

## PREDICCIONES DE LA DEMANDA DE LA ENERGÍA ELÉCTRICA CON DATOS DE LA ACTUAL CRISIS ECONÓMICA Y FINANCIERA. APLICACIÓN A LA REGIÓN CANARIA.

Gabriel Winter, Begoña González-Landín, Antonio Pulido-Alonso, Blas Galván-González y Mustapha Maarouf

Universidad de Las Palmas de Gran Canaria (ULPGC). Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI). División de Computación Evolutiva y Aplicaciones (CEANI).

Recibido: 7/ago/2015 - Aceptado: 15/sep/2015 - DOI: <http://dx.doi.org/10.6036/ES7782>

### *FORECASTING ELECTRICITY CONSUMPTION INCLUDING DATA OF THE PRESENT ECONOMIC AND FINANCIAL CRISIS. APPLICATION TO CANARY ISLANDS*

#### **ABSTRACT:**

The economic development is the most influential factor on the power consumption of each country and each region, in long term estimation. In years of economic and financial crisis like the current one, a great variability of Gross Domestic Product (GDP) and Consumer Price Index (CPI) is observed. Particularly, CPI is sensitive to changes in the price of energy and the establishment of monetary policy. Therefore, the improvement of including CPI, in addition to GDP and population, as an explanatory variable to forecast the electricity consumption is investigated. For electricity companies it is important to have efficient prediction techniques to reduce uncertainty in the energy demand and obtain an optimal and realistic scheduling of the production of electricity. In pursuit of more objective conclusions, estimates are made using prediction methods of different nature, such as Multiple Linear Regression and Multiple Logarithmic Regression, which are classical statistical techniques, Support Vector Machine, which is a statistical learning technique, a Genetic Algorithm, which is an evolutionary computation techniques and an Artificial Neural Network, which is a machine learning technique. As a case study, the prediction of electricity demand in the Canary Islands is considered. It is of great interest for being an insulated electric system. The best prediction results are obtained with techniques which possess a greater capability to emulate nonlinear dependencies of the electricity demand in relation to population, GDP and CPI.

**Keywords:** Electricity Demand, Long-Term Prediction, Multiple Linear Regression, Multiple Logarithmic Regression, Support Vector Machine, Genetic Algorithms, Artificial Neural Networks, Insular Electric System.

#### **RESUMEN:**

La evolución económica es el factor que más influye sobre el consumo eléctrico de cada país y de cada región, en las estimaciones a largo plazo. En años de crisis económica y financiera, como la actual, se constata una gran variabilidad del Producto Interior Bruto (PIB) y del Índice de Precios de Consumo (IPC). Este último sensible a la evolución del precio de la energía y al establecimiento de políticas monetarias. Por ello, en este trabajo se investiga la mejora de incluir el IPC, además del PIB y la población, como variable explicativa en las estimaciones de la demanda del consumo eléctrico. Para las compañías eléctricas es importante disponer de eficientes técnicas de predicción, para poder reducir la incertidumbre de la demanda de energía y obtener una programación óptima y realista de la producción de energía eléctrica. Para obtener conclusiones más objetivas, se realizan estimaciones con métodos de predicción de distinta naturaleza, tales como la Regresión Lineal Múltiple y la Regresión Logarítmica Múltiple, que son técnicas estadísticas clásicas, una Máquina de Vectores Soporte, que es una técnica de aprendizaje estadístico, un Algoritmo Genético, que es una técnica de computación evolutiva, y una Red Neuronal Artificial, que es una técnica de aprendizaje automático. Como caso de estudio se considera la predicción de la demanda de energía eléctrica en la Región Canaria, de gran interés por ser un sistema eléctrico aislado. Se obtienen mejores resultados de predicción con las técnicas de mayor capacidad de emular dependencias no lineales de la demanda de energía eléctrica en relación con la población, el PIB y el IPC.

**Palabras clave:** Demanda de Energía Eléctrica, Predicción a Largo Plazo, Regresión Lineal Múltiple, Regresión Logarítmica Múltiple, Máquina de Vectores Soporte, Algoritmos Genéticos, Redes Neuronales Artificiales, Sistemas Eléctricos Insulares.

## 1.- INTRODUCCION

Son varios los tipos de empresas que requieren las estimaciones a largo plazo de la demanda de energía eléctrica (DEE). Si las predicciones resultan muy bajas podrían dar lugar a deficiencias en el suministro de energía, con inconvenientes para distintos sectores económicos, siendo los costes muy superiores a la energía no abastecida. Si por el contrario, las

estimaciones resultan excesivas, los costes de oportunidad incurridos serán elevados, presentándose una elevada inversión económica improductiva. En ambos casos encarecería el precio que paga el consumidor.

Eficientes técnicas de predicción ayudan a las compañías eléctricas a reducir la incertidumbre de la DEE y a obtener una programación óptima y realista de la producción de energía eléctrica. La evolución económica es el factor más influyente sobre el consumo eléctrico de cada región a largo plazo (uno o varios años, incluso décadas). Una disminución del precio de la energía, como a su vez políticas monetarias expansivas incentivan el crecimiento de la economía. La mejora en la situación general de las economías domésticas da lugar a un aumento del consumo de bienes y servicios. Son pues el PIB y el IPC dos factores que intervienen en el nivel de demanda de consumo de energía eléctrica, y ambos presentan gran variabilidad en años de crisis económica y financiera como la actual. En este trabajo se realiza una comparativa entre varios métodos utilizados en la literatura para la predicción de la demanda eléctrica [1], entre ellos métodos no lineales, no deterministas y de aprendizaje automático. Así, se han utilizado técnicas estadísticas clásicas (Regresión Lineal Múltiple (RLM) y *Regresión Lineal Múltiple Logarítmica* (RLog)), una técnica de aprendizaje estadístico (*Máquina de Vectores Soporte* (MVS)), una técnica de Computación Evolutiva (*Algoritmo Genético* (AG)), y una técnica de Aprendizaje Automático (*Red Neuronal Artificial* (RNA)). Además se investiga el efecto de añadir el IPC como variable explicativa. Como caso de estudio se ha estimado la DEE en la Región Canaria para los años 2015-2020.

En la Sección 2 se trata la importancia y la complejidad de estimar la DEE a largo plazo. En la Sección 3 se exponen los principales fundamentos y características de los métodos considerados en este trabajo. En la Sección 4 se presentan los resultados obtenidos, constatando el interés de incluir el IPC como variable en el ajuste funcional, y analizando una comparativa de la bondad de la predicción. Por último, se presentan las conclusiones alcanzadas.

## 2.- IMPORTANCIA Y CRECIENTE COMPLEJIDAD DE LA ESTIMACIÓN DE LA DEE A LARGO PLAZO

El mercado eléctrico presenta algunas particularidades respecto al resto de suministros energéticos. Al no ser almacenable de forma competitiva ni en cantidades apreciables, se debe producir en cada instante la misma cantidad que se consume más las pérdidas. Diversas técnicas de almacenamiento podrían modificar el panorama en las próximas décadas.

El suministro eléctrico con una elevada fiabilidad, es algo imprescindible en una economía moderna. Los sistemas eléctricos se hallan fuertemente conectados. España se halla interconectada con 4 países: Marruecos, Portugal, Francia y Andorra, intercambiando fluido eléctrico con cualquiera de ellos. Las Islas Canarias están formadas por 6 sistemas eléctricos independientes, y cada uno se autoabastece. Esto hace que la implicación de la predicción de carga en Canarias puede dar lugar a una situación mucho más crítica que en el resto del territorio español.

En el pasado, antes del actual periodo de crisis económica, sólo ha existido predominantemente una tendencia significativamente creciente en la DEE, por el mayor uso de variados dispositivos eléctricos que han facilitado la vida, sin importar el consumo. Actualmente, aparecen una serie de nuevos factores que habría que considerar, pues podrían irrumpir con fuerza, introduciendo grandes cambios en la evolución de la demanda futura:

- La incertidumbre económica tras la crisis del 2008.
- La pobreza energética de la población, agravada por subidas bruscas del precio de la electricidad, unido a la alta tasa de desempleo y precariedad laboral.
- El incremento en la eficiencia de los equipos e instalaciones.
- La incorporación del vehículo eléctrico.
- La implantación de sistemas de gestión de la demanda, que con la instalación de contadores inteligentes desplazará temporalmente el consumo eléctrico, pero podrá afectar también a la cuantía del mismo, al hacer tomar más conciencia de ello.
- La evolución de las bombas de calor aerotérmicas podría desplazar a algunos combustibles fósiles [2].

Antes de detallar los métodos, es de interés destacar algunas consideraciones:

- En la Fig. 1 se puede observar que mientras que en el periodo previo al inicio de la actual crisis económica y financiera había una fuerte relación lineal entre la DEE y el PIB en Canarias, durante la misma se ha acentuado el comportamiento no lineal entre ambas variables.
- En un estudio del pronóstico del precio de la energía eléctrica en Colombia [3], se menciona que los análisis estadísticos preliminares indicaban que el IPC, la tasa de cambio real (TRC) y el precio internacional de combustibles no afecta significativamente a la variabilidad de los precios de la energía eléctrica. El contexto de este trabajo viene a ser una función inversa a la nuestra, pues trata del precio de la energía en función de la DEE.
- En un estudio de la DEE de Nueva Zelanda [4] con el modelo de regresión lineal múltiple se observó una alta correlación entre el PIB y la población, generándose cierta multicolinealidad, no obstante se menciona que la eliminación de algunas de estas variables no produjo mejores modelos, por lo que se optó por incluirlas todas. Es importante señalar que el problema de multicolinealidad sólo constituye un problema en métodos de regresión lineal, que resuelvan sistemas de ecuaciones algebraicos que puedan ser mal condicionados.
- En [5] se usó como variable explicativa sólo una combinación de variables, por ejemplo:  $(\text{PIB}/\text{IPC}) \times \text{POB}$ , y se obtuvo una mayor correlación, comprobando en ese caso que cuando se combina dos o más variables, la nueva variable cuenta con mayor correlación con la DEE.

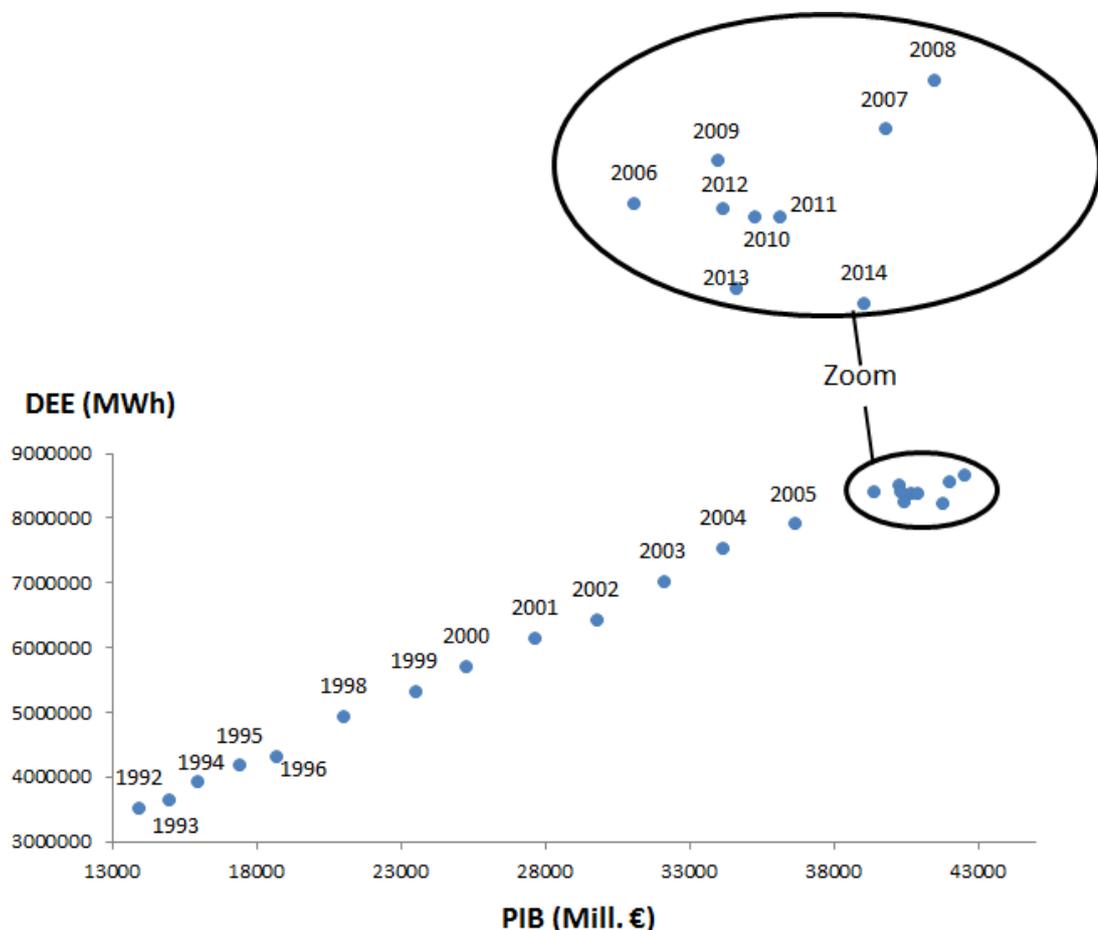


Figura 1. DEE respecto al PIB (Región Canaria, 1991-2014)

