Energy Conversion and Management 84 (2014) 133–139

Contents lists available at ScienceDirect

Energy Conversion and Management

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

An appraisal of wind speed distribution prediction by soft computing methodologies: A comparative study

Dalibor Petković [a,*], Shahaboddin Shamshirband [b,c], Nor Badrul Anuar [c], Hadi Saboohi [d], Ainuddin Wahid Abdul Wahab [c], Milan Protić [e], Erfan Zalnezhad [f], Seyed Mohammad Amin Mirhashemi [g]

[a] University of Niš, Faculty of Mechanical Engineering, Department for Mechatronics and Control, Aleksandra Medvedeva 14, 18000 Niš, Serbia
[b] Department of Computer Science, Chalous Branch, Islamic Azad University (IAU), 46615-397 Chalous, Mazandaran, Iran
[c] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[d] Department of Information System, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[e] University of Niš, Faculty of Occupational Safety, Čarnojevića 10A, 18000 Niš, Serbia
[f] Department of Mechanical Engineering, Faculty of Engineering Building, University of Malaya, 50603 Kuala Lumpur, Malaysia
[g] Institute of Asia Europe, University of Malaya, 50603 Kuala Lumpur, Malaysia

A R T I C L E I N F O

Article history:
Received 3 February 2014
Accepted 3 April 2014
Available online 4 May 2014

Keywords:
Wind turbine
Wind speed distribution
Weibull distribution
Support vector regression
Soft computing

A B S T R A C T

The probabilistic distribution of wind speed is among the more significant wind characteristics in examining wind energy potential and the performance of wind energy conversion systems. When the wind speed probability distribution is known, the wind energy distribution can be easily obtained. Therefore, the probability distribution of wind speed is a very important piece of information required in assessing wind energy potential. For this reason, a large number of studies have been established concerning the use of a variety of probability density functions to describe wind speed frequency distributions. Although the two-parameter Weibull distribution comprises a widely used and accepted method, solving the function is very challenging. In this study, the polynomial and radial basis functions (RBF) are applied as the kernel function of support vector regression (SVR) to estimate two parameters of the Weibull distribution function according to previously established analytical methods. Rather than minimizing the observed training error, SVR_poly and SVR_rbf attempt to minimize the generalization error bound, so as to achieve generalized performance. According to the experimental results, enhanced predictive accuracy and capability of generalization can be achieved using the SVR approach compared to other soft computing methodologies.

1. Introduction

Wind plays a central role in many applications, such as exploring wind energy and bridge construction. Hence, having knowledge of wind characteristics is of great importance. When the probability density of wind speed is known, characteristics such as mean, variance and power density can be easily determined. In recent years, the Weibull distribution has been a commonly applied, accepted and recommended distribution for the evaluation of wind energy potential addressed in literature.

In planning offshore wind farms, short-term wind speeds play a central role in estimating various engineering parameters, such as power output, extreme wind load, and fatigue load. Lacking wind speed time series of sufficient length, the probability distribution of wind speed serves as the primary substitute for data when estimating design parameters of wind farm [1].

The Weibull distribution [2] is a two-parameter function employed to fit the wind speed frequency distribution. Among several methods of estimating the parameters of Weibull wind speed distribution are the maximum likelihood and graphical methods. The results show that the wind distribution in the entire sitesbcan best be modelled using the Weibull distribution with the Rayleigh distribution being a close competitor [3]. Article [4] provided an extensive review of some discrete and continuous versions of Weibull distribution modifications. Six kinds of numerical techniques often utilized to estimate Weibull parameters are reviewed in [5], i.e. moment, empirical, graphical, maximum likelihood, modified maximum likelihood and energy pattern factor methods. Results from simulation tests of random variables indicate that the graphical method of estimating Weibull parameters performs the worst, followed by the empirical and energy pattern factor methods if the data number is smaller. Two mathematical models

* Corresponding author. Tel.: +381 643283048.
E-mail address: dalibortc@gmail.com (D. Petković).

http://dx.doi.org/10.1016/j.enconman.2014.04.010
0196-8904/

were proposed [6] that utilize Gaussian statistics and the Weibull distribution respectively, to accurately model the consequences on turbine productivity in turbulent environments. The main purpose of the study in [7] was to investigate the possibility of developing wind speed probability density functions with potentially better accuracy than the maximum entropy principle (MEP) and Weibull distribution. It was shown in [8] that the minimum cross entropy (MinxEnt) principle may serve as an alternative method for accurately estimating both wind distribution and wind power. In article [9], a generalized feed-forward type of neural network (GFNN) was used to predict the annual wind speed probability density distribution. Paper [10] proposed a methodology to estimate the parameters of the Weibull wind speed probability density distribution and its standard errors. Three of the most frequently applied methods to estimate the Weibull distribution parameters are revised and compared in [11]. In another study [12], a new technique was developed to estimate Weibull distribution parameters for wind energy applications, called the power density (PD) method. Estimating the energy output for small-scale wind power generators is the subject of article [13]. Article [14] presents the development of compressed wind speed data to be used in wind energy and performance calculations of standalone or hybrid wind energy systems.

