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*

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 six hours 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.

Keywords: Solar power; forecasting; smart grid; distributed sensors; smart metering; probabilistic; gradient boosting.

1. 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

*Correspondence to: Ricardo Bessa, INESC TEC, Campus da FEUP, Rua Dr. Roberto Frias, 4200 - 465 Porto Portugal. Telf: +351 22 209 4216. Fax: +351 22 209 4050. E-mail: ricardo.j.bessa@inesctec.pt

[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) six hours ahead (i.e., very short-term) and (ii) three days-ahead (i.e., short-term). For the first four hours 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 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 six hours-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 2 describes the information and communication infrastructure of the Smart Grid pilot in Portugal; section 3 describes the solar power point and probabilistic forecasting algorithms; the test case results are presented in section 4 and section 5 presents the conclusions.

2. 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 3 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.

3. Forecasting Framework

3.1 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 modelled 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:

$$\hat{p}_t^{cs} = \underset{\hat{p}_t^{cs}}{\arg\min} \sum_{i=1}^{N} K\left(h_t, doy_t, h_i, doy_i\right) \cdot \rho\left(\tau, e_i\right) \tag{1}$$

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

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

is the normalized kernel product of the two predictors that locally weights each observation, and

$$\rho\left(\tau, e_i\right) = \begin{cases} \tau \cdot e_i & , e_i \geq 0 \\ \left(1 - \tau\right) \cdot e_i , & e_i < 0 \end{cases} \tag{3}$$

is the quantile loss function [27]. Since both variables (h_t and doy_t) are circular, the following kernel is used:

$$K\left(x_t, x_i, \sigma\right) = e^{\frac{1}{\sigma} \cdot \cos\left[2\pi \cdot \frac{\left(x_t - x_i\right)}{d}\right]} \tag{4}$$

where *σ* is the smoothing parameter and *d* is the period of variable *x* (e.g., equal to 24 in the time of the day variable).

The output of the model from Eq. 1 is used to normalize the measured solar power (P_t) as follows:

$$p_t^{norm} = \frac{p_t}{\hat{p}_t^{cs}} \tag{5}$$

The model's parameters are the kernel bandwidths σ_h and σ_{doy}, as well as the quantile *τ* and are determined by trial-error experiences. It is expected to get a result of one for the normalized solar power p_t^{norm} in clear-sky days.

3.2 Vector Autoregressive (VAR) Model

A widely used class for univariate time series models is the autoregressive (AR) framework [28], in which the value of the target variable for time interval *t* is expressed as linear regression on past observations (or lags) of the time series. For one hour-ahead forecast, the AR model is:

$$\hat{p}_{t+1|t} = \alpha + \beta_1 \cdot p_t + \beta_2 \cdot p_{t-1} + \cdots + \beta_l \cdot p_{t-l} + e_{t+1|t} \tag{6}$$

where *β* are the model's coefficients, *α* a constant term, *l* the order of the AR model and $e_{t+1|t}$ is a white noise process with zero mean and constant variance σ_e^2.

This model can be extended with exogenous variables (such as NWP), forming an ARX model. However, for the very short-term time horizon, the main limitation of this model is that it only uses the past observations as predictors from the target variable.

In order to improve the forecast skill for this time horizon, a VAR model [29] is used to combine past observations from the solar power in each site with past values from neighbor sites, that is, it uses both time and spatial information. This consists in a multi-output linear regression model with *N* observations, *q*-dimensional response and *d*-lagged terms (or predictors).

In matrix format, for one step-ahead forecast, it is given by:

$$\hat{P}_{t+1|t} = \alpha + B_1 \cdot P_t + B_2 \cdot P_{t-1} + \cdots + B_d \cdot P_{t-d} + E_{t+1|t} \tag{7}$$

where $\hat{P}_{t+1|t}$ is a *q x 1* vector, P_{t-d} the *d-th* lag, B a coefficient matrix with dimension *q x q*, α a vector with *q x 1* intercept (or constant) terms, $E_{t+1|t}$ is a vector with dimension *q x 1* containing i.i.d. residuals with zero mean and constant covariance Σ_e.

Considering a case with two solar sites (i.e., two response variables) and a second order lag, Eq. 7 becomes:

$$\begin{aligned}\hat{p}_{t+1|t,1} &= \alpha_1 + \beta_{11} \cdot p_{t,1} + \beta_{12} \cdot p_{t-1,1} + \beta_{13} \cdot p_{t,2} + \beta_{14} \cdot p_{t-1,2} + e_{t+1|t,1} \\ \hat{p}_{t+1|t,2} &= \alpha_2 + \beta_{21} \cdot p_{t,1} + \beta_{22} \cdot p_{t-1,1} + \beta_{23} \cdot p_{t,2} + \beta_{24} \cdot p_{t-1,2} + e_{t+1|t,2}\end{aligned} \tag{8}$$

As shown in Eq. 8, the VAR consists of linear regression models, in which the $\hat{p}_{t+k|t}$ of each site depends on a constant term and lagged terms of the *q* response variables. Note that each regression equation takes the same predictors (P_{t-d}). In fact, Eq. 7 and 8 have the form of a Seemingly Unrelated Regression (SUR) model [29], where Ordinary Least Squares (OLS) can be applied independently to each regression equation, if the same predictors appear in every equation.

