energías sustitutivas, etc. resultarían clave.

## 2.4. Factores no predecibles

Debemos de considerar la existencia de una serie de factores que, siendo puramente aleatorios, pueden afectar sensiblemente a la demanda de energía eléctrica. Entre dichos factores destacamos, entre otros: Paros y huelgas generales, eventos deportivos, cierre de instalaciones industriales intensivas en electricidad, etc. En función del factor del que se trate podemos tratar el mismo como un *outlier* o dato atípico, siendo por tanto necesario realizar un filtrado de los datos de entrada.

También existen factores cuyo efecto sobre la demanda eléctrica puede ser considerado como un “*ruido blanco*”, ( $\mu = 0$ ,  $\sigma = k$ ), en cuyo caso no será necesario realizar filtrado alguno.

## 3. MODELO NEURONAL DE LA CURVA DE CARGA

### 3.1. Arquitectura de la red

La red que proponemos estará compuesta por tres capas: Una de entrada, una oculta<sup>9</sup> y una de salida. En lo referente a la capa de entrada, ubicaremos en la misma seis unidades que representarán cada uno de los factores explicativos seleccionados. La capa de salida solo contará con una unidad, la cual indicará el valor de la carga eléctrica horaria asociado a cada vector de inputs presentado a la red. Finalmente la capa oculta contará con un número variable de neuronas. Esquemáticamente nuestro modelo de red neuronal artificial para la curva de carga tendría la siguiente forma:

---

<sup>9</sup> En teoría las RNA con dos capas ocultas pueden representar funciones con cualquier forma. Hornik *et al.* (1994), White y Gallant (1992). Por tanto no existe razón alguna (al menos con base teórica) para utilizar redes con más de dos capas ocultas. De hecho se ha podido demostrar que en la mayor parte de los problemas reales es suficiente con aplicar redes de una sola capa oculta, siendo casos aislados aquellos problemas que requieren más de dos capas ocultas. Por ello hemos optado por incluir una única capa oculta, añadir capas adicionales supondría asumir un mayor coste computacional sin obtener a cambio mejoras razonables en la capacidad de predicción de nuestro modelo.

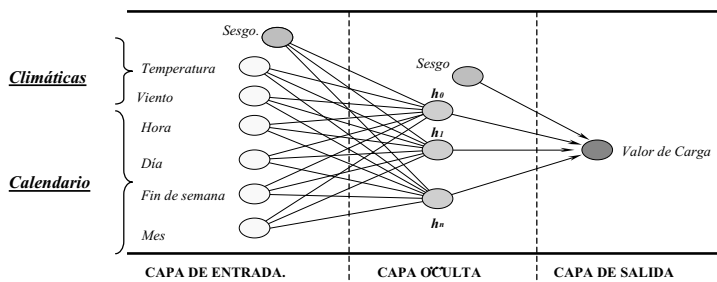


Figura 3.1. Representación esquemática de la RNA empleada.

La determinación del número óptimo de neuronas de la capa oculta se ha llevado a cabo mediante un proceso de ensayo y error en el que jugamos con el número de neuronas y el máximo error admisible. En todo caso, nuestro objetivo será dotar a la red de un número adecuado de neuronas en la capa oculta para que sea capaz de aprender las características de las posibles relaciones existentes entre los datos de la muestra. Demasiadas neuronas podrían llevarnos a un *sobre-ajuste*, la red memorizaría el conjunto de datos de entrenamiento y para valores pertenecientes al mismo su precisión sería casi absoluta. Sin embargo si le presentamos valores desconocidos fallaría estrepitosamente, dado que como consecuencia del sobre-ajuste sería incapaz de *generalizar*. Por otra parte si el número de neuronas en la capa oculta es reducido podremos llegar a una situación de *sub-ajuste*, en la cual la red no sería capaz de aprender ciertas relaciones existentes entre las variables. Aprovechando las propiedades atribuidas a las RNA Unidireccionales (*Teorema de Hornik-Stinchcombe-White*, Hornik et al., (1990)), nuestro objetivo será aproximar la curva de carga, dado el conjunto de datos disponible, para posteriormente poder realizar predicciones sobre valores futuros.

