Energy Conversion and Management 49 (2008) 3135–3142

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

# Monthly electric energy demand forecasting with neural networks and Fourier series

E. González-Romera *, M.A. Jaramillo-Morán, D. Carmona-Fernández

*School of Industrial Engineering, University of Extremadura, Avda de Elvas s/n, 06071 Badajoz, Spain*

ARTICLE INFO

*Article history:*
Received 17 January 2007
Received in revised form 26 October 2007
Accepted 8 June 2008
Available online 18 July 2008

*Keywords:*
Fourier series
Load forecasting
Neural network applications
Spectral analysis
Time series

ABSTRACT

Medium-term electric energy demand forecasting is a useful tool for grid maintenance planning and market research of electric energy companies. Several methods, such as ARIMA, regression or artificial intelligence, have been usually used to carry out those predictions. Some approaches include weather or economic variables, which strongly influence electric energy demand. Economic variables usually influence the general series trend, while weather provides a periodic behavior because of its seasonal nature. This work investigates the periodic behavior of the Spanish monthly electric demand series, obtained by rejecting the trend from the consumption series. A novel hybrid approach is proposed: the periodic behavior is forecasted with a Fourier series while the trend is predicted with a neural network. Satisfactory results have been obtained, with a lower than 2% MAPE, which improve those reached when only neural networks or ARIMA were used for the same purpose.



## 1. Introduction

Electric energy demand forecasting is a very useful tool for companies producing, delivering and reselling electric energy. Therefore many researchers have studied and applied different techniques to improve the prediction accuracy. Short-term load forecasting has engaged the attention of most researchers [1–14], because of its influence in the daily operation of generation and distribution systems. It has become a basic tool for companies that sell and buy electric energy in deregulated markets. This forecasting process tries to estimate the load or energy that will be demanded in the next hours or days. However, other time horizons are also interesting for power systems planning: the forecasting of the energy consumers will demand next weeks or months is very useful for the maintenance planning of grids and as a market research for electricity producers and resellers. This kind of prediction is known as medium-term forecasting. The progressive deregulation of power system operation in many countries has driven some researchers to study this time horizon [15–20], as it provides companies valuable information about the market needs of energy, which will help them to schedule fuel purchases and advantageously negotiate contracts with other companies, reducing financial risks. Finally, long-term forecasting pursues the prediction of the annual load peak or global energy that consumers will demand up to 20 years ahead, in order to schedule expansion planning strategies for production and distribution systems [21–23]. Both medium- and long-term load forecasting have attracted less attention than short-term forecasting in scientific literature.

As it has been previously pointed out, short- and medium-term load forecasting present a clear interest in deregulated markets, as the Spanish one. Short-term forecasting has been widely studied in literature, so medium-term is the time horizon selected in this paper. Data of the Spanish monthly electric demand from January 1975 to December 2002 (a total of 336 values) (Fig. 1) has been used. They have been obtained from statistical bulletins of REE (the Spanish Power System Manager, data available in www.ree.es).

Several techniques have been traditionally used for electric energy demand forecasting. Between them, the so-called classical techniques, ARIMA [15,16] and regression [1], were widely used in the earliest works. However, their low adaptability, mainly to solve problems with highly non-linear series, has promoted an increasing interest on artificial intelligence techniques: neural networks [2–5], expert systems [6] or hybrid methods, such as neural networks combined with fuzzy logic [22], wavelets [7] or genetic algorithms [8]. Neural networks have provided very reliable predictions because of their great ability to identify the behavior of non-linear systems. Many authors have written about the advantages of neural networks in forecasting and, specifically, in load prediction. These advantages could be summarized as follows: they present a great ability to model non-linear relations [21]; they do not need the supervision of human experts to make decisions [2] and they can adapt to changes in the series by self-learning [3]. So, they are one of the prediction tools to be used in this work.

* Corresponding author. Tel.: +34 924 289600; fax: +34 924 289601.
*E-mail address:* evagzlez@unex.es (E. González-Romera).

0196-8904/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.enconman.2008.06.004

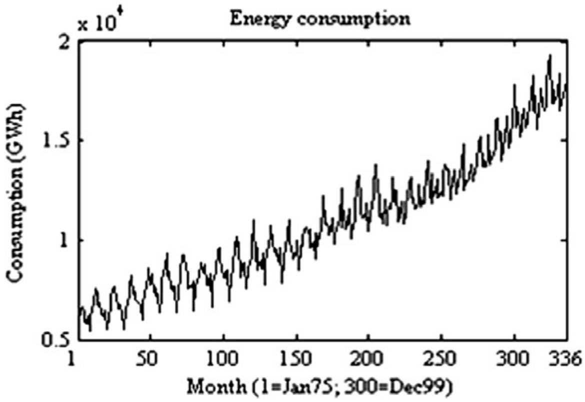


**Fig. 1.** Spanish monthly electric energy consumption from January 1975 to December 2002.

As electric demand is highly influenced by social and weather factors, they may be included in the forecasting process to improve the prediction accuracy [1,21]. So, in short-term forecasting, temperature, humidity, hour of the day or day of the week, must be taken into account along with past consumptions, because of their remarkable influence on demand. Therefore, a great amount of data is needed and complex models must be used to obtain accurate predictions. However, as medium-term forecasting deals with monthly data, the influence of the aforementioned factors is diluted in an overall value that represents the total monthly electric consumption, in which specific peaks or valleys are diluted. So, the forecasting model may use only past elements of the time series to obtain a prediction of the electric demand. Several authors have proved that inaccuracy in the prediction of weather variables, unavoidable in several months ahead forecasting, leads to a noticeable increase in demand forecasting errors [4,9–11]. Therefore, in this work, only past data of the demand time series will be used to forecast future ones.

