Renewable Energy 68 (2014) 550–559

Contents lists available at ScienceDirect

Renewable Energy

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

Wind speed forecasting using a portfolio of forecasters

Mario Graff [a,*], Rafael Peña [b], Aurelio Medina [a], Hugo Jair Escalante [c]

[a] División de Estudios de Postgrado, Facultad de Ingeniería Eléctrica, Universidad Michoacana de San Nicolás de Hidalgo, Mexico
[b] Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí, Mexico
[c] Computer Science Department, Instituto Nacional de Astrofísica, Óptica y Electrónica, Mexico

ARTICLE INFO

Article history:
Received 9 September 2013
Accepted 20 February 2014
Available online 15 March 2014

Keywords:
Algorithm portfolio
Algorithm selection problem
Time series forecasting
Power generation capacity
Wind turbines

ABSTRACT

This contribution presents the application of a portfolio of forecasters to the problem of wind speed forecasting. This portfolio is created using a single time series and it is based on a number of time series characteristics, previously proposed, and a set of novel time series features. The results show that the proposed portfolio produces accurate predictions, and, has better performance than the forecasters composing it. In addition to this, the forecast values are used to determine the power generation capacity of a wind turbine driven a permanent magnet synchronous generator.

1. Introduction

The study of trends and patterns of complex systems is of great interest since the results obtained from these studies support the decision-making process in many activities. For the case of wind farms, the wind speed is the variable that decides the amount of electrical energy that can be delivered by the farm. Given the fact that the wind speed is a random variable [1], it is necessary to develop techniques capable of forecasting this variable. The forecast results are in particular important to the optimal day-to-day operation of the power systems [2].

The forecast of the wind speed has been previously addressed and reported in the literature; for example in Refs. [3,4] an Auto-Regressive Integrate Moving Average (ARIMA) model was used, in Refs. [5,6] an Artificial Neural Network (ANN) was developed, closely related, in Refs. [7,8] an ensemble of ANN with empirical mode decomposition was proposed, in Ref. [9] a support vector machine (SVM) with wavelet transform was presented, a Bayesian structural model was used in Ref. [10], and, in Refs. [11,12], a Genetic Programming technique was applied, just to mention a few works. Clearly, there are other variables of interest to the electrical scientific community, such as electric load forecasting [2,13], economic forecasting [14], water flow [15], wave behaviour [16,17], maximum temperature of the next day [18], and solar radiation [19–23], among others. For a recent review in forecasting of wind power generation we refer the reader to Foley et al. [1].

The first problem to face, when forecasting a time series, is to decide among all the different forecasters which one to use. This problem is known as the *algorithm selection problem* [24], the idea is to select among a set of different algorithm which one would perform the best given a particular problem. In this context, the problem is to select the forecaster that would perform the more accurate predictions. Closely related to this problem is the algorithm portfolio. An *algorithm portfolio* is a collection of algorithms that are run in parallel or in sequence to solve a particular problem. That is, the portfolio would perform more accurately predictions than any of the forecasters included in it.

Recently, an algorithm portfolio composed by traditional forecasters was developed in Refs. [25], namely ARIMA (see Refs. [26,27]), and the Exponential smoothing state space model with Box–Cox transformation, Auto-Regressive Moving Average (ARMA) errors, Trend and Seasonal components (BATS) (see Ref. [28]). This portfolio was tested on different time series, and, for each time series, it selects the forecaster that would perform the best for that particular case. Clearly, the approach reported in Ref. [25] is not adequate when there is only one time series. This is because the portfolio would be used only once, consequently, the cost of building it would not be justified.

The idea of automatically selecting a forecaster has been studied previously; one of the first attempts to automate the process of selecting and/or tuning a forecaster was presented in Ref. [29]. The authors proposed a number of rules that were used to combine the forecasts performed by four techniques: random walk, linear

* Corresponding author.
E-mail addresses: mgraffg@dep.fie.umich.mx, mgraffg@gmail.com (M. Graff), rafael.pena@uaslp.mx (R. Peña), amedina@dep.fie.umich.mx (A. Medina), hugojair@inaoep.mx (H.J. Escalante).

http://dx.doi.org/10.1016/j.renene.2014.02.041
0960-1481/

regression, Holt's linear exponential smoothing [30], and Brown's linear exponential smoothing [31]. This work was then extended and improved in Refs. [32,33], where the authors reduced the number of rules and proposed an automatic process to identify the most prominent time series features.

The first fully automated selection procedure was presented in Ref. [34] (and later extended in Ref. [35]). In these works, an induction-based expert system was used to select the most promising forecasting technique based on time series features. Almost during the same period, a discriminant analysis was used to select the most appropriate forecasting technique (see Ref. [36]). It is interesting that neither of them used the terms meta-learning or an algorithm selection problem (a common issue across different domains, further discussed in Ref. [37]).

More recently, this problem has been addressed by different researchers (see Refs. [38–41]) using a variety of machine learning techniques, including linear combination of features, and decision trees, among others; and proposing novel time series characteristics that aim to enrich the set of features describing the time series. All these works have in common that all the features used to tackle the algorithm selection problem are based on time series characteristics. That is, these features are tailored specifically to describe time series, a few examples of these are: trend, seasonality, serial correlation, and periodicity, among others.

The aforementioned procedures, including our previous work [25], perform the next steps to decide which forecaster will be used given a time series, i.e., to solve the algorithm selection problem. Firstly, these techniques require a set of different time series, for each of them their characteristics are computed. Secondly, a machine learning technique is trained to relate the characteristics with the performance of the different forecasters composing the portfolio. Finally, given a new time series, the machine learning technique, trained previously, decides which forecaster will be used to predict the next points of the given time series.