### 3.2. Datos de entrada

Teniendo en cuenta lo explicado en el Apartado 2 alimentaremos el modelo con la información aportada por las siguientes variables:

CATEGORÍA	VARIABLE	UNIDADES	TIPO
CLIMÁTICAS	Temperatura	C°	Numérica. Representada por su valor actual.
	Viento	m /sg.	Numérica. Representada por su valor actual.
CALENDARIO	Hora	-	Numérica. Representada por su valor actual. Rango = (0-23)
	Día	-	Catagórica. Codificada: Lunes = 0, Martes = 1,..., Domingo = 6.
	Fin de Semana	-	Numérica binaria. Fin de semana ó festivo = 0 / Día laborable = 1
	Mes	-	Catagórica. Codificada: Enero = 0, Febrero = 1,..., Diciembre = 11
OBJETIVO	Carga Eléctrica	GW/Hora <sup>10</sup>	Numérica. Representada por su valor actual.

Dado el nivel de desagregación requerido (valores horarios), la obtención las series de datos referidas a la carga eléctrica y a las variables climáticas consideradas (temperatura y de velocidad del viento), ha sido una tarea especialmente laboriosa. La mayor parte de las bases de datos públicas recogen esta información, pero agregada en medias mensuales y anuales. En cuanto a las compañías privadas, si bien registran y almacenan toda esta información, se muestran reacias a facilitarla. Después de un intenso proceso de búsqueda pudimos obtener la información relativa a la carga eléctrica horaria en la página web de una compañía norteamericana de distribución eléctrica con base en Delaware y que opera en la zona sur de la costa este estadounidense. Una vez en posesión de la serie de carga eléctrica, la obtención de las series climáticas (temperatura y velocidad del viento) tampoco resultó tarea fácil. En primer lugar por el elevado nivel de desagregación necesario (series horarias), y en segundo lugar por el hecho de que las mismas tenían que estar referidas a la misma zona que los datos de carga. Finalmente conseguimos esta información en las páginas web de NOAA (National Oceanic and Atmospheric Administration) y de ATDD (Atmospheric Turbulence and Diffusion Division).

<sup>10</sup> Gigawatios por hora.

### 3.3. Descripción de las series

Todas las series están en medias horarias, abarcando el período comprendido entre las cero horas del uno de Enero de 1999 y las 24 horas del día 15 de Enero de 2000. Esto hace que dispongamos de un total de 9120 observaciones, de las cuales 8760 (el 96%, un año natural) se emplearán en el entrenamiento de la red, quedando reservadas las últimas 360<sup>11</sup> observaciones para la validación del modelo.

DISTRIBUCIÓN DE LA MUESTRA	
<b>Total de observaciones</b>	<b>9120</b>
<b>Datos para entrenamiento</b>	8760
<b>Datos reservados para predicción</b>	360

Tabla 3.2. Distribución de los datos muestrales según su uso en el modelo.

### 3.4. Preparación de la información de entrada: Filtrado, Codificación y cambio de escala

Uno de los aspectos más relevantes en el éxito de cualquier modelo neuronal es el conjunto de datos empleado en su entrenamiento. Incluso el más simple de los modelos puede ofrecer buenos resultados si es alimentado con datos que, tras ser preprocesados adecuadamente, revelan de forma clara y precisa la información relevante, Beltratti *et al.* (1996). Asimismo, el mejor de los modelos fracasará si es entrenado con datos de baja calidad. Es por tanto necesario llevar a cabo un pre-procesado de la información muestral para de este modo aumentar la capacidad de aprendizaje de la red. El pre-procesado constará de las siguientes operaciones:

- ✓ **Filtrado:** Se trataría de repasar las series de datos buscando posibles outliers. Hemos encontrado datos atípicos, concretamente valores perdidos, tanto en la serie de carga eléctrica como en las series climáticas. Creemos que estos datos atípicos se corresponden con cortes en el suministro eléctrico y con fallos en los instrumentos de medida respectivamente. En la serie de carga hemos remplazado los outliers con la media aritmética de los valores anterior y posterior al perdido. En cuanto a las

---

<sup>11</sup> 15 primeros días del mes siguiente, Enero de 2000.

series climáticas, al encontrarnos ya no con valores perdidos aislados sino con tramos de la serie (de longitud variable pero en ningún caso superior a las 24 observaciones) en los que no existían datos, preferimos cubrir estos huecos con valores históricos de las mismas series. Al hacer esto asumimos que variables como la temperatura y la velocidad del viento presenta un comportamiento muy similar dentro de un mismo mes para los diferentes años.