Although the influence of social and weather factors in monthly demand may be assumed as diluted in the cumulative nature of a monthly value, it is clear that they must highly influence the evolution of the consumption time series. On the one hand, as wealth of most nations presents a rising trend and technological development provides society with more and more devices that need electric energy to work, electric demand will suffer a rising trend, which growth rate will not be constant as it follows the changes of the economical situation. On the other hand, as weather has a seasonal behavior, the mean values of the variables that directly influence electric consumption, as temperature, ought to have a more or less constant value for the same month of different years (modified, of course, by the aforementioned rising trend of demand). It is worth noting that modifications in the usual value of monthly weather variables are usually associated with global weather variations which affect several months in the same way (a rise or fall of mean temperature usually spreads over several months, although drought periods may even spread over several years). Therefore their influence will be embedded in the series trend. Taking into account the aforementioned comments it may be assumed that the time series of monthly electric demand may have a more or less periodic behavior embedded in a rising trend.

Many authors have detrended the demand series in order to reject its rising trend. So, a new stationary series is obtained which is easier to handle. Differentiation [12,13,15], curve fitting [17,24] and smoothing techniques [16,23] are the most used tools for this purpose. The prediction of the fluctuation series obtained has been carried out with general forecasting tools, such as regression, ARIMA, neural networks or fuzzy sets.

Few authors have tried to explore the periodic behavior of the fluctuation series, and most of them have dealt with short-term demand forecasting. In [14] a short-term forecasting was carried out with a Fourier series where the fundamental frequency was assumed to be that corresponding to a period of a week. In [17] a Fourier series was trained to follow the monthly electric demand in a country whose social events were ruled by the moon calendar. As social periodic events did not keep a close relation with weather periodicity, the result of the forecasting process was not very successful.

Taking into account all these ideas, it is reasonable to suppose that for those time series in which a more or less periodic behavior is embedded in a rising trend, as that of the Spanish monthly consumption used in this work (Fig. 1), trend and fluctuations may be split up in order to obtain two different time series. They may be independently processed to provide two predictions, which will be then added up to obtain the monthly electric energy demand forecasting. The independent prediction of trend and fluctuations ought to provide a more successful forecasting than the direct prediction of demand, as the tool used to forecast each series may be better adapted to learn each dynamical behavior. This fact has been proved in [25] where trend and fluctuations of the same time series used in this work (Fig. 1) were independently forecasted with neural networks. A more accurate prediction than that provided by one neural network which directly forecasted the demand time series was obtained.

In [25] the spectral analysis of the fluctuation series showed the presence of several well defined frequency peaks that suggested a periodic behavior of this series. The neural network trained to forecast it provided good predictions and no further analysis of that periodic behavior was carried out. Nevertheless the presence of those well defined frequency peaks suggests that a tool specifically developed to deal with periodic data could be used instead of a neural network to improve the prediction accuracy. In this work such a tool will be tested: a Fourier series. Some data of the series will be used to fit a Fourier series made up with the peak frequencies of the spectrum plot of the fluctuation series. This Fourier series will be used to forecast fluctuations different from those used for fitting. The demand forecasting will be then calculated by adding the prediction of the trend, which will be obtained in the same way as it was in [25], to that of fluctuations. As it will be proved bellow the prediction of fluctuations with the Fourier series provides a more accurate forecasting of demand than that obtained when a neural network is used to predict fluctuations.

The paper is organized as follows: Section 2 deals with the split of the demand series into trend and fluctuation; in Section 3 a Fourier series is adjusted to reproduce the periodic behavior of fluctuations and then used to forecast data different from those used for adjustment; in Section 4 a neural network is optimized to forecast the trend series; in Section 5 both fluctuation and trend predictions are added up to obtain the demand forecasting and the results obtained are compared with those provided by other forecasting models; finally, Section 6 presents the conclusions.

## 2. Trend and fluctuation series

The Spanish monthly demand time series (Fig. 1) presents a clear rising trend with a more or less periodic fluctuation embedded in it. So, it may be split into two series: one describing the trend and the other the fluctuations. As trend represents the overall series behavior it must be obtained from the demand series rejecting fluctuations. To do this a smoothing process must be applied to the demand series. Once the trend series has been provided by one of those tools the fluctuation one will be obtained by subtracting it from the demand series. Now as two series, trend and fluctuations, are available, two independent forecasting processes may be carried out. Then they will be added up to obtain the demand forecasting. Therefore the first step to be given in order to get the demand series split into trend and fluctuations is the selection of an appropriate smoothing tool.

Several smoothing techniques may be used to extract the series trend. An updated summary may be found in [26]. They may be classified into two groups: parametric and non-parametric. The former tries to fit a smooth function to the series data, while the latter tries to reject the fluctuations applying smoothing tools, such as filters or local regression. Non-parametric smoothing techniques are very suitable for load trend forecasting, because of their capability to adapt to local changes in the time series [26]. So, a simple non-parametric technique will be used in this work to smooth the consumption series: the moving average. It transforms every data in the original series into another one obtained as the weighted mean of a data set, named window, around that to be modified. Usually, only preceding values are used to define the window, because otherwise the last elements of the series could not be smoothed.

Four kinds of moving averages have been tested [27] depending on the way the weights have been defined.

