A hybrid PSO-ANFIS approach for short-term wind power prediction in Portugal

H.M.I. Pousinho a, V.M.F. Mendes b, J.P.S. Catalão a,c,*

a Department of Electromechanical Engineering, University of Beira Interior, R. Fonte do Lameiro, 6201-001 Covilha, Portugal
b Department of Electrical Engineering and Automation, Instituto Superior de Engenharia de Lisboa, R. Conselheiro Emídio Navarro, 1950-062 Lisbon, Portugal
c Center for Innovation in Electrical and Energy Engineering, Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

Received 27 October 2009; received in revised form 30 June 2010

Abstract

The increased integration of wind power into the electric grid, as nowadays occurs in Portugal, poses new challenges due to its intermittency and volatility. Wind power prediction plays a key role in tackling these challenges. The contribution of this paper is to propose a new hybrid approach, combining particle swarm optimization and adaptive-network-based fuzzy inference system, for short-term wind power prediction in Portugal. Significant improvements regarding forecasting accuracy are attainable using the proposed approach, in comparison with the results obtained with five other approaches.

Keywords: Wind power; prediction; swarm optimization; neuro-fuzzy

1. Introduction

Wind-driven power resources have become increasingly important in the planning and operation of electric power systems. In Portugal, the wind power goal foreseen for 2010 was established by the government as 3750 MW, representing about 25% of the total installed capacity in 2010 [1]. This value has been raised to 5100 MW by the most recent governmental goals for the wind sector. Hence, Portugal has one of the most ambitious goals in terms of wind power and in 2006 was the second country in Europe with the highest wind power growth.

The wind energy is free, so all wind-generated electric energy is accepted as it comes, i.e. as it is available. However, the availability of the power supply generated from wind energy is not known in advance. Hence, the integration of a large share of wind power in an electricity system leads to some important challenges [2]. Wind power prediction plays a key role in tackling these challenges [3].
Short-term wind power prediction is an extremely important field of research for the energy sector, as the system operators must handle an important amount of fluctuating power from the increasing installed wind power capacity. The time scales concerning short-term prediction are in the order of some days (for the forecast horizon) and from minutes to hours (for the time-step) [4].

Wind as a natural phenomenon, lends itself to statistical analysis and physical forecasting [5]. In the technical literature, physical and statistical methods to predict wind power have been reported. The physical method requires a lot of physical considerations to reach the best prediction precision. For a physical model the input variables will be the physical or meteorology information. The statistical method aims at finding the relationship of the on-line measured power data. For a statistical model the historical data of the wind farm may be used. Physical method has advantages in long-term prediction while statistical method does well in short-term prediction [6].

The conventional statistical models are time-series-based models, including auto regressive (AR), and auto regressive integrated moving average (ARIMA) [7] models. The persistence models are considered as the simplest time-series models. They can surpass many other models in very short-term prediction. In spite of the unstable forecasting efficiency, they have been widely used in practice [6]. The persistence approach has proven to be a useful first approximation for short-term wind power prediction and provides a benchmark against which to compare alternative techniques. The Mycielski approach was proposed in [8], which performs a prediction using the total exact history of the data samples. Nevertheless, it has been reported that artificial-based models outperformed others in short-term prediction [6].

In the recent years, some new methods based on artificial intelligence are catching researcher’s attention. In [9], the wind farm power prediction models are built with five different data mining algorithms. A comparison of two advanced statistical short-term wind-power forecasting systems, based on artificial neural networks, is presented in [10]. In [11], NN presented better results for short-term wind power prediction, in comparison with ARIMA and persistence models. Since fuzzy logic [12] and NN are natural complementary tools, a neuro-fuzzy approach was proposed in [13], which can improve the NN results. Evolutionary algorithms [14] and other hybrid methods [15] have also been applied to short-term wind power prediction. Still, the accurate comparison of all the methods is quite difficult because these methods depend on different situations and the data collection is a formidable task.
Taking into account the available literature, novel methodologies are still required in order to improve forecasting accuracy and reduce the uncertainty in wind power predictions, while maintaining an acceptable computation time. This objective lead to the new hybrid approach proposed in this paper for accomplishing a successful short-term wind power prediction in Portugal.

