Renewable Energy 75 (2015) 301–307

Contents lists available at ScienceDirect

Renewable Energy

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

Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information

G.J. Osório [a], J.C.O. Matias [a], J.P.S. Catalão [a, b, c, *]

[a] University of Beira Interior, R. Fonte do Lameiro, 6201-001 Covilha, Portugal
[b] INESC-ID, R. Alves Redol, 9, 1000-029 Lisbon, Portugal
[c] IST, University of Lisbon, Av. Rovisco Pais, 1, 1049-001 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 21 January 2014
Accepted 11 September 2014
Available online

Keywords:
Forecasting
Wind power
Evolutionary particle swarm optimization
Neuro-fuzzy system
Mutual information
Wavelet transform

ABSTRACT

The non-stationary and stochastic nature of wind power reveals itself a difficult task to forecast and manage. In this context, with the continuous increment of wind farms and their capacity production in Portugal, there is an increasing need to develop new forecasting tools with enhanced capabilities. On the one hand, it is crucial to achieve higher accuracy and less uncertainty in the predictions. On the other hand, the computational burden should be kept low to enable fast operational decisions. Hence, this paper proposes a new hybrid evolutionary-adaptive methodology for wind power forecasting in the short-term, successfully combining mutual information, wavelet transform, evolutionary particle swarm optimization, and the adaptive neuro-fuzzy inference system. The strength of this paper is the integration of already existing models and algorithms, which jointly show an advancement over present state of the art. The results obtained show a significant improvement over previously reported methodologies.

1. Introduction

Recently, with the new paradigm shift in the energy sector, and the impositions for a gradual reduction of greenhouse gas emissions, producers are faced with delivering electricity using clean energy sources, in competitive deregulated electricity markets [1,2].

In this context, wind power sources have had the biggest jump in exploration and implementation within the electricity grid [3,4], in comparison with other clean energy technologies [5]. This worldwide expansion of wind energy has occurred due to the ratio between production and implementation costs, maintenance costs, the maturity of technology, and increasing production capacity [6]. However, due to the stochastic characteristic of wind power sources [7–9], its integration is responsible for the introduction of more variability, volatility, and uncertainty in system operation, which complicates the proper management of all production sources [10,11].

The behavior of wind farms depends on the quality and variation of wind speed, the weather conditions, total wind power capacity connected to the electricity grid, scheduled maintenance [12,13], and the wind power acceptance in electrical framework when it is available [14].

Portugal is one of the countries with the fastest growth in wind power production, and by 2020 it hopes to achieve an installed capacity of 8500 MW [15]. Thus, it becomes important to minimize the volatility and intermittent impacts of wind power [16,17], which can be accomplished by the scientific community in presenting new ideas for predicting wind power behavior [18–20]. Wind power forecasting tools represent a very important field of research for system operators, in order to reduce fluctuating power and optimize the installed wind power resources [21].

Wind power forecasting can be classified by time-scales, that is: very short-term, short-term and long-term (of the order of multiple days) [22]. Several wind power forecasting methodologies have been developed and described in the technical literature in recent years, which can be split into physical and statistical methodologies [23].

Physical methodologies need an extensive number of physical specifications, and their inputs are also physical variables, such as orography, pressure and temperature, presenting advantages in

* Corresponding author. University of Beira Interior, R. Fonte do Lameiro, 6201-001 Covilha, Portugal. Tel.: +351 275 329914; fax: +351 275 329972.
E-mail address: catalao@ubi.pt (J.P.S. Catalão).

http://dx.doi.org/10.1016/j.renene.2014.09.058
0960-1481/

long-term forecasting [24]. Statistical methodologies try to establish inherent relationships within the measured data, which can have advantages in short-term forecasting [25,26]. Some statistical methodologies are based on auto regressive techniques, i.e., auto regressive integrated moving average (ARIMA) [27]. Persistence and new reference model (NRM) [28] are also time-series models that can provide a valuable first approximation, and inclusively are able to beat numerical weather prediction (NWP) models for very short-term horizons (between few seconds till 6 h ahead).

Soft computing methodologies have become very popular recently, using an auto learning process from historical sets to identify future patterns, such as neural networks (NNs) [29,30], NNs with wavelet transform (WT), i.e., NNWT [31]; adaptive WT with NN (AWNN) [32], neuro-fuzzy (NF) systems [33,34], evolutionary algorithms [35], and some hybrid methods, such as wavelet-neuro-fuzzy (WNF) and particle swarm optimization (PSO)-WT-NF (WPA) [36].