- ✓ **Codificación:** La codificación se realizará siguiendo los criterios ya explicados.
- ✓ **Cambio de escala.** Todas las variables consideradas en el modelo serán linealmente escaladas (*Linear Scaling*) entre 0 y 1. Este procedimiento consiste simplemente en modificar la escala de las variables atendiendo a su valor mínimo y máximo. Como resultado obtendremos nuevas series de datos más homogéneas, las cuales oscilarán entre 0 y 1.

La información en su estado original tendría el siguiente formato:

Datos de entrada en su formato original							
FECHAS	CARGA	Temperatura	Viento	Hora	Día	Fin de Semana	Mes
01-01-1999	5.2873	-3.3	0.2	0	0	0	0
02-01-1999	4.8385	-4	0.2	1	0	0	0
03-01-1999	5.2598	-4.4	0.2	2	0	0	0

Tabla 3.3. Información de entrada en su formato original.

Una vez preparada su nuevo formato sería<sup>12</sup>:

Tabla 3.4. Información de entrada en nueva escala [0,1].

Datos de entrada en su formato original							
FECHAS	CARGA	Temperatura	Viento	Hora	Día	Mes	Fin de Semana
01-01-1999	0.5901	-0.1375	0.0277	0.0000	0.0000	0	0
02-01-1999	0.5328	-0.166	0.0277	0.0416	0.0000	0	0
03-01-1999	0.5866	-0.1833	0.0277	0.0833	0.0000	0	0

#### 4. ENTRENAMIENTO DE LA RED

<sup>12</sup> En la tabla únicamente se muestran tres valores de la serie a modo de ejemplo. Estos valores se corresponden con la carga observada en tres horas pertenecientes a un mismo día, por ello las variables "Día", "Fin de semana" y "Mes" toman un valor constante.

#### 4.1. Soporte informático para las RNA

Uno de los problemas que encontramos a la hora aplicar las RNA es la escasez de herramientas informáticas específicas; esto ha llevado a los investigadores a desarrollar sus propias aplicaciones para poder experimentar y validar las teorías enunciadas; Son programas informáticos que están hechos a medida, teniendo en cuenta las necesidades específicas de sus creadores, lo que hace difícil su aplicación fuera del proyecto en que nacieron. De cualquier modo existen una serie de programas de carácter más genérico que incorporan librerías o cajas de herramientas especializadas en las redes neuronales artificiales. Este es el caso de Matlab y su *Neural Networks Toolbox*. También hemos encontrado una serie de programas altamente especializados que ofrecen un sinfín de opciones para definir y manipular modelos neuronales mediante interfaces gráficas de fácil manejo como Neuro-Solutions, Brain-Maker, etc. En nuestro caso optamos por combinar el uso de dos aplicaciones. Por una parte empleamos un conjunto de *macros*<sup>13</sup> programadas en *Visual Basic* ejecutables desde la hoja de cálculo *Excel*. La principal razón por la que preferimos utilizar una macro como herramienta con la que llevar a cabo el entrenamiento estriba en el hecho de que la misma, una vez implementada, puede ser utilizada desde Excel, siendo por ello de uso más sencillo y accesible. Esta herramienta resultó sumamente útil y eficaz en una primera fase de la investigación, en la que entrenamos la red con datos pertenecientes al mismo mes, por lo que la amplitud de las series presentadas nunca excedía las 721 observaciones; si bien esta aproximación al problema nos permitió llegar a interesantes resultados, presentaba importantes carencias. En primer lugar al realizarse el entrenamiento dentro del mismo mes, la estacionalidad mensual desaparece, con lo que la variable “mes del año”, sería innecesaria, contradiciendo el modelo teórico. Por otra parte, la red solo sería capaz de reconocer patrones similares al del mes utilizado en el entrenamiento, con lo que fuera de esas fechas su capacidad de predicción se vería seriamente mermada. Todo esto nos llevo a repetir el entrenamiento aplicando el mismo modelo teórico pero con series de datos anuales, cuya longitud superaba las 8700 observaciones. La nueva escala del