3.3 Recursive Least Squares for Point Forecasts

The communication infrastructure of a Smart Grid generates a large volume of data streams that must be handled online with low data storage requirements. The coefficients of the AR and VAR models described in the previous section can be time-adaptive by using the recursive least squares (RLS) method with a forgetting factor (extensively described in [30]). This method overcomes the problem of handling large volumes of data since it is not necessary to store historical time series data for fitting (or re-fitting) the model, in each time step *t* only B_t, K_t and Q_t have to be stored in memory. Furthermore, the RLS method, with a forgetting factor λ, copes with concept drifting/shifting problems, such as loss of performance due to dust in PV panels or changes in the surrounding environment.

Since both VAR and AR can be fitted with OLS, the RLS method can also be applied to this model and it is of great importance since the relation between distributed PV sites is very dynamic and requires time-varying coefficients.

The update of the VAR model coefficients, using the example from Eq. 8 for site 1, is performed with the RLS method as follows for time step *t*:

$$[\alpha_1,\beta_{11},\beta_{12},\beta_{13},\beta_{14}]_{t+1}=[\alpha_1,\beta_{11},\beta_{12},\beta_{13},\beta_{14}]_t+K_{t+1}\cdot\left[p_{t+1,1}-\left(\alpha_1+\beta_{11}\cdot p_{t,1}+\beta_{12}\cdot p_{t-1,1}+\beta_{13}\cdot p_{t,2}+\beta_{14}\cdot p_{t-1,2}\right)\right] \quad (9)$$

where K_{t+1} is given by,

$$K_{t+1}=Q_{t+1}\cdot\left[p_{t,1},p_{t-1,1},p_{t,2},p_{t-1,2}\right] \quad (10)$$

and Q_{t+1} by

$$Q_{t+1}=\frac{1}{\lambda}\cdot\left[Q_t-\frac{Q_t\cdot\left[p_{t,1},p_{t-1,1},p_{t,2},p_{t-1,2}\right]\cdot\left[p_{t,1},p_{t-1,1},p_{t,2},p_{t-1,2}\right]^T\cdot Q_t}{\lambda+\left[p_{t,1},p_{t-1,1},p_{t,2},p_{t-1,2}\right]^T\cdot Q_t\cdot\left[p_{t,1},p_{t-1,1},p_{t,2},p_{t-1,2}\right]}\right] \quad (11)$$

$[\alpha_1,\beta_{11},\beta_{12},\beta_{13},\beta_{14}]_{t-1}$ are the coefficients from time step *t-1* and $[\alpha_1,\beta_{11},\beta_{12},\beta_{13},\beta_{14}]_t$ the updated coefficients after receiving the observation $p_{t+1,1}$.

A forgetting factor equal to 1 leads to a recursive estimation of the coefficients, while a smaller value discounts old data with an exponential decay. Initial values for B_0 and Q_0 are required. A simple and robust approach, purely as experimental observation, is to initialize B_0 with zeros and Q_0 as a diagonal matrix with a large constant value.

3.4 Gradient Boosting for Probabilistic Forecasts

Boosting is an ensemble machine-learning algorithm for classification and regression, which combines base learners [31]. It conducts numerical optimization, via steepest-descent, in function space by using a user-defined base learner recurrently on modified data that is the output from the previous iterations. Following the optimization phase, the final solution *F(x)* is a linear combination of the base learners, as follows:

$$\hat{F}(x) = \hat{f}_0(x) + \sum_{m=1}^{M} \hat{f}_m(x) \tag{12}$$

where $\hat{f}_0(x)$ is an initial guess, $\hat{f}_m(x)$ are the base learners, *x* the covariates and *M* is the maximum number of "boosts" (i.e., model's parameter). This algorithm allows different loss functions *ρ*, and for the probabilistic forecast problem the choice is the quantile loss function:

$$\rho\left(y, \hat{F}(x)\right) = \begin{cases} \tau \cdot \left(y - \hat{F}(x)\right) & , \quad y - \hat{F}(x) > 0 \\ -(1-\tau) \cdot \left(y - \hat{F}(x)\right), & y - \hat{F}(x) \le 0 \end{cases} \tag{13}$$

where *y* is the observed value and *τ* is the quantile nominal proportion.

The base learner for the model in Eq. 6 and 7 is a linear effect of a continuous predictor. The base learner of the component-wise gradient boosting (GB) algorithm proposed by Bühlmann in [32] selects only one predictor among all the d-predictors. The algorithm performs automatic variable selection and coefficients shrinkage. Like in section 3.3, the GB technique is also separately applied to each equation, meaning that, for each solar site, the most relevant lagged terms are selected automatically by the following algorithm.

