Solar Energy 143 (2017) 10–21

Contents lists available at ScienceDirect

Solar Energy

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

A copula method for simulating correlated instantaneous solar irradiance in spatial networks

Joakim Munkhammar [a,*], Joakim Widén [a], Laura M. Hinkelman [b]

[a] Built Environment Energy Systems Group (BEESG), Department of Engineering Sciences, Uppsala University, SE-751 21 Uppsala, Sweden
[b] Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, 3737 Brooklyn Ave. NE, Seattle, WA 98105, United States

ARTICLE INFO

Article history:
Received 7 October 2016
Received in revised form 9 December 2016
Accepted 12 December 2016
Available online 3 January 2017

Keywords:
Instantaneous solar irradiance
Clear-sky index
Aggregate clear-sky index
Copula

ABSTRACT

This paper presents a method for generating correlated instantaneous solar irradiance data for an arbitrary set of spatially dispersed locations. Based on the empirical clear-sky index distribution for one location and the cross-correlation between clear-sky index data at all location pairs, a copula is used to represent the dependence between locations. The method is primarily intended for probabilistic simulations of electricity distribution grids with high penetrations of photovoltaic (PV) systems, in which solar irradiance data for nodes in the grid can be sampled from the model. The method is validated against a 10-s resolution solar irradiance data set for 14 locations, dispersed within an array of approximately 1 km × 1.2 km, at the Island of Oahu, Hawai'i, USA. The results are compared with previous results for along- and cross-wind pairs of locations, and with models for adjacent (completely correlated) and dispersed (completely uncorrelated) locations. It is shown that the copula approach performs better than the adjacent model for a majority of all location pairs and for all but one pair of locations separated more than 500 m. It outperforms the dispersed model for all pairs of locations. In conclusion, the proposed method can generate correlated data and estimate the aggregate clear-sky index for any set of locations based only on the distribution of the clear-sky index for a single location.

1. Introduction

Over the last decades, grid-connected photovoltaic (PV) capacity has increased heavily in several countries. In Germany, for example, the contribution of PV to the total electricity production in 2015 was 6%, or 39 TW h, from a total installed capacity of 40 GW (IEA-PVPS, 2016). Worldwide, the total installed PV power continues to increase, with more than 50 GW added in 2015, bringing the total global capacity to almost 230 GW (REN21, 2016). With increasing generation capacities, often from systems dispersed and connected to medium-voltage (MV) and low-voltage (LV) distribution grids, comes the need for improved tools and approaches for taking PV into account in grid design, operation and control (Smith et al., 2015).

Analyses of power grids are typically based on power flow simulations, where the power flow in the grid is solved numerically, yielding currents in all lines and voltages in all buses (Grainger and Stevenson, 1994). To accurately simulate grid impacts of distributed PV systems, it is necessary to have unique PV power generation data for individual nodes of a grid. For example, in a LV network, this could be as detailed as the production profiles of individual customers. For a large MV network, it would at least be necessary to have data on the aggregate solar power in MV/LV substations.

Often, such studies have used solar irradiance or PV power generation data for one point location as representative for a larger area. Thomson and Infield (2007) and Watson et al. (2016) are two representative examples, although being solid and detailed studies in other respects. This is a rough method and does not take into account that the solar irradiance, especially on days with broken clouds, differs between points in the grid, or that the aggregate irradiance over an area can be considerably smoothed out (Perez and Hoff, 2013; Widén et al., 2015). The ideal would be to have unique data for each grid-connected PV system that would differ depending on the distance between sites, resulting in various amounts of smoothing of the aggregate power output from the systems.

One option for obtaining spatially resolved irradiance is to use satellite-based data (Miller et al., 2013; Perez et al., 2013). However, the spatial resolution (typically several km) is still too low to capture differences within LV and MV grids. Ruf et al. (2016) evaluated satellite data in comparison to ground-measured data from a point location for modeling the power output from a fleet

* Corresponding author.
E-mail address: joakim.munkhammar@angstrom.uu.se (J. Munkhammar).

http://dx.doi.org/10.1016/j.solener.2016.12.022
0038-092X/

of PV systems in a LV grid. As the satellite data gave a slightly lower root-mean-square error, it should be preferred over ground-measured data, but to capture spatial differences in near-instantaneous irradiance in distribution grids other methods would be required.

Another option is to use a model for generating spatially distributed irradiance data. Generating this type of data, that incorporates distance-dependent correlations and the dispersion-smoothing effect, is a more or less complex task. There exist models, e.g., for smoothing out single-site time series to represent the profile of a distributed PV fleet or a large PV plant (Lave et al., 2013). However, to our knowledge, there is no model available that generates data for individual sites that have a varying degree of similarity depending on site separation, and when aggregated yield realistic smoothed-out data.

