Enhanced Forecasting Approach for Electricity Market Prices and Wind Power Data Series in the Short-Term

Gerardo J. Osório 1, Jorge N. D. L. Gonçalves 2, Juan M. Lujano-Rojas 1,3 and João P. S. Catalão 1,2,3,*

1 C-MAST, University of Beira Interior, Covilhã 6201-001, Portugal; gjosilva@gmail.com (G.J.O.);
   lujano.juan@gmail.com (J.M.L.-R.)
2 INESC TEC and the Faculty of Engineering of the University of Porto, Porto 4200-465, Portugal;
   ee10191@fe.up.pt
3 INESC-ID, Instituto Superior Técnico, University of Lisbon, Lisbon 1049-001, Portugal
* Correspondence: catalao@fe.up.pt; Tel.: +351-220-413-295

Academic Editor: Frede Blaabjerg
Received: 31 July 2016; Accepted: 25 August 2016; Published: 31 August 2016

Abstract: The uncertainty and variability in electricity market price (EMP) signals and players’ behavior, as well as in renewable power generation, especially wind power, pose considerable challenges. Hence, enhancement of forecasting approaches is required for all electricity market players to deal with the non-stationary and stochastic nature of such time series, making it possible to accurately support their decisions in a competitive environment with lower forecasting error and with an acceptable computational time. As previously published methodologies have shown, hybrid approaches are good candidates to overcome most of the previous concerns about time-series forecasting. In this sense, this paper proposes an enhanced hybrid approach composed of an innovative combination of wavelet transform (WT), differential evolutionary particle swarm optimization (DEEPSO), and an adaptive neuro-fuzzy inference system (ANFIS) to forecast EMP signals in different electricity markets and wind power in Portugal, in the short-term, considering only historical data. Test results are provided by comparing with other reported studies, demonstrating the proficiency of the proposed hybrid approach in a real environment.

Keywords: adaptive neuro-fuzzy inference system (ANFIS); differential evolutionary particle swarm optimization (DEEPSO); electricity market prices (EMP); forecasting; short-term; time series; wavelet transform (WT); wind power

1. Introduction

In competitive and deregulated electricity markets, potential integration of renewables, especially wind power, which naturally introduces its stochastic, volatile, and uncertain behaviour, is totally reflected in the market players’ strategies and presents more difficulties for a sustainable and robust management of the power framework. Even more when the renewable potential is introduced very widely, yielding higher production costs, inflexibility, and unnecessary penalties due to wrong strategies by players or an increment in emissions caused by conventional producers filling the gaps, especially when the renewable resources suddenly fail or do not cover the required demand [1]. Moreover, with the growing need for smart grids, for example to meet the growing interest in electric vehicles and their integration, the above concerns may be even more pronounced without the use of innovative tools or mechanisms to ensure the quality, safety, and robustness of the electrical system [2].

One of the approaches discussed nowadays in the scientific domain to mitigate some of the problems described above and to achieve a profitable and sustainable management of the electrical framework involves the integration of energy storage systems, which makes the electrical system more flexible due to the increased exploitation of potential usage of renewables, especially under peak loads,
reducing the operational cost or curtailment events; however their implementation is still highly costly and in experimental phases in some cases [3].

An alternative way to tackle the aforementioned concerns in power systems and in competitive electricity markets, which are by nature more economical and useful for all agent players, is through the use of innovative forecasting tools to determine the future behaviour of the renewable potential or electricity market price (EMP) signals; making the creation of sets of possible market strategies suitable, considering other important indicators such as social behaviour, environmental factors, electrical constraints, and the behaviours of other electricity agents; in other words, the forecasting tools may be used as a first stage of defense for all market players [4]. In the last years, massive efforts, supported by the scientific community, have been made to propose more viable and reliable solutions, allowing mitigation of the countless concerns regarding power systems, which are reflected in widespread techniques and forecasting approaches for EMPs or wind power behaviour, considering statistical or physical models in soft or hard computing, as shown for instance in [5–7], considering very short-, short-, and long-term horizon forecasting [8,9].

Regarding EMP forecasting tools, since 2005, models such as autoregressive integrated moving average (ARIMA) combined with wavelet transform (WT) [10] can be found. This model belongs to the family of hard computing tools, which require a large amount of physical information and an exact modelling of the system, resulting in high computational complexity, and in this sense, will not be considered in this review of the state of the art. However, soft computing models, such as fuzzy neural network (FNN) [11] or hybrid intelligent system (HIS) [12], are among the soft computing models, which require the usage of any auto learning process from historical sets to identify future patterns and therefore require less computational complexity or information to model the problem. In this regard, several examples can be found such as neural network (NN) models [13], adaptive wavelet NN (AWNN) [14], cascaded neuro-evolutionary algorithm (CNEA) [15], cascaded NN (CNN) [16], the hybrid neuro-evolutionary system (HNES) [17], and some hybrid forecasting models, such as those presented in [18], or a combination of WT with particle swarm optimization (PSO) and the adaptive neuro-fuzzy inference system (ANFIS) (WPA) [19] and other hybrids [20], the hybrid fundamental-econometric model [21], or two-stage approaches such as those reported in [22,23]. Furthermore, more approaches considering singular spectrum analysis [24], informative vector machine [25], or even new genetic algorithms such as Levenberg-Marquardt and cuckoo search algorithms [26] and genetic regression of relevance vector machines [27] can be found for different EMP prices analyses, considering the Spanish, Pennsylvania-New Jersey-Maryland (PJM), Australian National Electricity Market (ANEM), and other liberalized electricity markets around the world as real case studies.