Even though a number of new mathematical functions have been proposed for modeling wind speed probability density distributions, the Weibull function continues to be the most popular model in literature. Aimed at determining such functions to ensure optimal unit operation, many soft computing techniques are presently used, such as fuzzy logic (FL) [15–17], artificial neural networks (ANNs) [18], neuro-fuzzy [19,20] and support vector machines (SVMs) [21,22]. Soft computing is fundamentally made to collect input/output data pairs and learn the proposed network from this data.

Artificial neural networks (ANNs) are being more extensively applied in various areas to overcome the problem with nonlinear relationships and predictions [23]. Support vector machines (SVMs) has recently gained importance in forecasting problems related to the environment [24]. Support vector machines fall under two main categories, namely support vector classification (SVC) and support vector regression (SVR). SVM is a learning system that uses a high-dimensional feature space [25]. Support vector regression (SVR) algorithms, specifically developed for regression problems, are appealing for solving a large variety of regression problems, since they not only take into account data error approximation, but also the model's generalization, i.e., capability to improve the model's prediction when new data is being evaluated [26,27]. SVR is based on a statistical learning theory and structural risk minimization principle, and has successfully been applied in nonlinear system modeling [28,29].

The SVR scheme is for the estimation of annual wind speed probability density distribution. For the presently developed neural network, the same parameters as those required by the Weibull function are used as input for predicting the density distributions in this study. The SVR models were designed based on three methods of estimating the Weibull wind speed distribution parameters: two variations of the maximum likelihood scheme as well as the popular graphical method. In other words, the SVR models should estimate the average two-parameter function of the Weibull distribution based on the existing methods. SVR_rbf and SVR_poly were examined. The first is a radial basis function and the second is a polynomial function. These represent kernel functions utilized to form qualified functions for SVM. Hence, the RBF and polynomial functions are applied to estimate the wake effect on a wind farm in this study.

The objective of this investigation is to establish an SVM to estimate two Weibull function parameters. An attempt is made to retrieve the correlation between effective wind speed and Weibull parameters by SVR methodology. This system should be able to forecast the Weibull parameters. The experimental training data is extracted with three analytical methods, thus quantifying the Weibull two-parameter function.

2. The Weibull distribution

This family of Weibull curves is widely applied in statistical analysis. In wind energy analysis it is used to represent the wind speed probability density function, commonly referred to as the wind speed distribution. The Weibull distribution function is given by

$$P(v < v_i < v + \mathrm{d}v) = P(v > 0)\left(\frac{k}{c}\right)\left(\frac{v_i}{c}\right)^{k-1}\exp\left[-\left(\frac{v_i}{c}\right)^{k}\right]\mathrm{d}v \qquad (1)$$

where c is the Weibull scale parameter, with units equal to the wind speed units; k is the unit-less Weibull shape parameter; v is wind speed; v_i is a particular wind speed; dv is an incremental wind speed; $P(v < v_i < v + \mathrm{d}v)$ is the probability that the wind speed is between v and $v + \mathrm{d}v$; and $P(v > 0)$ is the probability that the wind speed exceeds zero.

Eq. (1) along with the other equations in this paper that refer to probability can be applied equally well regardless of whether probability is interpreted as relative (fractional or percent) or absolute (number of data points). For example, $P(v > 0)$ in Eq. (1) can be interpreted as the fractional probability that wind speed exceeds zero or the number of hours per year that wind speed exceeds zero.

The cumulative distribution function is given by

$$P(v < v_i) = P(v \geqslant 0)\left\{1 - \exp\left[-\left(\frac{v_i}{c}\right)^{k}\right]\right\} \qquad (2)$$

where $P(v < v_i)$ is the probability that the wind speed is less than v_i, and $P(v \geqslant 0)$ is the probability that the wind speed equals or exceeds zero.

The two Weibull parameters and average wind speed are related by

$$\bar{v} = c \cdot \Gamma\left(1 + \frac{1}{k}\right) \qquad (3)$$

where $\bar{v}$ is the average wind speed and $\Gamma()$ is the gamma function.

2.1. Wind speed data

Measured wind speed data is normally available in time-series format, in which each data point represents either instantaneous sample or average wind speed over a time period. In some instances, wind speed data may instead be available in frequency distribution format. In this case, the frequency at which the wind speed falls within various ranges (bins) is given. The methods described in the following section can be used to estimate the Weibull parameters given wind speed in either time-series or frequency distribution format.

2.2. Determining the Weibull parameters

Three methods of estimating the parameters of Weibull wind speed distribution are used in this study: two variations of the maximum likelihood method as well as the popular graphical method.

2.2.1. The maximum likelihood method

The Weibull distribution can be fitted to time-series wind data using the maximum likelihood method. The shape factor k and scale factor c are estimated using the following two equations: