Solar Power Forecasting for Smart Grids Considering ICT Constraints

Ricardo J. Bessa
Center for Power and Energy Systems
INESC TEC
Porto, Portugal
ricardo.j.bessa@inesctec.pt

***Abstract*—The solar power capacity connected to the distribution grid is growing across several countries and the deployment of the Smart Grid concept can help to maximize its penetration. The communication infrastructure currently in place might impose constraints in terms of data transfer from the smart meters, which imposes limitations in new smart grid functions, such as very short-term solar power forecasting. This paper describes a new forecasting framework, based on autoregressive models and gradient boosting, that overcomes these limitations and produces point and probabilistic solar power forecasts for MV/LV substations. The results, for a smart grid test-pilot in Évora (demonstration site of the EU Project SuSTAINABLE), Portugal, show that the proposed methodology accomplishes a forecasting accuracy comparable to the one obtained if the time series with the total observed power is available in real-time.**

Keywords; Solar power, forecasting, micro-generation, smart grid, probabilistic, smart meters, communication.

I. INTRODUCTION

In several countries, the installed solar power capacity is growing, particularly connected to the medium and low voltage levels (MV and LV) of the distribution grid [1]. Distribution System Operators (DSO) are developing and testing information and communication technology (ICT) to boost the Smart Grid concept [2][3]. This new ICT will support and enable the maximization of renewable energy integration in distribution grids and requires new computation tools, such as solar power forecasting and voltage control for MV and LV grids.

The literature about solar power forecasting is mainly oriented to three main topics:

a) exploring the use of satellite frames in order to improve very short-term forecasts (up to six lead-times ahead);

b) using machine learning algorithms to convert weather variables (e.g., global horizontal irradiance) to solar power forecasts;

c) using machine learning and time series models to produce very short-term forecasts based on past values of the time series.

For the first topic, Hammer *et al.* [4] present an algorithm based on cloud-index images that are predicted with motion vector fields derived from two consecutive images. For the second topic, neural networks and support vector machines are used together with NWP [5]. An alternative approach to produce forecasts from NWP is to use an electrical representation of the PV system [6].

The list of algorithms for the third topic is vast. Some examples are ARIMA, neural networks and k-nearest neighbour [7]. An important feature of these algorithms is that past values of the time series are the key input and this information must be available when a new forecast is about to be generated. Bessa *et al.* [8] describe a new forecasting framework, based on vector autoregression theory, that combines spatial-temporal data collected by smart meters and distribution transformer controllers to produce six-hour-ahead forecasts at the residential solar PV and secondary substation (i.e., MV/LV substation) levels.

Different DSO are deploying and operating their own ICT infrastructure, using communication protocols, such as power-line communication (PLC), which might not allow 15 minutes data exchange between a remote terminal unit (RTU) installed at the secondary substation level and all connected smart meters [9]. Presently, in most Smart Grid test-pilots, it is either technically unfeasible to get, every 15 minutes, measured data from the smart meters to an upstream level (e.g., DSO control centre) or the communication costs are very high. Issues related to data transfer delays or failures might also contribute to having unavailable time series data.

The availability of time series data with, at least, a 15-minute update frequency is a key input for very short-term solar power forecasting since it improves the forecast error during the first four lead-times compared to information from Numerical Weather Prediction (NWP) systems [10]. Moreover, this information can also improve the net-load forecast under high penetration of photovoltaic (PV) panels in commercial and residential buildings [11]. Accurate solar power forecasts are an important requirement for several distribution grid management functions, such as voltage control [12] and storage management optimization [13].

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). It was also co-financed by the 7th RTD Framework Programme within the SuSTAINABLE (contract no. 308755) project.

In this paper, a feasible forecasting approach, which considers the communication constraints of the current ICT infrastructure, is proposed. The method work is as follows:

a) a subset of the smart meters associated with micro-generation is selected to have General Packet Radio Service (GPRS) communication protocol (or equivalent);

b) this subset is used to extrapolate the observed value of total solar generation in that secondary substation (using historical values from all smart meters that are generally available with a delay of 24 hours);

c) past observations from the extrapolated total solar generation together with past values from the GRPS-based smart meters are used to forecast the total solar generation for the next six lead-times.

Since the uncertainty associated to solar power generation is high, mainly due to cloud effects, probabilistic forecasts, represented by a set of quantiles ranging between 5% and 95%, are also generated. Component-wise gradient boosting [14] is used in step (a) to select the "reference" smart meters and also in step (c) to generate probabilistic forecasts.