The proposed approach is based on the combination of particle swarm optimization (PSO) and adaptive-network based fuzzy inference system (ANFIS). Our HPA (hybrid PSO-ANFIS) approach is compared with persistence, ARIMA, NN, NN in combination with wavelet transform (NNWT), and wavelet-neuro-fuzzy (WNF) approaches, to demonstrate its effectiveness regarding forecasting accuracy and computation time.

This paper is organized as follows. Section 2 presents the proposed approach to predict wind power. Section 3 provides the different criterions used to evaluate the forecasting accuracy. Section 4 provides the numerical results from a real-world case study. Finally, concluding remarks are given in Section 5.

2. Proposed approach

The proposed approach is based on a combination of particle swarm optimization (PSO) and adaptive-network based fuzzy inference system (ANFIS). The PSO is used to improve the performance of ANFIS, tuning the membership functions required to achieve a lower error. The ANFIS forecasts allow reconstructing the future behavior of the wind power series and therefore to predict wind power.

2.1 Particle swarm optimization

Particle swarm optimization is a heuristic approach first proposed by Kennedy and Eberhart in 1995 [16] as an evolutionary computational method developed for dealing with the optimization of continuous and discontinuous function decision making. The PSO algorithm is based on the biological and sociological behavior of animals such as schools of fish and flocks of birds searching for their food [17].

PSO is a population-based search method where each potential solution is represented as a particle in a population (called swarm). Particles change their position in a multidimensional search space until equilibrium or optimal state has been reached or until computation limitations are exceeded.

Empirical evidence has been accumulated to show that the algorithm is a useful tool for optimization [18]. PSO has been applied to many optimization problems, for instance [19].
Consider an optimization problem of $D$ variables. A swarm of $N$ particles is initialized in which each particle is assigned a random position in the $D$-dimensional hyperspace such that each particle’s position corresponds to a candidate solution for the optimization problem. Let $x$ denote a particle’s position (coordinate) and $v$ denote the particle’s flight velocity over a solution space. Each individual $x$ in the swarm is scored using a scoring function that obtains a fitness value representing how good it solves the problem.

The best previous position of a particle is $P_{best}$. The index of the best particle among all particles in the swarm is $G_{best}$. Each particle records its own personal best position ($P_{best}$), and knows the best positions found by all particles in the swarm ($G_{best}$). Then, all particles that fly over the $D$-dimensional solution space are subject to updated rules for new positions, until the global optimal position is found. Velocity and position of a particle are updated by the following stochastic and deterministic update rules:

$$v_i(t) = \omega v_i(t-1) + \rho_1(x_{p_{best}} - x_i(t)) + \rho_2(x_{g_{best}} - x_i(t)) \quad (1)$$

$$x_i(t) = x_i(t-1) + v_i(t) \quad (2)$$

where $\omega$ is an inertia weight, $\rho_1$ and $\rho_2$ are random variables. The random variables are defined as $\rho_1 = r_1 c_1$ and $\rho_2 = r_2 c_2$, with $r_1, r_2 \sim U(0,1)$, and $C_1$ and $C_2$ are positive acceleration constants.

Fig. 1 illustrates a search mechanism of a PSO technique using the velocity update rule (1) and the position update rule (2).

"See Fig. 1 at the end of the manuscript".

Acceleration constants $C_1$ and $C_2$ represent the weights of the stochastic acceleration terms that push a particle toward $P_{best}$ and $G_{best}$, respectively. Small values allow a particle to roam far from target regions. Conversely, large values result in the abrupt movement of particles toward target regions. In this work, constants $C_1$ and $C_2$ are both set at 2.0, following the typical practice in [20]. Suitable correction of inertia $\omega$ in (2) provides a balance between global and local explorations, thereby reducing the number of iterations when finding a sufficiently optimal solution. An inertia correction function called “inertia weight approach (IWA)” is used in this work [20]. During the IWA, the inertia weight $\omega$ is modified according to the following equation:

$$\omega = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{I_{max}} I_{t} \quad (3)$$
where \( \omega_{\text{max}} \) and \( \omega_{\text{min}} \) are the initial and final inertia weights, \( I_{\text{tr}} \) is the maximum number of iteration, and \( I_{\text{tr}} \) is the current number of iteration.

2.2 ANFIS

Neural networks (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 appropriated number of hidden units and enough computational resources available. Also, NN have the well-known advantages of being able to approximate any nonlinear function and being able to solve problems where the input-output relationship is neither well defined nor easily computable, because NN are data-driven. Multi-layered feedforward NN are specially suited for forecasting, implementing nonlinearities using sigmoid functions for the hidden layer and linear functions for the output layer [21].

Just like NN, a fuzzy logic system is a nonlinear mapping of an input vector into a scalar output, but it can handle numerical values and linguistic knowledge.

In general, a fuzzy logic system contains four components: fuzzifier, rules, inference engine, and defuzzifier. The fuzzifier converts a crisp input variable into a fuzzy representation, where membership functions give the degree of belonging of the variable to a given attribute. Fuzzy rules are of the type “if–then”, and can be derived from numerical data or from expert linguistic. Mamdani and Sugeno inference engines are two of the main types of inference mechanisms.

The Mamdani engine combines fuzzy rules into a mapping from fuzzy input sets to fuzzy output sets, while the Takagi–Sugeno type relates fuzzy inputs and crisp outputs. The defuzzifier converts a fuzzy set into a crisp number using the centroid of area, bisector of area, mean of maxima, or maximum criteria.

NN have the advantage over the fuzzy logic models that knowledge is automatically acquired during the learning process. However, this knowledge cannot be extracted from the trained network behaving as a black box. Fuzzy systems, on the other hand, can be understood through their rules, but these rules are difficult to define when the system has too many variables and their relations are complex [22].

A combination of NN and fuzzy systems has the advantages of each of them. In a neuro-fuzzy system, neural networks extract automatically fuzzy rules from numerical data and, through the learning process, the membership functions are adaptively adjusted.
ANFIS is a class of adaptive multi-layer feedforward networks [23], applied to nonlinear forecasting where past samples are used to forecast the sample ahead. ANFIS incorporates the self-learning ability of NN with the linguistic expression function of fuzzy inference [24].

The ANFIS architecture is shown in Fig. 2. The ANFIS network is composed of five layers. Each layer contains several nodes described by the node function. The node function is described next. Let \( O_i^j \) denote the output of the \( i \)th node in layer \( j \).

"See Fig. 2 at the end of the manuscript".

In layer 1, every node \( i \) is an adaptive node with node function:

\[
O_i^1 = \mu A_i (x), \quad i = 1, 2
\] (4)

or

\[
O_i^1 = \mu B_{i-2} (y), \quad i = 3, 4
\] (5)

where \( x \) (or \( y \)) is the input to the \( i \)th node and \( A_i \) (or \( B_{i-2} \)) is a linguistic label associated with this node.

Thus, \( O_i^1 \) is the membership grade of a fuzzy set \( A (= A_1, A_2, B_1, \text{ or } B_2) \) and it specifies the degree to which the given input \( x \) (or \( y \)) satisfies the quantifier \( A \). The membership functions for \( A \) and \( B \) are usually described by generalized bell functions, e.g.:

\[
\mu A_i (x) = \frac{1}{1 + \left( \frac{x - r_i}{p_i q_i} \right)^2}
\] (6)

where \((p_i, q_i, r_i)\) is the parameter set. As the values of these parameters change, the bell-shaped function varies accordingly, thus exhibiting various forms of membership functions on linguistic label \( A_i \).