In wind power forecasting, widespread use of forecasting models for the very short and short term can be found in specialized literature considering soft computing and statistical models. In this sense, several examples are usually found, such as an evolutionary algorithm using an artificial intelligence model [28], NN [29,30], ridgelet NN [31], hybrid approaches composed of WT and a neuro-fuzzy network (NF) [32], WT with NN [33], WT with ANFIS (WAN) [34], or WPA [35]. Also, wind power forecasting can be tackled by considering a combination of WT with support vector machine (SVM) and statistical analysis [36], adaptive WT combined with feed-forward NN (AWNN) [37], WT combined with ARTMAP [38], and optimized SVM using a genetic algorithm [39]. More recently some proposals have considered a principal component analysis algorithm [40], hybrid WT, PSO, and NN [41], multi-layer artificial NN improved with simplified swarm optimization [42], and WT combined with NN, trained by an improved clonal selection algorithm [43]. All of the aforementioned models were run considering real cases with data from wind farms or historical data collected from the public domain in different locations around the world.

In this paper, in accordance with the features demonstrated by hybrid forecasting models briefly presented above, a new approach to forecast the EMP or wind power performance in the short term (from a few to 168 h ahead) is proposed.
Specifically, in the case of EMP forecasting, the proposed approach will perform a forecast for the next 168 h ahead with a time step of 1 h, considering only historical data available from the public domain, without considering the inclusion of exogenous data such as load and other energy prices, among others, to allow a fair and clean comparison with other already published methodologies. In the case of wind power forecasting, the proposed forecasting approach will perform the forecasting for a range of 3 h ahead with a time step of 15 min, refreshing the system (input data and forecast results) until completion of the forecasting results for 24 h ahead. As in the previous case study, in wind power forecasting the proposed approach does not consider the inclusion of exogenous data such as wind profile and atmospheric data, among others, in order to make a fair and clean comparison with previously published approaches.

Furthermore, the proposed approach is composed of an innovative combination of WT as the pre-processing tool, which provides a smoothing effect of all inputs, providing more flexibility and more convergence to forecast the future behaviour, differential evolutionary particle swarm optimization (DEEPSO), which is itself a hybrid method and will be responsible for augmenting the performance of ANFIS (which is by nature a hybrid tool) by tuning the ANFIS membership functions to attain a lower forecasting error. Finally, the inverse WT will be used to introduce again the smoothing information collected at the beginning, providing the final forecasting signal. In this sense, hereafter the proposed approach will be called the hybrid WT+DEEPSO+ANFIS (HWDA) approach. In all case studies, the real historical data used will be comparable to those data used in reported and published models [44,45]. The remainder of the manuscript is organized as follows: Section 2 describes the concepts used to create the HWDA approach, the algorithm used for EMP or wind power forecasting, and the criteria used to validate and compare the capabilities of the proposed HWDA approach with previous and published methodologies. Section 3 describes the historical data used to carry out the forecasting considering the EMP or wind power, the detailed results, and the comparison carried out; finally, Section 4 presents the main conclusions drawn in this paper.

2. Proposed Approach

The HWDA approach results from the successful combination of WT, DEEPSO, and ANFIS. The WT is employed as a pre-processing step to decompose the historical sets of EMP or wind power into new constitutive sets with better behaviour. Then, the forthcoming values of those constitutive sets are the feeding sets of ANFIS responsible for creating the forecast results. DEEPSO augments the performance of ANFIS by tuning the ANFIS membership functions, resulting in lower forecasting error. In comparison with its ancestor, evolutionary particle swarm optimization (EPSO), the underlying evolutionary and differential concepts make real differences in terms of robustness, convergence, and computational time. So the combination of DEEPSO features with the adaptive characteristics of ANFIS means that they complement each other in positive way. Finally, the inverse WT is used to reconstruct the forecasting signal, and thus the final forecasting results are obtained.

2.1. Wavelet Transform

As reported in most of the previously described works on the state of the art, the application of WT in forecasting approaches is important for overcoming the limitations of non-stationary time series such as EMP or wind power; however, it may be applied in other engineering fields, since it enables the analysis of time series in their natural state. WT is used as a pre-processing tool for understanding non-stationary or time varying data [46], with sensibility to the irregularities of input data. In this sense, WT is especially useful for showing different aspects that constitute the data without losing the real signal content [47]. Despite the problems related with continuous WT (CWT) analysis, discrete WT (DWT) was created to give, in an effective way, a description relative to CWT, which is widely used to decompose the time series under study:
\[ DWT(m_{wt}, n_{wt}) = 2^{-(m_{wt}/2)} \sum_{h=0}^{H} p(t_{wt}) \varphi \left( \frac{t_{wt} - b}{a} \right) \]  
(1)

where \( H \) represents the length \( p(t_{wt}) \), and the parameters of scaling \( (a) \) and translation \( (b) \) are changed to integer variables \( a_{wt} = 2^m_{wt} \) and \( b_{wt} = n_{wt}2^m_{wt} \) respectively, with a time-step \( t_{wt} \), i.e.,:

\[ DWT(m_{wt}, n_{wt}) = 2^{-(m_{wt}/2)} \sum_{r=0}^{H} p(t_{wt}) \varphi \left( \frac{t_{wt} - n_{wt}2^m_{wt}}{2^m_{wt}} \right) \]  
(2)