The component-wise GB applied to the VAR model, using the example from Eq. 8 for site 1, works as follows:

1. initialize $\hat{f}_0\left(p_{t,1}, p_{t-1,1}, p_{t,2}, p_{t-1,2}\right)$ with the mean value of the solar power in site 1 (response variable);
2. for each *m*, compute the negative gradient *u[i]=-∂ρ/∂F* of the loss function and evaluate it at $\hat{F}_{m-1}\left(p_{t,1}[i], p_{t-1,1}[i], p_{t,2}[i], p_{t-1,2}[i]\right)$, where *i* takes values between 1 and N (i.e., number of fitting samples). For the quantile loss function, the negative gradient is given by:

$$u[i] = -\frac{\partial\rho\left(p_{t+1,1}[i], \hat{F}_{m-1}[i]\right)}{\partial F} = \begin{cases} \tau \quad , & p_{t+1,1}[i] - \hat{F}_{m-1}[i] > 0 \\ \tau - 1, & p_{t+1,1}[i] - \hat{F}_{m-1}[i] \le 0 \end{cases} \tag{14}$$

3. using the negative gradient *u[i]* calculated in step (2) as the response variable, estimate the coefficients associated to each candidate base learner:

$$\beta_j = \frac{\sum_{i=1}^{N} u[i] \cdot x^{(j)}[i]}{\sum_{i=1}^{N} \left(x^{(j)}[i]\right)^2} \tag{15}$$

   where *j* takes values between 1 and *d* (number of predictors or lagged terms) and *x* is the set of *d* predictors;

4. Determine the *s-th* predictor or base learner (from a set with *d* candidates) that minimizes the quadratic loss function:

$$s = \underset{1 \le j \le d}{\arg\min} \sum_{i=1}^{N} \left(u[i] - \beta_j \cdot x^{(j)}[i] \right)^2 \tag{16}$$

which gives the following selected base leaner: $\hat{f}_m\left(x^{(s)}\right) = \beta_s \cdot x^{(s)}$;

5. update function $\hat{F}_m(\cdot)$ as follows:

$$\hat{F}_m(\cdot) = \hat{F}_{m-1}(\cdot) + \nu \cdot \hat{f}_m\left(x^{(s)}\right) \tag{17}$$

where *v* is a shrinkage parameter;

6. stop when *m*=*M* and the final estimator is obtained.

A model is fitted for each quantile τ ranging between 5% and 95% with 5% increments.

The GB method has two parameters that need to be set: maximum number of boosting iterations (*M*); shrinkage parameter (*v*). The value of M is estimated through 5-fold cross-validation, where the value with the lowest square error is selected for each lead-time. The value of *v* does not influence significantly the results if set to be a low value [32]; a value of 0.15 is used in this problem.

In contrast to the RLS algorithm described in section 3.3., the GB selects the most relevant predictors (i.e., lagged terms), which contributes to increase the sparsity of the coefficients matrices since some predictions will never be selected in step (4).

3.5 VAR Applied to Solar Power Forecast

In order to apply the forecasting techniques from Eq. 6 and 7, in a first phase, it is necessary to normalize the solar power time series with the clear-sky model from section 3.1. The normalized solar power values outside the period between 7h00 and 19h00 (i.e., the average period with almost no solar generation during the whole year in Portugal) are removed from all sites. An alternative strategy is to remove all the hours with zenith angle greater than 90º.

In a second phase, the normalized solar power values are used to fit the AR and VAR models with RLS and GB. Both models are applied to forecast the solar power for each DTC and EB. Furthermore, the AR and VAR models are only fully specified after determining the number of lagged terms. The following classical model was adopted: (i) the autocorrelation plot of the normalized time series is analyzed to make a rough estimation of the relevant lags; (ii) the autocorrelation plot of the residuals is analyzed to check if the residuals are white noise (i.e., no autocorrelation remains). Note that the autocorrelation of the residuals can be removed by increasing the order of the model, by adding more lagged terms [33].

For instance, **Figure 1** depicts the autocorrelation function plot (ACF) of the residuals obtained with an AR model that includes lags *t*, *t-1* and *t-23* for one DTC. As depicted, the residuals are almost uncorrelated in time for a 95% confidence interval, which corroborates the choice of the lagged terms.

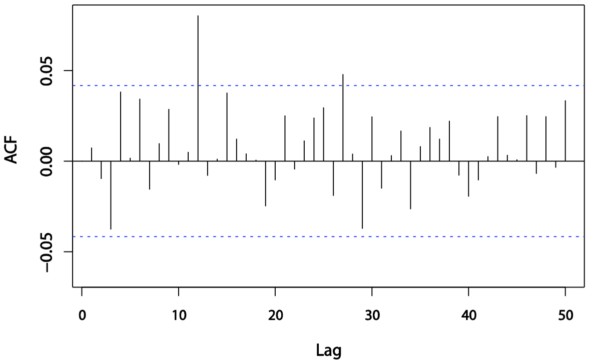

Figure 1: Autocorrelation plot of the AR model's residuals

Since the goal is to produce multi-step-ahead forecasts (in particular six hours-ahead), a different AR and VAR model is fitted for each lead-time. For instance, for lead-times 2 and 6, the VAR model has the following form:

$$\hat{P}_{t+2|t} = \alpha_2 + B_1 \cdot P_t + B_2 \cdot P_{t-1} + B_3 \cdot P_{t-22} + E_{t+2|t} \tag{18}$$

$$\hat{P}_{t+6|t} = \alpha_6 + B_1 \cdot P_t + B_2 \cdot P_{t-1} + B_3 \cdot P_{t-18} + E_{t+6|t} \tag{19}$$

where the terms P_t and P_{t-1} remain the same, and the seasonal lag associated to the previous day changes with the lead-time. The RLS and GB algorithms described in sections 3.3 and 3.4 are now used to estimate the coefficients for each lead-time.