In fact, any continuous and piecewise differentiable functions, such as triangular-shaped membership functions, are also qualified candidates for node functions in this layer [25]. Parameters in this layer are referred to as premise parameters.

In layer 2, each node \( \prod \) multiplies incoming signals and sends the product out:

\[
O_i^2 = w_i = \mu A_i (x) \mu B_i (y), \quad i = 1, 2
\] (7)

Hence, each node output represents the firing strength of a rule.
In layer 3, each node $N$ computes the ratio of the $i$th rule’s firing strength to the sum of all rules’ firing strengths:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \quad (8)$$

The outputs of this layer are called normalized firing strengths.

In layer 4, each node computes the contribution of the $i$th rule to the overall output:

$$O_i^4 = \bar{w}_i \ z_i = \bar{w}_i \ (a_i \ x + b_i \ y + c_i), \quad i = 1, 2 \quad (9)$$

where $\bar{w}_i$ is the output of layer 3 and $(a_i, b_i, c_i)$ is the parameter set. Parameters of this layer are referred to as consequent parameters.

In layer 5, the single node $\sum$ computes the final output as the summation of all incoming signals:

$$O_i^5 = \sum \bar{w}_i \ z_i = \sum \frac{w_i \ z_i}{\bar{w}_i} \quad (10)$$

Thus, an adaptive network is functionally equivalent to a Sugeno-type fuzzy inference system.

In this paper, ANFIS employs PSO method to adjust the parameters of the membership functions, as in [26]. The PSO techniques have the advantage of being less computationally expensive for a given size of network topology. The membership functions considered in this study are triangular-shaped.

### 3. Forecasting accuracy evaluation

To evaluate the accuracy of the HPA (hybrid PSO-ANFIS) approach in forecasting wind power, different criterions are used. This accuracy is computed in function of the actual wind power 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 defined as follows:

$$MAPE = \frac{100}{N} \sum_{h=1}^{N} \left| \frac{\hat{p}_h - p_h}{p} \right| \quad (11)$$

$$\bar{p} = \frac{1}{N} \sum_{h=1}^{N} p_h \quad (12)$$
where $\hat{p}_h$ and $p_h$ are respectively the forecasted and actual wind power at hour $h$; $\overline{p}$ is the average wind power of the forecasting period and $N$ is the number of forecasted hours.

The SSE criterion is given by:

$$SSE = \sum_{h=1}^{N} (\hat{p}_h - p_h)^2$$  \hspace{1cm} (13)

The SDE criterion is given by:

$$SDE = \sqrt{\frac{1}{N} \sum_{h=1}^{N} (e_h - \overline{e})^2}$$  \hspace{1cm} (14)

$$e_h = \hat{p}_h - p_h$$  \hspace{1cm} (15)

$$\overline{e} = \frac{1}{N} \sum_{h=1}^{N} e_h$$  \hspace{1cm} (16)

where $e_h$ is the forecast error at hour $h$ and $\overline{e}$ is the average error of the forecasting period.

A measure of the uncertainty of a model is the variability of what is still unexplained after fitting the model, which can be measured through the estimation of the variance of the error. The smaller this variance, the more precise is the prediction [27].

Consistent with definition (11), daily error variance can be estimated as:

$$\sigma^2_{e,\text{day}} = \frac{1}{N} \sum_{h=1}^{N} \left( \frac{\hat{p}_h - p_h}{p} - (e_{\text{day}}) \right)^2$$  \hspace{1cm} (17)

$$e_{\text{day}} = \frac{1}{N} \sum_{h=1}^{N} \left| \frac{\hat{p}_h - p_h}{p} \right|$$  \hspace{1cm} (18)

4. Numerical results

The HPA (hybrid PSO-ANFIS) approach has been applied for wind power prediction in Portugal. Historical wind power data are the main inputs for training. For the sake of clear comparison, no exogenous variables are considered.

