Renewable Energy 83 (2015) 362–374

Contents lists available at ScienceDirect

Renewable Energy

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

Assessment of large scale wind power generation with new generation locations without measurement data

Jussi Ekström [a, *], Matti Koivisto [a], Ilkka Mellin [b], John Millar [a], Eero Saarijärvi [a], Liisa Haarla [a]

[a] Aalto University School of Electrical Engineering, Department of Electrical Engineering, P.O. Box 13000, FI-00076 Aalto, Finland
[b] Aalto University School of Science, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076 Aalto, Finland

ARTICLE INFO

Article history:
Received 17 September 2014
Accepted 21 April 2015
Available online 14 May 2015

Keywords:
Autoregressive model
Monte Carlo simulation
Power system planning
Weibull distribution
Wind power
Wind speed

ABSTRACT

Large amounts of new wind power are currently under construction or planning in many countries. The constantly increasing percentage of wind power in the electricity generation mix has to be taken into consideration when planning power systems. This paper introduces a Monte Carlo simulation based methodology that can be used to assess the effects (e.g. need for new transmission lines, reserves, wind curtailment or demand side management) of large amounts of existing and planned wind power generation on the power system. The presented methodology is able to assess new wind power scenarios spread over a wide geographical area, comprising numerous existing and planned wind generation locations. The Monte Carlo simulation results are verified against measured aggregated wind power generation in Finland from 2008 to 2014. In addition, case studies of future scenarios with 232 individual wind generation locations are presented to show the applicability of the methodology as a tool in power system planning.

1. Introduction

The amount of wind power generation has been increasing rapidly and new wind farms are under construction or in planning in several countries. Therefore, as the amount of wind power in the generation mix rises, it becomes increasingly vital to assess the contemporaneous generation from a large number of wind generation locations covering large geographical areas. To be able to evaluate future scenarios, locations where no wind speed or wind power measurements exist (non-measured locations) have to be analyzed. However, existing methodologies that assess the contemporaneous generation of non-measured locations are limited.

This paper adds to the literature by presenting a Monte Carlo simulation based methodology to analyze future wind generation scenarios with a large number of non-measured locations. This methodology can be used as a flexible tool for power system planners to assess the impact of different wind power generation scenarios, e.g. the need for reserves, new transmission lines, wind curtailment or demand side management. The methodology is able to investigate the impact and feasibility of several generation locations or even large systems in multiple countries.

A major part of the presented methodology involves Monte Carlo based simulations and probability integral transformations. Time series models are created to enable the implementation of Monte Carlo simulations in the assessment of wind power generation in multiple locations. These simulations have also been used with wind power generation in Ref. [1–4]. Probability integral transformations are employed to separate the analysis of the marginal distributions of wind speeds or wind power in individual generation locations from the analysis of the dependence structures of multiple locations [1–4]. We use a similar approach to separate the wind speed distributions of the individual locations from their spatial and temporal dependencies.

To simulate wind speed time series, a vector autoregressive (VAR) model can be employed [4–7]. The VAR model assesses both temporal and spatial dependencies between all locations [4]. However, as shown in Refs. [4], new non-measured locations cannot be straightforwardly added to the model. Therefore, a VAR model was not considered a feasible approach. Thus, we use a time series model that follows the basic approach presented in

* Corresponding author. gsm +358 44 0592699.
E-mail addresses: jussi.ekstrom@aalto.fi (J. Ekström), matti.koivisto@aalto.fi (M. Koivisto), ilkka.mellin@aalto.fi (I. Mellin), john.millar@aalto.fi (J. Millar), eero.saarijarvi@aalto.fi (E. Saarijärvi), liisa.haarla@aalto.fi (L. Haarla).

http://dx.doi.org/10.1016/j.renene.2015.04.050
0960-1481/

Refs. [4,8]: individual univariate time series simulated by autoregressive (AR) models are transformed to multivariate cross-correlated time series with the Cholesky decomposition of a correlation matrix. This combination is referred to as the ARC model [4]. However, the ARC model presented in Ref. [4] assumes that the error terms of the model are normally distributed. In this paper, we have improved the model and present a refined ARC model with t-distributed error terms. This extended ARC model (with t-distributed error terms) allows a better analysis of the volatility of the wind generation, as the error terms of the AR models, which model the variability, are modeled more accurately.

The monthly changing diurnal variations are assessed by removing the diurnal variations from the data before estimation and adding them back in simulation, as was done in Refs. [4,5].

Two distributions are used for the marginal distributions: the Weibull distribution and a piecewise semi-parametric distribution with a non-parametric density estimator with Gaussian kernels for the majority of the data and a generalized Pareto distribution for the upper tail [4,9–11]. This piecewise distribution is referred as the KGP distribution.