To the best of our knowledge, a procedure, in a portfolio, that solves the algorithm selection problem using a single time series has not been proposed in the literature. This contribution tries to fill this gap by extending our previous work [25] to the case of a single time series. The idea is to solve the algorithm selection problem given a time series, and forecast n points ahead of that time series.

For a completeness of this contribution, the wind speed forecast is used to determine power generation capacity of a stand-alone wind turbine driven by a Permanent Magnet Synchronous Generator (PMSG) in the next hours. We have focused on this research topic due to the fact that wind power is today's most rapidly growing renewable energy source. This as a result of the generalized growing concern regarding the collaboration between fossil fuels and environmental pollution, the depletion of fossil fuels, and in consequence the increase of fossil fuels prices.

From the different types of wind turbines available in the market, in this paper a PMSG wind turbine is used, because this type of wind turbines have drawn great interest to wind turbine manufactures due to the advance of power electronics technology, improved designs and fabrication procedures [42,43]. The model used in this contribution was implemented in the software Simulink/Matlab [44], this with the aim of having a more accurate power generation representation.

The rest of this paper is organized as follows: Section 2 presents the time series characteristics proposed by Wang et al., Lemke et al., and our own. These time series features are extended in Section 3 to be used in the portfolio when there is only one time series. Section 4 presents the details of the PMSG wind generation system. In Section 5, the results obtained in the forecast of the wind speed time series are presented and discussed; finally, Section 6 draws the main conclusions of this research work.

2. Portfolio of forecasters

In general, an algorithm portfolio is built using a machine learning technique that selects the algorithm used to solve the problem at hand. Commonly, supervised learning algorithm [45] is used, and the portfolio is treated as a classification problem or as many regression problems. In a classification problem each different algorithm in the portfolio is given a unique label, and the classification algorithm is trained using a set of pairs of $T = \{(\mathbf{x}_i, y_i) \mid i = 1 \ldots n\}$ where $\mathbf{x}_i$ is a vector of features – this is related to the problems being solved – and y_i corresponds to the algorithm that solved the best the i-th problem.

On the other hand, when the portfolio uses a regression algorithm to make the selection, one needs to solve as many regression problems as algorithms are in the portfolio. More specifically, for the j-th algorithm in the portfolio, a regression problem is created using T, as previously stated, with the difference that y_i is the performance of the j-th algorithm on the i-th problem, e.g., mean square error. This process can be seen as modelling performance of an algorithm (see Ref. [46]); and, for now on, we refer to the trained regression algorithm as being a model of performance, all these models are used to estimate the performance of all the algorithms in the portfolio given a new problem, then the algorithm selected is the one that has the best predicted performance. This approach is followed in this contribution.

As it can be observed, the first step to create a portfolio is to build T. T is build using the pairs $(\mathbf{x}, y)$, where y is given by the performance of the algorithms, and $\mathbf{x} \in \Re^d$ is the vector of features that are related to the problem being solved.

To the best of our knowledge, all previous work designing portfolio of forecasters have followed the following procedure to compute $\mathbf{x}$. Let x_i be the i-th component of $\mathbf{x}$, then $x_i = f_i(\mathbf{y})$ where $\mathbf{y}$ is the time series of length ℓ, and $f_i: \Re^\ell \rightarrow \Re$. Consequently $\mathbf{x}$, is computed using a set of functions $F = \{f_i \mid i = 1 \ldots d\}$ and a time series $\mathbf{y}$. Finally T, is built by applying F on a set of different time series, i.e. $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$.

In the rest of this section, we describe different time series characteristics, i.e. f, that have been proposed previously by other researchers and our own previous works. Subsection 2.1 shows the features proposed by Wang et al. [40]. The time series characteristics proposed by Lemke et al. are shown in Subsection 2.2, and the features proposed in our previous research work [25,46,47] are described in Subsection 2.3.

2.1. Wang et al.'s time series characteristics

Wang et al. [40] proposed 13 time series features used to characterize univariate time series. Let us denote by w_i for $i = 1$ to 13 these time series characteristics, and w has the same signature than f.

Let $w_1 = 1 - \sigma^2(\mathbf{y}_f)/\sigma^2(\mathbf{y}_p - \mathbf{y}_s)$ and $w_2 = 1 - \sigma^2(\mathbf{y}_f)/\sigma^2(\mathbf{y}_p - \mathbf{y}_t)$, where $\mathbf{y}_f = \mathbf{y}_p - \mathbf{y}_t - \mathbf{y}_s$, $\mathbf{y}_p$ are the time series after the Box–Cox transformation[1] [49] $\mathbf{y}_t$ corresponds to the trend of $\mathbf{y}_p$, $\mathbf{y}_s$ is the seasonal component of $\mathbf{y}_p$, and $\sigma^2(\bullet)$ is the variance w_3; is the periodicity of $\mathbf{y}_p - \mathbf{y}_t$ which is computed using the autocorrelation. The next two features, w_4 and w_5, correspond to the serial correlation of $\mathbf{y}$ and $\mathbf{y}_f$, respectively w_6; and w_7 are the nonlinear autoregressive structure measured from $\mathbf{y}$ and $\mathbf{y}_f$, respectively, w_8 and w_9 are the skew of $\mathbf{y}$ and $\mathbf{y}_f$, respectively; and w_{10} and w_{11} are the kurtosis of $\mathbf{y}$ and $\mathbf{y}_f$,

[1] We used the BoxCox transformation implemented in Ref. [48]. λ is obtained using the loglik method with -1 and 1 as its limits.