Neural networks and wavelet transform for short-term electricity prices forecasting

J P S Catalão*,†, H M I Pousinho*, V M F Mendes‡

* Department of Electromechanical Engineering, University of Beira Interior, R. Fonte do Lameiro, 6201-001 Covilha, Portugal. E-mail: catalao@ubi.pt
† Center for Innovation in Electrical and Energy Engineering, Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal.

This paper proposes neural networks in combination with wavelet transform for short-term electricity prices forecasting. In the new deregulated framework, producers and consumers require short-term price forecasting to derive their bidding strategies to the electricity market. Accurate forecasting tools are required for producers to maximize their profits and for consumers to maximize their utilities. The accuracy of the price forecasting attained with the proposed approach is thoroughly evaluated, reporting the numerical results from a real-world case study based on the electricity market of mainland Spain.

Keywords: Electricity Market; Neural Networks; Price Forecasting; Wavelet Transform

1. INTRODUCTION

The electricity industry has undergone significant transformations since the advent of electricity generation in 1882 at Pearl Street Power Station, New York. The electricity industry was organized as regulated and vertically integrated, joining generation, transmission and distribution of electricity in government owned monopolistic companies [1].

When electricity markets were regulated, predicting future prices involved matching regional electricity demand to regional electricity supply. Future regional demand was estimated by escalating historical data, and regional supply was determined by stacking up existing and announced generation units in order of their variable operating costs [2]. Hence, in the regulated framework, the electricity industry’s attention mainly focused on load forecasting, existing little need for tools hedging against price risk given the deterministic nature of electricity prices.

Electricity has been turned into a traded commodity in nowadays, to be sold and bought at market prices. Two ways of trading are usually assumed: the pool trading and bilateral contracts trading. In the pool trading, producers and consumers submit bids respectively for selling and buying electricity on established intervals, typically on an hourly basis.
Finally, a market operator clears the market by accepting the appropriate selling and buying bids, giving rise to the electricity prices.

The new electricity industry deregulated framework was intended to encourage competition among companies in order to decrease the cost of electricity. However, occurrences seldom happening in the regulated framework, such as outages and blackouts, are now subject of increasing concern. Moreover, deregulation brings electricity prices uncertainty, placing higher requirements on forecasting [3]. In particular, accuracy in forecasting these electricity prices is very critical, since more accuracy in forecasting reduces the risk of under/over estimating the revenue for producers and provides better risk management [4].

Short-term electricity prices forecast has become a very helpful tool for producers and consumers. A producer needs to forecast electricity prices to derive its bidding strategy into the pool and to optimally schedule its energy resources [5]. In the regulated framework, traditional generation scheduling of energy resources was based on cost minimization. In the new deregulated framework, since generation scheduling of energy resources, such as hydro resources [6], is now based on profit maximization, electricity prices forecasting has become essential for developing negotiation skills in order to achieve better profits. Consumers need short-term electricity prices forecast for the same reasons as producers.

In the technical literature, several techniques to predict electricity prices have been reported [7], namely hard and soft computing techniques.

The hard computing techniques include auto regressive integrated moving average (ARIMA) [8], wavelet-ARIMA [9], and mixed model [10] approaches. The soft computing or artificial intelligence techniques include neural networks (NN) [11], fuzzy neural networks (FNN) [12], weighted nearest neighbors (WNN) [13], adaptive wavelet neural network (AWNN) [14], and hybrid intelligent system (HIS) [15] approaches.

NN are simple, but powerful and flexible tools for forecasting, provided that there are enough data for training, an adequate selection of the input-output samples, an appropriate number of hidden units and enough computational resources available. Also, NN have the well-known advantages of being able to approximate nonlinear functions and being able to solve problems where the input-output relationship is neither well defined nor easily computable, because NN are data-driven. Successful applications of NN have been reported in the technical literature [16–19]. Three-layered feedforward NN are specially suited for forecasting, implementing nonlinearities using sigmoid functions for the hidden layer and linear functions for the output layer.