### 2.1. Constant weights

The $i$th element of the smoothed series is obtained from the expression

$$T[i] = \frac{1}{n}(C[i] + C[i-1] + \cdots + C[i-(n-1)]) \tag{1}$$

where $T[i]$ is a smoothed datum, $C[i]$ a datum of the original series and $n$ the window size.

### 2.2. Exponentially decreasing weights

The smoothed series is obtained with the expression

$$T[i] = \frac{1}{r}C[i] + \left(1 - \frac{1}{r}\right)T[i-1], \quad r = \frac{2}{n+1} \tag{2}$$

It is equivalent to calculate a weighted mean with exponentially decreasing weights.

### 2.3. Linearly decreasing weights

The smoothed series is obtained from the expression

$$T[i] = \frac{2}{n+1}\sum_{j=1}^{n}\frac{j}{n}C[i-n+j] \tag{3}$$

### 2.4. Modified moving average

The smoothed series is obtained with the expression

$$T[i] = T[i-1] + \left[\frac{C[i] - T[i-1]}{n}\right] \tag{4}$$

As the aim of the series smoothing is to obtain a new one representing its trend, it is necessary to define some parameters to measure the performance of the smoothing process. Two parameters are usually used to find out the best option: the goodness of fit ($R^2$) and the smoothness ($S$) of the resulting series [26].

The first one is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(C[i] - T[i])^2}{\sum_{i=1}^{N}(C[i] - \widehat{C})^2} \tag{5}$$

where $C[i]$ and $T[i]$ are, respectively, the $i$th elements of the original and the smoothed series, $\widehat{C}$ is the original series mean and $N$ the number of data of the series. $R^2$ will be close to one if the smoothed curve fits accurately the original one, and will decrease to zero as the goodness of fit gets worse.

The smoothness ($S$) is defined as

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta^2 T[i])^2} \tag{6}$$

where $\Delta^2 T[i]$ is the discrete second derivative of the series $T$ in the $i$th element. The smoother the series, the lower $S$, with a zero value for the linear function.

The four kinds of moving averages have been applied to the consumption series with several window sizes ranging from 3 to 36 data to find out the configuration performing the best. A moving average of constant weights with a 12 data window has provided the best equilibrium between both parameters with $R^2 = 0.946$ and $S = 45.269$. Fig. 2 presents the smoothed series (thick line) and the original one (thin line).

For the sake of comparison splines [28] have been also tested although the results obtained were slightly worse. Therefore, only moving average has been presented in this work.

As it has been previously pointed out the fluctuation series may be easily obtained by subtracting the trend series from the demand one. It may be seen in Fig. 3. Now two series are available to carry out two independent forecasting processes. Before performing the prediction the tools to be used must be conveniently trained. So, the available data of both series must be split into two sets. The

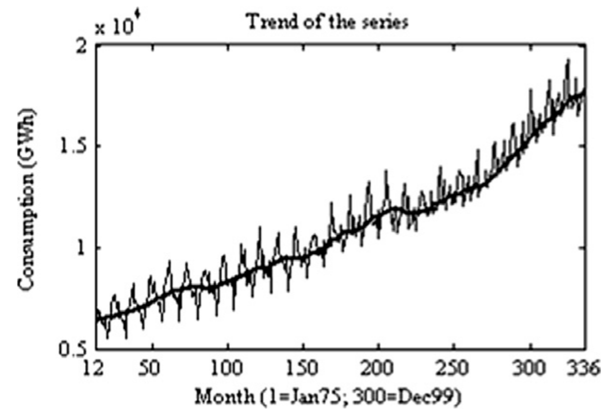


**Fig. 2.** Trend series (thick line) obtained with a moving average of 12 months with constant weights and consumption series (thin line).

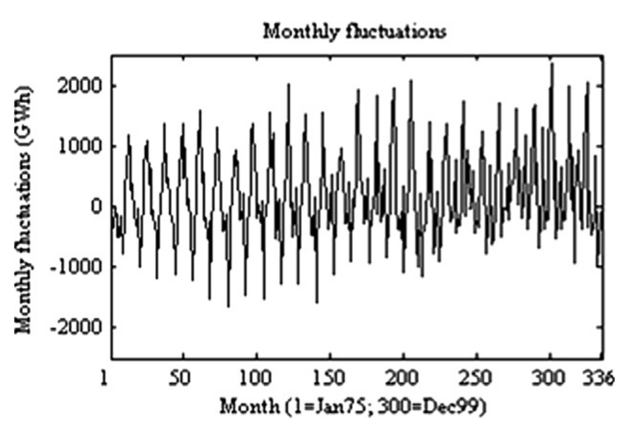


**Fig. 3.** Monthly fluctuation of the electric energy consumption.

first one will be used to train the corresponding tool while the second will be used to check their performance. So, the available data (336 values) have been split into two sets: 276 for training and the remaining 60 for validation.

To measure the performances of both trend and fluctuation predictions the following error indices will be used:

- MAPE: mean absolute percentage error between the actual ($S$) and the forecasted series ($S_f$):

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{S[i]-S_f[i]}{S[i]}\right|100 \tag{7}$$

- Percentile 95 of the absolute percentage errors (P95).
- Percentile 90 of the absolute percentage errors (P90).
- Maximum negative and positive errors (MNE and MPE).

