ARTICLE IN PRESS

Electrical Power and Energy Systems xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Probabilistic solar power forecasting in smart grids using distributed information

R.J. Bessa [a,*], A. Trindade [a,b], Cátia S.P. Silva [c], V. Miranda [a,b]

[a] INESC TEC – INESC Technology and Science, Porto, Portugal
[b] FEUP – Faculty of Engineering, University of Porto, Portugal
[c] Computational NeuroEngineering Lab, University of Florida, Gainesville, FL, USA

ARTICLE INFO

Article history:
Received 1 February 2015
Accepted 16 February 2015
Available online xxxx

Keywords:
Solar power
Forecasting
Smart grid
Distributed sensors
Probabilistic
Gradient boosting

ABSTRACT

The deployment of Smart Grid technologies opens new opportunities to develop new forecasting and optimization techniques. The growth of solar power penetration in distribution grids imposes the use of solar power forecasts as inputs in advanced grid management functions. This paper proposes a new forecasting algorithm for 6 h ahead based on the vector autoregression framework, which combines distributed time series information collected by the Smart Grid infrastructure. Probabilistic forecasts are generated for the residential solar photovoltaic (PV) and secondary substation levels. The test case consists of 44 micro-generation units and 10 secondary substations from the Smart Grid pilot in Évora, Portugal. The benchmark model is the well-known autoregressive forecasting method (univariate approach). The average improvement in terms of root mean square error (point forecast evaluation) and continuous ranking probability score (probabilistic forecast evaluation) for the first 3 lead-times was between 8% and 12%, and between 1.4% and 5.9%, respectively.

Introduction

Presently, the economics of photovoltaic (PV) solar power are attractive due to a high reduction in market prices of PV panels [1]. Across several countries, the installed solar power capacity is increasing, either consisting of medium/large solar parks connected to the medium voltage network or small PV installations at the building level (i.e., low voltage network) [2].

In this context, the power system management tasks of both Transmission System Operator (TSO) and Distribution System Operators (DSO) require PV power forecasting. In the DSO case, the advanced communication and monitoring capabilities of the Smart Grid infrastructure [3] creates conditions for advanced grid management tools, such as security-constrained unit commitment with demand response [4], voltage control [5] and probabilistic power flow [6]. Furthermore, solar power can be combined with storage, as a virtual power plant, for energy trading and/or providing system support services [7].

One key requirement of these advanced grid management tools is solar power forecasts covering two time horizons [8]: (i) 6 h ahead (i.e., very short-term) and (ii) three days-ahead (i.e., short-term). For the first 4 h ahead, the relevant inputs consist of past observations of the time series, after those lead-times information from Numerical Weather Predictions (NWP) is more relevant [9].

In the solar power forecasting literature, it is possible to find a vast number of works devoted to apply machine learning and statistical based algorithms to extrapolate solar power from NWP [9], which covers the short-term horizon. Fernandez-Jimenez et al. [10] convert NWP information to solar power using different statistical learning algorithms, as Auto-Regressive Integrated Moving Average, *k*-nearest neighbors, neural networks, and adaptive neuro-fuzzy models; Zamo et al. [11] compares several regression algorithms (e.g., random forests, boosting, support vector machines) that take NWP as input to produce solar power forecasts; Bacher et al. [12] proposes an autoregressive model with exogenous inputs (ARX) whose feature vectors are a combination of observations of solar power and NWP.

From a probabilistic forecasting point of view, Lorenz et al. [13] describes a method for computing situation-dependent predictions intervals for a single solar park. Also, Bacher et al. [12] uses weighted quantile regression conditioned to a clearness index (or normalized solar power) to produce probabilistic forecasts.

For the very short-term time horizon, the present literature is driven by statistical and time series models that use past values of the same power time series. For instance, Pedro and Coimbra [14] compare the performance of different statistical learning

* Corresponding author at: INESC TEC, Campus da FEUP, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal. Tel.: +351 22 209 4216; fax: +351 22 209 4050.
E-mail address: ricardo.j.bessa@inesctec.pt (R.J. Bessa).

http://dx.doi.org/10.1016/j.ijepes.2015.02.006
0142-0615/© 2015 Published by Elsevier Ltd.

algorithms (i.e., Auto-Regressive Integrated Moving Average, k-nearest neighbors and neural networks adjusted by genetic algorithms), which only use past observations of the time series as inputs. An important characteristic of these algorithms is that information from distributed time series data sources are not included in the model, i.e., only past values of the local response variable are used. Yet, an interesting development, proposed by Hammer et al. [15], uses cloud-index images to produce solar power in the very short-term horizon with motion vector fields derived from two consecutive frames. However, this information must be available in almost real-time and it may be more complex and expensive to operationalize such forecasting services. Moreover, in contrast to models for wind power forecast [16], all these approaches do not provide probabilistic forecasts of solar power.

The forecasting framework presented in this paper addresses the very short-term horizon and the key idea, in a Smart Grid context is to explore information from spatially distributed smart meters (or sensors). We should make sure it is readily available in acceptable "real-time" while it might be additionally combined with satellite information.

