Applied Energy 157 (2015) 95–110

Contents lists available at ScienceDirect

Applied Energy

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

An analog ensemble for short-term probabilistic solar power forecast

S. Alessandrini [a,*], L. Delle Monache [a], S. Sperati [b], G. Cervone [c]

[a] National Center for Atmospheric Research, Boulder, CO, USA
[b] Ricerca Sistema Energetico, Milano, Italy
[c] The Pennsylvania State University, University Park, PA, USA

HIGHLIGHTS

- A novel method for solar power probabilistic forecasting is proposed.
- The forecast accuracy does not depend on the nominal power.
- The impact of climatology on forecast accuracy is evaluated.

ARTICLE INFO

Article history:
Received 2 March 2015
Received in revised form 29 June 2015
Accepted 5 August 2015

Keywords:
Analog ensemble
Short-term solar power forecasting
Probabilistic predictions
Uncertainty quantification
Ensemble verification

ABSTRACT

The energy produced by photovoltaic farms has a variable nature depending on astronomical and meteorological factors. The former are the solar elevation and the solar azimuth, which are easily predictable without any uncertainty. The amount of liquid water met by the solar radiation within the troposphere is the main meteorological factor influencing the solar power production, as a fraction of short wave solar radiation is reflected by the water particles and cannot reach the earth surface. The total cloud cover is a meteorological variable often used to indicate the presence of liquid water in the troposphere and has a limited predictability, which is also reflected on the global horizontal irradiance and, as a consequence, on solar photovoltaic power prediction. This lack of predictability makes the solar energy integration into the grid challenging. A cost-effective utilization of solar energy over a grid strongly depends on the accuracy and reliability of the power forecasts available to the Transmission System Operators (TSOs). Furthermore, several countries have in place legislation requiring solar power producers to pay penalties proportional to the errors of day-ahead energy forecasts, which makes the accuracy of such predictions a determining factor for producers to reduce their economic losses. Probabilistic predictions can provide accurate deterministic forecasts along with a quantification of their uncertainty, as well as a reliable estimate of the probability to overcome a certain production threshold. In this paper we propose the application of an analog ensemble (AnEn) method to generate probabilistic solar power forecasts (SPF). The AnEn is based on an historical set of deterministic numerical weather prediction (NWP) model forecasts and observations of the solar power. For each forecast lead time and location, the ensemble prediction of solar power is constituted by a set of past production data. These measurements are those concurrent to past deterministic NWP forecasts for the same lead time and location, chosen based on their similarity to the current forecast and, in the current application, are represented by the one-hour average produced solar power.

The AnEn performance for SPF is compared to a quantile regression (QR) technique and a persistence ensemble (PeEn) over three solar farms in Italy spanning different climatic conditions. The QR is a state-of-the-science method for probabilistic predictions that, similarly to AnEn, is based on a historical data set. The PeEn is a persistence model for probabilistic predictions, where the most recent 20 power measurements available at the same lead-time are used to form an ensemble. The performance assessment has been carried out evaluating important attributes of a probabilistic system such as statistical

Abbreviations: T2M, air temperature at 2 m above ground; AnEn, analog ensemble; AZ, azimuth angle; BS, Brier score; BSS, Brier skill score; CRPS, continuous ranked probability score; CC, cloud cover; GHI, global horizontal irradiance; NP, nominal power; MAE, mean absolute error; MP, mean power; MRE, missing rate error; NN, neural network; NWP, numerical weather prediction; PeEn, persistence ensemble; PV, photovoltaic power; PDF, probability density function; QR, quantile regression; RAMS, regional atmospheric modeling system; RMSE, root mean squared error; EL, solar elevation; SPF, solar power forecasting; TSOs, transmission system operators; WPF, wind power forecast.

* Corresponding author at: National Center for Atmospheric Research, PO Box 3000, Boulder, CO 80307-3000, USA.
E-mail address: alessand@ucar.edu (S. Alessandrini).

http://dx.doi.org/10.1016/j.apenergy.2015.08.011

consistency, reliability, resolution and skill. The AnEn performs as well as QR for common events, by providing predictions with similar reliability, resolution and sharpness, while it exhibits more skill for rare events and during hours with a low solar elevation.