---

<sup>13</sup> Una macro no es más que un conjunto de instrucciones que automatizan un determinado proceso liberando, por tanto, de trabajo al usuario. En nuestro caso estas instrucciones consisten en operaciones matemáticas llevadas a cabo sobre celdas de un libro de Excel. El resultado de dichas operaciones será precisamente el entrenamiento de la red. Como algoritmo de entrenamiento utilizamos el Algoritmo de Retrogradación de Errores (BP). En cuanto a las funciones de activación empleamos funciones logísticas en ambas capas.



problema resultó desmesurada para el algoritmo implementado en Visual Basic que, si bien procesaba, lo hacía a un ritmo desalentador. Los requerimientos computacionales de la nueva situación nos llevaron a cambiar de herramienta informática, encontrando en Matlab y en su Toolbox el perfecto sustituto. La potencia de motor del cálculo de Matlab, muy superior al de Excel, permitió superar fácilmente el inconveniente asociado a la escala del problema. Este software también permitió modificar algunos parámetros del entrenamiento, como el criterio de parada. En vez de fijar el número máximo de iteraciones, determinado de manera arbitraria, establecimos un umbral para el error cuadrático medio marginal. Si este umbral no era alcanzado en el transcurso de 25 iteraciones se podía asumir que cada nueva iteración no aportaba nada al entrenamiento, con lo que el mismo podía darse por concluido. Este nuevo criterio de parada basado en el descenso por la función de error cuadrático medio, además de ser técnicamente más adecuado, evita la posibilidad de llegar a un sobre-entrenamiento de la red. En cuanto a las funciones de transferencia sustituimos la función logística por la tangente hiperbólica.

#### 4.2. Resultados del entrenamiento

Tal y como ya comentamos, el entrenamiento se ha llevado a cabo utilizando una muestra de 8760 observaciones y para llegar a determinar el valor óptimo de los parámetros que definen la red hemos recurrido al método de ensayo y error:

- ✓ **Número óptimo de iteraciones:** La aportación de cada nueva iteración en términos de reducción del error cuadrático medio es despreciable a partir de las 200 iteraciones.
  
- ✓ **Número óptimo de neuronas en la capa oculta:** Pude ser determinado mediante la aplicación de una serie de fórmulas heurísticas, ó alternativamente mediante un proceso de ensayo y error. Nosotros hemos optado por la segunda opción. Para llevar a cabo el experimento fijamos la tasa de aprendizaje y el valor del momento, de este modo pudimos aislar el efecto que tiene el número de neuronas ocultas sobre el error cuadrático medio del entrenamiento. Si bien cada neurona adicional aportará mayor capacidad explicativa a la red, esta aportación será cada vez menor hasta llegar a un punto en que la capacidad explicativa del modelo no se ve afectada con la inclusión de

nuevas neuronas. Los experimentos realizados revelaron que el número óptimo de neuronas de la capa oculta estaría situado en torno a 8. Teniendo en cuenta el número de vectores con el que se han realizado los entrenamientos (8760), este número de neuronas nos parece razonable.

✓ **Valor óptimo de la tasa de aprendizaje:** Fijando todos los parámetros que inciden en el entrenamiento de la red salvo la tasa de aprendizaje, llegamos a la conclusión de que el valor óptimo de este parámetro se situaría en torno a 0.7.

✓ **Valor óptimo del momento:** Repitiendo una vez más el mismo proceso ya explicado, obtuvimos como valor óptimo del momento un valor situado en torno a 0.6.. En resumen, los valores óptimos que definen el entrenamiento de nuestro modelo neuronal serían:

<b>Valores óptimos de los parámetros del entrenamiento<sup>14</sup></b>	
<b>Iteraciones</b>	200
<b>Neuronas Ocultas</b>	8
<b>Tasa de Aprendizaje</b>	0.6-0.8
<b>Momento</b>	0.4-0.6

Tabla 4.1. Valores óptimo de los parámetros del entrenamiento.