### 3.- MÉTODOS ESTADÍSTICOS, DE COMPUTACION EVOLUTIVA Y DE APRENDIZAJE AUTOMÁTICO

Los métodos de estimación pueden clasificarse de distintas formas, una de ellas es en dos grandes grupos: los denominados de inteligencia artificial y los paramétricos. Entre los métodos de inteligencia artificial se puede citar

principalmente: AG, RNA, máquinas de soporte vectorial, lógica difusa y sistemas expertos. Los métodos paramétricos se basan en modelos econométricos, principalmente: regresión múltiple, regresión logarítmica o bien series temporales. Los métodos basados en la inteligencia artificial presentan como principal ventaja la habilidad de emular las relaciones no lineales. Entre los más populares y de demostrada eficiencia en resolver problemas altamente no lineales están los AG y las RNA.

A fin de facilitar la discusión de los resultados y las conclusiones del trabajo, es preciso contextualizar los distintos métodos en relación a sus fundamentos y operatividad, aflorando en un contexto de conocimiento multidisciplinar información precisa que facilite el uso práctico y aplicación de los mismos, y nos permite ver con claridad las ventajas e inconvenientes de unos métodos respecto a otros.

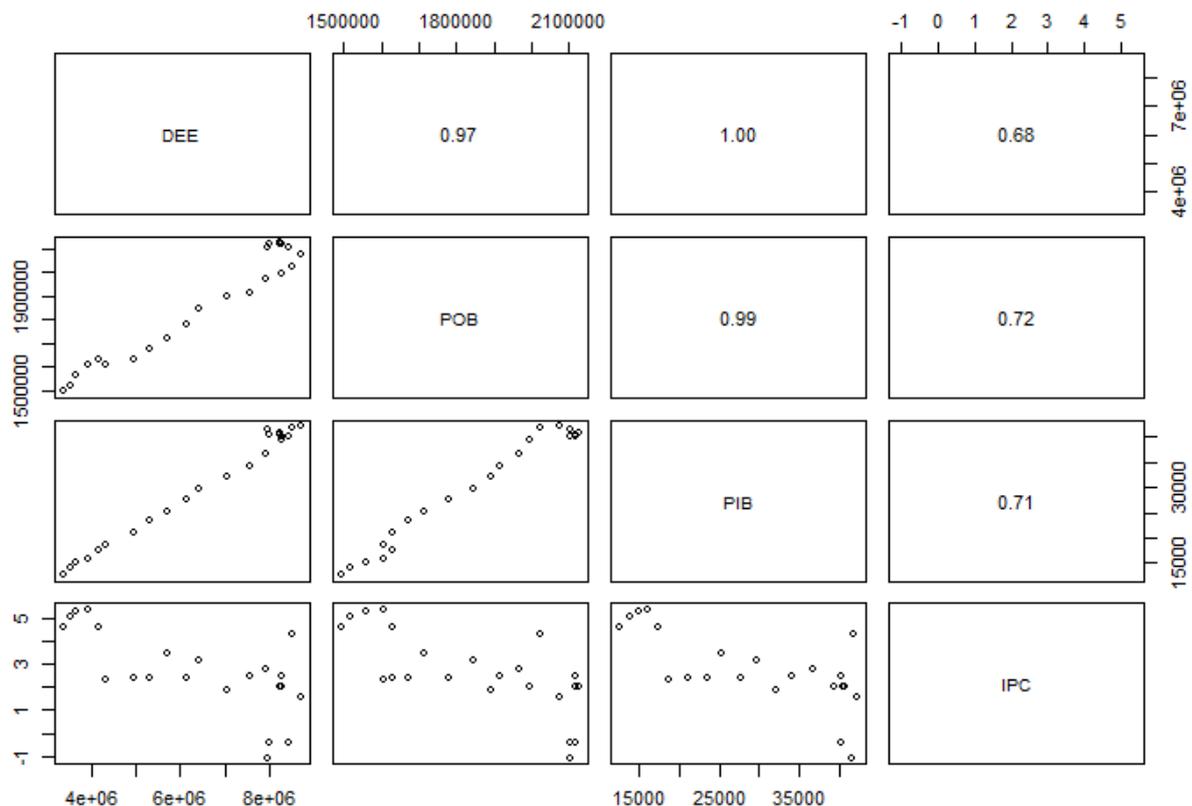


Figura 2. Gráfico matricial de dispersión de las variables de estudios en el periodo 1991-2014. Región Canaria.

### 3.1.- METODOS ESTADÍSTICOS CLÁSICOS DE REGRESIÓN

El análisis de regresión es una técnica estadística utilizada para analizar la dependencia entre una variable, la dependiente, y una variable (regresión simple) o más variables (regresión múltiple) adicionales, las explicativas.

En este apartado, se aplica regresión lineal múltiple a la función:

$$DEE = \beta_1 POB + \beta_2 PIB + \beta_3 IPC + \beta_4 \quad (1)$$

donde POB es la población y los parámetros  $\beta_i, i = 1,2,3,4$ , son coeficientes que se determinan con el ajuste mínimos cuadrados.