## 3. Fluctuations forecasting with Fourier series

The fluctuation series (Fig. 3) presents an oscillatory behavior that ought to be studied in order to find out whether it is periodic. In that case, the forecasting process might be carried out with a tool that takes advantage of this fact, instead of using a standard forecasting tool, as it was done in [25], where a neural network was used to forecast fluctuations. So, a spectral analysis must be performed. Fig. 4 shows the power spectrum of the fluctuation series, where the $X$-axis represents the normalized frequency, whose higher value ($f = 1$) is the so-called Nyquist frequency, defined as half the sampling frequency [29].

As it is shown in Fig. 4, high and sharp peaks appear in frequencies 0.16, 0.33, 0.67 and 1, corresponding to periods of 12, 6, 3 and 2 months, respectively, and lower peaks in 0.5 and 0.83, corresponding to periods of 4 and 2.4 months. A continue component also appears as the smoothing process was not able to retain the whole non-oscillatory behavior of the consumption series. As all these frequencies are multiples of a fundamental one (that corresponding to a 12-month period), a six-term Fourier series can fit the fluctuation series:

$$F(n) = a_0 + \sum_{k=1}^{6}[a_k\cos(k\omega n) + b_k\sin(k\omega n)] \tag{8}$$

In this expression $\omega$ is the fundamental frequency of the series, and $a_0$, $a_k$ and $b_k$ the fitting coefficients.

Once the Fourier series has been defined it may be used to forecast future values of fluctuation as the series presents a clear periodic behavior.

Several strategies may be used to adjust the parameters in (8) depending on the number of data used to do it. They all use the least squares (LS) method to adjust them.

A first option is the use of the whole training data set. This option presents a clear drawback: it describes the global behavior of the series throughout 22 years (264 months) and provides the curve with a certain obsolescence. To avoid the curve fitting becomes obsolete, new data may be added to the training set, repeating the training process of Eq. (8). So, a new strategy may be defined in which all the data preceding that to be forecasted are used to adjust the parameters in (8). In this way many fitting processes are necessary, so this option will spend more time although it can better adapt to changes in the consumer's behavior.

Another strategy may be defined taking into account the fact that, as (8) has 13 parameters to be adjusted, only 13 data are necessary to carry out that adjustment. So, when a new value is to be forecasted the 13 data preceding it are used to obtain the parameters of (8). The function so obtained would exactly fit the training data set avoiding the obsolescence of the information used to obtain the parameters. Nevertheless as small data sets are used, this technique might lose generalization capability, because the presence of outliers will highly influence the resultant function.

A balanced approach between the previous strategies may be defined avoiding the presence of obsolete information and retaining a generalization capability: a new data will be forecasted with a curve whose parameters will be adjusted with a set of past data higher than 13 but smaller than the whole data set preceding it. The number of data to be considered must be defined in a trial and error process.

On the other hand, as fluctuations evolve around zero, the way the MAPE is obtained must be modified to avoid a division by zero: the error between a prediction and its corresponding actual value is divided by the actual consumption instead of the corresponding fluctuation.

Fig. 5 shows the forecasting MAPEs of the fluctuation series obtained with the validation data for different sizes of the data set used to fit the Fourier series. The best MAPE was obtained with a set of five years.

Table 1 shows the error indices of the fluctuation forecasting obtained with the validation set with the four fitting strategies described previously.

## 4. Trend forecasting with neural networks

As the trend series presents significant variations in its rising rate, classical forecasting tools are not well-suited to forecast it, as it will be seen bellow. So, an artificial intelligent tool will be a

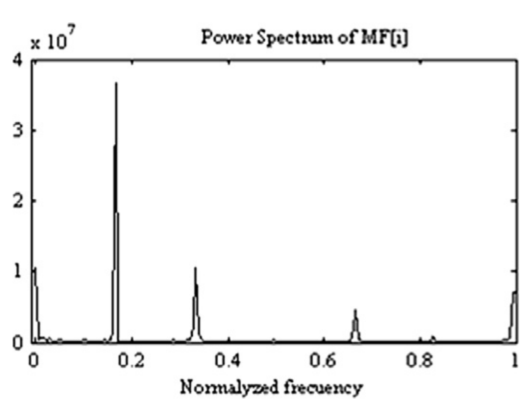


**Fig. 4.** Power spectrum of the monthly fluctuation series.

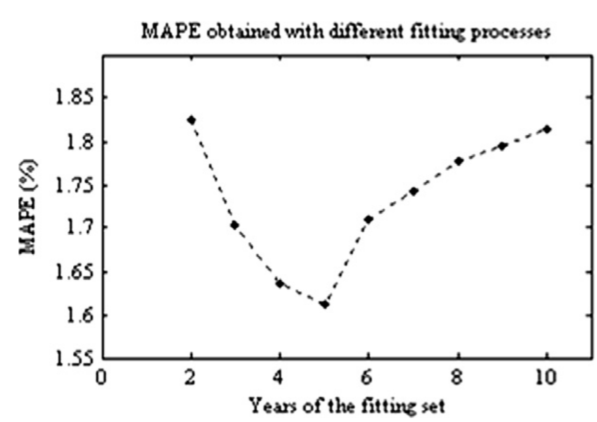


**Fig. 5.** MAPE obtained with different data sizes of the training set used to fit the parameters of (8).

**Table 1**
Fluctuation forecasting errors indices reached with the proposed fitting techniques

| | MAPE (%) | P95 (%) | P90 (%) | MNE (%) | MPE (%) |
|---|---|---|---|---|---|
| Training set fitting | 2.283 | 5.553 | 4.498 | −4.116 | 6.675 |
| All previous data fitting | 2.188 | 5.271 | 4.243 | −3.990 | 6.429 |
| 13-data fitting | 1.762 | 4.503 | 4.072 | −5.493 | 4.302 |
| 5-year fitting | 1.609 | 4.051 | 3.798 | −4.166 | 4.790 |

better choice. As it has been previously stated, in [25] a neural network was used to forecast the trend of the same demand series used in this work, and a very accurate prediction was obtained. So, this tool will be also used in this work.