The DWT is performed by multi-resolution analysis, where a “father wavelet”, responsible for the low-frequency series, is used with a complementary “mother wavelet”, which is responsible for the high-frequency series components [38]. In this paper and following the description cited in [44,45] the Daubechies of fourth order, or Db4, was used as the mother-wavelet function. The Db4 has asymmetrical and continuous proprieties, where a higher order level will create a higher level oscillation, which is desirable in forecasting [38,47]. The coefficients of approximations \( A_n \) and details \( D_n \) are expressed as:

\[ A_n = \sum_{n} DWT(m_{wt}, n_{wt}) \varphi_{mn} (t) \]  
(3)

\[ D_n = \sum_{n} DWT(m_{wt}, n_{wt}) \psi_{mn} (t) \]  
(4)

where \( \varphi_{mn} (t_{wt}) \) is the father-wavelet and \( \psi_{mn} (t_{wt}) \) is the mother-wavelet, and \( DWT(m_{wt}, n_{wt}) \) are the coefficients obtained from Equation (2) [33]. Furthermore, the Db4 is chosen as the mother-wavelet function due to a better trade-off between smoothness and length [19]. Also, the DWT used in this paper was created on four filters divided into two groups: the decomposition group, composed of low-pass and high-pass filters, and the reconstruction group, composed of low-pass and high-pass filters as described in [44,45]. Figure 1 shows a general decomposition model of WT, where approximation steps \( A_n \) are able to analyse the universal information of original sets; that is, the low-frequency representation and description of the high-frequency component and the detailed steps \( D_n \) are able to describe the difference between the successive approximations.

\[ \text{Original Data} \rightarrow A_1 \rightarrow \ldots \rightarrow A_n \rightarrow D_1 \rightarrow \ldots \rightarrow D_n \]

**Figure 1.** Universal \( n \) level decomposition model of WT.

### 2.2. Differential Evolutionary Particle Swarm Optimization

DEEPSO is a successful hybrid combination of the EPSO model [44,45], which is itself a hybrid combination of its ancestor model, namely PSO, where weight factors have self-adaptive features, with evolutionary programming, which brings self-adaptive operators [48], and a differential evolution algorithm, which provides a new solution from the current particle of the swarm by adding a fraction difference between two other points found from the previously evaluated swarm [49]. The DEEPSO schema is similar to EPSO [50]; however, the movement rule Equation (6) has new notation:

\[ X_{i}^{new} = X_{i} + V_{i}^{new} \]  
(5)

\[ V_{i}^{new} = w_{i0}^{*} V_{i} + w_{i1}^{*} \left( X_{r1} - X_{i}^{*} \right) + P \omega_{2}^{*} \left( b_{i}^{*} - X_{i} \right) \]  
(6)
where the weights $w^*_{ik}$ (inertia, memory and cooperation) are defined as:

$$w^*_{ik} = w_{ik} + \tau N(0, 1)$$  \hspace{1cm} (7)

and the global position is defined as:

$$b^*_g = b_g (1 + w_g N(0, 1))$$  \hspace{1cm} (8)

From Equation (6), components $X^i_{r1}$ should be any pair of different particle already tested from the swarm, and ordered to minimize at the end of respective iteration, i.e.,:

$$f \left( X^i_{r1} \right) < f \left( X^i_{r2} \right)$$  \hspace{1cm} (9)

From Equations (5)–(9), $X^{new}_i$ is the new position of the particle, $V^{new}_i$ is the new velocity found, $P$ is a diagonal binary matrix with a value of 1 when the probability is $p$ and 0 when the probability is $1 - p$, $w^*_i$ are the mutated weights of inertia, memory, and cooperation of the swarm, given by a learning parameter $\tau$ (fixed or mutated), and $N(0, 1)$ is a random Gaussian variable with 0 mean and variance 1.

Also, $b^*_g$ is the global position provided by the new weight $w^*_g$, which is collected from a diagonal matrix, having a self-adaptive feature, and in this sense, it is a mutated element [48,49]. Components $X^i_{r1}$ and $X^i_{r2}$ guarantee that a suitable extraction really happens, considering macro-gradient points in a descending direction depending on the structured comparison of $f \left( X^i_{r1} \right)$ and $f \left( X^i_{r2} \right)$. In this sense, component $X^i_{r2}$ is assumed to be as $X^i_{r2} = X_i$, and component $X^i_{r1}$ is sampled from the set of best ancestors from the swarm of n particles, that is, $S_{hA} = \{b_1, b_2, \ldots, b_n\}$ [50–52]. The main idea underlying DEEPSO movement is briefly illustrated in Figure 2.

![Figure 2](image-url)

**Figure 2.** Brief illustration of DEEPSO (differential evolutionary particle swarm optimization) particle movement rule.

2.3. Adaptive Neuro-Fuzzy Inference System

ANFIS is a well-known hybrid combination of NN and fuzzy algorithms combining useful features such as low computational requirements, the possibility of dealing with a large number of data, and high response features. Furthermore, it has self-learning capabilities provided by the NN, which help it to self-adjust its parameters due to fuzzy capabilities [19,45]. The general ANFIS structure is based on several layers, which provide the fuzzification, rules, normalization data, desfuzzification, and data reconstruction process as described in [35,44]. Figure 3 briefly describes the multi-layer feed-forward network ANFIS structure. Mathematically, each of the five layers $l_{nk}$ used is:

$$\begin{align*}
\left\{ \begin{array}{ll}
l_{1k} = \mu A_k(x), & k = 1, 2 \\
l_{3k} = \mu B_{k-2}(y), & k = 3, 4
\end{array} \right. \hspace{1cm} (10)
\end{align*}$$
From Equation (10), all nodes $k$ are adaptive nodes with node function $l_{1,k}$, where $x$ and $y$ are the input of the $k_{th}$ node and $A_k$ and $B_{k-2}$ are the membership function, also called the linguistic label, associated with these nodes. In this paper, a triangular membership function is normally used \[44,45\], where \{\(p_k, q_k, r_k\)\} are parameter sets, because it is a continuous and piecewise differentiable function, described in Equation (11), which represents the first layer. In Equation (12), all output nodes represent the firing strength of the rule $w_k$, where each node signal is multiplied by the previous inputs signals, representing the second layer. In Equation (13), the third layer, every node computes the ratio of firing strength rules $k_{th}$ to the sum of all firing strength rules. Equation (14) represents the computation of all nodes’ contribution to $k_{th}$ rule with global output, where \{\(a_k, b_k, c_k\)\} are parameter sets, and $\overline{w}_k$ is the layer output (fourth layer). Finally, Equation (15) defines the ANFIS output node, that is, the fifth layer where the summation $\Sigma$ is made. As reported in \[19,35\], in this paper, the ANFIS structure follows the least-squares and back-propagation gradient descent method, considering the Takagi-Sugeno approach.

