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 and 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).

Received: 7/aug/2015 - Accepted: 15/sep/2015 - DOI: http://dx.doi.org/10.6036/ES7782

PREDICTIONS OF ELECTRICITY DEMAND, INCLUDING DATA OF THE PRESENT ECONOMIC AND FINANCIAL CRISIS. APPLICATION TO THE 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, Genetic Algorithm, which is an evolutionary computation technique, and 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.- INTRODUCTION

Different types of companies require long-term forecasts of electric energy demand (EED). Very low predictions could lead to deficiencies in the power supply, with disadvantages for different economic sectors, and costs could be higher than the non-supplied energy. By contrast, very high estimates could lead to high incurred opportunity costs, and a high unprofitable economic investment. In both cases, the power price would increase for the consumer.

Efficient prediction techniques help power companies reduce the uncertainty of the EED, and obtain an optimal and realistic scheduling of electricity production. Economic development is the most influential factor on the power consumption of each region in the long term (one or more years, even decades). A decrease in energy prices along with expansive monetary policies, incentivize the growth of the economy. The improvement in the general situation of household economies leads to an increase in consumption of goods and services. Therefore, GDP and CPI are two factors that intervene in the level of electricity consumption, and both show great variability on years of economic and financial crisis like the present. This paper presents a comparison between several methods used in literature to predict the electricity demand [1], including nonlinear, not deterministic and machine learning methods. Thus, classical statistical techniques (Multiple Linear Regression (MLR) and Logarithmic Multiple Linear Regression (LogR)), statistical learning techniques (Support Vector Machine (SVM)), Evolutionary Computation techniques (Genetic Algorithm (GA)), and Machine Learning techniques (Artificial Neural Network (ANN)) have been used. Furthermore, the effect of adding the CPI as an explanatory variable is investigated. As a case study, the DEE in the Canary Islands has been estimated for the years 2015-2020.

Section 2 shows the importance and increasing complexity of estimating the long-term EED. Section 3 describes fundamental principles and characteristics of the methods considered in this paper. Section 4 presents the interest of including the CPI as a variable in the functional fit, the results obtained, and a comparative analysis of the prediction goodness. Finally, the conclusions are presented.

2.- IMPORTANCE AND COMPLEXITY OF ESTIMATING THE LONG-TERM EED

The power market has some peculiarities compared to other energy supplies. At every moment, the same amount of electricity consumed plus losses must be produced, because it cannot be stored competitively or in appreciable quantities. Different storage techniques could change the landscape in the coming decades.

High reliability in power supply is essential in a modern economy. European electrical systems are strongly connected. Spain is interconnected with four countries: Morocco, Portugal, France and Andorra; and can exchange electrical power with any of them. In the Canary Islands, there are six independent power systems, and each one of them is self-sufficient. Therefore, a “bad” demand prediction in the Canary Islands is much more critical than in the rest of the Spanish territory.

In the past, before the current period of economic and financial crisis, there has mainly been a significantly increasing trend in EED, due to increased use of varied electric devices, which have made life easier, regardless of consumption. Nowadays, new factors should be considered, because they could burst with force, introducing major changes in the evolution of future demand:

- The economic uncertainty created by the 2008 crisis.
- Energy poverty of the population, aggravated by sharp rises in the price of electricity, coupled with high rate of unemployment and precarious working conditions.
- Increased efficiency of equipment and facilities.
- The incorporation of electric vehicles.
- The implementation of demand management systems, which temporarily will displace the electrical consumption by installing smart meters, and can also affect the amount thereof, by making become more aware of it.
- The evolution of aerothermal heat pumps could displace some fossil fuels [2].

Before describing the methods, it is interesting to highlight some points:

➢ Fig. 1 shows that, while prior to the current economic and financial crisis there was a strong linear relationship between EED and GDP in the Canary Islands, during the crisis the nonlinear behavior between both variables has been accentuated.

Publicaciones DYNA SL -- c) Mazarredo nº69 - 4º -- 48009-BILBAO (SPAIN)
Tel +34 944 237 566 – www.dyna-energia.com - email: info@dyna-energia.com

- Jorge et al. [3], in a study about the forecast of the electricity price in Colombia, say that preliminary statistical analysis showed that CPI, the real exchange rate and the international fuel price do not significantly affect the prices of electricity. The context of that study becomes an inverse function of the treated in this paper, because it refers to the energy price according to the EED.
- Mohamed and Bodger [4], in a study about the EED in New Zealand, observed that there was a high correlation between GDP and population when they applied the multiple linear regression model, generating some multicollinearity. However, they concluded that the removal of some of these variables did not produce better fits, so all were included. It is also important to note that the multicollinearity problem is only a problem when the method used to adjust has to solve an algebraic equation system that is unconditioned.
- Wagdy et al. [5] used a combination of variables as the only explanatory variable and obtained a higher correlation. For example: $\left(\frac{GDP}{CPI}\right) POP$, where POP is population.