Finally, in addition to the AR and VAR models described in the previous sections, a VAR with exogenous variables (VARX) is also proposed and evaluated. The model consists in adding exogenous variables to Eq. 7, which are the solar power values collected by each (or a subset of) EB.

The desired output is to have the EB measurements (P_t^{EB}) improving the solar power forecast at the DTC level. The VARX for lead-time *t+1* has the following structure:

$$\hat{P}_{t+1|t} = \alpha + B_1 \cdot P_t + B_2 \cdot P_{t-1} + B_3 \cdot P_{t-23} + B_4 \cdot P_t^{EB} + B_5 \cdot P_{t-1}^{EB} + E_{t+1|t} \tag{20}$$

Note that, for the EB observations, only the past observations *t* and *t-1* are included in the model since the goal is to use the EB as distributed sensors that characterize the current atmospheric conditions (in terms of solar power) across the region.

4. Test Case Results

4.1 Description

The dataset used as test case is from the city of Évora, which presently represents a large-scale Smart Grid pilot with more than 30 000 EB and 300 DTC (all costumers and substations), in order to have the entire municipality covered (an area of 1307 km^2).

The dataset consists of time series from 44 EB associated to residential PV (rated power ranging between 1.1 kWp and 3.7 kWp). Moreover, these EB are also related to 10 different DTC, and the total values of each DTC are also forecasted.

The parameters of the clear-sky model determined by trial-error tests are: σ_h=0.01, σ_{doy}=0.02, τ=85%. The forgetting factor λ for both AR and VAR was found to be 0.999.

The original data was sampled in 15 minutes, but it was resampled to hourly values. The model's fitting period was between 1 February 2011 and 31 January 2012, and the test period was between 1 February 2012 and 6 March 2013.

The point forecast results are evaluated with the Root Mean Square Error (RMSE) calculated for the kth lead-time [12]:

$$\mathrm{RMSE}_k = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{p}_{t+k|t} - p_{t+k}\right)^2} \tag{21}$$

The RMSE is to be normalized with the solar peak power.

The probabilistic forecast results are evaluated with Continuous Ranking Probability Score (CRPS) modified for quantile forecasts [34]:

$$\mathrm{C}RPS_k = 2 \cdot \int_0^1 \rho\left(p_{t+k}, \hat{p}^{\tau}_{t+k|t}\right) d\tau \tag{22}$$

where ρ is the average of the quantile loss function (Eq. 13) calculated on the test period dataset and $\hat{p}^{\tau}_{t+k|t}$ is the quantile forecast. The integral is calculated through numerical integration with the Simpson's rule.

The forecast skill of two models (VAR and VARX) is evaluated by computing the improvement over AR in terms of RMSE and CRPS:

$$\mathrm{Imp}_k = \frac{CRPS\,(\text{or } RMSE)_{k,AR} - CRPS\,(\text{or } RMSE)_{k,VAR}}{CRPS\,(\text{or } RMSE)_{k,AR}} \cdot 100\% \qquad (23)$$

4.2 Point Forecast Results

The improvement of the VAR and VARX over the AR model for each lead-time is plotted in **Figure 2** for two DTC and for the RMSE_k calculated with the complete dataset of DTC forecast errors.

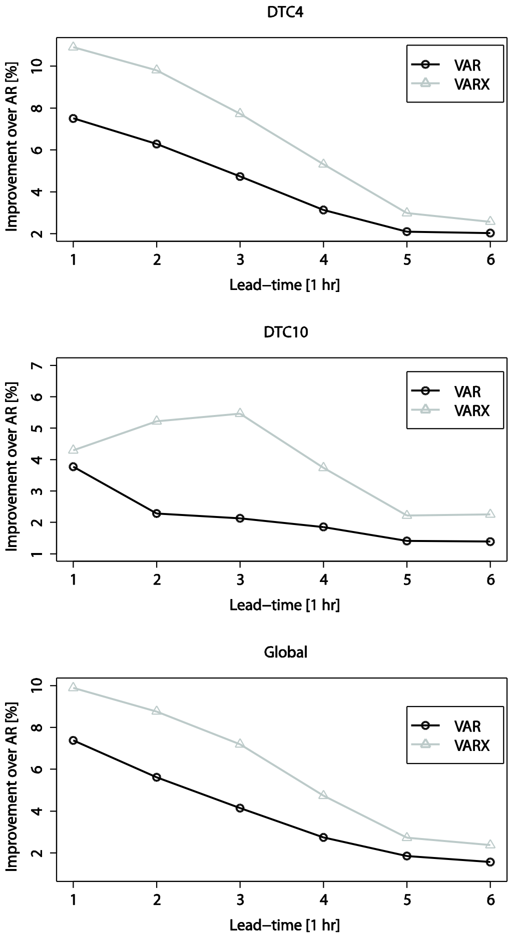

Figure 2: Imp_k of the VAR and VARX models for two DTC and for the RMSE calculated using the entire dataset of forecast errors.

These plots clearly show that the VARX model achieves the highest improvement. From the full set of DTC, number 4 is the one with the highest overall improvement, reaching a value around 11% for the first lead-time and around 6% for the sixth lead-time. Number 10 is the one with the lowest improvement, ranging between 3% and 4%.

The VAR model also achieves a positive improvement in all lead-times, but lower than VARX. This shows that information from surrounding smart meters (or EB), explored as distributed sensors, can improve the forecast skill at the DTC (or MV/LV substation) level. This is corroborated by a global improvement of the VARX between 5.5% and 10%.