The marginal distributions are used to transform the measured wind speed data to data with normally distributed marginal distributions to be used as an input to the ARC model with t-distributed error terms. This combination is referred to as the transformed ARC model with t-distributed error terms. The ARC model and the marginal distributions are both fitted to the measured wind speed data. The transformed ARC model with t-distributed error terms can analyze new wind generation locations without measured data and can be utilized to simulate hundreds or even thousands of generation locations.

Wind speed time series are simulated for each location with the ARC model with t-distributed error terms. A turbine model is specified for each location and is used to transform the simulated wind speed time series to a wind power time series through a piecewise power curve specified by turbine specific parameters. A third degree function is considered for the non-linear part of the power curve [12]. The presented methodology also assesses the wake effect inside wind farms [13], the changing installed aggregate generation capacity and the availability of the turbines. The presented model (built from individual locations) is verified against aggregated wind power generation data in Finland from 2008 to 2014. Case studies comprising six future wind power generation scenarios and an evaluation of the European Wind Energy Association (EWEA) high 2020 target scenario [14] are presented to illustrate the applicability of the proposed methodology.

2. The wind speed model

This section presents the data used to estimate and verify the model and considers the marginal distributions, introduces the transformed ARC model with t-distributed error terms and shows how non-measured wind generation locations are added to the model.

2.1. The data

Low and high altitude wind speed data, and wind power generation and capacity data are used in the analyses in this paper. The low altitude wind speed data is from 19 locations in Finland. The measurement data is hourly and the measurement height is approximately 15 m above the surrounding ground level. The data has been measured in all of the 19 locations, from July 2008 to July 2011. The low altitude wind speed data are obtained from the Finnish Meteorological Institute.

The high altitude wind speed data are from 12 different locations in Finland and the time resolution of the measurement data is 1 h. Seven of the 12 locations are tower measurements and the measurement height varies from location to location, from 74 m to 120 m. Five of the measurements were made with SODAR technology and in these cases measurements from 75, 100, 125 and 150 m above the surrounding ground level are considered. The measurement lengths and dates vary between the different high altitude locations. The high altitude wind speed data are obtained from the Finnish Meteorological Institute.

For both high and low altitude locations the wind speed data is denoted as $\boldsymbol{y}_t = [y_{1,t}, \ldots, y_{k,t}]'$, where $t = 1, \ldots, T$ is time and k is the number of locations.

The wind power generation and capacity data from Finland are from January 2008 to January 2014. The generation data are aggregated from all wind generators in Finland. The resolution of the measured aggregate generation data is hourly. Both the generation and capacity data are obtained from the Finnish Energy Industries.

The Finnish Wind Atlas database [15] is used to obtain the Weibull parameters for the marginal distributions of the non-measured locations. As the Finnish Wind Atlas provides Weibull parameters for different coordinates and altitudes, the ARC model can be applied to different geographical locations and all realistic hub heights of the turbines. The terrain roughness and other geographical factors are already considered in the Weibull parameters, as the Finnish Wind Atlas utilizes geographical information in the estimation of the Weibull parameters.

2.2. The marginal distributions

This paper uses two different distributions as marginal distributions for wind speeds. The KGP distribution consists of two parts, as presented in Ref. [4]. First, a non-parametric density estimator with Gaussian kernels is fitted for the wind speed data below the highest 10% of the measurements. Second, the distribution of the highest 10% of the measured wind speeds is fitted with a generalized Pareto distribution [4,9–11].

The KGP distribution is a very flexible distribution, but the non-parametric density estimator cannot be used in locations without measurement data. Accordingly, KGP distributions are only used for the measured locations employed in the estimation of the model. The estimated KGP distributions were capable of depicting the measured data more accurately than, or at least as accurately as, the Weibull distributions for all the measured locations [4].

The Weibull parameters for the new non-measured locations, however, are available from various databases, such as the Finnish Meteorological Institute's Finnish Wind Atlas [15]. Thus the Weibull distribution, widely used to describe wind speed distributions, can be used with the non-measured locations [12,16].

2.3. The transformed ARC model with t-distributed error terms

This section introduces the transformed ARC model with t-distributed error terms and the estimation process of the model. The flowchart of the estimation is presented in Fig. 1. First, the marginal distributions are estimated for the measured wind speed data $\boldsymbol{y}_t$ and then the estimated marginal distributions are transformed to the standard normal. The normality of the data is not requisite with AR models but it is desirable [5].

The probability integral transformation from wind speeds to data with normally distributed marginal distributions is