Moreover, in grid simulations a time series approach is not always ideal. Probabilistic load flow (PLF) approaches, first proposed in 1974 (Borkowska, 1974), have become increasingly more common for distribution systems with distributed generation. In PLF, power consumption and generation data for nodes in a grid are drawn from probability distributions and the power flow is typically solved numerically with Monte Carlo methods (Bollen and Hassan, 2011). To improve these kinds of simulations, probabilistic models for the instantaneous solar irradiance are required, from which it should be possible to randomly draw correlated values for an arbitrary number of sites. Existing studies on probabilistic power flow have commonly used elaborate methods for correlating different simulation variables, for example PV systems with wind turbines and load (Ran and Miao, 2016), but not so much taken into account the spatial correlations within the same power source.

In this paper, we describe and test the validity of a model for spatially distributed instantaneous solar irradiance, using copula modeling (Nelsen, 2006). Since the model is primarily intended for PLF simulations of instantaneous system states (e.g., the probability for overvoltage in a distribution grid at a specific time), it only considers correlation in the spatial domain. The model is based on random sampling from joint probability distributions describing the instantaneous solar irradiance at sites in a spatial network. The dependence between the sites is represented by the station-pair correlation, which is used to generate correlated samples from the distributions. With this approach we connect three methodologies that have so far not been combined in the field of solar irradiance modeling:

1. *Correlation modeling for spatial solar irradiance networks*. Perez et al. (2011), Hoff and Perez (2012) and Perez et al. (2012) were the first to describe station-pair correlations for solar irradiance over short distances and high time resolutions. Isotropic models for this correlation and its dependence on the station-pair distance were later suggested by Perez and Hoff (2013) and Lave and Kleissl (2013). Hinkelman (2013) showed with empirical measurements that the correlation is anisotropic, i.e. highly dissimilar for along-wind station pairs and cross-wind station pairs. Anisotropic correlation models have later been proposed, e.g., by Lonij et al. (2013) and Arias-Castro et al. (2014). So far, correlations have been mainly evaluated for step changes and not for instantaneous radiation, as in this paper.
2. *Probabilistic modeling of instantaneous solar irradiance for point locations*. The nature of solar irradiance for individual locations has been studied thoroughly, starting with initial studies by Ångström (1922), Ångström (1924), Ångström (1956) relating the duration of bright sunshine to the average solar irradiance by a linear fit. Improved measurement resolution enabled more advanced statistical analysis, yielding various probability distribution models of the instantaneous clear-sky index (Suehrcke and McCormick, 1988; Suehrcke, 1994; Jurado et al., 1995; Tovar and Olmo, 1998). These include models based on mixture distributions (Hollands and Huget, 1983; Hollands and Suehrcke, 2013) and of convolution type (Munkhammar and Widén, 2016). No attempt has so far been made to generalize this to multiple locations, as we do in this paper, apart from the special case of completely uncorrelated locations by Munkhammar et al. (2015b).
3. *Copula modeling of joint probability distributions to describe dependence between dispersed sites*. A copula is an increasingly common method for modeling correlation between stochastic variables (Nelsen, 2006). It is based on the notion that any N stochastic variables, which have a dependency that is reflected in the correlation between variables, can be modeled via a joint distribution function (the so-called copula) for all stochastic variables considered. The methodology has been used for similar purposes as in this paper in modeling of aggregated wind power production, e.g., by Hagspiel et al. (2012), but in the field of solar irradiance modeling it has so far only been used to relate monthly mean solar irradiance and number of sunshine hours (Bazrafshan et al., 2014) and to correlate diffuse and beam radiation for individual sites (Munkhammar and Widén, 2016).

Related models, used for forecasting solar irradiance at both observed and unobserved locations, have been proposed by Yang et al. (2014) and Aryaputera et al. (2015), and it is worthwhile to point out some differences between their approach and the one presented here. Yang et al. (2014) devise a methodology to obtain stationary, symmetrical and separable spatio-temporal covariance fields, and use time-forward kriging based on these to forecast solar irradiance at unobserved locations. Aryaputera et al. (2015) extend the analysis to include four different parametric correlation functions and forecasting based on spatio-temporal kriging using the same solar irradiance network as in this paper. First, it is important to once again note that we do not require the temporal dimension, which simplifies the modeling in comparison to these studies. Second, the methods applied serve different purposes. We are interested in estimating representative distributions for solar irradiance at correlated, arbitrarily located sites. The studies above aim at forecasting the solar irradiance at unobserved locations, in relation to existing data points, over a specific time horizon. Although the methodologies are similar, kriging is essentially a method for interpolation, while copula modeling is a method for joining probability distributions.

The paper is structured as follows. Section 2 presents the methodology, including probabilistic model definitions, data and simulation procedure. Results are presented in Section 3 and a concluding discussion is held in Section 4.

2. Methodology

This study is based on modeling the clear-sky index (normalized solar irradiance) at N stations in a spatial network as probability distributions. A copula (a multivariate probability distribution), correlated according to the cross-correlations of all station pairs, is then used to obtain correlated samples from the distributions. The model parameters are estimated using 10-s resolution solar irradiance data for 14 dispersed locations at the Island of Oahu, Hawai'i, USA, and the method is validated by comparing the simulated aggregate clear-sky index for different sets of stations to the empirical ones given by the measured data.

The probabilistic model is presented in Section 2.1, the copula modeling is presented in Section 2.2, model estimation and simulation is outlined in Section 2.3 and special cases of the model are