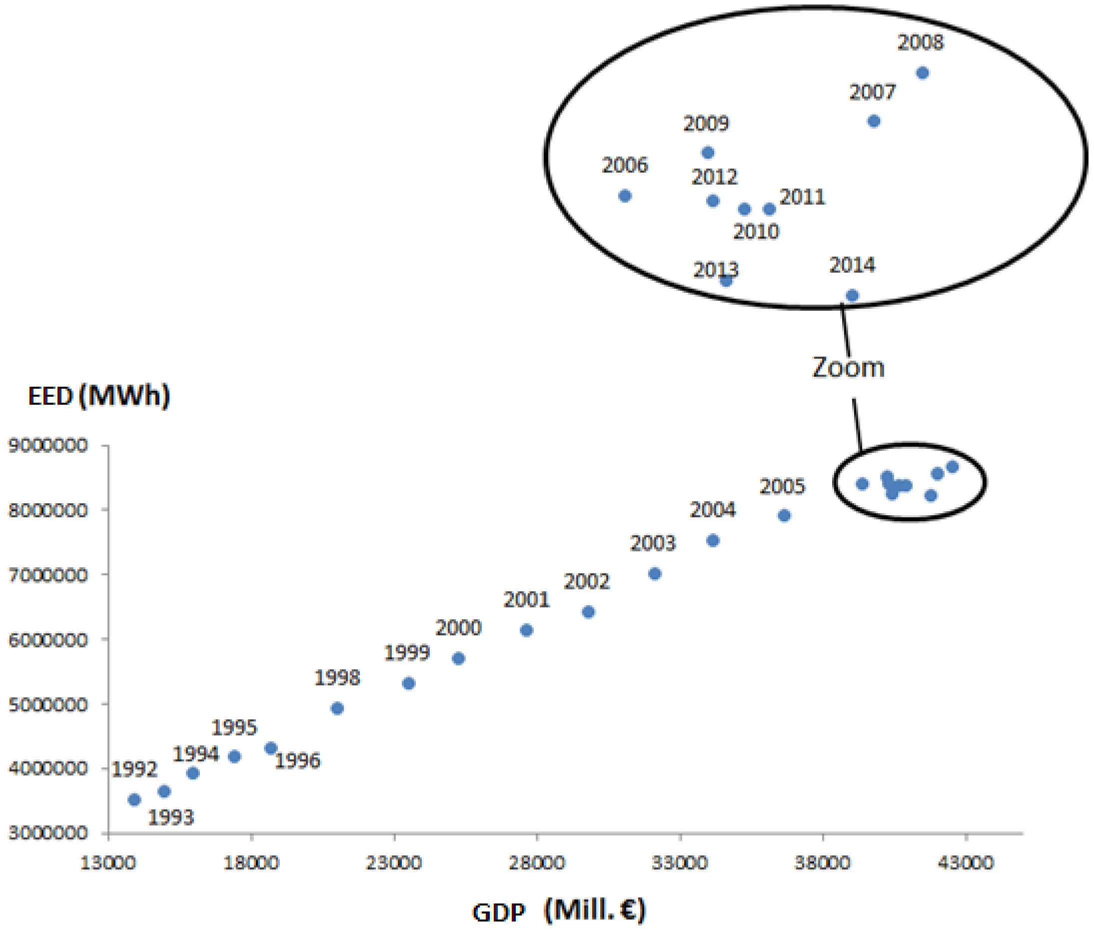

Figure 1. Relationship between EED and GDP (Canary Islands, 1991-2014)

3.- STATISTICAL METHODS, EVOLUTIONARY ALGORITHMS AND AUTOMATIC LEARNING TECHNIQUES

Forecasting methods can be classified into two broad categories: parametric methods and artificial intelligence–based methods. Among the artificial intelligence methods are: Genetic Algorithms (GA), Artificial Neural Networks (ANN), Support Vector Regression (SVM), Fuzzy Logic and Expert Systems. The parametric methods are based on relating electricity demand to its affecting factors by a mathematical model, for example: Multiple Linear Regression (MLR), Logarithmic Multiple Linear Regression (LogR), Time Series, etc. The main advantage of methods based on artificial intelligence is the ability to emulate nonlinear relationships. GA and ANN are among the most popular and proven efficiency in solving highly nonlinear problems.

In order to facilitate the discussion of the results and conclusions of this work, it is necessary to contextualize the different methods in relation to its fundamentals and operability, also in the context of multidisciplinary knowledge, it is necessary to provide accurate information to facilitate their practical use and implementation, that clearly shows the advantages and disadvantages of one method over another.