\[
\begin{align*}
\mu A_k(x) &= \frac{1}{1 + \left| \frac{x - r_k}{p_k} \right|^{2q_k}} \\
l_{2_k} &= w_k = \mu A_k(x) \mu B_k(y), \quad k = 1, 2 \\
l_{3_k} &= \overline{w}_j = \frac{w_k}{w_1 + w_2}, \quad k = 1, 2 \\
l_{4_k} &= \overline{w}_k z_k = \overline{w}_k (a_k x + b_k y + c_k), \quad k = 1, 2 \\
l_{5_k} &= \sum_{k} \overline{w}_k z_k = \frac{\sum_{k} w_k z_k}{\sum_{k} w_k}
\end{align*}
\]

Figure 3. Brief illustration of ANFIS (adaptive neuro-fuzzy inference system) structure.

2.4. Hybrid Proposed Approach

As stated before, the HWDA approach results from a combination of WT, DEEPSO, and ANFIS. The WT is employed as a pre-processing step to decompose the historical sets. The DEEPSO augments the ANFIS performance by tuning the ANFIS membership functions. Finally, the inverse WT is used to reconstruct the forecasting signal, and then the final forecasting results are obtained. Figure 4 shows the HWDA flowchart. In detail, HWDA follows the following steps:

- Step 1: Initialize the HWDA approach with a historical data matrix of EMP or wind power, respectively, considering the forecasting time-scale of each forecast field;
- Step 2: Choose a set of historical data of the previous step to run the pre-processing process carried out by the WT tool. This step is performed by a backtracking process, in order to attain a smaller error at the end by choosing the best set of candidates. Also, the approach considered in this paper uses $A_3$, $D_3$, and $D_1$ steps as inputs for the next step;
- Step 3: Train the ANFIS tool with the previous sets of constitutive historical data obtained from WT. The optimization process of the ANFIS membership function parameters will be achieved with the DEEPSO method. All parameters considered from all methods are summarized in Table 1.
As in [44,45], the ANFIS inference rules are obtained by considering the automatic ANFIS mode, due to the nature of the data, which requires a large number of inference rules, and thus additional improvement is achieved.

**Figure 4.** HWDA (hybrid WT+DEEPSO+ANFIS) forecasting approach flowchart.

**Table 1.** DEEPSO (differential evolutionary particle swarm optimization) and ANFIS (adaptive neuro-fuzzy inference system) parameters used for EMP (electricity market price) and wind power forecasting.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Parameters</th>
<th>Type or Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>Decomposition Direction</td>
<td>Row</td>
</tr>
<tr>
<td></td>
<td>Level of Decomposition</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mother-Wavelet Function</td>
<td>Db4</td>
</tr>
<tr>
<td></td>
<td>Denoising Methods</td>
<td>“sqtwolog”–“minimaxi”</td>
</tr>
<tr>
<td></td>
<td>Multiplicative Thresholds Rescaling</td>
<td>“one”–“sln”</td>
</tr>
<tr>
<td>DEEPSO</td>
<td>Communication Probability</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Final Inertia Wight</td>
<td>0.01–0.15</td>
</tr>
<tr>
<td></td>
<td>Initial Inertia Weight</td>
<td>0.50–0.90</td>
</tr>
<tr>
<td></td>
<td>Initial Population Size</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Initial Sharing Acceleration</td>
<td>0.50–2.00</td>
</tr>
<tr>
<td></td>
<td>Initial Swarm Learning Process</td>
<td>1.00–2.00</td>
</tr>
<tr>
<td></td>
<td>Initial Swarm Sharing Process</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>Learning Parameter</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Maximum Value of New Position</td>
<td>Set of Max. Inputs</td>
</tr>
<tr>
<td></td>
<td>Minimum Value of New Position</td>
<td>Set of Min. Inputs</td>
</tr>
<tr>
<td></td>
<td>Necessary iterations</td>
<td>100–1000</td>
</tr>
<tr>
<td>ANFIS</td>
<td>Structure Type</td>
<td>Takagi-Sugeno</td>
</tr>
<tr>
<td></td>
<td>Style of Membership Function</td>
<td>Triangular</td>
</tr>
<tr>
<td></td>
<td>Number of Inference Rules</td>
<td>Automatic</td>
</tr>
<tr>
<td></td>
<td>Membership Functions</td>
<td>2–15</td>
</tr>
<tr>
<td></td>
<td>Number of Epochs</td>
<td>2–50</td>
</tr>
<tr>
<td></td>
<td>Number of Nodes</td>
<td>3–9</td>
</tr>
<tr>
<td></td>
<td>Number of Inputs / Outputs</td>
<td>2–5/1</td>
</tr>
</tbody>
</table>
• Step 4: until the best results are obtained or convergence is reached:
  o Step 4.1: Jump to Step 4 in the case of EMP if convergence is not reached;
  o Step 4.2: Jump to Step 2 in the case of wind power forecasting, refreshing the historical
data matrix.