Another interesting observation is that the improvement decays with the lead-time, meaning that the distributed information is more relevant for the first three hours. This makes sense since the forecasting model in this test case only includes information from a small municipality. If solar power data from neighboring municipalities and regions is included in the model, a higher improvement for lead-times between 4 and 6 is expected.

Figure 3 shows the improvement obtained with the VAR model for two EB and for the RMSE_k calculated with the full dataset of EB forecast errors.

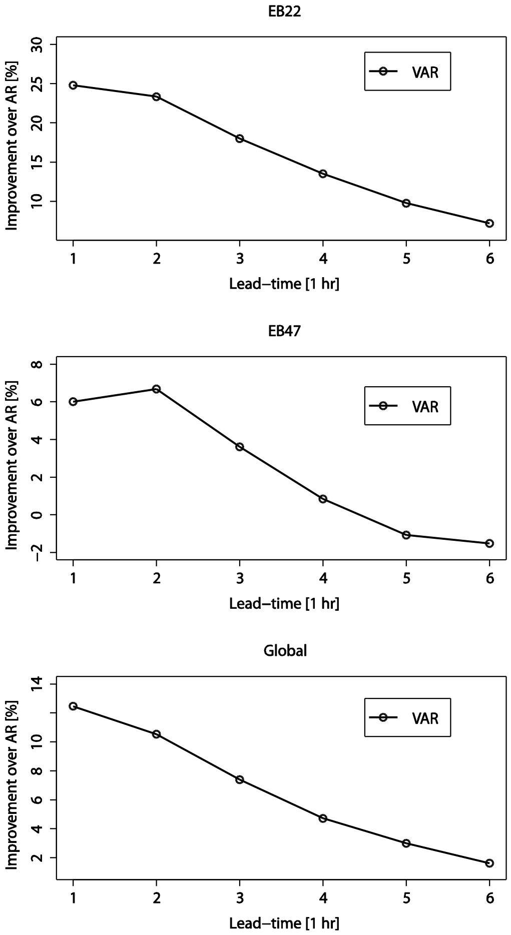

Figure 3: Imp_k of the VAR model for two EB and for the RMSE calculated using the entire dataset of forecast errors.

The VAR model attained the highest improvement for EB number 16, with a value around 25% for lead-time 1 and around 11% for lead-time 6. The lowest improvement was attained for EB number 39, with 6.5% for lead-time 1 and a negative value of around -1% for lead-time 6. The global improvement for the EB dataset varies between 0.1% and 12.5%. Compared to the DTC results, the improvement obtained for the EB dataset is higher for the first two lead-times.

Figure 4 depicts the average, minimum and maximum values of the normalized RMSE for the EB and DTC datasets, calculated from the individual RMSE_k values of each EB and DTC. For the DTC, the forecast errors are from the VARX model, while for the EB are from the VAR model. The RMSE magnitude is consistent with the state of the art (see [12] and [10]).

The average RMSE_k of the EB and DTC is similar for the first two lead-times, but the difference increases for the other lead-times. The main difference is in the minimum and maximum values, with higher amplitude for the EB. For instance, there are EB with an RMSE below 10% (even for lead-time 6), but also EB with an RMSE close to 20%. Since the variability in solar power due to the clouds is smoothed by the aggregation of EB in one DTC, this is an expected result. Nevertheless, this also shows that the forecasts for some EB are significantly improved with the distributed information (e.g., EB number 16 in **Figure 3**).

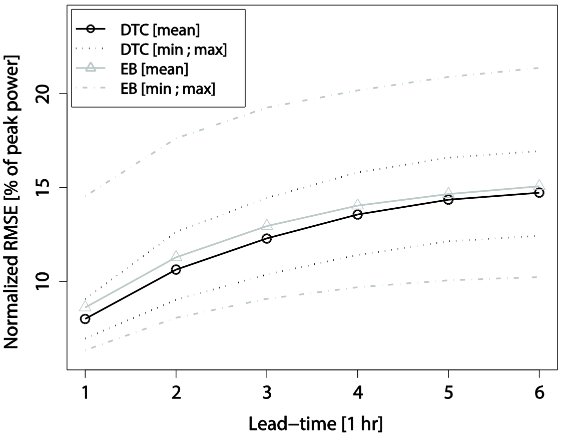

Figure 4: Average, minimum and maximum normalized RMSE_k calculated with the individual RMSE_k obtained for each EB and DTC.

4.3 Probabilistic Forecast Results

The improvement in terms of Continuous Ranking Probability Score (CRPS) of the VAR and VARX over the AR model for each lead-time is plotted in **Figure 5** for two DTC and averaged over the entire dataset of DTC. These three plots clearly show that the VAR framework improves over the AR in all lead-times. In DTC number 5 the improvement goes up to 12%, while the overall improvement is between 1.4% and 5.9%. The DTC number 10 is the one that presents the lowest improvement, with an average value around 0.1%. The results for DTC number 5 and 10, as well as for the global dataset show that the VARX achieves the highest improvement for lead-time between 2 and 6 (i.e., 16.4% for lead-time 3 in DTC 5).

Similarly to the point forecast results, these results confirm that the distributed information from DTC and EB improve the probabilistic forecast skill.

