Renewable Energy 91 (2016) 11–20

Contents lists available at ScienceDirect

Renewable Energy

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

Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest

David P. Larson, Lukas Nonnenmacher, Carlos F.M. Coimbra*

Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, Center of Excellence in Renewable Energy Integration and Center for Energy Research, University of California San Diego, La Jolla, CA 92093, USA

A R T I C L E I N F O

Article history:
Received 2 July 2015
Received in revised form 30 November 2015
Accepted 8 January 2016
Available online xxx

Keywords:
Forecasting
Solar power output
Photovoltaics
Numerical weather prediction
Interannual performance variability

A B S T R A C T

A forecasting method for hourly-averaged, day-ahead power output (PO) from photovoltaic (PV) power plants based on least-squares optimization of Numerical Weather Prediction (NWP) is presented. Three variations of the forecasting method are evaluated against PO data from two non-tracking, 1 MWp PV plants in California for 2011–2014. The method performance, including the inter-annual performance variability and the spatial smoothing of pairing the two plants, is evaluated in terms of standard error metrics, as well as in terms of the occurrence of severe forecasting error events. Results validate the performance of the proposed methodology as compared with previous studies. We also show that the bias errors in the irradiance inputs only have a limited impact on the PO forecast performance, since the method corrects for systematic errors in the irradiance forecast. The relative root mean square error (RMSE) for PO is in the range of 10.3%–14.0% of the nameplate capacity, and the forecast skill ranges from 13% to 23% over a persistence model. Over three years, an over-prediction of the daily PO exceeding 40% only occurs twice at one of the two plants under study, while the spatially averaged PO of the paired plants never exceeds this threshold.

1. Introduction

High renewable energy penetration grids are challenging to balance due to inherently variable generation weather-dependent energy resources. Solar and wind forecasting are proven methods for mitigating resource uncertainty and reducing the need for scheduling of ancillary generation. Several forecasting methodologies have been developed to target different forecast time horizons [1].

Short-term forecasting ($t < 1$ h) is mainly based on sky imaging techniques and time-series models, while satellite-based forecasts achieve usable results for time horizons of 1–6 h [2–4]. Usually, horizons larger than 6 h require numerical weather prediction models (NWP) to generate accurate results, although there are exceptions such as [5]. Recent advances in solar forecasting have mainly covered intra-day resource forecasts, driven by advances in sky image techniques and time-series modeling [6–16].

While intra-day forecasts are important for grid stability, day-ahead forecasts are critical for market participation and unit commitment. Current market regulations in many world regions require day-ahead forecasts [17–19] and there are financial incentives to produce accurate forecasts [18]. Besides market participation, day-ahead forecasts can also be useful for energy storage dispatch [20,21]. In this paper, we focus on day-ahead forecasting of power output from photovoltaic power plants in the American Southwest.

1.1. Previous work

Most NWP models generate global horizontal irradiance (GHI) at the ground level as one of the outputs, with some newer generation models now including direct normal irradiance (DNI). Previous studies evaluated the accuracy of this GHI variable and suggested ways to improve it, which have mostly focused on spatial averaging and bias removal [22–28]. [22,27,29] showed that forecast errors for all sky conditions can be reduced by averaging the GHI forecasts from all NWP grid points within a set distance of the target site. In addition [22], showed that the forecast performance could be further improved through bias removal using a polynomial regression based on the solar zenith angle and clear sky index. [27] showed that a similar improvement can be achieved

* Corresponding author.
E-mail address: ccoimbra@ucsd.edu (C.F.M. Coimbra).

http://dx.doi.org/10.1016/j.renene.2016.01.039
0960-1481/

through Kalman filtering.

In contrast to our knowledge of the performance of day-ahead GHI predictions, the application of these forecasts directly to prediction of day-ahead power output (PO) of PV plants is poorly understood. There are currently less than a dozen published studies that cover the subject of day-ahead PV PO forecasts, the majority of which were published in the last 5 years. We summarize these papers below and in Table 1. It is important to note that this lack of knowledge is partly due to the difficulty of obtaining data from operational PV plants, due to security restrictions and lack of data infrastructure. However, data access should improve in the coming years due to energy policies that require PO forecasting and therefore necessitate the collection of PO data.

[30] applied the Danish Meteorological Institute's High Resolution Limited Area Model (HIRLAM) to forecast PO of 21 PV systems in Denmark. The PV systems had a total rated power capacity of 1–4 kWp each and one year of data was used to evaluate the forecasts. HIRLAM GHI forecasts were used as the main input to autoregressive-based PO models. [31] also used HIRLAM to forecast PO of a 6.2 kWp test site in Spain.

[29] and [32] used the European Center for Medium-Range Weather Forecasts Model (ECMWF) to generate regional, PO forecasts in Germany. [29] only evaluated two months (April and July 2006) of forecasts for 11 PV systems in Southern Germany, while [32] tested on 10 months (July 2009–April 2010) of PO data from approximately 400 representative PV systems.