When the best result is found or convergence is reached, the wind power data are forecasted for
the next 3 h until the forecast for the next 24 h ahead is complete.

• Step 5: Apply the inverse WT. The output of the proposed HWDA approach is attained; that is,
the forecasted EMP or wind power results are ready to be presented;
• Step 6: Compute the forecasting errors of EMP or wind power results with different criteria to
validate the proposed HWDA approach and show the results.

2.5. Forecasting Error Evaluation

To compare the proposed approach with other methodologies for EMP or wind power forecasting
previously published in the specialized literature, the mean absolute percentage error (MAPE) criterion
is used. This criterion is given as [44,45]:

$$\text{MAPE} = \frac{100}{N} \sum_{n=1}^{N} \frac{|\hat{p}_n - p_n|}{\bar{p}}$$

(16)

$$\bar{p} = \frac{1}{N} \sum_{n=1}^{N} p_n$$

(17)

where $\hat{p}_n$ is the data forecasted at hour $n$, $p_n$ is the real data at hour $n$, $\bar{p}$ is the average value for the
forecasting time horizon, and $N$ has the length value of observed points. Following the same concept
from the MAPE criterion, the uncertainty of the HWDA model is evaluated using the error variance,
described as [19,35]:

$$\sigma^2_{e,n} = \frac{1}{N} \sum_{n=1}^{N} \left( \frac{|\hat{p}_n - p_n|}{\bar{p}} - e_n \right)^2$$

(18)

$$e_n = \frac{1}{N} \sum_{n=1}^{N} \left| \frac{\hat{p}_n - p_n}{\bar{p}} \right|$$

(19)

Moreover, for wind power forecasting, the normalized mean absolute error (NMAE) criterion is
used [35,45]:

$$\text{NMAE} = \frac{100}{N} \sum_{n=1}^{N} \left| \frac{\hat{p}_n - p_n}{P_{\text{installed}}} \right|$$

(20)

where $P_{\text{installed}} = 2700 \text{ MW}$, which corresponds to the total wind power capacity installed in
accordingly to [53]. Furthermore, the normalized root mean square error (NRMSE) is also used
and is described as [45]:

$$\text{NRMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left( \frac{\hat{p}_n - p_n}{P_{\text{installed}}} \right)^2} \times 100$$

(21)
3. Case Studies and Results

3.1. Electricity Market Prices Forecasting

As briefly stated before, the HWDA approach is used first to forecast EMP for the next 168 and 24 h considering the historical data from the Spanish market available in [54].

As mentioned in [10,21], this market has features that are difficult to forecast due to influences from dominant players, which are reflected in historical data. The EMP historical data used for the Spanish market date back to the year 2002, allowing a clear and fair comparison with the already published results from other proposed methodologies, considering the same four test weeks of the year 2002, which are consistent with the four seasons. As stated before, only EMP historical data sets were used, for the reasons stated above, otherwise a correct comparative study would not be possible.

The HWDA approach forecasts the next 168 h of EMP considering the previous 1008 h (six weeks), which are used as input sets. In order to avoid over-training during the learning process, very large training sets are not used. The output of the HWDA approach results in a set of 168 points representing the forecasting horizon. For day-ahead forecasting, the same idea may be followed; that is, the HWDA approach has as its input the previous six days, considering the historical data from the same market for the year 2006, which were analysed by the case studies reported in [44].

Furthermore, the HWDA approach is tested for the PJM market, forecasting the EMP for the next 24 and 168 h ahead. The historical data of electricity prices are available in [55]. Similarly to the Spanish market, no exogenous data were considered for the same reason as described above.

3.1.1. Spanish Market Results

The results obtained with the HWDA approach are provided in Figures 5–8 for the four test weeks (168 h ahead) of 2002, where the solid and dash-dot black lines represent the actual and forecasted EMP, respectively, while the blue line at the bottom of each figure represents the resulting errors as absolute values. Tables 2 and 3 shows the comparative MAPE criterion and weekly error variance criterion results, respectively, between the HWDA approach and ten previous published methodologies, namely NN [13], FNN [11], AWNN [14], HIS [12], CNEA [15], CNN [16], WPA [19], mutual information with composite NN (MI+CNN) [22], and hybrid evolutionary algorithm (HEA) [44], indicating the enhancements as the percentage evolution between the HWDA approach and the respective comparative methodology under analysis.

![Figure 5. Winter week 2002 results for the Spanish market.](image-url)
Figure 6. Spring week 2002 results for the Spanish market.

Figure 7. Summer week 2002 results for the Spanish market.

