Applied Energy 191 (2017) 653–662

Contents lists available at ScienceDirect

Applied Energy

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

Short-term wind speed forecasting by spectral analysis from long-term observations with missing values

Hüseyin Akçay *, Tansu Filik

Department of Electrical and Electronics Engineering, Anadolu University, Eskişehir 26555, Turkey

HIGHLIGHTS

- Novel wind speed forecast framework introduced.
- Framework building on de-trending, subspace, identification, and Kalman filtering.
- Intermittently or sequentially missing measurements allowed.
- Persistence forecast method outperformed by the proposed method.

ARTICLE INFO

Article history:
Received 9 October 2016
Received in revised form 21 January 2017
Accepted 27 January 2017
Available online 10 February 2017

Keywords:
Wind energy
Wind speed forecasting
Time-series
Auto-regressive moving average
Kalman filter
Spectrum estimation
Missing data

ABSTRACT

In this paper, we propose a novel wind speed forecasting framework. The performance of the proposed framework is assessed on the wind speed measurements collected from the five meteorological stations in the Marmara region of Turkey. The experimental results show that trimming of the diurnal, the weekly, the monthly, and the annual patterns in the measurements significantly enhances the estimation accuracy. The proposed framework builds on data de-trending, covariance-factorization via a recently developed subspace method, and one-step-ahead and/or multi-step-ahead Kalman filter prediction ideas. The data sets do not have to be complete. In fact, as in sensor failures, intermittently or sequentially missing measurements are permitted. The numerical experiments on the real data sets show that the wind speed forecasts, in particular the multi-step-ahead forecasts, outperform the benchmark values computed with the persistence forecasting models by a clear difference.

1. Introduction

Many countries today are re-thinking their energy portfolio and developing diverse sources of renewable and clean energy due to environmental concerns and supply uncertainties. The worldwide goal is to generate cost-effective energy without major negative impact on the environment. The European Union, for example, has set an ambitious target for the year 2020 to reduce the greenhouse gas emission by 20%, to increase the renewable energy to 20% of the energy supply, and to reduce the overall energy consumption by 20% via the improved energy efficiency in comparison to the year 1990 [1].

Wind is a clean and renewable energy source and has the potential to supply a significant portion of the demand for electricity in favourable regions. For example, in Denmark, Portugal, and Spain the wind energy supplied between 16 and 21% of the electricity demand in 2010. Currently in Turkey, the wind energy is supplying about 5% of the electricity demand; but by the year 2023, it is targeted to reach to 30% [2].

The benefits of the wind energy are, however, accompanied by several challenges, *i.e.,* the limited predictability, the high variability, the nonstorability, and the limited dispatchability. A power system balances the electricity supply and the demand at a minimum cost subject to the transmission network constraints and the contingency plans. In addition, it regulates the frequency and meets the reserve requirements for the reliable and secure system operations. High uncertainty in the wind energy, on the other hand, reduces stability and reliability of power systems; thereby operation costs increase. Instant wind fluctuations make also difficult to mechanically control wind turbines. Thus, it is imperative to accurately forecast short-term wind speeds.

In the last decades, reducing uncertainty in wind speed has been the focus of research and new developments. Numerous methods, generally divided as the physical and the statistical

* Corresponding author.
E-mail addresses: huakcay@anadolu.edu.tr (H. Akçay), tansufilik@anadolu.edu.tr (T. Filik).

http://dx.doi.org/10.1016/j.apenergy.2017.01.063
0306-2619/

approaches, have been proposed in the literature to forecast the wind speed. The physical approaches use the atmospheric whether prediction models [3–5], which produce poor estimation results when used for the short-term wind speed forecasting [6]. As a result, they are suitable only for the long-term predictions, *i.e.*, at least for three-hour predictions [7]. The statistical approaches, on the other hand, typically utilize the linear-in-parameters time-series models [8–15], the nonlinear-in-parameters artificial neural network models [16–18], the wavelet transform models [19,20], and the space-time models, *i.e.,* spatio-temporal, [21–25]. The Kalman filter based forecasting methods were also proposed [26–30]. The hybrid methods unifying the auto-regressive moving average (ARMA) or the auto-regressive integrated moving average (ARIMA) model structures, the Kalman filters, the artificial neural networks, the wavelet transforms, the support vector machines, and the artificial bee colony algorithms were proposed in [31–37]. Time resolution of the short-term forecasting models is usually from several minutes-ahead to several hours-ahead. The fractional-ARIMA models were used in [38] to model and forecast the wind speeds on one day-ahead and two days-ahead horizons. The ARMA models were used in [39] to forecast the wind speed and the wind direction simultaneously.