The following original contributions are produced:

a) a forecasting framework that overcomes the ICT constraints is proposed in order to produce solar power forecasts at the secondary substation level;

b) the value of combining information from distributed sensors (i.e., PV panels with real-time measurements) is demonstrated;

c) probabilistic forecasts based on component-wise gradient boosting are produced.

The paper is organized as follows: section II describes the smart grid infrastructure in Portugal and the potential limitations in terms of time series data transfer; section III describes the method for selecting the reference smart meters; section IV presents the forecasting framework; the test case results are presented in section V; section VI presents the conclusions and future work.

II. Smart Grid Infrastructure

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 [15]. This resulted in a large-scale demonstration pilot in the city of Évora in Portugal, named InovCity [16], which is also one of the demonstration sites of the EU Project SuSTAINABLE [17].

The main components of this infrastructure, depicted in Fig. 1, are the EDP Box (EB) and the Distribution Transformer Controller (DTC). The EB is a smart meter with load and generation management functions, located at each delivery point. Load and generation are metered separately. It can interact with other devices through a home area network. The DTC is located at the secondary substation level, comprising modules for measurement, remote control and communication actions. It collects data from the EB and the secondary substation. Both the EB and the DTC are part of a hierarchical control and communication architecture. Each EB has a bi-directional communication with the corresponding DTC through GPRS (General Packet Radio Service) or PLC (Power Line Communications), and the DTC communicates with the SCADA/DMS through a wide area network based on GPRS [16]. Other communication technologies, such as radio frequency mesh modules, can be explored in this framework.

At the primary substation level (HV/MV), a Smart Substation Controller (SSC) is installed. The SSC is responsible for aggregating and managing the operational data from EB and DTC, and for applying demand-side management, self-healing and generation management strategies. Therefore, the MV grid is managed by the SSC. On the top of these technologies, there are services capable of handling large volumes of data and, at the same time, providing an overview of all existing devices.

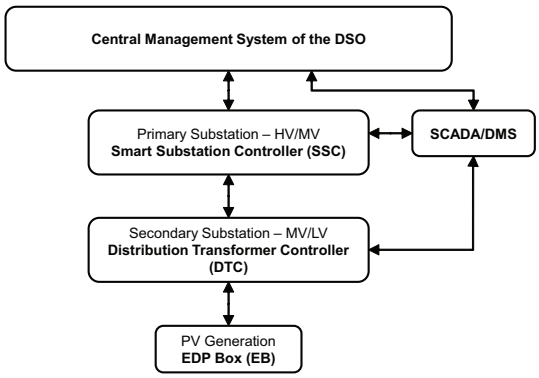

Figure 1. Smart grid infrastructure of the InovCity test pilot (demo site of the EU Project SuSTAINABLE).

In this architecture, the forecasting tools are installed at the central management level and produce inputs used in the voltage control tools [9].

Presently, the current smart grid test pilots have limitations in terms of data transfer capability or costs [18][19]. In some cases, the time series data measured by the smart meters is only available at the DSO control centre (or automatic metering management system) with a minimum delay of six hours. An alternative approach would be a fully decentralized paradigm, self-governed by autonomous devices. However, the costs associated to this solution would be very high nowadays. Regarding ICT constraints, the following overview is provided in [9]:

a) *between primary substation and SCADA/DMS*: optical fibre infrastructure with no particular limitations in terms of data latency and amounts of data being transferred;

b) *between secondary substation and SCADA/DMS or primary substation*: if optical fibre exists, there is no particular limitation. Otherwise, the latencies and bandwidths associated to the available communication protocols, such as GPRS or other, have to be taken into consideration, namely the frequency of data retrieval and the amount of retrieved data, and the maximum delay that is acceptable depending on the functionality to be executed.

c) *between smart meter and secondary substation*: regardless of the technology, it is the one that

introduces stricter constraints, both on volume of data to be transferred, available bandwidth and data latency. Current technologies, such as PLC, may impose limitations on the amount of data that can be retrieved or the frequency of data retrieval.

Moreover, the use of the same communications infrastructure to transfer different types of data, such as data for metering and operation, may introduce more constraints.

Since the equipment installed at the secondary substation level (e.g., DTC, RTU) only measures the net-load, in order to measure the total solar generation it is necessary to sum the individual generation of each smart meter. This means performing a query to hundreds of smart meters at the same time and with a frequency below one hour (in order to feed the forecasting algorithms with recent past observations). In the following sections, a forecasting framework will be described to overcome these constraints.

III. Reference Smart Meters Selection

The basic idea of the forecasting framework is depicted in Fig. 2.