In this paper, a new hybrid evolutionary-adaptive (HEA) methodology is tested for forecasting wind power, based on MI-mutual information, WT, EPSO-evolutionary particle swarm optimization, and ANFIS-adaptive neuro-fuzzy inference system. The HEA methodology is tested on a real case study using wind power data from Portugal. The object of the study is short-term prediction in wide area forecasting. To prove its superior forecasting accuracy and reduced computational burden, a comparison study will take into account persistence, NRM, ARIMA, NN, NNWT, NF, WNF, and WPA methodologies. This paper is organized in five sections: the proposed methodology (Section 2), forecasting accuracy validation (Section 3), case study (Section 4), and finally conclusions (Section 5).

2. Proposed methodology

The HEA methodology results from the innovative combination of MI, WT, EPSO and ANFIS. The MI is used to eliminate the randomness in the selection of wind power series as inputs, increasing the robustness of the methodology and helping to decrease the final forecasting error [37]. MI is a nonlinear feature selection technique that is more adequate for wind power series than a correlation analysis [23,38]. MI-based techniques in Ref. [23] outperform correlation analysis, which is a linear feature selection method, while wind power is a nonlinear mapping function of its input variables. The WT is employed to decompose the sets of wind power into new constitutive sets with better behavior. Then, the forthcoming values of those constitutive sets are predicted with the ANFIS. EPSO brings on augmented ANFIS performance by tuning their membership functions to attain a lesser error. Comparatively to a classical PSO, the evolutionary concepts behind of EPSO can make a real difference in terms of convergence properties. EPSO is self-adaptive, more robust and less sensitive to parameter initialization, comparatively to classical PSO. The evolutionary characteristics of EPSO and the adaptive characteristics of ANFIS complement each other perfectly. Finally, the inverse WT is used to reconstruct the signal, obtaining then the final forecasting results.

2.1. Mutual information

The MI is based on the concept of entropy. In the case where variable X is a random discrete variable, for example, $(X_1, \ldots, X_n)$, with distribution probabilities $P(X_n)$, the entropy $H(X)$ is given by Refs. [39,40]:

$$H(X) = -\sum_{i=1}^{n} P(X_i)\log_2(P(X_i)) \tag{1}$$

The conditional entropy is defined as:

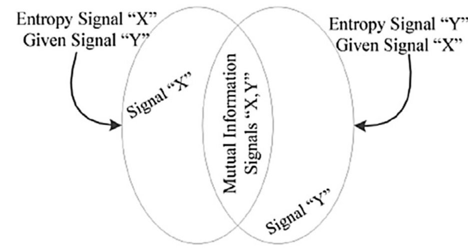

Fig. 1. Simplified MI representation.

$$H(Y/X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i, Y_j)\log_2(P(Y_j/X_i)) \tag{2}$$

The conditional entropy $H(Y/X)$ quantifies the remaining uncertainty of Y when X is known. The joint and conditional entropies are related by:

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y) \tag{3}$$

The MI measures the level of information between a set of information data. The discrete expression is defined as:

$$\mathrm{MI}(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i, Y_j)\log_2\left(\frac{P(X_i, Y_j)}{P(X_i)P(Y_j)}\right) \tag{4}$$

The MI may be given as:

$$\mathrm{MI}(X,Y) = \mathrm{MI}(Y,X) = H(X) - H(X/Y) \tag{5}$$

To ensure the convergence of the HEA methodology, the bounds of MI are very important to guarantee the best performance of the ANFIS. MI helps to determine the best sets of candidates that will be inputs for training the ANFIS tool [41]. Fig. 1 shows a simplified representation about MI.

2.2. Wavelet transform

Non-stationary behavior in a time series arises from instability in the mean and variance of the series. The WT is used in non-stationary or time varying sets [42], being sensitive to the irregularities of input sets [43]. WT tools are capable of illustrating different aspects in the sets without losing the signal [44], reducing the noise of the sets without degradation. The discrete wavelet transform (DWT) is defined [13] as:

$$W(m,n) = 2^{-(m/2)}\sum_{t=0}^{T-1} f(t)\varphi\left(\frac{t-b}{a}\right) \tag{6}$$

In (6), the variable T represents the signal length $f(t)$, the parameters of scaling and translation of φ are given by $a = 2^m$ and $b = n2^m$, respectively, and the time step is given by t. The DWT algorithm used in this work is based on four filters divided into two groups: the decomposition in low and high pass filters, and the reconstruction in low and high filters. The approximations (A_n) and details (D_n) of the

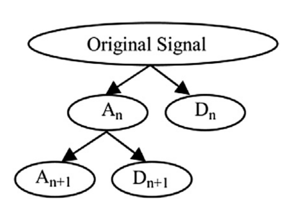

Fig. 2. Level decomposition model of WT.