## 5 PREDICCIÓN

Una vez concluido el entrenamiento, si ha resultado exitoso, podemos asumir que la red ha sido capaz de aprender las relaciones entre los datos presentados, con lo que la misma estará preparada para llevar a cabo tareas de predicción<sup>15</sup>. La predicción podrá realizarse sobre datos ya conocidos por la red, utilizados en su entrenamiento, ó alternativamente sobre datos nunca vistos. En el primer caso contrastaremos la eficacia del aprendizaje, es decir, en que medida la red es capaz de replicar la serie objetivo al

---

<sup>14</sup> Óptimos en el sentido que hacen caer el error cuadrático medio más rápidamente, es decir, modifican los pesos para que la función de error “caiga” siguiendo su gradiente.

<sup>15</sup> Antes de comprobar la capacidad de predicción de la red hemos verificado la estacionariedad de la serie de datos a través del test de raíces unitarias de Dickey-Fuller.

presentarle de nuevo los mismos datos con los que fue entrenada. En cuanto al segundo caso, podremos contrastar la capacidad de generalización de la red, y por tanto validar su capacidad como instrumento de predicción.

### 5.1. Predicción con datos conocidos por la red

En este caso utilizaremos como datos de entrada los mismos que utilizamos en el entrenamiento, es decir emplearemos series de 8760 observaciones.

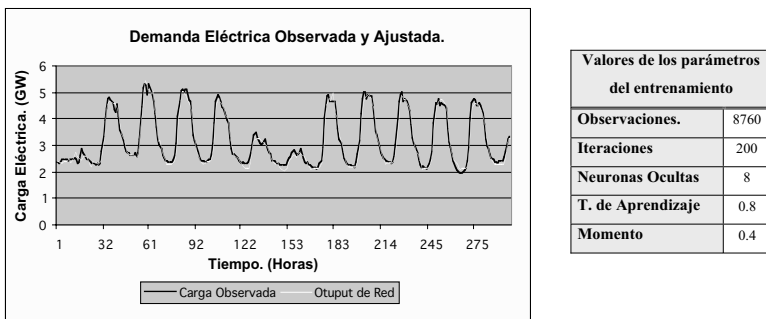


Gráfico 5.1. Curva de carga observada y la ajustada por el modelo en el entrenamiento.

Atendiendo al gráfico se puede apreciar como la red neuronal reproduce con gran precisión la curva de carga original, con lo que podemos intuir que el entrenamiento ha resultado un éxito. Esta misma idea es corroborada por el error cuadrático medio, el cual toma un valor del 0.6% ; por tanto podemos concluir que el entrenamiento ha sido un éxito. Veamos ahora como responde la red ante datos que le son desconocidos.

## 5.2. Predicción con datos no conocidos por la red

A continuación presentaremos a la red el conjunto de datos que habíamos reservado previamente. La red procesará esta nueva información teniendo en cuenta las relaciones aprendidas en el entrenamiento y que guarda (recuerda) en forma de vector de pesos sinápticos. La salida de la red no será más que la predicción. Comparando esta predicción con el valor real de la carga podremos contrastar la capacidad predictiva del modelo.

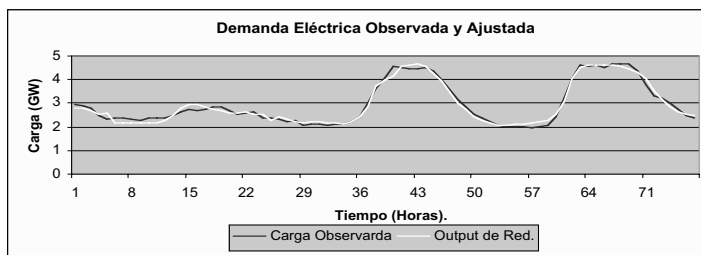


Gráfico 5.2. Representación de la curva de carga observada y la ajustada por el modelo con datos desconocidos.

Atendiendo a la representación gráfica se puede apreciar como el perfil de la curva de carga generada por la red se aproxima a la curva real con una precisión aceptable. En cuanto al error cuadrático medio de la predicción (0,92%), tal y como cabía esperar es ligeramente superior a su homólogo en el entrenamiento, siendo lo suficientemente reducido como para admitir la capacidad predictiva del modelo.

Fase	Datos	Neuronas	$\lambda^{16}$	$\mu^{17}$	Iteraciones	$\sqrt{ECM}$
<b>Entrenamiento</b>	8760	8	0.8	0.5	200	7,73%
<b>Predicción</b>	360	8	0.8	0.5	200	9,64%

Tabla 5.1. Resultados de la predicción.