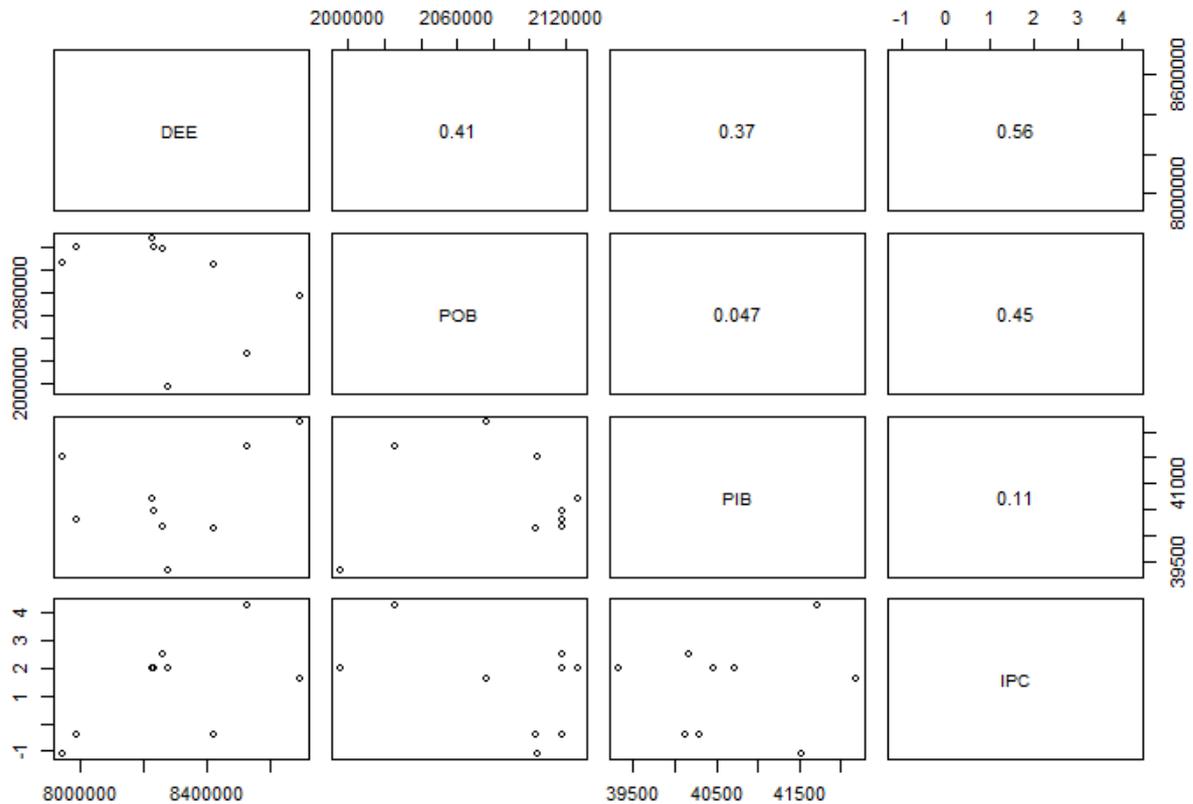


Figura 3. Gráfico matricial de dispersión de las variables de estudios en el periodo 2006-2014. Región Canaria.

Si bien, si consideramos todos los datos del periodo de estudio se observa una fuerte dependencia lineal no sólo de la DEE con la población y el PIB sino también entre estas dos variables explicativas (ver Fig. 2), si consideramos los datos del periodo 2006-2014, esta colinealidad desaparece, tal y como se puede apreciar en la Fig. 3. Es decir, en el periodo con años de la crisis económica se acentúa la dependencia no lineal entre la DEE y las variables explicativas. Por ello, se decide aplicar también regresión lineal múltiple logarítmica (RLog):

$$\log(DEE) = \beta_1 \log(POB) + \beta_2 \log(PIB) + \beta_3 \log(IPC) + \beta_4 \quad (2)$$

En ocasiones se emplean los logaritmos como simple estrategia de transformación matemática tendente a reducir la dispersión original del conjunto de datos. Tiene la ventaja de reducir el problema de la heterocedasticidad, al reducirse la escala de medida de los datos.

Si bien la Ec. (2) equivale a una expresión no lineal exponencial al despejar la DEE, el método de cálculo usado para la obtención de los coeficientes es una regresión lineal en el plano transformado, considerando las variables logaritmo de las variables explicativas como las variables del cálculo.

### 3.2.- TÉCNICAS DE COMPUTACIÓN EVOLUTIVA Y DE APRENDIZAJE AUTOMÁTICO

Los AG, las RNA y las MVS se caracterizan por ser capaces de modelar automáticamente relaciones no lineales. A continuación se destacan algunas características generales de los modelos aplicados a la estimación de la DEE y se dan pautas para el uso eficientes de los métodos aplicados al caso de estudio.

### 3.2.1.- Capacidades y buen uso de los AG

Los Algoritmos Evolutivos (AE) son métodos de búsqueda global, inspirados en procesos de la eficiente evolución de los seres vivos. Su simplicidad y la habilidad de buscar eficientemente las mejores soluciones son características que motivan, sin lugar a dudas, el uso de los mismos. Los AE surgen a principios de los 60 con la Programación Evolutiva [6], empleando únicamente el operador mutación. En 1965, Schwefel [7] y Rechenberg [8] plantean las bases de las Estrategias Evolutivas. Entre ellos los AG han sido los más aplicados para resolver problemas de optimización en muchas disciplinas de la ciencia e ingeniería [9]. Fue John Holland [10], quien puso en escena internacional a los AG incorporando a su vez la recombinación o cruce genético para generar diversidad de nuevas soluciones candidatas a ser las óptimas en el símil de obtención de descendientes.

Los AG, como los AE, operan en una dinámica de poblaciones, evolucionando en el tiempo una población de soluciones candidatas a ser las óptimas, bajo un proceso iterativo de paso de una generación a otra, que hace evolucionar favorablemente las soluciones aproximadas hasta convertirlas en las óptimas. Son métodos heurísticos de búsqueda, pues en cada iteración hay operaciones, principalmente cruce y mutación, que se realizan con probabilidades asignadas.

Se considera en este trabajo un AG que usa operadores bien conocidos, simples de implementar, resultando un AG que con poco esfuerzo de calibración en sus parámetros generan buenas soluciones y de forma robusta (una vez fijados los parámetros, cada ejecución del mismo da lugar a obtener soluciones muy similares):

- operador de selección por torneo [11] de fácil implementación y de calibración de su presión de selección. Realiza, al inicio de cada iteración del AG, sucesivos torneos entre dos o tres de las soluciones candidatas (individuos) de la población, con degradación en sus evaluaciones en la función objetivo por uso de la función de compartición de recursos (sharing function). Así se penaliza la aglomeración de soluciones candidatas cercanas a óptimos locales (símil a debilitarse los individuos en mayor medida por compartir en mayor número y con más cercanía entre ellos los recursos).
- operador de cruce o recombinación denominado antitético [12], que a partir de dos soluciones candidatas genera por combinación lineal entre ellas otras dos. Dos padres  $p_1$  y  $p_2$  generan los descendientes:

$$h_1 = ap_1 + (1 - a)p_2$$

$$h_2 = (1 - a)p_1 + p_2$$

Donde  $a$  es un número aleatorio (peso) comprendido en el intervalo  $[-0,5; 0,5]$ . Así, actúa interpolando cuando está en el intervalo  $[0,1]$ , o extrapolando cuando está en los intervalos  $[-0,5; 0]$  ó  $[1; 1,5]$ .