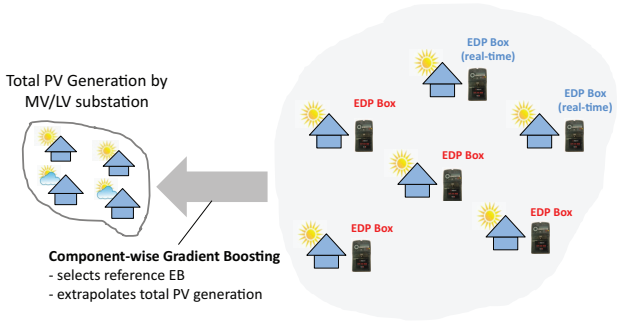

Figure 2. Selection of reference EDP Box (EB) to estimate total solar power at the secondary substation level.

Basically it consists in selecting, from a total of *k* EB, a subset of *s* EB (or smart meters) that should have real-time communication capability, while only the historical time series data is used from the remaining EB (*k*-*s*). The total PV generation in the MV/LV substation is estimated from the EB subset. This approach somehow resembles the upscaling method used for wind power forecast [10]. Since most of the relations between total and individual power are linear (see Fig. 3), the following multi-regression linear model can be defined to extrapolate the total power from a subset with *s* EB:

$$\hat{p}_t^{sub} = \beta_0 + \beta_1 \cdot p_{1,t}^{EB} + \beta_1 \cdot p_{2,t}^{EB} + \cdots + \beta_s \cdot p_{s,t}^{EB} + \varepsilon_t \qquad (1)$$

where β are the model's coefficients and ε_t is a white noise process with zero mean and constant variance.

The model's coefficients are estimated using historical time series data from the *s* EB and the total solar power (p_t^{sub}) is the sum of the *k* EB individual values. The component-wise gradient boosting (GB) method [14] conveys automatic variable selection capability, which makes it very useful to calculate the coefficients and select the most relevant variables (i.e., reference EB).

Boosting is an ensemble machine-learning algorithm for classification and regression, which combines base learners. 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:

$$\hat{F}(x) = \hat{f}_0(x) + \sum_{m=1}^{M} \hat{f}_m(x) \qquad (2)$$

where $f_0(x)$ is an initial guess, $f_m(x)$ are the base learners, *x* the covariates (or explanatory variables) and *M* is the maximum number of "boosts" (which is a model's parameter). The empirical risk function, $\rho\left(p^{sub}, \hat{F}(x)\right)$, is the squared loss and the base learner is a linear effect of a continuous predictor.

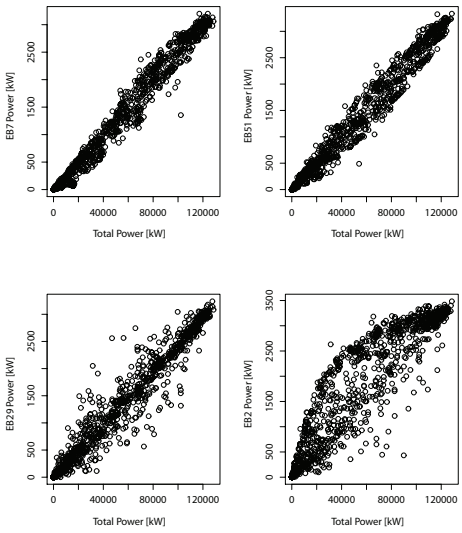

Figure 3. Scatter plot between total and individual (by EB) solar power.

The component-wise GB works as follows:

1. initialize $\hat{F}_{m-1}(x)$ with the mean value of the p_t^{sub} response variable. Note that in Eq. 1, *x* corresponds to the vector of explanatory variables $p_{s,t}^{EB}$;
2. for each *m*, given $\hat{F}_{m-1}(x)$, compute the negative gradient $u[i] = -\partial\rho/\partial F$ of the loss function and evaluate it at $\hat{F}_{m-1}(x[i])$, where *i* takes values between 1 and *N* (i.e., number of training samples);
3. using the negative gradient *u[i]* calculated in step 2 as the response variable, produce an estimate of the coefficients associated to each candidate base learner (i.e., $\beta_j \cdot x^{(j)}$):

$$\beta_j = \frac{\sum_{i=1}^{N} u[i] \cdot x^{(j)}[i]}{\sum_{i=1}^{N} \left(x^{(j)}[i]\right)^2} \qquad (3)$$

4. determine the s^{th} predictor or base learner (from a set with *d* candidates) that minimizes the squared loss function:

$$s = \arg\min_{1\le j\le d} \sum_{i=1}^{N} \left(u[i] - \beta_j \cdot x^{(j)}[i]\right) \quad (4)$$

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

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

$$\hat{F}_m(x) = \hat{F}_{m-1}(x) + \nu \cdot \hat{f}_m\left(x^{(s)}\right) \quad (5)$$

where v is a shrinkage parameter;

6. stop when $m=M$.