<sup>16</sup> Tasa de aprendizaje.

<sup>17</sup> Momento.

### 5.3. Comparación con otros modelos

Finalmente, una vez superadas las fases de entrenamiento y validación, podemos comparar los resultados obtenidos a partir del modelo neuronal con los ofrecidos por otras técnicas más convencionales. De este modo podremos determinar en que medida las RNA superan a las técnicas tradicionales como instrumento de modelización y predicción. A tal efecto hemos considerado los siguientes modelos:

- ✓ **Modelo ingenuo:** Modelo autoregresivo de primer orden,  $y_t = y_{t-1} + \varepsilon$   $\varepsilon \approx N(0, \sigma)$ .
- ✓ **ARMA:** Hemos identificado un modelo  $ARMA(2,2) \times ARMA_s(24,24)$
- ✓ **Alisado Exponencial:** Alisado exponencial de Holt-Winters con  $\alpha=0.5$

Los resultados obtenidos aparecen recogidos en la siguiente tabla:

	RNA	Modelo ingenuo	ARMA	Alisado exponencial
$R^2$ del modelo	0.64	0.90	0.93	0.94
$\sqrt{ECM}$ <sup>18</sup>	9,64%	29.91%	11.95%	14%

Tabla 5.2. Comparación con otros modelos.

Vemos como si bien el  $R^2$  asociado al modelo neuronal es inferior al obtenido con las técnicas convencionales, en la predicción el mismo resulta ser más eficaz. Esto puede deberse al hecho de que las variables que explican la evolución de la demanda eléctrica en el tiempo se encuentran vinculadas mediante relaciones de carácter no lineal. En la medida en que pretendamos capturar dichas relaciones empleando especificaciones lineales (tipo ARMA), estaremos obviando la parte no-lineal de la dinámica del fenómeno, cometiendo por tanto un importante error en el ajuste. Mediante el uso de RNA, al emplear funciones de transferencia no-lineales, estaremos dotando al modelo de la capacidad necesaria para capturar las relaciones no-lineales que determinan la dinámica del fenómeno estudiado.

<sup>18</sup> La raíz del error cuadrático medio está referida a la predicción obtenida con dichos modelos.

## 6. CONCLUSIONES

1) El modelo neuronal propuesto ha sido capaz de predecir la evolución horaria de la demanda eléctrica con una precisión del 98%. Las RNA tipo Perceptrón Multicapa pueden ser empleadas como modelos *no-paramétricos* de predicción económica

2) Queda también demostrada la incidencia que tienen variables como la temperatura o el calendario en la evolución de la curva de carga eléctrica. De cualquier modo no podemos conocer el tipo de relación que vincula dichas variables con la demanda eléctrica. En este sentido la red puede ser entendida como una especie de *caja negra* que procesa la información de entrada para generar una salida que se ajusta a un determinado criterio o norma.

3) Dada la incertidumbre existente en torno al valor que debe de asignarse a los parámetros que definen la red neuronal (número de unidades ocultas, tasa de aprendizaje, momento, etc), se podría estudiar la posibilidad de optimizar dichos valores mediante el uso de algoritmos genéticos. En la misma línea también resultaría interesante investigar en que medida este tipo de procedimientos reduce el riesgo de los mínimos locales asociado a los métodos de entrenamiento basados en el gradiente.

4) Finalmente, resultaría sumamente interesante estudiar la aplicación conjunta de los modelos neuronal y las transformadas wavelet. La idea sería la siguiente: filtrar la serie de datos mediante el empleo “wavelets” para obtener una serie libre de ruido. Sobre la nueva serie aplicaríamos el modelo neuronal correspondiente para poder predecir la evolución futura del fenómeno objeto de estudio.