The statistical approaches are usually preferred for the short-term wind speed forecasting. Nevertheles, they have also found applications in atmospheric whether prediction models. For example, in [27–29] the Kalman filtering was applied as a posterior data-processing procedure. A linear-in-parameters time-series model, see for example [10], does not require long data records for model building and analysis. Moreover, it is easy to implement and applies to a broad range of dynamic systems. A problem with this model structure is that its performance severely deteriorates when some measurements are missing due to sensor failures and the nature of the wind. Degradation in the estimation accuracy when some measurements are missing is usually offset by collecting data over long horizons. The non-linear model structures, which also need long data records, have been introduced to ameliorate shortcomings of the linear models. The non-linear artificial neural network model proposed in [18], for example, uses nine years of wind speed measurements for the assessment of probability distributions. Likewise, the spatio-temporal estimation algorithms proposed in [23,24] utilize longer prediction horizons over multiple sensor locations in order to build accurate models.

In this paper, for the short-time wind speed forecasting we use the wind speed measurements supplied by the Turkish Meteorological Service and averaged over one-hour segments. The short-term forecasting needs significant penetration levels of the wind energy, which may take a long time to develop, as pointed out in [40]. While many forecasting studies have alluded to the annual and the diurnal affects, only in a few works [41,42,23] the diurnal and the annual patterns were explicitly modelled. Any data acquisition procedure is inherently subject to missing values due to sensor failures. In this paper, we propose a novel wind speed forecasting framework. The ingredients of the proposed framework are the data de-trending for periodic components, the covariance-factorization via a recently developed subspace method, and one-step-ahead and/or multi-step-ahead Kalman filter prediction. In this framework, intermittently or sequentially missed measurements are easily handled, without a significant deterioration in the estimation accuracy.

The contents of this paper are as follows. In Section 2, the data sets consisting of the wind speed and the wind direction measurements at the five meteorological stations in the Marmara region of Turkey are described. These measurements cover a period of over six years and have intermittently and sequentially missing values. In Section 3, we present our forecasting scheme and apply it to the experimental data in several stages. Firstly, in Section 3.1 a detailed modelling approach to capture the diurnal, the weekly, the monthly, and the annual patterns is outlined. This stage is necessary in order to represent the residuals as the outputs of ARMA filters driven by white noise. Secondly, in Section 3.2 the auto-correlation coefficients of the ARMA filters are estimated from the modified ensemble averages that accommodate the missing values. Thirdly, in Section 3.3 the subspace-based algorithm developed in [43] are used to identify the ARMA models in the innovation form. A key property of this algorithm is that the identified transfer functions are guaranteed to be positive-real. Next, in Section 3.4, the missing wind speed values and the wind speed forecasts are estimated by the one-step-ahead and the multi-step-ahead predictors. In Section 3.5, the proposed predictors are validated on the unused data sets. A well-known competitor, the family of persistence predictors is applied to the same wind speed data and the estimation results are validated in Section 4. The comparison of the proposed forecasting scheme with the persistence prediction method and other popular methods for the wind speed estimation and the discussion of the numerical results are presented in Section 5. The proposed multi-step-ahead predictors are observed to outperform the persistence predictors by a clear difference. Section 6 concludes the paper.

2. Data sets

The data used in this paper were collected from the five stations of the Turkish Meteorological Service around the Marmara region, shown schematically in Fig. 1. This region is known as having one of the highest wind energy potential in Turkey. The five stations were selected arbitrarily from the available measurement stations in this region. Each measurement station is isolated from the rest and is not at the proximity of the Bosphorus and Çanakkale Straits. From the five meteorological stations, the wind speed and the direction measurements were collected over a period of 6 and 2/3 years between 2008 and 2014. The wind speed was measured 10 m above the ground. The data are available on the web site [44]. In Fig. 2, the wind speed measurements over 58,969 h are plotted for the five stations. The data records are not complete. At Stations BOZ, IPS, GON, BAN, and SIL shown in Fig. 1, 1.68%, 2.04%, 3.70%, 1.80%, and 1.62% of the wind speed and the direction measurements are missing. We will consider the wind speed forecasting from the wind speed measurements with missing values. The wind speed forecasts can be converted into the wind power forecasts by using the deterministic power curves supplied by the wind turbine manufacturer.

3. Short-term wind speed forecasting by spectral analysis

In this section, in several stages we will build input/output models for the wind speed measurements plotted in Fig. 2. Since the data sets are not complete and have been collected over a long period of time, standard identification methods cannot be applied directly. Inspection of Fig. 2 reveals that the measurements vary wildly; thus, making short-term forecasting difficult. On the other hand, except for magnitudes, each data set follows a similar pattern. By focusing on large data segments, we aim to utilize tools developed for stationary random processes. The first-stage of the proposed scheme is the removal of trends from the measurements.

3.1. De-trending of the measurements

In several works [45,23], it has been suggested to remove a diurnal pattern

$$y_d[k] = c_0 + \sum_{i=1}^{n_d} c_{1i} \cos\left(\frac{2\pi ik}{24}\right) + \sum_{i=1}^{n_d} c_{2i} \sin\left(\frac{2\pi ik}{24}\right)$$