[33] used an artificial neural network (ANN) to predict GHI in Trieste, Italy. The predicted GHI was then mapped directly to PO using the efficiency data of the studied 20 kWp PV system. Unfortunately, PO forecast results were only reported for four consecutive clear sky days in May 2009. Similarly [34], evaluated an autoregressive-moving-average model with exogenous inputs (ARMAX) forecast model that did not use NWP data as input. However [34], only forecasted the day-ahead mean daily PO, rather than day-ahead hourly values, for a 2.1 kWp PV system in China.

[27] forecasted PO of three small PV systems (6.72, 19.8 and 45.6 kWp) in mainland Canada, using the Canadian Meteorological Centre's Global Environmental Multiscale Model (GEM). The GEM's GHI forecasts were validated against ground measurements from the SURFRAD Network in the United States. Spatial averaging of 300 km–600 km and bias removal via Kalman filtering were used to improve the GEM forecast performance. Reported RMSE values were in the range of 6.4%–9.2% of the rated power of the PV systems for a 1 year testing set (April 2008–March 2009).

[35] and [36] generated regional forecasts for Japan using Support Vector Regression (SVR) together with inputs from the Japan Meteorological Agency's Grid Point Value-Mesoscale Model (GPV-MSM). [35] used PO data from approximately 450 PV plants from four regions (Kanto, Chubu, Kansai, Kyushu) that had a net rated power of approximately 15 MWp while [36] used data from 273 PV plants spread over two regions (Kanto and Chubu), with an approximate total power of 8 MWp.

[37] and [38] forecasted PO of 28 PV plants in mainland France using Meteo France's Action de Recherche Petite Echelle Grande Echelle (ARPEGE) model. [37] presented a deterministic forecast that achieved a RMSE of 8–12% of the plant capacity over two years of testing data (2009–2010) while [38] presented a probabilistic forecast. Both methods used 31 variables from the ARPEGE, including GHI as well as environmental conditions, e.g., temperature, humidity and precipitation. Unfortunately, no actual values were reported for the PV plant power ratings or other technical details, which limits analysis into the applicability of the results to other PV plants and regions.

Most recently [39], forecasted the PO of five vertical-axis tracking PV plants in Spain, using a nonparametric model based on the Weather Research and Forecasting Model (WRF). The PV plants ranged in size from 775 to 2000 kWp and the forecasts were evaluated with data from 2009 to 2010. Quantile Regression Forests, a variation of random forests, was used to generate the PO forecasts, with WRF variables such as GHI, temperature and cloud cover as the inputs.

Although these studies all presented day-ahead PO forecasting for PV plants, further research is still required, especially for sites in the United States. In this study we seek to provide the following contributions: (1) Introduction and evaluation of a PV PO forecast model for the American Southwest, a region with both high solar energy generation potential and a favorable political environment for solar, especially in California. (2) Investigation of the interannual forecast performance variability and (3) the occurrence of severe forecasting errors, as they relate to large-scale renewable energy integration. (4) Spatial smoothing through pairing of PV plants in proximity. To achieve these goals, we systematically test approaches and inputs to generate PO forecasts based on two operational PV plants in Southern California.

2. Data

2.1. Ground data

Two sites are used to evaluate day-ahead forecasts of GHI and PO: Canyon Crest Academy (CCA) and La Costa Canyon High School (LCC). CCA (32.959°, −117.190°) and LCC (33.074°, −117.230°) are both located in San Diego County, USA, and each feature 1 MW peak (MWp) of non-tracking photovoltaic (PV) panels. The PV panels for

Table 1
Summary of the current literature on day-ahead PO forecasting. “Type” refers to whether the forecast was for a single PV plant site or for an entire region while “Testing data” is how much data was used in the forecast model evaluation.

	Region	NWP model	Type	PV plants		
				No. of sites	Total capacity	Testing data
[30]	Denmark	HIRLAM	Point	21	~100 kWp	1 year
[31]	Spain	HIRLAM	Point	1	6 kWp	10 months
[29]	Germany	ECMWF	Regional	11	Unknown	2 months
[32]	Germany	ECMWF	Regional	383	~100 MWp	10 months
[35]	Japan	GPV-MSM	Regional	454	~10 MWp	1 year
[36]	Japan	GPV-MSM	Regional	273	~10 MWp	1 year
[27]	Canada	GEM	Point	3	~100 kWp	2 years
[37]	France	ARPEGE	Point	28	Unknown	2 years
[38]	France	ARPEGE	Point	28	Unknown	2 years
[39]	Spain	WRF	Point	5	~10 MWp	2 years
[33]	Italy	N/A	Point	1	20 kWp	4 days
[34]	China	N/A	Point	1	2 kWp	6 months