Among the different neural models, feedforward networks (FF) are the most widely used in time series forecasting because of their simplicity and proved ability to identify the non-linear dynamics of complex systems. Multilayer perceptron (MLP) and radial basis functions (RBF) are the most popular FF models because they have been proved to be universal approximators [30,31]. Both of them will be used in this work to forecast the trend series.

### 4.1. Multilayer perceptron (MLP)

This neural structure is comprised of an input layer, which is not an actual layer but the set of input data to the network, an output layer and one or several hidden layers. Each neuron performs a weighted sum of its inputs, which activates an output function of the form

$$y_k = \Psi\left(\sum_{j=0}^{n} w_{kj}x_j + \theta_k\right) \tag{9}$$

In this expression $x_j$ represents the $j$th input to the $k$th neuron (which may be a network input or a neural output of the previous layer), $w_{kj}$ the strength of the connections between this neuron and its $j$th input, $y_k$ the neuron output and $\theta_k$ a bias constant. The activation function $\Psi(\cdot)$ is usually the logistic function or the hyperbolic tangent, although linear functions are also used in the output layers. A logistic function in the hidden layers with a linear one in the output layer is the configuration selected in this work.

The neural network ability to approximate any system is provided by its learning capability: the neural network is trained to learn the system behavior with the adaptive modification of its weights. The way this adaptation process is carried out defines the learning strategy of the neural network. The backpropagation algorithm [32] is used to train multilayer perceptrons, where a function relating the output error to the neural weights is minimized.

### 4.2. Radial basis functions (RBF)

This neural model is comprised of an input layer (the input data), an output layer and only one hidden layer with Gaussian output functions of the form

$$y_k = e^{-\left(\frac{\|\mathbf{x}-\mathbf{c}_k\|}{\sigma}\right)^2} \tag{10}$$

where $\|\cdot\|$ represents the Euclidean distance, $\mathbf{c}_k$ is a vector parameter, named center of the neuron, $\mathbf{x}$ the input vector and $\sigma$ a parameter defined as $\sigma = \text{disp}/\sqrt{-\ln 0.5}$, with disp being a dispersion parameter. The output layer has the same form as (9) with the activation function $\Psi(\cdot)$ being a linear one.

The learning algorithm in this model is the orthogonal least squares [33], which is carried out in two steps: First, the centers of the neurons of the hidden layer are selected from the training input vectors so that the network output provides a minimum error, and then the weights of the output layer are adapted to minimize the sum-squared error.

### 4.3. Neural network structure

Before using a neural network to solve any problem it is necessary to normalize the input data in order to avoid saturation of the activation function of neurons. It is usual in time series forecasting when data spread over a large range and specially when they present a rising trend. Differentiation is the classic method to do it. Some authors have employed it for load forecasting [5]. A scaled first order differentiation will enclose the trend series data within an appropriate interval:

$$S_n[i] = \frac{S[i] - S[i-1]}{1000} \tag{11}$$

where $S[i]$ is any element of the series and $S_n[i]$ its corresponding normalized value. The scaling parameter (1000) has been obtained by trial and error.

A rescaling process must be also applied to the network output in order to obtain an actual value of the prediction. It will be carried out with an expression easily obtained from (11).

An important issue in the definition of the network structure is the length of the input vector. An intuitive option could be to select them taking into account the periodic behavior of the series. Although periodic components have already been removed from the trend series, weather and social factors can influence the trend behavior, and those factors usually repeat yearly. So, 12 past values of the trend series will form the input vectors. Other options have been tested [34] and this length provided a better performance. It is worth noting that, due to the way the trend extraction and the normalization have been carried out, the first 12 data of the training set are lost, so only 264 are now available for this purpose.

Once the input data structure has been defined different numbers of neurons for both one-layer MLP and RBF networks have been tested. Different values of disp have been also tested with the RBF. Some of the MAPE obtained in the validation stage are shown in Fig. 6. The best performance has been provided by a two neuron RBF with disp = 0.5. The forecasting error indices obtained with this configuration are: MAPE = 0.198%, P95 = 0.498%, P90 = 0.376%, MNE = −0.356% and MPE = 0.540%.

Two-layer MLPs have been also tested, but they have not provided better results than those obtained with a one-layer network, so they have been discarded as they demand a higher computational load.

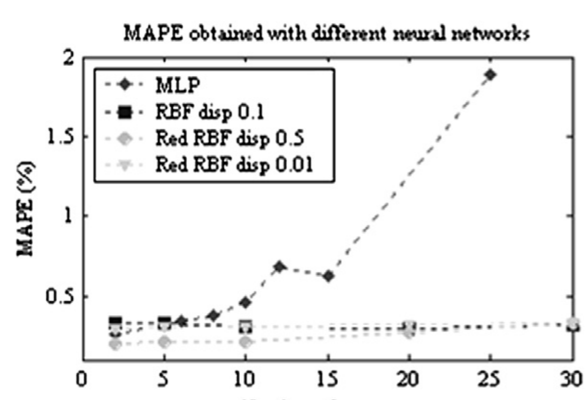


**Fig. 6.** MAPE obtained by neural networks with different number of neurons in the hidden layer.

## 5. Demand forecasting

As trend and fluctuations have been separately forecasted both predictions must be added up in order to obtain the demand forecasting. Then the actual performance of the proposed model may be obtained. The forecasting error indices of the validation set provided by a RBF network (disp = 0.5) with two neurons and the Fourier series (8) adjusted with the data of the five years preceding that to be forecasted, the structures which provided the best performances, may be seen in Table 2, where they have been also grouped by years to provide a more detailed description of the system performance. A plot of actual and forecasted trend, fluctuations and demand may be seen in Figs 7–9, respectively.