This paper proposes NN in combination with wavelet transform (WT) for short-term electricity prices forecasting. The proposed NNWT approach is examined on the electricity market of mainland Spain, commonly used as the test case in several price forecasting papers [8–15]. It has been concluded that the Spanish market has a hard nonlinear behavior and time variant functional relationship [9,12]. So, this market is a real world case study with sufficient complexity.

The proposed NNWT approach is compared with ARIMA, mixed-model, NN, wavelet-ARIMA, WNN, FNN, HIS and AWNN approaches, to demonstrate its effectiveness regarding forecasting accuracy and computation time.

This paper is structured as follows. Section 2 presents the NNWT approach. Section 3 describes the algorithm used to forecast electricity prices. Section 4 provides the importance of price in electricity markets and the main factors that influence it, as well as the different criterions used to assess the behavior of the proposed approach. Section 5 presents the numerical results from a real-world case study based on the electricity market of mainland Spain. Finally, Section 6 outlines the conclusions.

2. PROPOSED APPROACH

The proposed NNWT approach to forecast electricity prices is based on a combination of NN and WT. The WT is used to decompose the usually ill-behaved price series into a set of better-behaved constitutive series. Then, the future values of these constitutive series are forecasted using NN. In turn, the NN forecasts allow, through the inverse WT, reconstructing the future behavior of the price series and therefore to forecast prices.

The WT convert a price series in a set of constitutive series. These constitutive series present a better behavior than the original price series, and therefore, they can be predicted more accurately. The reason for the better behavior of the constitutive series is the filtering effect of the WT [9].

A brief summary of WT is presented hereafter. For the sake of simplicity, one-dimensional wavelets are considered to illustrate the related concepts. WTs can be divided in two categories: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The CWT \( W(a,b) \) of signal \( f(x) \) with respect to a mother wavelet \( \phi(x) \) is given by [14]:

\[
W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \phi \left( \frac{x-b}{a} \right) dx
\]

where the scale parameter \( a \) controls the spread of the wavelet and translation parameter \( b \) determines its central position. The DWT can be defined by:

\[
W(m,n) = 2^{-(m/2)} \sum_{t=0}^{T-1} f(t) \phi \left( \frac{t-n 2^m}{2^m} \right)
\]

where \( T \) is the length of the signal \( f(t) \). The scaling and translation parameters are functions of the integer variables \( m \) and \( n \) \( (a = 2^m, b = n \cdot 2^m) ; t \) is the discrete time index.

Many mother-wavelets are used for different applications. The most known are Haar and Daubechies wavelets. Of great importance are also the symlet, coiflet, Mexican Hat and biorthogonal wavelets. A discussion on the wavelet types is presented as follows.

- Haar: This wavelet is discontinuous, and resembles to a step function. It represents the same wavelet as Daubechies Db1.
- Daubechies: This wavelet has excellent properties of orthogonality and minimum compact support and for hav-
are considered, as shown in Figure 1, since it describes levels (a level of approximation, \(A_3\), and details \(D_1, D_2\) and 

\[
\phi(t)
\]

term electricity prices forecasting. Also, three decomposition wavelet offers an appropriate trade-off between wave-length 

\[\text{viated as } D_b_4\] is used as the mother wavelet 

original signal is broken down into lower resolution compo-

successive decomposition of the approximations, a multilevel 

tain “approximations” and “details” from a given signal. By 

Table 1 presents some properties of the mother-wavelets.

- Symlets: The symlets are nearly symmetrical wavelets proposed by Daubechies as modifications to the Db family. The properties of the two wavelet families are similar.

- Coiflets: Compactly supported wavelets with highest number of vanishing moments for both scaling and wavelet function for a given support width.

- Mexican Hat: This wavelet has no scaling function and is derived from a function that is proportional to the second derivative function of the Gaussian probability density function.

- Biorthogonal: Symmetry with FIR (Finite Impulse Response) filters, desirable properties for decomposition and reconstruction are split and nice allocation is possible. Main difficulty is that the orthogonality is lost.