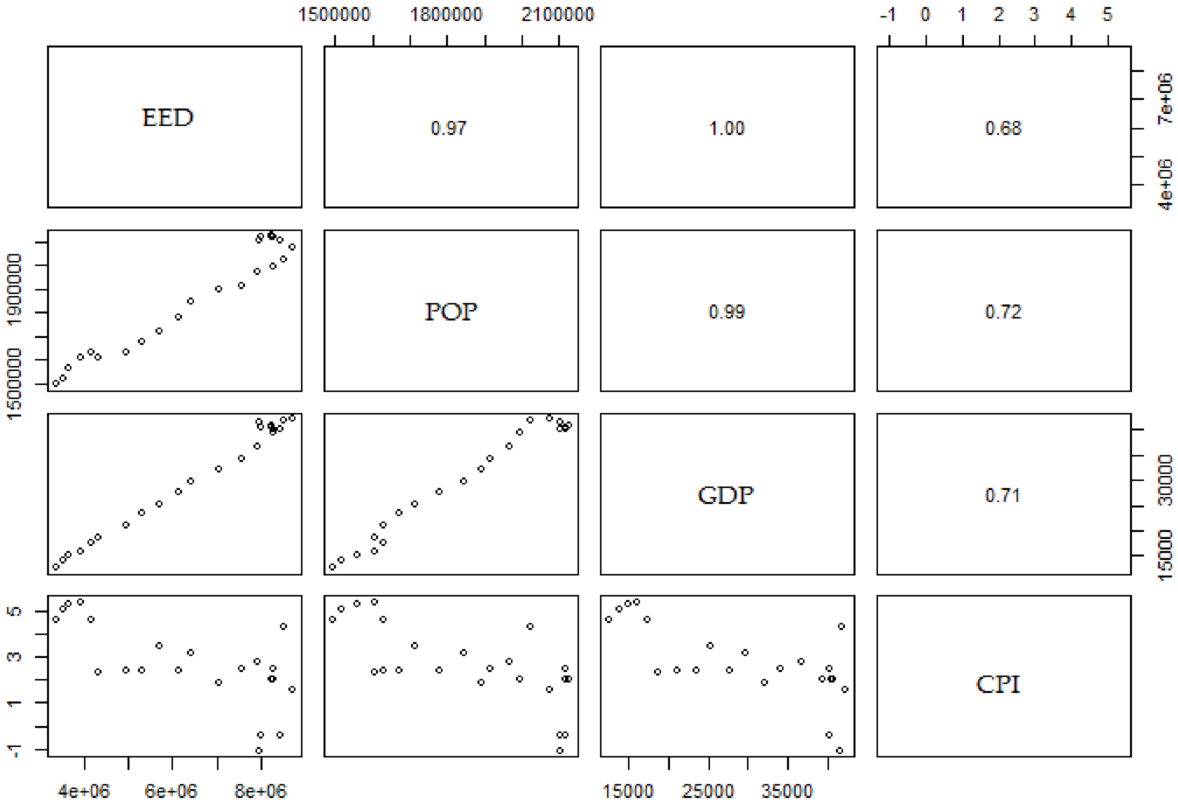

Figure 2. Scatter plot matrix (Canary Islands, 1991-2014).

3.1.- CLASSIC REGRESSION STATISTICAL METHODS

Regression analysis is a statistical technique used for modeling the relationship between a single variable, called the response or dependent variable, and one or multiple independent variable(s). The response variable must be a continuous variable, but the independent variables can be either continuous or discrete or categorical.

In this section, a multiple linear regression is applied to the function defined in Eq. (1)

$$EED = \beta_1 POP + \beta_2 GDP + \beta_3 CPI + \beta_4 \quad (1)$$

where POP is the population and $\beta_i,\ i = 1,2,3,4$, are coefficients that are determined by the least squares adjustment.