- La mutación considerada es una de las más simples. Consiste en sumar un número aleatorio comprendido entre -1 y 1 a la variable que se muta.
- Los parámetros de ejecución han sido los siguientes:
  - ✓ Codificación de las variables: real.
  - ✓ Número de soluciones candidatas: 100
  - ✓ Probabilidad de cruce: 0,55
  - ✓ El parámetro  $a$  se genera aleatoriamente en cada cruce.
  - ✓ Probabilidad de mutación: 0,45
  - ✓ Parámetro de compartición de recursos:  $\sigma_{share} = 10$ . Representa la distancia mínima entre dos soluciones candidatas a ser el mínimo global.

La única operación matemática relevante que realizan es simplemente evaluar la función objetivo del problema de optimización. En este trabajo, se ha considerado la siguiente función objetivo no lineal, por dar representaciones gráficas adaptables como caso de funciones exponenciales, y con el fin de aprovechar la ventaja de compresión de los valores numéricos y así del espacio de búsqueda:

$$DEE = \exp \left( \frac{a_1 \log(POB)}{(a_2 + \log(POB))} + \frac{a_3 \log(PIB)}{(a_4 + \log(PIB))} + \frac{a_5 IPC}{(a_6 + IPC)} + a_7 \right) \quad (3)$$

donde los  $a_i$ ,  $1 \leq i \leq 7$ , son coeficientes reales que serán los óptimos en la minimización de la función de la **DEE**.

El buen uso de un AG radica en que operen en la búsqueda de los óptimos con equilibrio entre lo que se denomina explotación y exploración del espacio de búsqueda. La explotación es la intensidad con la que se buscan buenas soluciones candidatas en cada iteración, y el mayor nivel de producción de diversidad de soluciones candidatas induce un comportamiento del AG con mayor exploración en el espacio de búsqueda.

### 3.2.2.- Capacidades y buen uso de las RNA

Las RNA son modelos que tratan de imitar las habilidades del cerebro humano, ofreciendo ventajas en cuanto a su capacidad de aprendizaje, con funcionamiento en paralelo y tolerancia a fallos por información imprecisa [13]. Se originaron en 1943, por el trabajo de McCulloch y Pitts, pero el trabajo de Rumelhart, Hinton y Williams en 1986 fue el que desencadenó un crecimiento exponencial de la literatura neuronal [13]. Algunas de estas técnicas se han aplicado al problema de la predicción de los precios de la energía en el mercado eléctrico español [14] dando lugar a resultados competitivos con los resultados obtenidos de la aplicación de otras técnicas.

Una RNA con una única capa oculta puede aproximar hasta el nivel deseado cualquier función continua en un intervalo, por lo tanto las RNA multicapa son aproximadores universales de funciones, operando este tipo de red con una analogía matemática al desarrollo en funciones, que recuerda a la expresión de series de Fourier, pero aquí es un desarrollo en funciones sigmoideas [15].

A continuación se describen los pasos que conlleva el diseño de una red neural para aproximar una función.

#### 3.2.2.1.- Definir la arquitectura

La selección de la arquitectura de la red consiste básicamente en determinar las entradas, preparación de los datos, tamaño de la red y función de activación. Una RNA de única capa oculta puede aproximar hasta el nivel deseado para cualquier función continua en un intervalo [16]. En este trabajo, la RNA contiene una capa oculta (Ver Fig. 4).

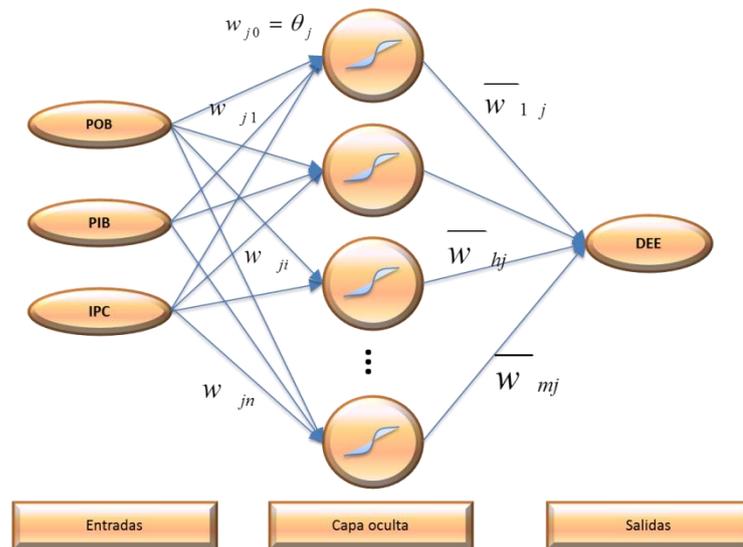


Figura 4. Estructura de la RNA

Los datos se normalizan dentro del intervalo  $[-1,1]$ :

$$X = \frac{2x - (x_{max} + x_{min})}{(x_{max} - x_{min})} \quad (4)$$

siendo  $x_{max}$  (resp.  $x_{min}$ ) el máximo de los  $x$  (resp. mínimo).

La ecuación que define la relación entre la salida deseada, en este caso la DEE, y las entradas: PIB, POB, IPC, es:

$$DEE = \sum_{s=0}^l \overline{w_{hs}} k_s = \sum_{s=0}^l \overline{w_{hs}} f(p_s) \quad (5)$$

donde  $l$  es el número de neuronas ocultas,  $p_s = w_{s1}POB + w_{s2}PIB + w_{s3}IPC$  es la entrada de la neurona oculta  $s$ ,  $1 \leq s \leq l$ ,  $f$  es la función de activación de la red,  $k_s = f(p_s)$  es la salida de la neurona oculta  $s$ , y los coeficientes  $w_{hs}$ ,  $\overline{w_{hs}}$  son los pesos de la RNA.

Se entrena la red para un número limitado de episodios, variando en cada entrenamiento el número de neuronas ocultas. El número de neuronas con una convergencia satisfactoria se toma como tamaño de la capa oculta.

Las funciones de salida de las neuronas ocultas deben ser senoidales [17-18]. Generalmente, la función más usada es la tangente hiperbólica (anti-simétrica) o la sigmoideal (simétrica), que se definen como:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (6)$$

$$\text{sigm}(x) = \frac{1}{e^{-x} + 1} \quad (7)$$

En este trabajo se usa la función (6) por ser la preferiblemente usada en la literatura [13, 19].

### 3.2.2.2.- Algoritmo de entrenamiento

Los algoritmos utilizados para ajustar los pesos de la red, con el fin de minimizar el error, se basan en métodos de optimización no lineales locales [13], aunque también se han utilizado técnicas de optimización global para el entrenamiento de las RNA [20]. Los métodos de optimización locales no lineales son algoritmos que se basan en el gradiente descendiente, como el algoritmo de retropropagación estándar [13], y los algoritmos de segundo orden utilizan el Hessiano [21]. En este trabajo se utiliza el algoritmo “Resilient Backpropagation” (Rprop) [22-23]. El Rprop es un método rápido de primer orden, y sus requisitos de memoria son menores que los métodos de segundo orden.