In order to test the performance of the model proposed in this work, its predictions have been compared with those obtained with other forecasting models. In the first of them the fluctuation forecasting is carried out by another RBF neural network instead of the Fourier series [25]. A trial and error process was performed

**Table 2**
Forecasting errors year-by-year in the validation period

| Year | MAPE (%) | P95 (%) | P90 (%) | MNE (%) | MPE (%) |
|---|---|---|---|---|---|
| 1998 | 1.571 | 4.212 | 4.187 | −0.504 | 4.216 |
| 1999 | 1.869 | 5.079 | 3.965 | −2.590 | 5.264 |
| 2000 | 1.326 | 4.485 | 2.907 | −2.119 | 4.746 |
| 2001 | 1.719 | 3.087 | 2.794 | −3.136 | 2.647 |
| 2002 | 2.218 | 4.389 | 4.309 | −4.402 | 1.071 |
| Total | 1.740 | 4.335 | 4.180 | −4.402 | 5.264 |

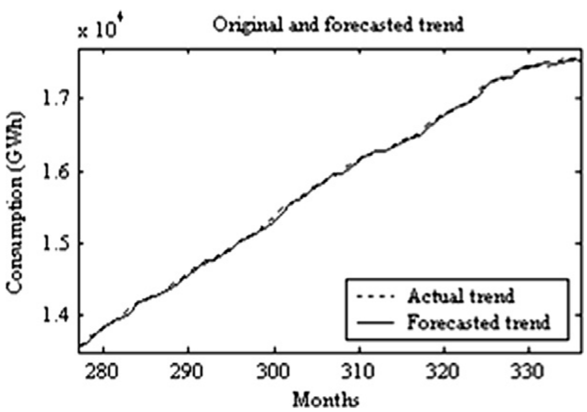


**Fig. 7.** Forecasted versus actual trend data in the validation period.

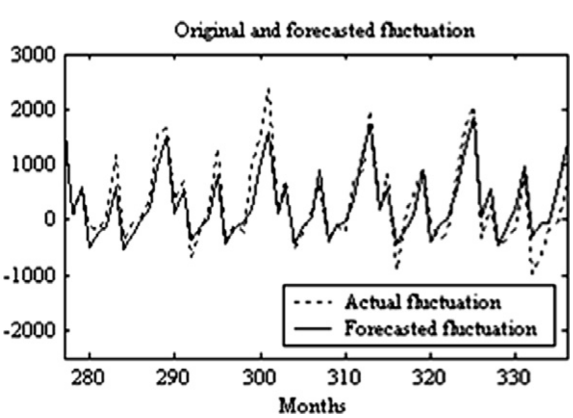


**Fig. 8.** Forecasted versus actual fluctuation data in the validation period.

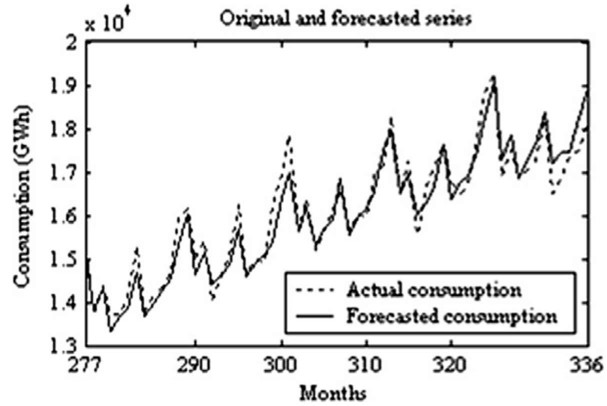


**Fig. 9.** Forecasted versus actual consumption data in the validation period.

to find out the neural structures performing the best. They were a network with two neurons with disp = 0.5 for the trend prediction, the same as that used in the present work, and a network with 50 hidden neurons and a dispersion parameter disp = 0.1 for the fluctuation forecasting. They both had a 12 data input vector. Trend and fluctuations series were obtained in the same way they have been in the present work and then scaled: trend as it has been done in this work and fluctuation by dividing each datum by the trend value of the previous month. The error indices obtained are shown in Table 3. As in [25] a slightly smaller MAPE was obtained when the trend was extracted with splines (MAPE = 1.894%), the trend has been also obtained with this tool and then forecasted with an optimized RBF network, while the fluctuation was predicted with the previously described Fourier series. The performance so obtained was worse than that obtained with the structure described in this work.

Another option is the prediction of the demand with only one neural network. Again a trial and error process was used to find out the best neural structure. It was a RBF network with 100 hidden neurons and a dispersion parameter disp = 0.1, with a 12 data input vector. The consumption series was normalized dividing each datum by the mean of that to be normalized and the preceding 11. The error indices obtained are also shown in Table 3.