In terms of absolute values, the CRPS averaged over the entire dataset of DTC ranges between 3.9% (normalized by peak power) and 6.2% for AR model, between 3.6% and 6.1% for the VAR model, between 3.7% and 6% for the VARX model.

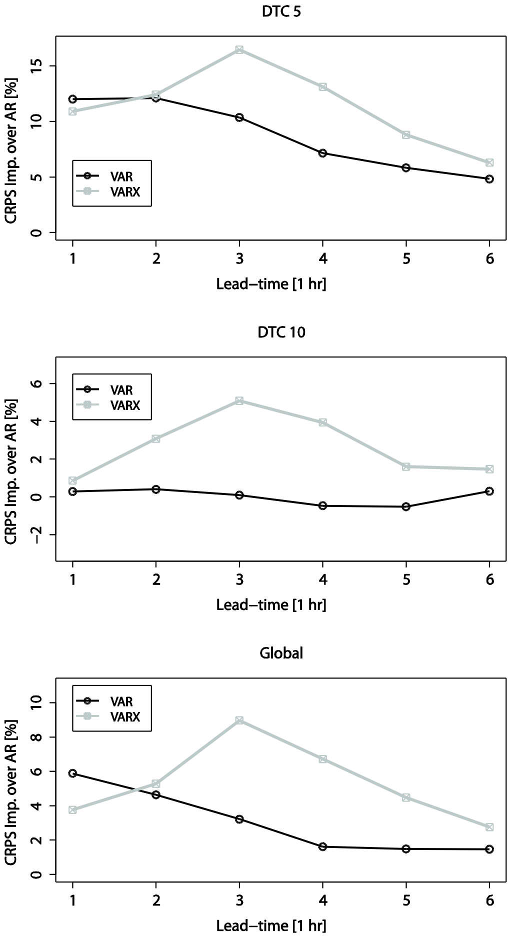

Figure 5: Imp$_k$ of the VAR and VARX models for two DTC and for the CRPS calculated using the entire EB dataset.

The CRPS improvement obtained with the VAR model is depicted in **Figure 6** for the EB dataset. In EB number 14 the improvement ranges between 7.3% and 21.4%, while in EB number 29 is negative between lead-times 2 and 6. The overall improvement, i.e., average of the individual improvements of the 44 EB, ranges between -2.8% and 4.6%.

In contrast to the DTC dataset, the improvement is negative for lead-times between 4 and 6. These results are very different from the point forecast improvement presented in section 4.3 and also the improvement obtained with the median, which ranges between 6.5% (lead-time 1) and 3.3% (lead-time 6). This negative improvement is mainly explained by a poor performance of VAR models in some quantiles. This is illustrated by **Figure 7** that shows the quantile loss calculated for each nominal probability and lead-times 1 and 5. For lead-time 1, the quantile loss of the VAR is lower than the one obtained with AR for all quantiles. Conversely, the skill of the VAR model in lead time 5 is lower than the AR model for the 5-20% quantiles. In this case, the information from distributed sensors decreases the forecast skill of some quantiles, which impacts the overall score.

This is an interesting result since in [16] is mentioned that, for the wind power problem, “forecast improvements mainly come from the space-time correction of the point forecasts”. However, in this paper, and for the solar power problem, an improvement in point forecast skill is not translated to an improvement in some quantile forecasts. A future development to improve this probabilistic forecast is the employment of a method that combines different quantile forecasts [35] (from the AR and VAR models in this case), which in turn may improve the individual score of each quantile. For instance, the VAR model outperforms AR in quantile 40%, while AR outperforms VAR in the 10% quantile.

Finally, the absolute values of CRPS are between 4.4% and 6.5% for the AR model, and between 4.2% and 6.65% for the VAR model.

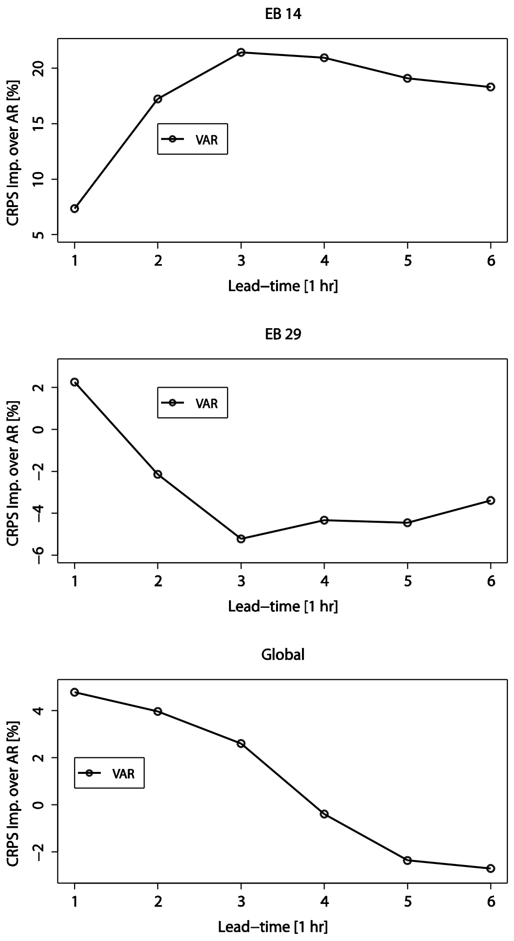

Figure 6: Imp$_k$ of the VAR model for two EB and for the CRPS calculated using the entire EB dataset.

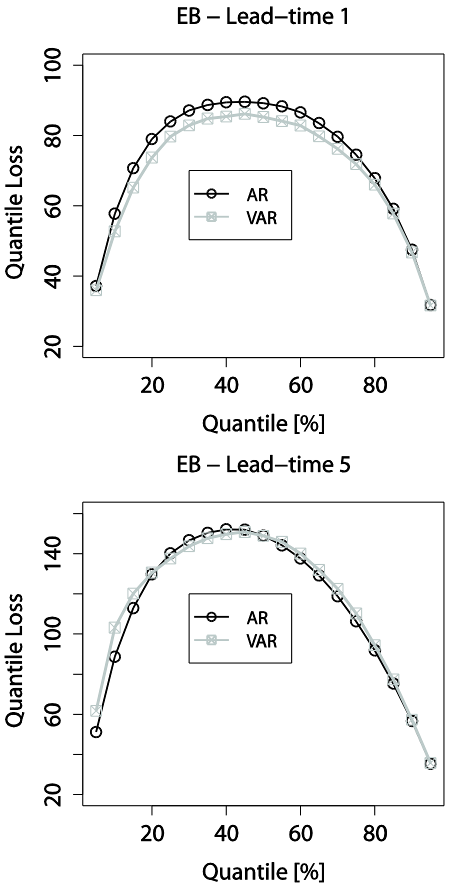

Figure 7: Quantile loss of the AR and VAR models for lead-times 1 and 5 and EB dataset.

5. Conclusions

This paper proposes a new framework for very short-term solar power probabilistic forecasts, which uses information collected by a Smart Grid infrastructure (e.g., smart meters, pyranometer sensors, remote terminal units). The results for real data from a Smart Grid test pilot, in the city of Évora, Portugal, show that information from distributed PV generation, when combined in a common forecasting framework, can improve the point forecast skill, compared to an univariate model (i.e., only uses past observations of a single time series), between 8% and 12% on average for the first three lead-times and the probabilistic forecast skill between 1.4% and 5.9%. Furthermore, time series observations collected by smart meters associated to residential PV improve both point and probabilistic forecast skill at the secondary substation level. Therefore, the adoption of multivariate (spatial-temporal) models, which fully explore the information collected by smart meters and intelligent electronic device, is recommended to create high quality forecasting methods that will support Smart Grid management functions. The main requirement is to have a quasi-real-time data flow between smart meters and central systems, which can be waived by considering upscaling techniques for solar power measurements [36].

These results open new avenues of future research, such as: combination of information from satellite frames and weather stations; apply new data mining and optimization techniques to enable an increase of the spatial coverage; develop regime-switching models.

Acknowledgements

This work was made in the framework of the BEST CASE project ("NORTE-07-0124-FEDER-000056") financed by the North Portugal Regional Operational Programme (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), and by national funds through Fundação para a Ciência e a Tecnologia (FCT), Portugal. It was also co-financed by the 7th RTD Framework Programme within the SuSTAINABLE project (contract no. 308755) and the FCT project «SusCity – MITP-TB/CS/0026/2013».

References

[1] M. Bazilian, I. Onyeji, M. Liebreich, I. MacGill, J. Chase, J. Shah, D. Gielen, D. Arent, D. Landfear, and S. Zhengrong, "Re-considering the economics of photovoltaic power," *Renewable Energy*, vol. 53, pp. 329-338, May 2013.

[2] G. Masson, M. Latour, M. Rekinger, I. Theologitis, and M. Papoutsi, "Global market outlook for photovoltaics: 2013-2017," Tech. Report, European Photovoltaic Industry Association, 2013.

[3] F. Vallée, V. Klonari, T. Lisiecki, O. Durieux, F. Moiny, and J. Lobry, "Development of a probabilistic tool using Monte Carlo simulation and smart meters measurements for the long term analysis of low voltage distribution grids with photovoltaic generation," *International Journal of Electrical Power & Energy Systems*, vol. 53, pp. 468-477, Dec. 2013.

[4] K. Chandrasekaran, S. P. Simon, and N. Prasad Padhy, "SCUC problem for solar/thermal power system addressing smart grid issues using FF algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 62, pp. 450-460, Nov. 2014.

[5] A.G. Madureira, J.A. Peças Lopes, "Ancillary services market framework for voltage control in distribution networks with microgrids," *Electric Power Systems Research*, vol.86, pp.1-7, May 2012.

[6] C. Delgado and J.A. Domínguez-Navarro, "Point estimate method for probabilistic load flow of an unbalanced power distribution system with correlated wind and solar sources," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 267-278, Oct. 2014.

[7] M. Peik-Herfeh, H. Seifi, M.K. Sheikh-El-Eslami, "Decision making of a virtual power plant under uncertainties for bidding in a day-ahead market using point estimate method," *International Journal of Electrical Power & Energy Systems*, vol. 44, pp. 88-98, Jan. 2013.

[8] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, "Wind power forecasting: state-of-the-art 2009," Report ANL/DIS-10-1, Argonne National Laboratory, November 2009.