El algoritmo Rprop funciona en procesamiento por lotes, es decir, los pesos se actualizan basándose en los valores del error correspondientes al error total sobre el conjunto de entrenamiento. Si la  $p^{th}$  salida deseada es  $t^{(p)}$  y la salida obtenida mediante la red es  $y^{(p)}$  para el conjunto de  $p^{th}$  entrada, y si hay  $P$  número de ejemplos de entrenamiento, entonces el error se calcula mediante el error cuadrático medio (ECM) definido como:

$$ECM = \frac{1}{P} \sum_{p=1}^P (t^{(p)} - y^{(p)})^2 \quad (8)$$

Para cada procedimiento de inicialización de la RNA, 30 redes diferentes (con distintos pesos aleatorios iniciales) son entrenadas durante un máximo de 200 iteraciones para evitar el sobre-entrenamiento.

El rendimiento sobre los datos de entrenamiento y los conjuntos de datos de prueba se calcula (para cada red entrenada) usando el ECM y se conserva la mejor red cuyo ECM es más pequeño.

### 3.2.3.- Capacidades y buen uso de las MVS

Las Máquinas de Vectores Soporte tienen su origen en una generalización del método “generalized portrait” propuesto por Vapnik y Lerner [24] para resolver problemas de clasificación linealmente separables [25]. Inicialmente se aplicó al problema de clasificación en dos clases, buscando el hiperplano que separara los puntos de una clase respecto de la otra, y que a su vez tuviera la máxima distancia (margen) con los puntos que están más cerca del hiperplano. Al vector formado por los puntos más cercanos al hiperplano se le llama vector de soporte. Las MVS surgen así como clasificadores lineales basados en teorías de aprendizaje estadístico y en el marco del principio de minimización del riesgo estructural, frente a los que usan el principio de minimización del riesgo empírico, como es el caso de las redes neuronales. Un coeficiente de regularización,  $C$ , pondera la minimización del riesgo estructural frente a la minimización

del riesgo empírico, otorgando a los MVS una buena capacidad de generalización a costa de cometer más errores en la fase de entrenamiento y facilitando formular las MVS como problemas de optimización cuadrática de funcional convexo y, por tanto, con un único óptimo global. A fin de superar las limitaciones del aprendizaje lineal, se recurre a formular los MVS con funciones denominadas Kernel, utilizando una proyección o transformación,  $\Phi$ , en general no lineal. Las funciones kernel se definen como el producto escalar de las imágenes de dicha transformación,  $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$ . La formulación de las MVS con las funciones kernel facilita la implementación computacional. Las soluciones de problemas de clasificación y regresión buscadas por algoritmos basados en el kernel, operan así con las funciones  $\Phi$  sólo implícitamente, sin ser preciso conocer sus expresiones. Tratan de encontrar fundamentalmente:

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) \quad (9)$$

donde  $\alpha_i$ ,  $1 \leq i \leq n$ , son coeficientes que se determinan con el modelo y  $n$  es el número de puntos de entrenamiento.

Entre los kernels más utilizados tenemos:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right) \quad (\text{Función Gaussiana; } \|\cdot\| \text{ es la norma euclídea}) \quad (10)$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_j^T \mathbf{x}_i \quad (\text{Función Lineal}) \quad (11)$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = (t + \mathbf{x}_j^T \mathbf{x}_i)^d \quad (\text{Función Polinómica}) \quad (12)$$

Una de las ventajas de las Máquinas de Vector Soporte es su eficiencia con pocos datos de entrenamiento.

## 4.- APLICACIÓN y RESULTADOS

En el problema se trata de encontrar buenos ajustes funcionales de la DEE. El objetivo es encontrar métodos con mayor bondad de ajuste, siendo lo más importante que den buenas predicciones en años futuros. A fin de evaluar en el presente el poder predictivo de un método, se reserva varios años históricos, entre ellos algunos recientes, como años-test. Éstos no son utilizados para construir el ajuste funcional del método. Es decir, el ajuste o entrenamiento se hace con los datos de los restantes años. Los años test (elegidos aleatoriamente) fueron: 1992, 2001, 2004, 2009 y 2014.

Señalar que, el hecho de que uno de los años-test sea 2009, el primero de la crisis económica que rompe de forma acusada la tendencia creciente de la DEE de muchos años anteriores (ver Fig. 1), no ayuda al ajuste funcional de los métodos en comparación a si hubiese sido utilizado como año de ajuste funcional o entrenamiento.

### 4.1.- CASO DE ESTUDIO

El estudio propuesto tiene por objeto estimar la DEE en Canarias. Los valores históricos de la misma se han obtenido del Instituto Canario de Estadística, ISTAC, <http://www.gobiernodecanarias.org/istac/>, tomando los datos del consumo de energía eléctrica. Sólo se dispone de datos a partir del año 1991.

A fin de poder realizar la comparación de los diferentes métodos detallados en la Sección 3, se han considerado como entradas a los mismos los valores, en las Islas Canarias, del PIB, del IPC y de la población (POB), desde el año 1991 hasta el año 2014 (disponibles en el Instituto Nacional de Estadística, INE, <http://www.ine.es>), y la variable de salida es la DEE. En el periodo 1991-2014 se observa colinealidad entre las variables explicativas POB y el PIB (ver Fig. 2), sin embargo, ésta desaparece si nos centramos en el periodo más reciente 2006-2014 (ver Fig. 3).

Los valores de las predicciones de la población se han tomado de la *Proyección de la población española* desde el año 2014 hasta 2064, elaborado por el INE. Respecto al IPC y el PIB, se ha empleado una estimación de la *Planificación energética indicativa* (Ley 2/2011, de 4 marzo de Economía Sostenible).

## 4.2.- RESULTADOS

El uso de métricas es necesario para la validez de las conclusiones y para efectuar una comparación eficaz entre los métodos.

Como medida del error, atendiendo a que el número de datos es limitado, se emplea el error medio absoluto porcentual conocida como MAPE:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{100\varepsilon_i}{DEE_i} \right| \quad (13)$$

siendo  $\varepsilon_i = DEE_i - \overline{DEE}_i$  ( $DEE_i$  es la demanda real y  $\overline{DEE}_i$  es la demanda estimada mediante el método en el año  $i$ ).

A continuación se procede mostrar los resultados de los modelos para estimar la DEE, teniendo dos casos de entradas:

- ✓ Primer caso: POB y PIB.
- ✓ Segundo caso: POB, PIB e IPC.