A multiplicative ARIMA $(p,d,q)(P,D,Q)_s$ model [35] has been also used to check the accuracy of the proposed model. In this case a $(0,1,1)\ (0,1,1)_{12}$ is used. Trend and fluctuations have been separately forecasted with the ARIMA model and their predictions added up. As the errors so obtained was too high they are not shown in this work. Such a bad result appeared because the ARIMA was unable to accurately forecast the trend series, due to the significant changes in the rising rate it presents, which suggest a non-linear behavior of the series. Nevertheless, fluctuations were predicted with an acceptable accuracy. So, the demand series was directly forecasted with the ARIMA model and a better accuracy was obtained. The yearly and total error indices are shown in Table 4. As it may be seen, not only these results are worse than those obtained with the model proposed in this work, a neural network with a Fourier series, but also they are worse than those provided by one or two neural networks as well.

The results obtained in this work may be compared with those obtained in other works with the same time scale of electric energy demand forecasting [15–20]. In [15,16] ARIMA models have proved to perform better than other methods, like regression or expert system models, to forecast the series under consideration. In [15] a MAPE of 3.8% was obtained with this method. In [16] the ARIMA model was combined with a high pass filter to improve the monthly demand forecasting in Lebanon. In [17] a curve fitting approach was used to forecast monthly demand in countries that

**Table 3**
Forecasting errors obtained when RBFs are used to predict both trend and fluctuation, and when demand is directly forecasted with only one RBF neural network

| Year | Two RBF networks for trend and fluctuation forecasting | | | | | Only one RBF network for the whole demand forecasting | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | MAPE (%) | P95 (%) | P90 (%) | MNE (%) | MPE (%) | MAPE (%) | P95 (%) | P90 (%) | MNE (%) | MPE (%) |
| 1998 | 1.959 | 5.525 | 4.875 | −1.390 | 5.634 | 2.396 | 5.780 | 5.222 | −0.430 | 5.873 |
| 1999 | 1.920 | 5.026 | 4.445 | −2.431 | 5.123 | 2.263 | 5.403 | 4.542 | −1.648 | 5.546 |
| 2000 | 1.788 | 4.614 | 3.679 | −3.212 | 4.770 | 2.002 | 5.760 | 4.734 | −3.119 | 5.931 |
| 2001 | 1.873 | 3.096 | 3.054 | −3.103 | 3.033 | 1.735 | 3.295 | 3.104 | −3.327 | 3.009 |
| 2002 | 2.231 | 4.672 | 4.213 | −4.748 | 2.666 | 2.159 | 3.974 | 3.867 | −3.992 | 3.814 |
| Total | 1.954 | 4.759 | 4.069 | −4.748 | 5.634 | 2.111 | 5.244 | 4.166 | −3.992 | 5.931 |

**Table 4**
Forecasting errors obtained with an ARIMA $(0,1,1)\,(0,1,1)_{12}$ model

| Year | MAPE (%) | P95 (%) | P90 (%) | MNE (%) | MPE (%) |
|---|---|---|---|---|---|
| 1998 | 1.877 | 3.778 | 3.679 | −2.825 | 3.782 |
| 1999 | 3.339 | 6.480 | 5.822 | −0.442 | 7.232 |
| 2000 | 4.863 | 6.974 | 6.746 | 2.616 | 7.223 |
| 2001 | 6.046 | 8.871 | 8.560 | 2.347 | 9.204 |
| 2002 | 4.828 | 6.729 | 6.407 | 1.814 | 7.115 |
| Total | 4.190 | 7.973 | 7.126 | −2.825 | 9.204 |

works with moon calendar, obtaining a MAPE of 12.84%. In [18–20] neural networks are used as forecasting tools. The first one uses both past consumption and actual meteorological data to forecast monthly demand in a fast developing country, reaching a MAPE of 2.03%. In [19], monthly demands in different distribution zones were forecasted with FIR neural networks. Those series showed a clearly linear trend, a fact that defines a little simpler problem than that presented in this work. As a result an average MAPE of 1.88% was reached. Finally, in [20] an MLP neural network was used to forecast a monthly electric consumption series, reaching a MAPE of 6.84%.

The technique proposed in this work provides a MAPE lower than those obtained in the aforementioned works, although the important differences between the demand time series studied in they all make that comparison of results must be done carefully and always from a relative viewpoint. Nevertheless, it has been proved that the split of the demand series into trend and fluctuations in order to forecast them separately by a neural network and a Fourier series provides better results than those obtained with other techniques like ARIMA or only neural networks (forecasting both trend and fluctuation or directly the demand) when they are applied to the same time series.

## 6. Conclusions

Neural networks and curve fitting are very common techniques used for time series forecasting and specifically for load forecasting. However, the combined use of neural networks and Fourier series in electric demand forecasting is a novel approach, which has been proved to provide an adequate performance: a MAPE of 1.740% has been obtained for the whole validation period, while a value of 1.326% was obtained for the best year prediction. The correct separation of trend and fluctuation and the optimization of the forecasting tools used to carry out the predictions are the key issues in the success of this technique.

Furthermore the combined use of a neural network and a Fourier series provides a simpler structure (2 neurons for the network and a Fourier series) than that of two neural networks (2 neurons for the trend and 50 for the fluctuation) [25], and, indeed, than that with only one neural network to carry out the whole prediction (100 neurons).

So it may be assumed that other series with similar characteristics (a more or less periodic fluctuation embedded in a general trend) could be accurately forecasted by using the method proposed in this work.