[9] R. H. Inman, H. Pedro, and C. Coimbra, "Solar forecasting methods for renewable energy integration," *Progress in Energy and Combustion Science*, vol. 39, no. 6, pp. 535-576, Dec. 2013.

[10] L.A. Fernandez-Jimenez, A. Muñoz-Jimenez, A. Falces, M. Mendoza-Villena, E. Garcia-Garrido, P.M. Lara-Santillan, E. Zorzano-Alba, P.J. Zorzano-Santamaria, “Short-term power forecasting system for photovoltaic plants,” *Renewable Energy*, vol. 44, pp. 311-317, August 2012.

[11] M. Zamo, O. Mestre, P. Arbogast, and O. Pannekoucke, “A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production,” *Solar Energy*, vol. 105, pp. 792-803, July 2014.

[12] P. Bacher, H. Madsen, H.A. Nielsen, “Online short-term solar power forecasting,” *Solar Energy*, vol. 83, no. 10, pp. 1772-1783, October 2009.

[13] E. Lorenz, J. Hurka, D. Heinemann, and H. Beyer, “Irradiance forecasting for the power prediction of grid-connected photovoltaic systems,” *IEEE Journal of Special Topics in Earth Observations and Remote Sensing*, vol. 2, pp. 2-10, 2009.

[14] H. Pedro, C. Coimbra, “Assessment of Forecasting Techniques for Solar Power Production with no Exogenous inputs,” *Solar Energy*, vol. 86, no. 7, pp. 2017-2028, July 2012.

[15] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, H.G. Beyer, “Solar energy assessment using remote sensing technologies,” *Remote Sensing of Environment*, vol. 86, no. 3, pp. 423-432, August 2003.

[16] J. Tastu, P. Pinson, P.-J. Trombe, and H. Madsen, “Probabilistic forecasts of wind power generation accounting for geographically dispersed information,” *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 480-489, 2014.

[17] V. Berdugo, C. Chaussin, L. Dubus, G. Hebrail, V. Leboucher, “Analog method for collaborative very-short-term forecasting of power generation from photovoltaic systems,” *Next Generation Data Mining Summit (NGDM '11)*, Athens, Greece, 4 September 2011.

[18] C. Yang, L. Xie, “A novel ARX-based multi-scale spatiotemporal solar power forecast model,” in *Processings of the North American Power Symposium (NAPS)*, USA, September 2012.

[19] V. Lonij, A. Brooks, A. Cronin, M. Leuthold, K. Koch, “Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors”, *Solar Energy*, vol. 97, pp. 58-66, 2013.

[20] R.J. Bessa, A. Trindade, A. Monteiro, C. Silva, V. Miranda, “Solar power forecasting in smart grids using distributed information,” in *Proc. of the PSCC 2014 - 18th Power Systems Computation Conference*, Wroclaw, Poland, Aug. 2014.

[21] R.J. Bessa, A. Trindade, V. Miranda, “Spatial–temporal solar power forecasting for Smart Grids,” IEEE Transactions on Industrial Informatics, In Press, 2015. DOI: 10.1109/TII.2014.2365703

[22] C. Gouveia, D. Rua, F.J. Soares, C. Moreira, P.G. Matos, J.A. Peças Lopes, “Development and implementation of Portuguese smart distribution system,” *Electric Power Systems Research*, vol. 120, pp. 150-162, March 2015.

[23] P. Lúcio, P. Paulo, H. Craveiro, “InovCity - Building smart grids in Portugal,” in *Proceedings of the 21nd International Conference on Electricity Distribution (CIRED)*, Frankfurt, Germany, 6-9 June 2011.

[24] www.sustainableproject.eu (accessed on September 2014)

[25] H.M. Diagne, M. David, P. Lauret, and J. Boland, “Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids,” *World Renew. Energ. For.*, USA, May 2012.

[26] P. Bacher, “Short-term solar power forecasting,” MSc Thesis, Technical University of Denmark, 2008.

[27] R. Koenker, G. Bassett, “Regression quantiles”, *Econometrica*, vol. 46, pp. 33-50, 1978.

[28] H. Madsen, *Time Series Analysis*, London: Chapman and Hall/CRC, 2006.

[29] R. Davidson, J.G. MacKinnon, *Econometric Theory and Methods*, New York: Oxford University Press, 2003.

[30] L. Ljung, T. Soderstrom, T*heory and Practice of Recursive Identification*, Cambridge: The MIT Press, 1983.

[31] J.H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Annals of Statistics*, vol. 29, no. 5, pp. 1189-1232, 2001.

[32] P. Bühlmann, “Boosting for high-dimensional linear models,” *Annals of Statistics*, vol. 34, no. 2, pp. 559-583, 2006.

[33] J.M. Wooldridge, Introductory Econometrics: A Modern Approach, 2nd Edition, South-Western College Pub, 2002.

[34] G. Anastasiades and P. McSharry, “Quantile forecasting of wind power using variability indices,” *Energies*, vol. 6, no. 2, pp. 662-695, 2013.

[35] K. Shan and Y. Yang, “Combining regression quantile estimators,” *Statistica Sinica*, vol. 19, pp. 1171-1191, 2009.

[36] R.J. Bessa, “Solar power forecasting for smart grids considering ICT constraints,” in *Proc. of the 4th International Workshop on Integration of Solar Power into Power Systems (SIW2014)*, Berlin, Germany, Nov. 2014.