Figure 8. Autumn week 2002 results for the Spanish market.
Table 2. MAPE (the mean absolute percentage error) comparison considering the year 2002 Spanish market case study for 168 h ahead.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
<th>Enhancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN [13], 2007</td>
<td>5.23</td>
<td>5.36</td>
<td>11.40</td>
<td>13.65</td>
<td>8.91</td>
<td>54.66%</td>
</tr>
<tr>
<td>FNN [11], 2006</td>
<td>4.62</td>
<td>5.30</td>
<td>9.84</td>
<td>10.32</td>
<td>7.52</td>
<td>46.28%</td>
</tr>
<tr>
<td>HIS [12], 2009</td>
<td>6.06</td>
<td>7.07</td>
<td>7.47</td>
<td>7.30</td>
<td>6.97</td>
<td>42.04%</td>
</tr>
<tr>
<td>AWNN [14], 2008</td>
<td>3.43</td>
<td>4.67</td>
<td>9.64</td>
<td>9.29</td>
<td>6.75</td>
<td>40.15%</td>
</tr>
<tr>
<td>CNEA [15], 2009</td>
<td>4.88</td>
<td>4.65</td>
<td>5.79</td>
<td>5.96</td>
<td>5.32</td>
<td>24.06%</td>
</tr>
<tr>
<td>CNN [16], 2009</td>
<td>4.21</td>
<td>4.76</td>
<td>6.01</td>
<td>5.88</td>
<td>5.22</td>
<td>22.61%</td>
</tr>
<tr>
<td>HNES [17], 2010</td>
<td>4.28</td>
<td>4.39</td>
<td>6.53</td>
<td>5.37</td>
<td>5.14</td>
<td>21.40%</td>
</tr>
<tr>
<td>MI+CNN [22], 2012</td>
<td>4.51</td>
<td>4.28</td>
<td>6.47</td>
<td>5.27</td>
<td>5.13</td>
<td>21.25%</td>
</tr>
<tr>
<td>WPA [19], 2011</td>
<td>3.37</td>
<td>3.91</td>
<td>6.50</td>
<td>6.51</td>
<td>5.07</td>
<td>20.32%</td>
</tr>
<tr>
<td>HEA [44], 2014</td>
<td>3.04</td>
<td>3.33</td>
<td>5.38</td>
<td>4.97</td>
<td>4.18</td>
<td>3.35%</td>
</tr>
<tr>
<td>HWDA</td>
<td>3.00</td>
<td>3.16</td>
<td>5.23</td>
<td>4.76</td>
<td>4.04</td>
<td>-</td>
</tr>
</tbody>
</table>

When the HWDA approach was used, the MAPE criterion reached an average value of 4.04%, which is significant, even when it is compared for each week independently or considering the improvements over all comparative methodologies. The weekly error variance criterion results obtained using the HWDA approach reached an average value of 0.0012, showing a notable accuracy compared with the other methodologies described and reported, even when its improvements are analysed independently.

3.1.2. PJM (Pennsylvania-New Jersey-Mary) Land Market Results

The HWDA approach was also used to forecast the EMP considering the historical data from the PJM market, available in [55], providing results for the next 24 and 168 h ahead. As in the previous case study, no exogenous data are taken into account. Figures 9–11 illustrate some results for some days and weeks tested considering the historical data of 2006 for the PJM market, and the same condition as described in [44] is applied to give a clear and fair comparison with other published methodologies. Moreover, in all figures, the solid and dash-dot black lines represent the actual and forecasted EMP, respectively, while the blue line at the bottom of each figure represents the resulting errors as absolute values.
Tables 4 and 5 shows the MAPE and error variance results, respectively, for the HWDA approach and four other methodologies. When using the HWDA approach, the MAPE criterion reached an average value of 3.16% and the error variance reached an average of 0.0011, which is notable for this competitive market.
3.2. Wind Power Forecasting

The HWDA approach was used to forecast the wind power for 3 h ahead with a time-step of 15 min until the forecast for the whole 24 h ahead was complete, considering the historical data of wind power in Portugal between 2007 and 2008 as described in [45,53] and considering the different seasons of the year. Also, as in the previous case studies, to allow a fair and clean comparison, only historical wind power data are considered, for the same reason as described above. Figures 12–15 show the numerical wind power results for winter, spring, summer, and autumn days, respectively, where solid and dash-dot black lines represent the actual and forecasted wind power, respectively, while the blue line in the bottom figures represents the errors as absolute values. For all results, it is possible to observe how the HWDA approach correctly forecasts the unexpected and abrupt changes of the wind power profile, that is, its uncertainty behaviour during the whole day of forecasting.

![Figure 12. Real and forecasted wind power results (15 min intervals) for the Winter day.](image)

**Table 4.** MAPE (the mean absolute percentage error) comparison considering the year 2006 PJM (Pennsylvania-New Jersey-Mary) land market case study for 24/168 h ahead.

<table>
<thead>
<tr>
<th></th>
<th>HNES [17], 2010</th>
<th>Hybrid [44], 2010</th>
<th>CNEA [15], 2009</th>
<th>HEA [44], 2014</th>
<th>HWDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan. 20</td>
<td>4.98</td>
<td>3.71</td>
<td>4.73</td>
<td>3.29</td>
<td>3.22</td>
</tr>
<tr>
<td>Feb. 10</td>
<td>4.10</td>
<td>2.85</td>
<td>4.50</td>
<td>2.80</td>
<td>2.71</td>
</tr>
<tr>
<td>Mar. 5</td>
<td>4.45</td>
<td>5.48</td>
<td>4.92</td>
<td>3.32</td>
<td>3.27</td>
</tr>
<tr>
<td>Apr. 7</td>
<td>4.67</td>
<td>4.17</td>
<td>4.22</td>
<td>3.55</td>
<td>3.42</td>
</tr>
<tr>
<td>May 13</td>
<td>4.05</td>
<td>4.06</td>
<td>3.96</td>
<td>3.43</td>
<td>3.40</td>
</tr>
<tr>
<td>Feb. 1–7</td>
<td>4.62</td>
<td>5.27</td>
<td>4.02</td>
<td>3.11</td>
<td>3.09</td>
</tr>
<tr>
<td>Feb. 22–28</td>
<td>4.66</td>
<td>5.01</td>
<td>4.13</td>
<td>3.08</td>
<td>3.02</td>
</tr>
<tr>
<td>Average</td>
<td>4.50</td>
<td>4.36</td>
<td>4.35</td>
<td>3.23</td>
<td>3.16</td>
</tr>
<tr>
<td>Enhancement</td>
<td>29.78%</td>
<td>27.52%</td>
<td>27.36%</td>
<td>2.17%</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 5.** Error variance comparison considering the year 2006 PJM (Pennsylvania-New Jersey-Mary) land market case study for 24/168 h ahead.