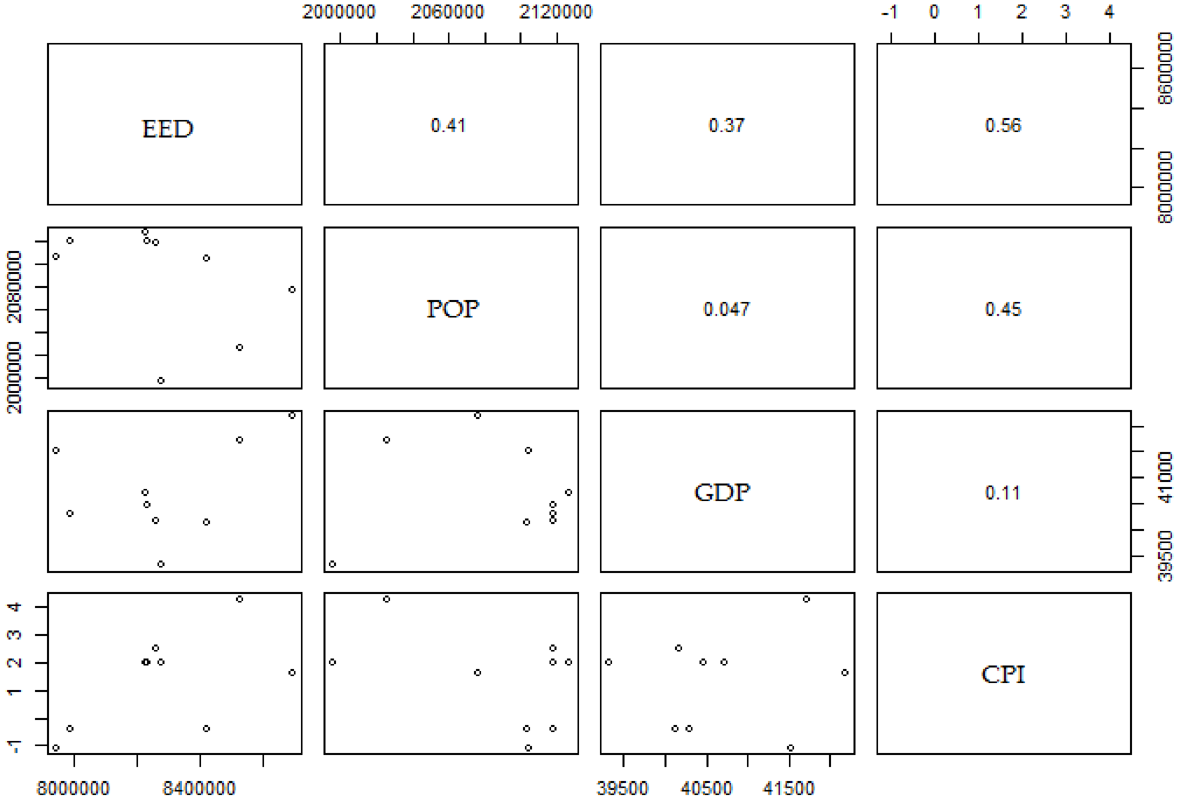

Figure 3. Scatter plot matrix (Canary Islands, 2006-2014)

If all the data of the study period are considered, a strong linear dependence not only between EED and population, and between EED and GDP, but also between population and GDP, is observed (see Fig. 2). However, considering the period 2006-2014, this collinearity disappears, as can be seen in Fig. 3. That is, in the period concentrated around the economic and financial crisis, the nonlinear dependence between EED and the explanatory variables is emphasized. Therefore a logarithmically transformed model (LogR) is also applied:

$$log(EED) = \beta_1\, log(POP) + \beta_2\, log(GDP) + \beta_3\, log(CPI) + \beta_4 \qquad (2)$$

Sometimes logarithms are used as a simple mathematical transformation strategy, aimed at reducing the original dispersion of the dataset. It has the advantage of reducing the heteroscedasticity problem, since the scale of measurement of the data is also reduced.

Although Eq. (2) is equivalent to a nonlinear exponential expression when EED is solved, the calculation method used to obtain the coefficients is a multiple linear regression on the transformed plane, considering the logarithm of the original explanatory variables as the independent variables of the calculation.

3.2.- EVOLUTIONARY COMPUTATION AND MACHINE LEARNING TECHNIQUES

GA, ANN and SVM are characterized by being able to automatically modeling nonlinear relationships. Some general characteristics of the models used to estimate EED in this paper are highlighted below, and guidelines for their efficient use are given.

3.2.1.- Capabilities and good use of GA

Evolutionary Algorithms (EA) are global search methods inspired by natural evolution and genetics. Their simplicity and the ability to find efficiently the best solutions are features that motivate, undoubtedly, the use thereof. EA theories emerged in the early 1960's, with Evolutionary Programming [6], using only the mutation operator. In 1965, Schwefel [7] and Rechenberg [8] laid the groundwork of Evolutionary Strategies. GA have been the AE mostly applied to solve optimization problems in many disciplines of science and engineering [9]. John Holland [10] put GA on the international scene incorporating recombination or genetic cross to generate diversity in the new solutions candidates to be optimal in the simile of obtaining descendants.

GA, as EA, operate in a dynamic of populations, evolving over time a population of candidate solutions to be optimal, under an iterative process of transition from one generation to another, which makes evolve favorably the approximate solutions until making them optimal. They are heuristic search methods, because there are operations, mainly crossover and mutation, which are made in each generation with assigned probabilities.