### Método RLM

En este caso se calculan los coeficientes de la Ec. (1) y se obtuvo la siguiente expresión:

$$\overline{DEE} = 218,41 \text{ PIB} - 1,84 \text{ POB} + 19.947,86 \text{ IPC} + 32.372.363 \quad (14)$$

El coeficiente de la población es negativo, lo que es indicativo de la colinealidad existente entre las variables POB y PIB (ver Fig. 2).

### Método RLog

Se calculan los coeficientes de la Ec. (2). Como en los años de crisis hay valores negativos para el IPC, todos mayores que  $-5$ , para poder aplicar este método se ha decidido trasladar los datos sumando 5 al IPC de cada uno de los años, de forma que ahora todos los valores ya son positivos.

Se ha obtenido la siguiente expresión:

$$\log(\overline{DEE}) = 0,86 \log(\text{PIB}) - 0,13 \log(\text{POB}) + 0,04 \log(\text{IPC} + 5) + 8,69 \quad (15)$$

### Algoritmo Genético

En este apartado se calculan los coeficientes de la Ec. (3), que es una expresión no lineal que se propone por primera vez en la literatura para la estimación de la demanda eléctrica. La expresión obtenida es:

$$DEE = \exp \left( \frac{-1.083,2 \log(\text{POB})}{(4.077,267 + \log(\text{POB}))} + \frac{-1.209,1 \log(\text{PIB})}{(-1.373,4 + \log(\text{PIB}))} + \frac{0,0058 \text{ IPC}}{(1,197 + \text{IPC})} + 10,368 \right) \quad (16)$$

donde el espacio de búsqueda último considerado, tras unas ejecuciones previas, ha sido:

$$\begin{aligned} -4.000 &\leq a_1, a_3, a_4 \leq -100 \\ 1.000 &\leq a_2 \leq 7.000 \\ -20 &\leq a_5, a_6, a_7 \leq 20 \end{aligned}$$

## Redes Neuronales Artificiales

Aquí los resultados se obtuvieron con un tipo de predicción directa con la RNA descrita en la Tabla 1, obtenida a partir del aprendizaje realizado con el algoritmo Rprop [22] aplicado a nuestra base de datos.

Estructura de la RNA	Configuración
Transformación datos iniciales	[-1,1]
Número de neuronas (1ª capa, 2ªcapa, 3ªcapa)	3-7-1
Función activación oculta / salida	Tanh / Lineal
Algoritmo	Rprop
Episodios (máx.)	200

Tabla 1. Estructura de la RNA utilizada

## Máquinas de Vector Soporte

En este caso los resultados se obtuvieron con la función *ksvm* de la librería **kernlab** en R [26], cuyos parámetros se recogen en la Tabla 2.

Parámetros	Valores
Formulations	$\epsilon$ -SVR
Kernels	Gaussian / Lineal
Optimizer	SMO
Model Selection	Hyper-parameter estimation for Gaussian kernels
$\epsilon$	0,1
C	3; 5; 10; 15; 20; 25; 30
$\sigma$	0,01; 0,05; 0,1; 0,2; 0,5; 0,7; 0,9

Tabla 2. Parámetros utilizados en la función *ksvm* de la librería *kernlab* en R.

En la Tabla 3 se presentan los resultados de error obtenidos. Añadir el IPC como variable independiente mejora el error de las predicciones de los datos test con los métodos AG, RNA y MVS (Gaussiana) que son métodos de ajuste funcional no lineal. A continuación se presentan únicamente los resultados de los métodos aplicados al segundo caso (con POB, PIB e IPC).

	Mape % (Datos de Ajuste)		Mape % (Datos Test)	
	POB, PIB	POB, PIB, IPC	POB, PIB	POB, PIB, IPC
RLM	1,52	1,25	3,04	3,41
RLog	1,46	1,46	1,66	2,46
MVS (Lineal)	1,45 (C=30)	2,06 (C=3)	3,42 (C=30)	3,63 (C=3)
MVS (Gaussiana)	2,06 (C=5, $\sigma=0,5$ )	2,29 (C=10, $\sigma=0,05$ )	2,38 (C=5, $\sigma=0,5$ )	2,09 (C=10, $\sigma=0,05$ )
AG	1,48	1,46	2,77	1,71
RNA	0,82	0,65	1,58	1,22

Tabla 3. Evaluación del error de los métodos con /sin IPC

En la Tabla 4 se presentan las estimaciones futuras de la demanda eléctrica desde 2015 hasta el año objetivo 2020, obtenidas con cada uno de los modelos considerados en el presente estudio, y a partir de las últimas predicciones de la

población, el PIB y el IPC, en dicho periodo para Canarias, publicadas por los organismos e instituciones citados en la Sección 4.1.

Con el AG se ha usado software propio programado en lenguaje ANSI C. En el caso de las RNA se ha empleado la librería libre **encog** (<http://www.heatonresearch.com/encog>). Y con el resto de los métodos se ha usado el software libre R (<https://www.r-project.org/>).

Año	POB	PIB	IPC	DEE (GWh)					
	Nº de habitantes	(Mill. €)	(%)	RLM	Rlog	MVS (Gaussiana) C = 10; $\sigma = 0,05$	MVS (Lineal) C = 3	AG	RNA
2015	2.127.891	42.437	1,8	8.654,230	8.609,778	8.572,495	8.725,847	8.612,651	8.605,344
2016	2.130.870	43.073	1,8	8.785,770	8.712,794	8.618,520	8.833,802	8.724,213	8.707,273
2017	2.133.853	43.719	1,8	8.921,383	8.822,908	8.671,168	8.950,171	8.837,880	8.809,547
2018	2.136.841	44.375	1,7	9.057,109	8.928,386	8.712,969	9.061,564	8.952,319	8.897,254
2019	2.138.123	45.041	1,7	9.200,124	9.042,180	8.762,588	9.180,825	9.070,874	8.994,387
2020	2.139.405	45.716	1,7	9.345,320	9.157,424	8.810,961	9.301,868	9.191,003	9.058,278

Tabla 4. Estimaciones futuras del 2015-2020 usando los distintos modelos considerados en el presente artículo, con las tres variables explicativas.

## 5.- DISCUSION

Motivado por las grandes oscilaciones del PIB e IPC durante la actual crisis económica, en este trabajo se analiza la mejora de la predicción de la demanda de energía eléctrica, en la Región Canaria, al incorporar el IPC, por sí sola, sin estar multiplicada o dividida por otras variables, como una variable independiente o explicativa en los modelos considerados. Se incorpora pues el IPC como una variable explicativa más, junto a las tradicionales: población y PIB; y los datos recopilados se ajustan con métodos de predicción de distinta naturaleza.