## References

[1] Papalexopoulos AD, Hesterberg TC. A regression-based approach to short-term load forecasting. IEEE Trans Power Systems 1990;5(4):1535–50.
[2] Peng TM, Hubele NF, Karadi GG. Advancement in the application of neural networks for short-term load forecasting. IEEE Trans Power Systems 1992;7(1):250–7.
[3] Park DC, El-Sharkawi MA, Marks II RJ, Atlas LE, Damborg MJ. Electric load forecasting using an artificial neural network. IEEE Trans Power Systems 1991;6(2):442–9.
[4] Vermaak J, Botha EC. Recurrent neural networks for short-term load forecasting. IEEE Trans Power Systems 1998;13(1):126–32.
[5] Chow TWS, Leung CT. Neural network based short-term load forecasting using weather compensation. IEEE Trans Power Systems 1996;11(4):1736–42.
[6] Rahman S, Hazim O. A generalized knowledge-based short-term load-forecasting technique. IEEE Trans Power Systems 1993;8(2):508–14.
[7] Gao R, Tsoulakas LH. Neural-wavelet methodology for load forecasting. J Intel Robot Systems 2001;31:149–57.
[8] Tzafestas S, Tzafestas E. Computational intelligence techniques for short-term electric load forecasting. J Intel Robot Systems 2001;31:7–68.
[9] Khotanzad A, Hwang RC, Abaye A, Maratukulam D. An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities. IEEE Trans Power Systems 1995;10(3):1716–22.
[10] Abdel-Aal RE. Short-term hourly load forecasting using abductive networks. IEEE Trans Power Systems 2004;19(1):164–73.
[11] Chen BJ, Chang MW, Lin CJ. Load forecasting using support vector machines: a study on EUNITE competition 2001. IEEE Trans Power Systems 2004;19(4):1821–30.
[12] Mbamalu GAN, El-Hawary ME. Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation. IEEE Trans Power Systems 1993;8(1):343–7.
[13] Charytoniuk W, Chen MS. Very short-term load forecasting using artificial neural networks. IEEE Trans Power Systems 2000;15(1):263–8.
[14] Fan JY, McDonald JD. A real-time implementation of short-term forecasting for distribution power systems. IEEE Trans Power Systems 1994;9(2):988–94.
[15] Abdel-Aal RE, Al-Garni AZ. Forecasting monthly electric energy consumption in Eastern Saudi Arabia using univariate time-series analysis. Energy 1997;22(11):1059–69.
[16] Saab S, Badr E, Nasr G. Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. Energy 2001;26(1):1–14.
[17] Barakat EH. Modeling of nonstationary time-series data. Part II. Dynamic periodic trends. Electr Power Energy Systems 2001;23:63–8.
[18] Islam SM, Al-Alawi SM, Ellithy KA. Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network. Electr Power System Res 1995;34:1–9.
[19] Soto CP, Vellasco MMBR, Pacheco MAC. Electric load time series forecasting using neural networks. In: VII International congress of electronic, electrical and system engineering (INTERCON 2000), Lima, Perú. 2000. p. 165–71.
[20] Yalcinoz T, Eminoglu U. Short term and medium term power distribution load forecasting by neural networks. Energy Conv Manage 2005;46:1393–405.
[21] Jia NX, Yokoyama R, Zhou YC, Gao ZY. A flexible long-term load forecasting approach based on new dynamic simulation theory – GSIM. Electr Power Energy Systems 2001;23:549–56.
[22] Padmakumari K, Mohandas KP, Thiruvengadam S. Long term distribution demand forecasting using neuro fuzzy computations. Electr Power Energy Systems 1999;21:315–22.
[23] Ringwood JV, Bofelli D, Murray FT. Forecasting electricity demand on short, medium and long time scales using neural networks. J Intel Robot Systems 2001;31:129–47.
[24] Barakat EH, Al-Qasem JM. Methodology for weekly load forecasting. IEEE Trans Power System 1998;13(4):1548–55.

[25] González-Romera E, Jaramillo-Morán MA, Carmona-Fernández D. Monthly electric energy demand forecasting based on trend extraction. IEEE Trans Power System 2006;21(4):1935–46.
[26] Zhao S, Wei GW. Jump process for the trend estimation of time series. Comput Statist Data Anal 2003;42:219–41.
[27] Achelis SB. Technical analysis from A to Z. McGraw-Hill; 1995.
[28] Kincaid D, Cheney W. Numerical analysis: mathematics of scientific computing. Belmont: Brooks/Cole Publishing Company; 1991.
[29] Mitra SK, Kaiser JF. Handbook for digital signal processing. New York: John Wiley and Sons; 1993.
[30] Hornik K, Stinchcombe M, While H. Multilayer feedforward networks are universal approximators. Neural Networks 1989;2:359–66.
[31] Hartman EJ, Keeler JD, Kowalski JM. Layered neural networks with Gaussian hidden units as universal approximators. Neural Comput 1990;2:210–5.
[32] Bishop CM. Neural networks for pattern recognition. Oxford University Press; 1995.
[33] Chen S, Cowan CFN, Grant PM. Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans Neural Networks 1991;2(2):302–9.
[34] Jaramillo MA, Carmona D, González E, Álvarez JA. Reliability of the forecasting of the monthly demand of electric energy with neural networks. In: International conference on renewable energy and power quality (ICREPQ-2003), Vigo, Spain. 2003. p. 111–2.
[35] Box GEP, Jenkins GM. Time series analysis: forecasting & control. Oakland: Holden-Day, Inc.; 1976.