In this paper a GA using well-known and simple to implement operators is considered. With little effort in the calibration of its parameters, it generates good solutions and robustly (once the parameters are set, in each run this GA gets very similar solutions):

- Tournament selection operator [11]. It is easy to implement and to calibrate its pressure of selection. At the beginning of each iteration, the GA performs successive tournaments between two or three candidate solutions (individuals) of the population, with degradation in their assessments in the objective function by using the sharing function. So the agglomeration of candidate solutions around local optima is penalized (simile to the further weakening of a population when resources are shared with a greater number of individuals and nearest among them).
- Recombination or crossover operator: antithetical crossover [12]. From the linear combination of two candidate solutions, two new solutions are generated. Two parents p_1 and p_2 generate offspring:
$$h_1 = ap_1 + (1-a)p_2$$
$$h_2 = (1-a)p_1 + p_2$$
where a is a random number in the interval $[-0{,}5; 0{,}5]$. Thus, it acts interpolating when a is in the interval $[0{,}1]$, or extrapolating when a is in the intervals $[-0{,}5; 0]$ or $[1; 1{,}5]$.
- This mutation considered is one of the simplest. It consists in adding a random number between -1 and 1 to the variable to be mutated.
- The execution parameters have been the following:
  - ✓ Encoding of variables: real.
  - ✓ Number of candidate solutions: 100
  - ✓ Crossover probability: 0,55
  - ✓ The parameter a is randomly generated at each recombination.
  - ✓ Mutation probability: 0,45
  - ✓ Resource sharing parameter: $\sigma_{share} = 10$. It represents the minimum distance between two candidate solutions to be the global minimum.

The only relevant mathematical operation performed by the GA is simply the evaluation of the objective function of the optimization problem. In this paper, the following nonlinear objective function, whose graphics are similar to those of exponential functions, has been considered. Also to take advantage of the compression of the numerical values, and so the compression of the search space:

$$EED = \exp\left(\frac{a_1 log(POP)}{(a_2 + log(POP))} + \frac{a_3 log(GDP)}{(a_4 + log(GDP))} + \frac{a_5 CPI}{(a_6 + CPI)} + a_7\right) \quad (3)$$

where $a_i, 1 \le i \le 7$, are real coefficients to be optimal during the minimization process of the objective function.
The good use of a GA lies in operating with a balance between exploitation and exploration of the search space, in the search for the optimal. Exploitation is the intensity with which good candidate solutions are sought in each iteration. A higher level of production of diversity among the candidate solutions induces a further exploration of the search space.

3.2.2.- Capabilities and good use of ANN

Publicaciones DYNA SL -- c) Mazarredo nº69 - 4º -- 48009-BILBAO (SPAIN)
Tel +34 944 237 566 – www.dyna-energia.com - email: info@dyna-energia.com

ANN are models that try to mimic the abilities of the human brain, offering advantages in terms of their learning ability, with parallel operation and tolerance to failures caused by inaccurate information [13]. McCulloch and Pitts (1943) is often cited as the starting point in neural network research, but the Rumelhart, Hinton and Williams' paper in 1986 was the trigger of an exponential growth of neuronal literature [13]. Some of these techniques have been applied to predict energy price in the Spanish electricity market [14], producing competitive results with those obtained with the application of other techniques.

A single hidden layer ANN can approximate any continuous function on an interval up to the desired level; so the multilayer ANN are universal methods that approximate functions, and operate with a mathematical analogy to the development of functions, which are reminiscent of the Fourier series expression, but here it is a development in sigmoid functions [15].

The steps involved in designing a neural network to approximate a function are described below.

3.2.2.1.- The network architecture

The selection of the network architecture basically consists of preparing the data, and defining the inputs, the size of the network and the activation function. In this work, a single hidden layer ANN is considered (See Fig. 4).