A fast DWT algorithm based on the four filters (decom-

position low-pass, decomposition high-pass, reconstruction low-pass, and reconstruction high-pass filters), similar to that of Mallat [20], is considered in this paper.

Multiresolution via Mallat’s algorithm is a procedure to ob-

tain “approximations” and “details” from a given signal. By successive decomposition of the approximations, a multilevel decomposition process (Figure 1) can be achieved where the original signal is broken down into lower resolution compo-

A wavelet function of type Daubechies of order 4 (abbreviated as Db4) is used as the mother wavelet \(\phi(t)\). This wavelet offers an appropriate trade-off between wave-length and smoothness, resulting in an appropriate behavior for short-term electricity prices forecasting. Also, three decomposition levels (a level of approximation, \(A_3\), and details \(D_1, D_2\) and 

\(D_3\)) are considered, as shown in Figure 1, since it describes the price series in a more thorough and meaningful way than the others [21].

NN are highly interconnected simple processing units de-

signed in a way to model how the human brain performs a particular task [22]. Each of those units, also called neurons, forms a weighted sum of its inputs, to which a constant term called bias is added. This sum is then passed through a transfer function: linear, sigmoid or hyperbolic tangent.

Multilayer perceptrons are the best known and most widely used kind of NN. The units are organized in a way that defines the network architecture. In feedforward networks, units are often arranged in layers: an input layer, one or more hidden layers and an output layer.

In order to find the optimal network architecture, several combinations should be evaluated. These combinations include networks with different number of hidden layers, different number of units in each layer and different types of transfer functions. The configuration chosen consists of a one hidden layer that uses a hyperbolic tangent sigmoid transfer function and a one unit output layer with a pure linear transfer function. This configuration has been proven to be a universal mapper, provided that the hidden layer has enough units [23]. On one hand, if there are too few units, the network will not be flexible enough to model the data well and, on the other hand, if there are too many units, the network may over-fit the data. Typically, the number of units in the hidden layer is chosen by trial and error, selecting a few alternatives and then running simulations to find out the one with the best results.

Forecasting with NN involves two steps: training and learning. Training of feedforward networks is normally performed in a supervised manner. One assumes that a training set is available, given by the historical data, containing both inputs and the corresponding desired outputs, which is presented to the network. The adequate selection of inputs for NN training is highly influential to the success of training. In the learning process a NN constructs an input-output mapping, adjusting the weights and biases at each iteration based on the minimization of some error measure between the output produced and the desired output. The error minimization process is repeated until an acceptable criterion for convergence is reached.

The most common learning algorithm is the backpropaga-

tion algorithm [24], in which the input is passed layer through layer until the final output is calculated, and it is compared to the real output to find the error. The error is then propagated back to the input adjusting the weights and biases in each layer.

The standard backpropagation learning algorithm is a steep-
est descent algorithm that minimizes the sum of square errors. However, the standard backpropagation learning algorithm is not efficient numerically and tends to converge slowly. An al-

algorithm that trains a NN 10 to 100 times faster than the usual backpropagation algorithm is the Levenberg-Marquardt algo-

While backpropagation is a steepest descent algorithm, the Levenberg-Marquardt algorithm is a variation of Newton’s method [25]. Hence, a three-layered feedforward NN trained by the Levenberg-Marquardt algorithm is considered in this paper.
3. **ALGORITHM DESCRIPTION**

In this section, the algorithm used to implement the proposed approach is described step-by-step. As depicted in Figure 2, wavelet techniques are implemented in the initial and final stages. The actual time-series (electricity price data) are first decomposed into a number of wavelet coefficient signals and one approximation signal. The decomposed signals are then fed into the NN at the second stage to predict the future patterns for each of the signals. Finally, the predicted signals are recombined in the last stage to form the final predicted price series.

1) **First step:** Form a matrix with a set of historical data on electricity prices, arranged in $C$ columns of a matrix thereof. Each column of the array has an associated profile of prices for a particular week where prices are known beforehand. In this first step the matrix has 6 columns, corresponding to the 6 previous weeks to the week whose prices are to be forecasted.