1. Introduction

Solar photovoltaic power (PV) generation has increased steadily in several countries in the last 10 years, becoming an important component of a sustainable solution of the energy problem. In Italy, for instance, thanks also to substantial government subsides over the past 5 years, the annual generation by PV has reached 19,418 GWh in 2013, corresponding to 7% of the total Italian energy demand. The rate of growth is significant considering that in 2008 the generation from PV was just 200 GWh (0.1% of total) [1]. During the last 10 years wind power installed capacity has also followed the same trend in Italy, going from 664 MW to 8551 MW in 2013 [2], which has led to an annual generation of about 4.7% of the total Italian energy demand. As a consequence more than 10% of the total Italian energy production is now made of wind and solar resources that have a variable nature and a limited predictability. Astronomical and meteorological factors are the main causes of PV variability. The former are the solar elevation and the solar azimuth, which are easily predictable without any uncertainty. The amount of liquid water met by the solar radiation inside the troposphere is the main meteorological factor influencing the solar power production. In fact, a fraction of short wave solar radiation is reflected back outside the atmosphere or absorbed by the cloud water particles and cannot reach the earth surface. A similar process involves other kind of particles generally defined as aerosols that have different origins both natural (marine salt, soil erosion) and anthropogenic (combustion). The interaction between solar radiation and aerosols increases the amount of the diffused component (that still reaches the earth surface contributing to PV generation) but alters only marginally the reflected and absorbed components, which become significant only in the case of particularly high aerosol concentrations, such those found in the proximity of a volcanic eruption or during episodes of high urban pollution. The PV panel energy efficiency is also related to air temperature (the efficiency decreases at higher temperatures) [3]. Hence, the total cloud cover (CC), the global horizontal irradiance (GHI) and air temperature at 2 m above the ground (T2M) are the meteorological variables directly related to PV production and their limited predictability is also reflected on PV predictions. Aerosol interactions with solar radiation are generally taken into account by meteorological models, but the aerosol distribution is based on climatological estimates.

Electricity is distributed over a region via a grid, which is an interconnected network from suppliers to consumers. To maintain grid stability at an effective cost, it has now become crucial to be able to predict with accuracy the renewable energy production which is combined with other more predictable sources (e.g., coal, natural gas) to satisfy the energy demand [4,5]. The efforts of providing an accurate solar power prediction have now been mandated by recent legislation in Italy for utility companies, which have to pay penalties proportional to the forecast errors (defined as the difference between the day-ahead planned energy and the actual production). This has now put a greater focus on solar power forecasting (SPF). Even though the first methods for wind power forecasting (WPF) and SPF were both developed in early 1980s [6,7], the number of publications focused on WPF is much higher.

Predictions can be categorized into deterministic and probabilistic forecasts. A deterministic forecast consists of a single predicted value of the variable for each prediction time, while probabilistic forecasting provides probability density functions (PDFs) from which probabilities of future outcomes can be estimated. Probabilistic forecasts also provide information about uncertainty in addition to the commonly provided single-valued (best-estimate) power prediction.

There are many examples of how probabilistic predictions can provide a higher value than deterministic ones. For instance, one example is estimating the optimal level of reserves that need to be allocated to compensate for wind and solar power variability and their limited predictability, as discussed in [8]. Another significant application is when renewable energy is traded in day-ahead electricity markets. In [9–11] it is shown that trading future wind energy production using probabilistic wind power predictions can lead to higher economic benefits than those obtained by using deterministic forecasts alone. Indeed, the maximum income for a producer is obtained by offering in the day-ahead market an amount of energy that can be different from the most expected one.

Two reviews of the status of forecasting GHI on different time scales for energy generation are reported in [12] and [13], while different forecasting techniques for PV power are evaluated and compared in [14], and a full review of SPF can be found in [15]. In this work we focus on short-range PV forecasting, i.e., 0–72 h ahead. An approach to forecasting on this time range can be based on applying statistical or machine learning techniques directly to historical time series of PV production data. Several applications of this kind can be found in the literatures [16–18]. In [19] it is shown that for forecasts up to 2 h ahead the most important input is the available observations of solar power, while for longer horizons numerical weather prediction (NWP) model output becomes crucial for a better accuracy. Even though there are several contributions on the topic of probabilistic WPF [20–23]for probabilistic PV forecasts only a few methods have been proposed in the literature. [24] and [25] propose probabilistic forecast systems for GHI only. In [24] the system is based on a stochastic differential equation framework together with NWP for modeling the uncertainty associated with the solar irradiance point forecast. In [25] the GHI forecast system is based on studying the correlation of uncertainty to local meteorological conditions describing synoptic-scale atmospheric flow. The direction and magnitude of geostrophic flow were used as an indicator of coastal cloud cover probability to produce regime-dependent forecast intervals.

In [26] PV energy probabilistic forecast for 1-h ahead are based on a Bayesian auto regressive time series model without using NWP. In [27] a normal distribution with zero mean and a standard deviation dependent on the solar zenith angle and the cloud situation is assumed to describe forecast errors of GHI. In this paper we propose the Analog Ensemble (AnEn) as a novel method for PV power forecasting over the 0–72 h lead time period. It has been originally proposed by [28] and [29] for deterministic and probabilistic meteorological forecasting, by [30] for WPF, and by [32] for wind resource assessment applications. The AnEn technique provides a set of likely PV predictions (i.e., an ensemble that is a Monte Carlo approximation of the PDF associated with future power production) using an historical dataset of observations and deterministic NWP. For each forecast lead time the ensemble