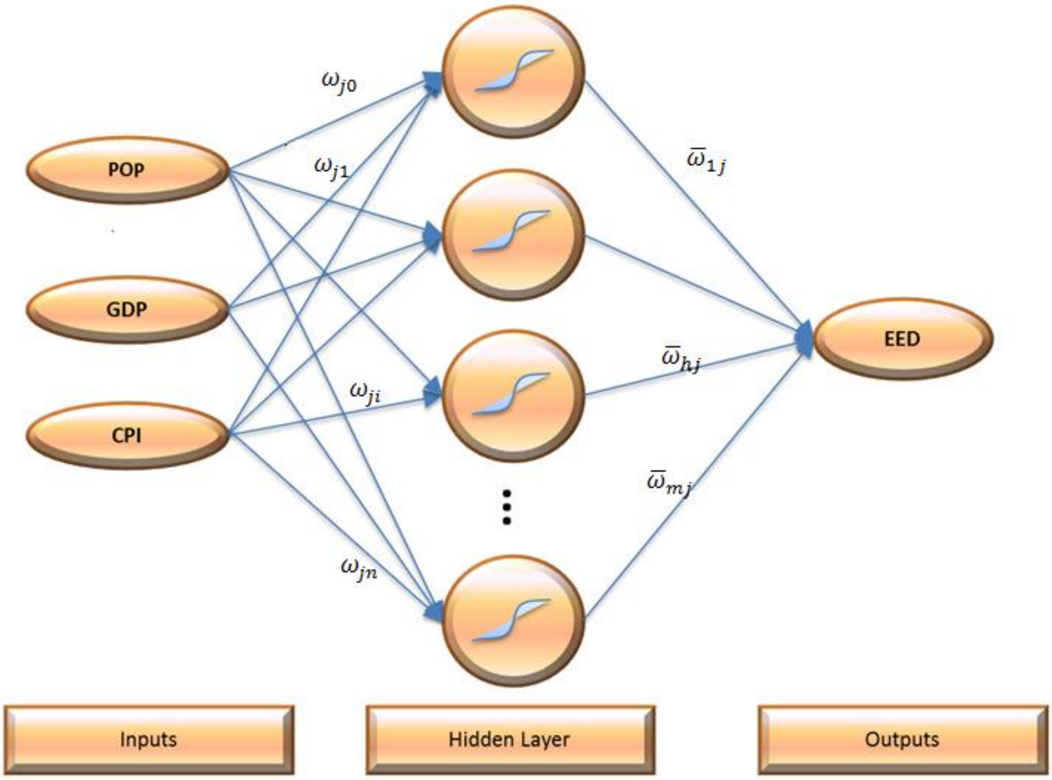

Figure 4. Network structure used for training the ANN

Data are normalized within the interval $[-1,1]$, as shown below:

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

where x_{max} and x_{min} are the maximum and minimum of x, respectively.

The equation that defines the relationship between the desired output (EED) and the inputs (GDP, POP and CPI) is defined as:

$$DEE = \sum_{s=0}^{l} \varpi_{hs}\, k_s = \sum_{s=0}^{l} \varpi_{hs}\, f(p_s) \tag{5}$$

where l is el number of hidden neurons, $p_s = \omega_{s1}POP + \omega_{s2}PIB + \omega_{s3}IPC$ is the input of the hidden neuron $s, 1 \leq s <= l$, f is the activation function of the network, $k_s = f(p_s)$ is the output of the hidden neuron s, and the coefficients ω, ϖ_{hs} are the weights of the ANN.

The network is trained for a limited number of episodes. The number of hidden neurons varies in each training process. The number of neurons with a satisfactory convergence is taken as the size of the hidden layer.

The output functions of the hidden neurons must be sinusoidal [17-18]. Generally, the most used functions are the hyperbolic tangent (anti-symmetric) and the sigmoidal (symmetrical), defined as:

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

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

Function (6) is used in this paper to be preferably used in the literature [13, 19].

3.2.2.2.- Training algorithm

The algorithms used to adjust the network weights, in order to minimize the error, are based on nonlinear local optimization methods [13]; though global optimization techniques for the training of ANN have also been used [20]. Nonlinear local optimization methods are algorithms which are based on gradient descent, as the standard back-propagation algorithm [13], and the second-order algorithms use the Hessian [21]. In this paper, the Resilient Back-propagation (Rprop) algorithm [22-23] is used. The Rprop is a fast method of first-order, and its memory requirements are lower than second-order methods.

The Rprop algorithm works in batch processing, that is, the weights are updated based on the error values corresponding to the total error on the training set. If the p^{th} desired output is $t^{(p)}$ and the output obtained via the network is $y^{(p)}$ for the p^{th} input set, and if there are P number of training examples, then the error is calculated using the mean squared error (MSE) defined as:

$$MSE = \frac{1}{P} \sum_{p=1}^{P} \left(t^{(p)} - y^{(p)}\right)^2 \tag{8}$$

For each ANN initialization procedure, 30 different networks (with different initial random weights) are trained for a maximum of 200 iterations, to avoid over-training.

Performance on the training data and test data sets is calculated (for each trained network) using the MSE, and the best network (with the smallest MSE) is preserved.

3.2.3.- Capabilities and good use of SVM

The algorithm implemented by Support Vector Machines is a nonlinear generalization of the Generalized Portrait algorithm proposed by Vapnik and Lerner [24] to solve linearly separable classification problems [25]. Initially it was applied to the two-class classification problem, looking for the best separation hyperplane, that is, the hyperplane that provides the highest margin distance between the nearest points of the two classes (called functional margin). This approach, in general, guarantees that the larger the margin is the lower is the generalization error of the classifier.