2) **Second step:** Select a number of columns of the previous array so that the set of values derived from it represents the real input data. Correlation analysis is used for feature selection of price forecasting. Hence, appropriate inputs are selected based on a correlation analysis. The candidate inputs with correlation coefficient greater than 0.8 are selected, corresponding to 4 of the 6 weeks with the highest correlation.

3) **Third step:** Decompose the input data using the WT tool available in MATLAB. The operation mode of this process is to decompose the vector with the input data selected. The decomposition is made from the choice of basis functions (wavelet family of functions), and the number of levels wanted to split the series. The signal is divided into three levels, namely, a level of approximation (A) and details (D).

   Figure 1 illustrates the decomposition process. The wavelet function used is the Db4 type, which offers a good approach and ability to use a relatively small number of coefficients, making the code faster. Subsequently, in the level of decomposition, the detail series (for high frequencies) obtained is analyzed, so that they make a selection of coefficients in this series. This selection procedure is known as thresholding, because the purpose is to eliminate the coefficients smaller than a given value, with the aim of improving signal quality by removing noise. Finally, there is the process of reconstruction of the series (from the series of approximate level with the N series about the modified thresholding process - levels 1 to N). The approximation (low-frequency component, A3) subseries, which represents a smoothed version of the price signal, constitutes the main component of the transform, while the details (high-frequency components, D1, D2 and D3) subseries provides “small” adjustments [9].

4) **Fourth step:** In this stage, the wavelet coefficients obtained from WT decomposition are fed into the NN to predict future price data. The approach developed in this paper uses A3, along with D3 and D1, as inputs for the NN.

5) **Fifth step:** A set of feedforward NNs are allocated to forecast the wavelet at different resolution levels. These networks contain only one hidden layer, which is adequate to approximate functions of any complexities. The Levenberg-Marquardt algorithm is used in training the NNs, due to its advantage of small computation time for a large NN size.

6) **Sixth step:** Use WT again to reconstruct the price series forecast given by NN. The final output corresponds to the prediction of our NNWT approach.

4. **ELECTRICITY PRICES FORECASTING**

The electricity price is of extreme importance in an electricity market to all the market players, and in particular for producers and consumers. A priori knowledge of the electricity price is important for risk management and may represent an advantage for a market player facing competition. For companies that trade in electricity markets, the ability to forecast prices means that the company is able to strategically set up bids for the spot market in the short-term.

Electricity price is influenced by many factors: historical prices and demand, bidding strategies, operating reserves, imports, temperature effect, predicted power shortfall and generation outages.

The daily average price in the electricity market of mainland Spain at 2002 is shown in Figure 3.

If all possible factors that influence the electricity price are considered, forecasting will be very accurate, which, however,
is very difficult to do in a real-world case study. Some factors are more important than others and practically only those more important should be considered. The amount of different types of reserves, power import and predicted power shortfall do not improve the forecast at all [26], the effect of the temperature can be incorporated in the demand, and unit outage information is generally proprietary thus not available to all market agents. Also, in the case of NN and ARIMA models, historical demand data does not significantly improve predictions [5]. Extremely high prices with no assessable reasons are the consequence of bidding strategies, which are confidential. Hence, it was decided to use only publicly available information, namely historical price data, to forecast the future prices. The historical prices are natural selections since history and future are correlated.

The shape of price profiles presents seasonality characteristics, usually day and week cycles. The price profile is modified from day to day and week to week, to reflect changes in the electricity market behavior. Typically, daily price profiles are classified as weekdays, from Monday to Friday, and weekend days, Saturday and Sunday, which are different. Another consideration besides weekend is public holiday, known as the calendar effect, since price profiles on non-holidays are particularly different from those on public holidays.

To evaluate the accuracy of the NNWT approach in forecasting electricity prices, different criterions are used. This accuracy is computed in function of the actual prices that occurred. The mean absolute percentage error (MAPE) criterion, the sum squared error (SSE) criterion, and the standard deviation of error (SDE) criterion, are defined as follows.