The forecast horizon is one day with a time-step of fifteen minutes. The following days are randomly selected: July 3, 2007, October 31, 2007, January 14, 2008, and April 2, 2008, corresponding to the four seasons of the year. Hence, days with particularly good wind power behavior are deliberately not chosen. This results in an uneven accuracy distribution throughout the year that reflects reality.
Numerical results with the HPA approach are shown in Figs. 3 to 6, respectively for the winter, spring, summer and fall days. Each figure shows the actual wind power, solid line, together with the forecasted wind power, dash-dot line.

"See Fig. 3 at the end of the manuscript".

"See Fig. 4 at the end of the manuscript".

"See Fig. 5 at the end of the manuscript".

"See Fig. 6 at the end of the manuscript".

Table 1 presents the values for the criterions to evaluate the accuracy of the HPA approach in forecasting wind power. The first column indicates the day, the second column presents the MAPE, the third column presents the square root of the SSE, and the fourth column presents the SDE.

"See Table 1 at the end of the manuscript".

Table 2 shows a comparison between the HPA approach and five other approaches (persistence, ARIMA, NN, NNWT and WNF), regarding the MAPE criterion.

"See Table 2 at the end of the manuscript".

The proposed approach presents better forecasting accuracy: the MAPE has an average value of 5.41%. Improvement in the average MAPE of the proposed approach with respect to the five previous approaches is 71.6%, 47.7%, 26.4%, 22.4% and 9.7%, respectively.

The absolute values of forecast errors, considering ARIMA, NNWT and HPA approaches, are shown in Figs. 7 to 10, respectively for the winter, spring, summer and fall days.

"See Fig. 7 at the end of the manuscript".

"See Fig. 8 at the end of the manuscript".

"See Fig. 9 at the end of the manuscript".

"See Fig. 10 at the end of the manuscript".

The ARIMA approach provides larger errors compared with NNWT and HPA approaches.

In addition to the MAPE, stability of results is another important factor for the comparison of forecast approaches.
Table 3 shows a comparison between the HPA approach and five other approaches (persistence, ARIMA, NN, NNWT and WNF), regarding the daily error variance.

"See Table 3 at the end of the manuscript".

Note that the average error variance is smaller for the HPA approach, indicating less uncertainty in the predictions. Improvement in the average error variance of the proposed approach with respect to the five previous approaches is 90.1%, 71.3%, 54.9%, 51.1% and 28.1%, respectively.

The HPA approach presents enhanced forecasting accuracy, outperforming the other approaches. Moreover, the average computation time is less than 5 seconds, using MATLAB on a PC with 1 GB of RAM and a 2.0-GHz-based processor. Hence, the proposed approach is both novel and effective for short-term wind power prediction.

5. Conclusions

A new hybrid approach is proposed in this paper for short-term wind power prediction. The proposed approach is based on the combination of particle swarm optimization and adaptive-network-based fuzzy inference system. The application of the proposed approach to wind power prediction is both novel and effective. The MAPE has an average value of 5.41%, outperforming five other approaches while the average computation time is less than 5 seconds. Hence, the presented results validate the proficiency of the proposed approach in short-term wind power prediction.

Acknowledgements

H.M.I. Pousinho thanks the Fundação para a Ciência e a Tecnologia (FCT) for a Ph.D. grant (SFRH/BD/62965/2009).