The vector formed by the points closest to the hyperplane is called a support vector. So the SVM arise as linear classifiers based on statistical learning theory and under the principle of structural risk minimization, as opposed to other methods which use the principle of empirical risk minimization, such as neural networks. A regularization coefficient, C, weighs the structural risk minimization versus the empirical risk minimization, giving to SVM a good generalization capability at the cost of obtaining greater errors in the training phase, and facilitating the formulation of SVM as quadratic optimization problems of convex functional and, therefore, with a single global optimum. In order to overcome the linear learning limitations, many author formulate SVM as kernel functions, using a non-linear transformation function Φ. The kernel functions are defined as the scalar product of the images of this transformation: $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$. The formulation of SVM with kernel functions makes easy their computational implementation. Solutions of classification and regression problems, sought by the kernel-based algorithms, operate with Φ functions only implicitly, without being necessary to know their expressions. They try to find mainly:

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

where: $\alpha_i, 1 \leq i \leq n$, are coefficients which are determined with the model and n is the number of training points.

Among the most used kernels are:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2 / \sigma^2\right) \quad \text{(Gaussian function; } \|\cdot\| \text{ is the Euclidean norm)} \tag{10}$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_j^T \mathbf{x}_i \quad \text{(Linear function)} \tag{11}$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = \left(t + \mathbf{x}_j^T \mathbf{x}_i\right)^d \quad \text{(Polynomial function)} \tag{12}$$

One of the advantages of SVM is their efficiency with few training data.

4.- APPLICATION AND RESULTS

The goal is to find good EED functional fittings. Different methods are used to check which of them have a greater goodness of fit, but the most important thing is to obtain good predictions in future years. In order to evaluate the predictive power of each method with the available data, several historical years were reserved as test-years. Thus, the adjustment or training phase was performed using the remaining data. The (randomly chosen) test years were: 1992, 2001, 2004, 2009 and 2014.

Note that the fact that 2009, the first year of the economic and financial crisis that markedly breaks the previous increasing linear trend of EED (see Fig. 1), is one of the test-years and not one of the training-years, makes the functional fitting not so good.

4.1.- CASE STUDY

The proposed study aims to estimate the EED in the Canary Islands. Historical data were downloaded from the Instituto Canario de Estadística, ISTAC, http://www.gobiernodecanarias.org/istac/, taking data from electricity consumption. There is only data from 1991.

As inputs to the methods detailed in Section 3, data of GDP, CPI and population (POP) in the Canary Islands, from 1991 until 2014 (available from the Instituto Nacional de Estadística, INE, http://www.ine.es), have been considered, and the output variable is EED. In the period 1991-2014 it is observed collinearity between explanatory variables and POP GDP (see Fig. 2), however, it disappears if we focus on the most recent period from 2006 to 2014 (see Fig. 3).

Collinearity between explanatory variables GDP and POP is observed in the period 1991-2014 (see Fig. 2) however it disappears if we focus on the most recent period 2006-2014 (see Fig. 3).

Values of predictions of the population have been taken from the “Proyección de la población española desde el año 2014 hasta 2064”, by INE. The CPI and GDP estimations made for the "Planificación energética indicativa" (Ley 2/2011, de 4 marzo de Economía Sostenible) have been used.

4.2.- RESULTS

The use of metrics is required for the validity of the conclusions and to make an effective comparison between the methods.

As a measure of error, considering that the data number is limited, the mean absolute percentage error, known as MAPE, is used:

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

being $\varepsilon_i = EED_i - \overline{EED_i}$ (where EED_i is the real demand and $\overline{EED_i}$ is the estimated demand in year *i*).

Two cases are considered:

- ✓ First case: POP-GDP.
- ✓ Second case: POP-GDP-CPI.

MLR method

In this case, the coefficients of Eq. (1) are calculated. The following expression is obtained:

$$\overline{\mathrm{EED}} = 218.41\, GDP - 1.84\, POP + 19{,}947.86\, CPI + 32{,}372{,}363 \tag{14}$$

The coefficient of the population is negative. This is indicative of colineality between the POP and GDP variables (see Fig. 2).

LogR method

Now, the coefficients of Eq. (2) are calculated. As in the crisis years there are negative values for the CPI, all higher than –5, the CPI data have been moved by adding 5, in order to can apply this method.

The following expression was obtained:

$$log(\overline{\mathrm{EED}}) = 0.86\, log(GDP) - 0.13\, log(POP) + 0.04\, log(CPI+5) + 8.69 \tag{15}$$

GA method

In this case, the coefficients of Eq. (3) are calculated. This equation is a nonlinear expression that is proposed for the first time in the literature to estimate the EED. The expression obtained is:

$$EED = exp\left(\frac{-1{,}083.2\, log(POP)}{(4{,}077.267 + log(POP))} + \frac{-1{,}209.1\, log(GDP)}{(-1{,}373.4 + log(GDP))} + \frac{0.0058\, CPI}{(1.197 + CPI)} + 10.368\right) \tag{16}$$

where the last search space considered, after some previous runs, has been:

$$-4{,}000 \le a_1, a_3, a_4 \le -100$$
$$1{,}000 \le a_2 \le 7{,}000$$
$$-20 \le a_5\, a_6, a_7 \le 20$$

ANN method

Here the results were obtained from the ANN described in Table 1, derived from the learning with the Rprop algorithm [22] applied to our database.