The MAPE criterion is given by:

$$MAPE = \frac{100}{N} \sum_{h=1}^{N} \frac{|\hat{p}_h - p_h|}{\bar{p}}$$  (3)

where $\hat{p}_h$ and $p_h$ are respectively the forecasted and actual electricity prices at hour $h$, $\bar{p}$ is the average price of the forecasting period and $N$ is the number of forecasted hours.

Electricity price can rise to tens or even hundreds of times of its normal value at particular hours, and it may drop to zero at other hours. Hence, the average price is used in (3) to avoid the adverse effect of prices close to zero [27].

The SSE criterion is given by:

$$SSE = \sum_{h=1}^{N} (\hat{p}_h - p_h)^2$$  (5)

The SDE criterion is given by:

$$SDE = \sqrt{\frac{1}{N} \sum_{h=1}^{N} (e_h - \bar{e})^2}$$  (6)

$$e_h = \hat{p}_h - p_h$$  (7)

$$\bar{e} = \frac{1}{N} \sum_{h=1}^{N} e_h$$  (8)
where $e_h$ is the forecast error at hour $h$ and $\bar{e}$ is the average error of the forecasting period.

5. CASE STUDY

The proposed NNWT approach is applied to forecast next-week prices in the electricity market of mainland Spain. Price forecasting is computed using historical data of year 2002 for the Spanish market.

It should be noted that the electricity market of mainland Spain is a duopoly with a dominant player, resulting in price changes related to the strategic behavior of the dominant player, which are hard to predict [9].

For the sake of simplicity and clear comparison, no exogenous variables are considered. Also, for the sake of a fair comparison, the same test weeks as in [8–15] are selected, which correspond to the four seasons of year 2002. To build the forecasting model, the hourly historical price data of the 42 days previous to the day of the week whose prices are to be forecasts have been considered.

Numerical results with the proposed NNWT approach are shown in Figures 4–7 respectively for the winter, spring, summer and fall weeks. Each figure shows the actual prices, solid line, together with the forecasted prices, dashed line.

Table 2 presents the values for the criterions to evaluate the accuracy of the proposed NNWT approach in forecasting electricity prices. The first column indicates the week, the second column presents the MAPE, the third column presents the square root of the SSE, and the fourth column presents the SDE.

A good accuracy of the proposed NNWT approach was ascertained. The MAPE for the Spanish market has an average value of 6.65%.

All the cases have been run on a PC with 1 GB of RAM and a 2.0-GHz-based processor.
Table 3 Comparative MAPE results.

<table>
<thead>
<tr>
<th>Method</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet-ARIMA [9]</td>
<td>4.78</td>
<td>5.69</td>
<td>10.70</td>
<td>11.27</td>
<td>8.11</td>
</tr>
<tr>
<td>WNN [13]</td>
<td>5.15</td>
<td>4.34</td>
<td>11.40</td>
<td>10.83</td>
<td>8.05</td>
</tr>
<tr>
<td>FNN [12]</td>
<td>4.62</td>
<td>5.30</td>
<td>9.84</td>
<td>10.32</td>
<td>7.52</td>
</tr>
<tr>
<td>AWNN [14]</td>
<td>3.43</td>
<td>4.67</td>
<td>9.64</td>
<td>9.29</td>
<td>7.52</td>
</tr>
<tr>
<td>NNWT</td>
<td>3.61</td>
<td>4.22</td>
<td>9.50</td>
<td>9.28</td>
<td>6.65</td>
</tr>
</tbody>
</table>

Figure 7 Fall week: actual prices, solid line, together with the forecasted prices, dashed line, in euro per megawatt hour.

The proposed NNWT approach presents better forecasting accuracy over the other approaches. Moreover, the average computation time is less than 5 seconds.

6. CONCLUSIONS

A NNWT approach is proposed for electricity prices forecasting on the Spanish market. The application of the proposed approach to price forecasting is both novel and effective. The MAPE has an average value of 6.65%, while the average computation time is less than 5 seconds. Hence, the proposed approach presents a good trade-off between forecasting accuracy and computation time, taking into account results previously reported in the technical literature.

REFERENCES