In the scientific community, only three publications explore the use of information from neighboring solar sites to improve power forecast. Berdugo et al. [17] proposes an analog searching algorithm for similar local and global current states as neighbor sites. However, the main goal is not to produce the forecast with minimum error, instead, it is to efficiently handle large volumes of streaming data and keep power measurements' confidentiality. Yang et al. [18] proposes an ARX model for each solar site where the exogenous variables are measurements from neighbor sites. Lonij et al. [19] combines solar power observations from a network of data loggers and wind speed information from a NWP model to estimate cloud edge velocity and infer solar power from this information. Considering the reviewed literature, this paper presents three main contributions for 6 h-ahead solar power forecasts:

- Probabilistic forecasting method based on vector autoregression framework (VAR), which combines information from the distributed PV panels collected by the Smart Grid infrastructure. The residential PV and secondary substation (i.e., MV/LV) levels are covered by this approach.
- Improve of probabilistic forecast skill at the secondary substation level by introducing exogenous variables (i.e. observations from micro-generation units with smart meters) to the VAR model.
- Probabilistic forecasting approach based on gradient boosting technique, which is the main contribution compared to [20,21].

This paper is organized as follows: Section 'Smart Grid pilot in Portugal' describes the information and communication infrastructure of the Smart Grid pilot in Portugal; Section 'Forecasting framework' describes the solar power point and probabilistic forecasting algorithms; the test case results are presented in Sections 'Test case results' and 'Conclusions' presents the conclusions.

Smart Grid pilot in Portugal

The Portuguese DSO promoted the development of new ICT technology and computational tools for automating network management in order to create a full smart distribution grid [22]. This resulted in a large-scale demonstration pilot in the city of Évora in Portugal, named InovCity [23], which is also one demonstration site of the EU Project SuSTAINABLE [24].

The main Smart Grid equipment of this infrastructure is the following: EDP Box (EB), the Distribution Transformer Controller (DTC) and Smart Substation Controller (SSC). This architecture covers the whole distribution grid and the control layers correspond to the main voltage levels from the HV grid down to the LV consumers that define the hierarchy of control and management.

The SSC, housed in the HV/MV distribution substation level, is responsible for managing the MV grid and includes local intelligence (e.g., self-healing and control of distributed generation connected to MV network) and several operational functionalities. Regarding the LV grid, it is controlled by a DTC located at the secondary (or MV/LV) substation level that will be responsible for managing the distributed energy resources at the LV level. The DTC comprises modules for monitoring and remote control. At this level – LV level – the smart meters (EB) associated to consumers and microgeneration units also have monitoring and management functions, interacting with other devices through a home area network.

In this architecture, the SSC is responsible for aggregating and managing the operational data from EB and DTC, using a GRPS Wide Area Network. The DTC collects data from the EB through a Local Area Network with GPRS or PLC technology.

The forecasting system that is described in Section 'Forecasting framework' requires a centralized data flow topology and can be installed at the DSO control center level, SSC, and DTC. In this paper, without loss of generality, it is assumed to be installed at the central management level (i.e., in the DMS). Point and probabilistic forecast outputs are generated for each DTC and EB. For the forecast at the DTC or secondary substation level, the EB measurements can be used as distributed sensors to better capture the influence of clouds and therefore improve the forecast skill, which in turn increase the amount of transferred data. Note that even if the system is operating at the EB level, if a centralized topology is created with a peer-to-peer communication channel between smart meters [17], it is possible to explore distributed information within a neighboring area.

Finally, driven by the high uncertainty associated to renewable energy, the trend in grid management functions falls into a stochastic paradigm. Therefore, the probabilistic forecasts are an important input to management functions.

Forecasting framework

Calculation of clear-sky generation

The solar power time series presents a seasonal pattern dependent on the time of the day and day of the year [12]. This deterministic variation of the solar irradiance can be modeled with different physical and statistical methods, which can be found in [25].

In [12], a statistical model based on quantile regression with varying coefficients is described. It computes the clear-sky power for a given solar power time series. The method is presented as a statistical normalization of solar power, capable of generating a stationary time series suitable for classical models, such as AR and VAR.

The work described in [12] and [26], indicates that the most relevant predictors of the clear-sky model are the time of the day (h) and day of the year (doy). The clear-sky generation ($\hat{p}_t^{cs}$) is estimated as a local constant model and the quantile regression with varying coefficient for quantile τ can be expressed as:

$$\hat{p}_t^{cs} = \arg\min_{\widehat{p_t^{cs}}} \sum_{i=1}^{N} K(h_t, doy_t, h_i, doy_i) \cdot \rho(\tau, e_i) \quad (1)$$

with $e_i = p_t - \hat{p}_t^{cs}$, and

$$K(h_t, doy_t, h_i, doy_i) = \frac{K(h_t, h_i, \sigma_h) \cdot K(doy_t, doy_i, \sigma_{doy})}{\sum_{i=1}^{N}[K(h_t, h_i, \sigma_h) \cdot K(doy_t, doy_i, \sigma_{doy})]} \quad (2)$$

Please cite this article in press as: Bessa RJ et al. Probabilistic solar power forecasting in smart grids using distributed information. Int J Electr Power Energ Syst (2015), http://dx.doi.org/10.1016/j.ijepes.2015.02.006