Con los métodos no lineales, el incluir o no el IPC como variable explicativa, es significativo. De hecho, su inclusión mejora los resultados de predicción de los métodos no lineales.

El AG se usó para hallar los coeficientes óptimos para una expresión matemática no lineal. Dicha expresión matemática es la primera vez que se propone en la comunidad científica y tecnológica.

## BIBLIOGRAFÍA

- [1] Suganthi L, Anand AS. "Energy models for demand forecasting - A review". *Renewable and Sustainable Energy Reviews*, 2012, Vol.16-2 p.1223-1240. (DOI: <http://dx.doi.org/10.1016/j.rser.2011.08.014>)
- [2] Blokhuis E, Brouwers B, van der Putten E, et al. "Peak loads and network investments in sustainable energy transitions". *Energy Policy*, 2011. Vol. 39-10, p.6220-6233. (DOI: <http://dx.doi.org/10.1016/j.enpol.2011.07.021>. ISSN 0301-4215)
- [3] Jorge BM, Edwin R, Esteban V, et al. "Modelo para el pronóstico del precio de la energía eléctrica en Colombia". *Lecturas de Economía*, 2012, N° 77p. 91-127. (ISSN-e 0120-2596)
- [4] Mohamed Z, Bodger P. "Forecasting electricity consumption in New Zealand using economic and demographic variables". *Energy*, 2005, Vol. 30-10 p.1833-1843. (DOI: <http://dx.doi.org/10.1016/j.energy.2004.08.012>)
- [5] Wagdy M, Moenes M, Mahmoud H, et al. "Long Term Load Forecasting for the Egyptian Network using ANN and Regression Models". *21st International Conference on Electricity Distribution*. CIREN. Frankfurt, 6-9 June 2011. Paper 0043.
- [6] Fogel LJ, Owens AJ, Walsh MJ. *Artificial Intelligence through Simulated Evolution*. John Wiley and Sons, Chichester, UK. 1966.
- [7] Schwefel, HP. "Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie". *Interdisciplinary systems res*Birkhaeuser, 1977.
- [8] Rechenberg I. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Frommann-Holzboog Verlag.1973.
- [9] Goldberg, DE. "Genetic Algorithms in Search, Optimization and Machine Learning". First Edition. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1989. ISBN: 0201157675
- [10] Holland JH. "Genetic algorithms and the optimal allocations of trials". *Siam Journal on Computing – SIAMCOMP*. 1973, Vol.2-2. p.88-105. (DOI: <http://dx.doi.org/10.1137/0202009>)

- [11] Brindle A. "Genetic algorithms for function optimization". Technical Report TR81-2, Department of Computer Science, University of Alberta, Edmonton.
- [12] Michalewicz Z. "Genetic Algorithm + Data Structure = Evolution Programs". Third Edition. Berlin Heidelberg, New York: Springer-Verlag, 1996. ISBN: 978-3-662-03315-9
- [13] Haykin S. "Neural Networks: A Comprehensive Foundations". Prentice Hall PTR, Upper Saddle River, NJ, USA. 1998. ISBN 81-7808-300-0
- [14] Cybenko G. "Approximation by superpositions of a sigmoidal function". *Math. Control, Signals, and Systems*, 1989. Vol.2-4. p.303-314. (DOI: <http://dx.doi.org/10.1007/BF02551274>)
- [15] Bonifacio M, Alfredo SM. *Redes neuronales y Sistemas Borrosos*. 3ª edición. RA-MA EDITORIAL. 2006, p.65. ISBN: 978-84-7897-743-7.
- [16] Funahashi K. "On the approximate realization of continuous mapping by neural networks". *Journal of Neural Networks*, 1989. Vol. 2-3. p.183-192. (DOI: [http://dx.doi.org/10.1016/0893-6080\(89\)90003-8](http://dx.doi.org/10.1016/0893-6080(89)90003-8))
- [17] Duch W, Jankowski N. "Survey of neural transfer functions". *Neural Computing Surveys*, 1999. Vol.2. p.163-212. <http://www.icsi.berkeley.edu/jagota/NCS>
- [18] Chandra P. "Sigmoidal function classes for feedforward artificial neural networks". *Neural Processing Letters - NPL*. 2003. Vol.18-3. p.205-215. (DOI: <http://dx.doi.org/10.1023/B:NEPL.0000011137.04221.96>)
- [19] LeCun Y, Bottou L, Orr GB, et al. "Efficient backprop". In *Neural Networks: Tricks of the trade*, ser. LNCS:1524, Eds. Berlin: Springer. 1998. p.9-50. (DOI: [http://dx.doi.org/10.1007/3-540-49430-8\\_2](http://dx.doi.org/10.1007/3-540-49430-8_2). [http://dx.doi.org/10.1007/3-540-49430-8\\_2](http://dx.doi.org/10.1007/3-540-49430-8_2))
- [20] Han F, Zhu JS. "Improved particle swarm optimization combined with backpropagation for feedforward neural networks". *International Journal of Intelligent Systems*, 2013. Vol. 28-3. p.271-288. (DOI: <http://dx.doi.org/10.1002/int.21569>)
- [21] Rumelhart DE, Hinton GE, Williams RJ. "Learning internal representations by error propagation". In *Parallel Distributed Processing Foundations*. Eds. Cambridge: MIT Press. 1987. Vol. 1. p.318-362. ISBN 0-262-68053-X
- [22] Riedmiller M, Braun H. "A direct adaptive method for faster backpropagation learning: The RPROP algorithm". *Neural Networks, IEEE International Conference on*. 1993. Vol.1. p.586-591. (DOI: <http://dx.doi.org/10.1109/ICNN.1993.298623>)
- [23] Riedmiller M. "Advanced supervised learning in multi-layer perceptrons from backpropagation to adaptive learning algorithms". *Computer Standards & Interfaces*, 1994. Vol.16-3. p.265 - 278. (DOI: [http://dx.doi.org/10.1016/0920-5489\(94\)90017-5](http://dx.doi.org/10.1016/0920-5489(94)90017-5))
- [24] Vapnik, V. N. and Lerner, A. "Pattern recognition using generalized portrait method". *Automation and Remote Control*, 1963, Vol.24-6, pp.774-780.
- [25] Vapnik V. *The Nature of Statistical Learning Theory*. Springer, 1995. ISBN: 978-1-4419-3160-3
- [26] Karatzoglou A, Smola A, Hornik K, Zeileis A. "kernlab - An S4 Package for Kernel Methods in R." *Journal of Statistical Software*, Vol.11-9 pp.1-20. ( URL <http://www.jstatsoft.org/v11/i09/>)