## BIBLIOGRAFÍA

- BAO, JIE (2000): "Short-term Load Forecasting based on Neural Network and Moving Average". Artificial Intelligence Lab, Dept of Computer Science Iowa State University.
- BARTON, ALLEN H.(1969): "Communities in Disaster: A Sociological Analysis of Collective Stress Situations", Nueva York, Garden City, *Doubleday and Company, Inc.*
- BELTRATTI, A.; MARGARITA, S.; TERNA, P. (1996): "Neural Networks for economic and financial modelling". *International Thomson*.
- C.M., KUAN; WHITE, H.: "Artificial Neural Networks: An Econometric Perspective," *Econometric Reviews*, pp. 13, 1-92.
- DYNES, RUSSELL R. (1970): "Organized Behavior in Disaster", Lexington, Massachusetts: *Heath Lexington Books*.
- FARAGÓ, A.; LUGOSI, G. (1993), "Strong Universal Consistency of Neural Network Classifiers", *IEEE Transactions on Information Theory*, 39, 1146-1151.
- GARCÍA DELGADO, J. LUIS; MYRO, RAFAEL; MARTÍNEZ SERRANO, J.A (1997): "Lecciones de Economía Española". (3ª Edición). *Editorial Civitas*.
- GARCÍA ESTÉVEZ, PABLO (2001): "Aplicaciones de las redes neuronales en las finanzas", *Documento de Trabajo*, Universidad Complutense de Madrid.
- GALLANT, A.R.; WHITE, H. (1992): "On Learning the Derivatives of an Unknown Mapping with Multilayer Feedforward Networks", *Neural Networks*, 5, 129-138.
- GRANGER, C.W.J.: (1991): "Developments in the Nonlinear Analysis of Economic Series," *Scandinavian Journal of Economics*, 93(2), 263-76.
- HEINAMANN, M. (2000): "Adaptative Learning of Rational Expectations using Neural Networks", *Journal of Economic Dynamics and Control*, 24 (5-7), 1007-1026.
- HORNIK, K. (1989): "Multilayer feedforward networks are universal approximators", *Neural Networks*, Vol. 2., pp. 359-366.
- HORNIK, K.; M. STINCHCOMBE; H. WHITE (1990): "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks". *Neural Networks* 3 551- 560.

- JONES, L.K. (1990): "Constructive approximation for neural networks by sigmoidal functions", *Proc. IEEE*, 78, 1586-1589.
- KRÖSE, B; SMAGT, P. (1996): "An introduction to Neural Networks." University of Amsterdam.
- MEDSKER, L.; TURBAN, E.; TRIPPI: "Neural Network Fundamentals for Financial Analysts", *Neural Networks in Finance and Investing* edited by Trippi and Turban, Irwin, USA, Capítulo 1, pp. 329-339, ISBN 1-55738-919-6, 1996.
- MILLER III, W.T.; SUTTON, R.S.; ERBOS, P.J. (1990): "Neural Networks for Control" Cambridge, *Massachusetts Institute of Technology Press*.
- MINSKY, M.; PAPERT, S. (1969): "Perceptrons: An introduction to Computational Geometry", *The MIT Press*. Cambridge, MA.
- MURTO, PAULI (1998): "Neural Network Models for short-term load forecasting". Department of Engineering, Physics and Mathematics. Helsinki University of Technology.
- RIPLEY, B.D. (1994): "Neural Networks and Related Methods for Classification". *Journal of the Royal Statistical Society*, N° 3, pp 409-456.
- RODRÍGUEZ VILLAGARCÍA, CARMEN (Dirección) (1998): "Atlas de la demanda eléctrica española. Proyecto Indel". *Red Eléctrica de España*.
- RUMELHART, R.E.; MC CLELLAND, J.L. (1986): "Parallel Distributed Processing", Cambridge, MA: *MIT Press*.
- SHACHMUROVE, Y. (2000): "Applying Neural Networks to Business, Economics and Finance", Working Paper, pp. 1-7. Department of Economics. The City College of the City University of New York and, The University of Pennsylvania.
- TAN, CLARENCE N.W. (2001): "An Artificial Neural Networks Primer with Financial Applications Examples in Financial Distress Predictions and Foreign Exchange Hybrid Trading System", pp. 51-80. School of Information Technology, Bond University, Gold Coast, QLD 4229, Australia.
- WEIGEND, A.S.; GERSHENFELD, N.A. (1994): "Time Series Predictions: Forecasting the future and Understanding the past", *Addison-Wesley*.
- WHITE, H. (1990): "Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings", *Neural Networks*, 3, 535-550.