<table>
<thead>
<tr>
<th></th>
<th>CNEA [15], 2009</th>
<th>Hybrid [44], 2010</th>
<th>HNES [17], 2010</th>
<th>HEA [44], 2013</th>
<th>HWDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan. 20</td>
<td>0.0031</td>
<td>0.0010</td>
<td>0.0020</td>
<td>0.0010</td>
<td>0.0010</td>
</tr>
<tr>
<td>Feb. 10</td>
<td>0.0036</td>
<td>0.0015</td>
<td>0.0012</td>
<td>0.0009</td>
<td>0.0008</td>
</tr>
<tr>
<td>Mar. 5</td>
<td>0.0042</td>
<td>0.0033</td>
<td>0.0015</td>
<td>0.0011</td>
<td>0.0010</td>
</tr>
<tr>
<td>Apr. 7</td>
<td>0.0022</td>
<td>0.0013</td>
<td>0.0018</td>
<td>0.0011</td>
<td>0.0011</td>
</tr>
<tr>
<td>May 13</td>
<td>0.0027</td>
<td>0.0015</td>
<td>0.0013</td>
<td>0.0012</td>
<td>0.0012</td>
</tr>
<tr>
<td>Feb. 1–7</td>
<td>0.0044</td>
<td>0.0037</td>
<td>0.0016</td>
<td>0.0012</td>
<td>0.0011</td>
</tr>
<tr>
<td>Feb. 22–28</td>
<td>0.0035</td>
<td>0.0025</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0016</td>
</tr>
<tr>
<td>Average</td>
<td>0.0034</td>
<td>0.0021</td>
<td>0.0016</td>
<td>0.0012</td>
<td>0.0011</td>
</tr>
<tr>
<td>Enhancement</td>
<td>67.65%</td>
<td>47.62%</td>
<td>45.45%</td>
<td>8.33%</td>
<td>-</td>
</tr>
</tbody>
</table>
Tables 6 and 7 provide a comparative study between the HWDA approach using MAPE and the daily error variance criterion and five other previously published methodologies, namely NN [29],
NF [32], WNF [34], WPA [35], and HEA [45], respectively. When the HWDA approach is used, the MAPE criterion has an average value of 3.37%, representing an enhancement of 11.28% compared to the HEA methodology, which is again significant.

**Table 6.** MAPE (the mean absolute percentage error) comparison for wind power forecasting.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
<th>Enhancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN [29]</td>
<td>9.51</td>
<td>9.92</td>
<td>6.34</td>
<td>3.26</td>
<td>7.26</td>
<td>53.58%</td>
</tr>
<tr>
<td>NF [32]</td>
<td>8.85</td>
<td>8.96</td>
<td>5.63</td>
<td>3.11</td>
<td>6.64</td>
<td>49.25%</td>
</tr>
<tr>
<td>WNF [34]</td>
<td>8.34</td>
<td>7.71</td>
<td>4.81</td>
<td>3.08</td>
<td>5.99</td>
<td>43.74%</td>
</tr>
<tr>
<td>WPA [35]</td>
<td>6.47</td>
<td>6.08</td>
<td>4.31</td>
<td>3.07</td>
<td>4.98</td>
<td>32.33%</td>
</tr>
<tr>
<td>HEA [45]</td>
<td>5.74</td>
<td>3.49</td>
<td>3.13</td>
<td>2.62</td>
<td>3.75</td>
<td>11.28%</td>
</tr>
<tr>
<td>HWDA</td>
<td>5.08</td>
<td>3.19</td>
<td>2.96</td>
<td>2.27</td>
<td>3.37</td>
<td></td>
</tr>
</tbody>
</table>

Furthermore, the daily error variance obtained using the HWDA approach has an average value of 0.0013%, presenting lower uncertainty in the forecasts done, and again, in all results the HWDA approach shows better accuracy in comparison with analyses of the same real case by all other previously published methodologies.

Finally, Tables 8 and 9 show a comparison of the results obtained with the HWDA approach according to the NMAE and NRMSE criteria, respectively. In all cases analysed, it is possible to observe that the HWDA approach gave better results than the other published methodologies considering the same cases studies. The proposed HWDA approach was performed on a standard PC equipped with an Intel Core i7-3537U, 2 GHz CPU and 4 GB of RAM with Windows 10 and the MATLAB® 2016a platform. The authors used the ANFIS and WT structure functions available in MATLAB toolboxes, while DEEPSO was programmed from scratch in MATLAB considering the information available in [49–52].

**Table 7.** Daily error variance comparison for wind power forecasting.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
<th>Enhancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN [29]</td>
<td>0.0044</td>
<td>0.0106</td>
<td>0.0043</td>
<td>0.0010</td>
<td>0.0051</td>
<td>76.47%</td>
</tr>
<tr>
<td>NF [32]</td>
<td>0.0041</td>
<td>0.0086</td>
<td>0.0038</td>
<td>0.0008</td>
<td>0.0043</td>
<td>72.09%</td>
</tr>
<tr>
<td>WNF [34]</td>
<td>0.0046</td>
<td>0.0051</td>
<td>0.0021</td>
<td>0.0011</td>
<td>0.0032</td>
<td>62.50%</td>
</tr>
<tr>
<td>WPA [35]</td>
<td>0.0021</td>
<td>0.0035</td>
<td>0.0016</td>
<td>0.0011</td>
<td>0.0021</td>
<td>42.86%</td>
</tr>
<tr>
<td>HEA [45]</td>
<td>0.0019</td>
<td>0.0015</td>
<td>0.0010</td>
<td>0.0008</td>
<td>0.0013</td>
<td>7.69%</td>
</tr>
<tr>
<td>HWDA</td>
<td>0.0017</td>
<td>0.0016</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0012</td>
<td></td>
</tr>
</tbody>
</table>