ANN structure	Configuration
Range of normalization	[-1,1]
Number of neurons (1st layer 2º layer, 3º layer)	3-7-1
Activation function: hidden / output	Tanh / Linear
Algorithm	Rprop
Iterations (max.)	200

Table 1. Structure of the ANN used

SVM method

The results were obtained with the *ksvm* function of the library **kernlab** in R [26], whose parameters are shown in Table 2.

Parameter	Values
Formulations	ε-SVR
Kernels	Gaussian / Linear
Optimizer	SMO
Model Selection	Hyper-parameter estimation for Gaussian kernels
ε	0.1
C	3; 5; 10; 15; 20; 25; 30
σ	0.01; 0.05; 0.1; 0.2; 0.5; 0.7; 0.9

Table 2. Parameters used in the ksvm function of the library kernlab *in R.*

Table 3 shows the errors obtained with all the methods considered. Add the CPI as an independent variable improves the error in the predictions of the test data with GA, ANN and SVM (Gaussian kernel) methods, which are nonlinear functional adjustment methods.

	MAPE %(Training data)		MAPE %(Test data)	
	POP, GDP	POP, GDP, CPI	POP, GDP	POP, GDP, CPI
MLR	1.52	1.25	3.04	3.41
LogR	1.46	1.46	1.66	2.46
SVM (Linear)	1.45 (C=30)	2.06 (C=3)	3.42 (C=30)	3.63 (C=3)
SVM (Gaussian)	2.06 (C=5, σ=0,5)	2.29 (C=10, σ=0,05)	2.38 (C=5, σ=0,5)	2.09 (C=10, σ=0,05)
GA	1.48	1.46	2.77	1.71
ANN	0.82	0.65	1.58	1.22

Table 3. Error of the methods with/without CPI

Table 4 shows estimates of EED from 2015 until 2020. They were obtained with each of the methods considered in this paper, based on the latest predictions of population, GDP and CPI, in that period for the Canary Islands, published by agencies and institutions referred to in Section 4.1.

The GA has been programmed in ANSI C language. The ANN was run with the free ENCOG library (http://www.heatonresearch.com/encog). The other methods were run with the free software R (https://www.r-project.org/).

Publicaciones DYNA SL -- c) Mazarredo nº69 - 4° -- 48009-BILBAO (SPAIN)
Tel +34 944 237 566 – www.dyna-energia.com - email: info@dyna-energia.com

	POP	GPD	CPI	EED (GWh)					
Year	population	(Mill. €)	(%)	MLR	LogR	SVM (Gaussian) C = 10; σ = 0,05	SVM (Linear) C = 3	GA	ANN
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	8,921383	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

Table 4. Estimates of EED from 2015 until 2020, using different methods considered in this article, and with three explanatory variables.

5.- DISCUSSION

Motivated by large oscillations of the GDP and CPI during the current economic and financial crisis, this paper discusses the improvement of the electricity energy demand prediction in the Canary Islands, by incorporating the CPI (without being multiplied or divided by other variables) as an independent or explanatory variable in the models considered. Thus, CPI is added as an explanatory variable along with the traditional ones: population and GDP. Moreover, the collected data are adjusted with different methods.

With nonlinear methods, the effect of including or not the CPI as an explanatory variable is significant. In fact, its inclusion improves the results of prediction.

ANN have shown one again their ability to find and learn the relationship among the input variables to obtain better estimates of EED. A GA was used to find the optimal coefficients for a non-linear mathematical expression. This mathematical expression is proposed, in the scientific and technological community, for the first time.

BIBLIOGRAPHY

[1] Suganthi L, Anand AS. “Energy models for demand forecasting - A review”. *Renewable and Sustainable Energy Reviews*, 2012, Vol.16-2, pp.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, pp.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, Nº 77, pp. 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, pp.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*. CIRED. 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 resBirkhaeuser , 1977.
[8] Rechenberg I. *Evolutiansstrategie: 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, pp.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. Berling 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, pp.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, pp.183–192. (DOI: 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, pp.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, pp.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, pp.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)

[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, pp.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, pp.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, pp.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. pp.265 – 278. (DOI: 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/)

Publicaciones DYNA SL -- c) Mazarredo nº69 - 4º -- 48009-BILBAO (SPAIN)
Tel +34 944 237 566 – www.dyna-energia.com - email: info@dyna-energia.com