References

Figure captions

Fig. 1. Updating the position mechanism of PSO.
Fig. 2. ANFIS architecture.
Fig. 3. Winter day: actual wind power, solid line, together with the forecasted wind power, dash-dot line, in megawatt.
Fig. 4. Spring day: actual wind power, solid line, together with the forecasted wind power, dash-dot line, in megawatt.
Fig. 5. Summer day: actual wind power, solid line, together with the forecasted wind power, dash-dot line, in megawatt.
Fig. 6. Fall day: actual wind power, solid line, together with the forecasted wind power, dash-dot line, in megawatt.
Fig. 7. Winter day: absolute value of forecast errors considering ARIMA (dashed line), NNWT (dash-dot line) and HPA (solid line) approaches.
Fig. 8. Spring day: absolute value of forecast errors considering ARIMA (dashed line), NNWT (dash-dot line) and HPA (solid line) approaches.
Fig. 9. Summer day: absolute value of forecast errors considering ARIMA (dashed line), NNWT (dash-dot line) and HPA (solid line) approaches.
Fig. 10. Fall day: absolute value of forecast errors considering ARIMA (dashed line), NNWT (dash-dot line) and HPA (solid line) approaches.
Tables

Table 1
Statistical analysis of the daily forecasting error

<table>
<thead>
<tr>
<th>Day</th>
<th>MAPE</th>
<th>$\sqrt{\text{SSE}}$</th>
<th>SDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>6.71</td>
<td>432.57</td>
<td>26.86</td>
</tr>
<tr>
<td>Spring</td>
<td>7.22</td>
<td>382.98</td>
<td>25.34</td>
</tr>
<tr>
<td>Summer</td>
<td>4.59</td>
<td>168.80</td>
<td>12.44</td>
</tr>
<tr>
<td>Fall</td>
<td>3.13</td>
<td>179.80</td>
<td>11.29</td>
</tr>
</tbody>
</table>
Table 2
Comparative MAPE results

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence</td>
<td>13.89</td>
<td>32.40</td>
<td>13.43</td>
<td>16.49</td>
<td>19.05</td>
</tr>
<tr>
<td>ARIMA</td>
<td>10.93</td>
<td>12.05</td>
<td>11.04</td>
<td>7.35</td>
<td>10.34</td>
</tr>
<tr>
<td>NN</td>
<td>9.51</td>
<td>9.92</td>
<td>6.34</td>
<td>3.26</td>
<td>7.26</td>
</tr>
<tr>
<td>NNWT</td>
<td>9.23</td>
<td>9.55</td>
<td>5.97</td>
<td>3.14</td>
<td>6.97</td>
</tr>
<tr>
<td>WNF</td>
<td>8.34</td>
<td>7.71</td>
<td>4.81</td>
<td>3.08</td>
<td>5.99</td>
</tr>
<tr>
<td>HPA</td>
<td>6.71</td>
<td>7.22</td>
<td>4.59</td>
<td>3.13</td>
<td>5.41</td>
</tr>
</tbody>
</table>
Table 3
Daily forecasting error variance

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence</td>
<td>0.0074</td>
<td>0.0592</td>
<td>0.0085</td>
<td>0.0179</td>
<td>0.0233</td>
</tr>
<tr>
<td>ARIMA</td>
<td>0.0025</td>
<td>0.0164</td>
<td>0.0090</td>
<td>0.0039</td>
<td>0.0080</td>
</tr>
<tr>
<td>NN</td>
<td>0.0044</td>
<td>0.0106</td>
<td>0.0043</td>
<td>0.0010</td>
<td>0.0051</td>
</tr>
<tr>
<td>NNWT</td>
<td>0.0055</td>
<td>0.0083</td>
<td>0.0038</td>
<td>0.0012</td>
<td>0.0047</td>
</tr>
<tr>
<td>WNF</td>
<td>0.0046</td>
<td>0.0051</td>
<td>0.0021</td>
<td>0.0011</td>
<td>0.0032</td>
</tr>
<tr>
<td>HPA</td>
<td>0.0026</td>
<td>0.0038</td>
<td>0.0023</td>
<td>0.0006</td>
<td>0.0023</td>
</tr>
</tbody>
</table>