**Table 8.** NMAE (the normalized root mean square error) comparison for wind power forecasting.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
<th>Enhancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN [29]</td>
<td>5.22</td>
<td>3.72</td>
<td>2.35</td>
<td>2.15</td>
<td>3.36</td>
<td>84.23%</td>
</tr>
<tr>
<td>NF [32]</td>
<td>4.86</td>
<td>3.36</td>
<td>2.09</td>
<td>2.05</td>
<td>3.09</td>
<td>82.85%</td>
</tr>
<tr>
<td>WNF [34]</td>
<td>4.58</td>
<td>2.89</td>
<td>1.78</td>
<td>2.03</td>
<td>2.82</td>
<td>81.21%</td>
</tr>
<tr>
<td>WPA [35]</td>
<td>3.56</td>
<td>2.28</td>
<td>1.60</td>
<td>2.02</td>
<td>2.37</td>
<td>77.64%</td>
</tr>
<tr>
<td>HEA [45]</td>
<td>2.73</td>
<td>1.48</td>
<td>0.74</td>
<td>1.10</td>
<td>1.51</td>
<td>64.90%</td>
</tr>
<tr>
<td>HWDA</td>
<td>0.94</td>
<td>0.49</td>
<td>0.28</td>
<td>0.39</td>
<td>0.53</td>
<td></td>
</tr>
</tbody>
</table>

**Table 9.** NRMSE (the normalized root mean square error) comparison for wind power forecasting.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
<th>Enhancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEA [45]</td>
<td>3.60</td>
<td>3.18</td>
<td>1.78</td>
<td>2.07</td>
<td>2.66</td>
<td>39.47%</td>
</tr>
<tr>
<td>HWDA</td>
<td>2.19</td>
<td>1.27</td>
<td>1.81</td>
<td>1.18</td>
<td>1.61</td>
<td></td>
</tr>
</tbody>
</table>
4. Conclusions

An enhanced HWDA approach was proposed in this paper for short-term EMP and wind power forecasting considering real cases studies, specifically the analyses from Spanish and PJM markets, as well as the wind power behavior in Portugal. The innovative and successful combination of WT, DEEPSO and ANFIS provided interesting and valuable results. The main findings resulting from this study are related to the lower forecasting errors attained while providing an acceptable computational time. The MAPE criterion reached an average value of 4.04% for the Spanish Market, surpassing all other methodologies, and for the PJM market reached an average value of 3.16%. Regarding the wind power forecasting results, the MAPE criterion had an average value of 3.37%. Lower error variances were also obtained in all cases. Moreover, the computational time required for HWDA approach was less than two min, on average, for the EMP results, and for wind power forecasting took less than one min per iteration. Hence, the overall results obtained with the HWDA approach provided an excellent trade-off between computational time and accuracy, which is crucial for real-life and real-time applications.


Author Contributions: All authors have worked on this manuscript together and all authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

\[ a \] WT scaling integer variable

\[ A_k \] ANFIS linguistic label

\[ a_k \] ANFIS contribution parameter set

\[ A_n \] WT approximation coefficient

\[ b \] WT translation integer variable

\[ b_g \] DEEPSO actual global position

\[ b_g^* \] DEEPSO global position provided by a new weight \( w_g \)

\[ B_k \] ANFIS linguistic label

\[ b_k \] ANFIS contribution parameter set

\[ c_k \] ANFIS contribution parameter set

\[ D_n \] WT detail coefficient

\[ DWT \] Discrete wavelet transform set

\[ e_n \] Error at hour \( n \)

\[ \varphi_{wt} \] WT father-wavelet function

\[ H \] WT length of set \( p \ (t_{wt}) \)

\[ i \] DEEPSO integer time-step from global search space

\[ k \] ANFIS number of nodes

\[ k_{th} \] ANFIS output node

\[ l_{ns} \] ANFIS layer

\[ MAPE \] Mean absolute percentage error

\[ m_{wt} \] WT integer scaling parameter

\[ N \] Length of observed values points

\[ N (0, 1) \] DEEPSO random Gaussian variable with 0 mean and variance 1

\[ NMAE \] Normalized mean absolute error

\[ NRMSE \] Normalized root mean square error

\[ n_{wt} \] WT integer translation parameter
Average value for the forecasting horizon

DEEPSO probabilistic diagonal binary matrix

Data forecasted at hour $n$

Total wind power capacity installed

ANFIS parameter set of membership function

Real data at hour $n$

WT mother-wavelet function

WT signal input

ANFIS parameter set of membership function

ANFIS parameter set of membership function

Error variance from the forecasting horizon

DEEPSO learning parameter

WT time-step

DEEPSO actual velocity

DEEPSO new velocity of the particle

DEEPSO new weight with self-adaptive features

DEEPSO mutated weights of inertia, memory and cooperation

ANFIS firing strength

ANFIS output firing strength

ANFIS input data

DEEPSO actual position

DEEPSO new position of the particle

DEEPSO set of best ancestors from the swarm

DEEPSO set of recorded positions of the swarm

ANFIS input data

ANFIS defuzzification parameters data

References


5. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecast.* 2014, 30, 1030–1081. [CrossRef]


49. Miranda, V.; Alves, R. Differential evolutionary artificial particle swarm optimization (DEEPSO): A successful hybrid. In Proceedings of the 1st BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence (BRICS-CCI and CBIC), Recife, Brazil, 8–11 September 2013.

© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).