The final estimator is obtained in step 6. Since the squared loss is used as empirical risk function, the negative gradient in step 2 corresponds to the residuals given by:

$$u[i] = p_t^{sub}[i] - \hat{F}_{m-1}(x[i]) \quad (6)$$

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 m value with the lowest square error is selected for each lead-time. According to [14], the value of v does not influence significantly the results if set to be a low value (e.g., 0.1). The R package "mboost" is used in this paper for the component-wise GB [20].

In step 4 of the GB algorithm, only one base learner, which is a coefficient multiplied by one covariate, is selected and used to update the model in step 5. This means that if one variable is never selected during step 4, it will have a null coefficient and can be removed from the model. This characteristic enables the selection of reference EB. For instance, Fig. 4 depicts the squared error as a function of the number of coefficients with non-null value. Each point in the curve is an iteration of the GB algorithm. The data used in this plot is from the test case described in section V, where p_t^{sub} corresponds to one substation with 44 microgeneration EB connected. This plot shows that after including 12 non-zero coefficients (i.e., 11 EB measurements $p_{S,t}^{EB}$ plus the constant term β_0) the squared error almost stabilizes. Therefore, the total power from this substation can be extrapolated with lower error using measurements from a subset with 11 EB.

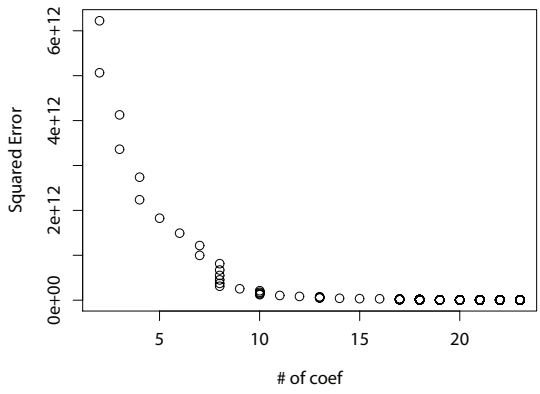

Figure 4. Curve relating the squared error and the number of coefficients with non-null coefficients in each GB iteration.

IV. Forecasting Framework

The forecasting framework is based on the ARIMA process, which is a well-known class for univariate time series models [21]. A special model of this class is the autoregressive (AR), in which the value of the response variable for time interval t can be interpreted as a regression on past observations (or lags) of the time series. For one hour-ahead forecast (and substation values), the AR model is the following:

$$\hat{p}_{t+1|t}^{sub} = \beta_0 + \beta_1 \cdot \hat{p}_t^{sub} + \beta_2 \cdot \hat{p}_{t-1}^{sub} + \cdots + \beta_l \cdot \hat{p}_{t-l}^{sub} + \varepsilon_{t+1|t} \quad (7)$$

where β are the model's coefficients, l the order of the AR model and $\epsilon_{t+1|t}$ is a white noise process with zero mean and constant variance. Note that $\hat{p}^{sub}$ is the total solar power extrapolated using the method described in section III.

Since the subset with s EB provides time series data in real-time, the past observations (or lags) of the time series can be added to Eq. 7. This model is named autoregressive with exogenous variables (ARX) and is given by:

$$\begin{aligned}\hat{p}_{t+1|t}^{sub} = \beta_0 + \beta_1 \cdot \hat{p}_t^{sub} + \beta_2 \cdot \hat{p}_{t-1}^{sub} + \cdots + \beta_l \cdot \hat{p}_{t-l}^{sub} + \\ \theta_1 \cdot p_t^{EB[1]} + \theta_2 \cdot p_{t-1}^{EB[1]} + \cdots \\ + \theta_n \cdot p_{t-1}^{EB[2]} + \beta_{n+1} \cdot p_{t-1}^{EB[2]} + \cdots + \varepsilon_{t+1|t}\end{aligned} \quad (8)$$

where θ are the model's coefficients of the EB observations lags.

The coefficients of the models in Eq. 7 and 8 are estimated with the Recursive Least Squares (RLS) method with forgetting factor described in the appendix section.

The solar power time series are pre-processed in order to remove the seasonal component with the weighted quantile regression model (or clear-sky model) described in [22]. Since this model is unable to completely remove the seasonal component of the model, a diurnal term corresponding to lag *t-23* (for hourly time series) and *t-92* (for 15 min time series) is included in the models of Eq. 7 and 8 (also for the p_t^{EB} variables). Following the results in [8], the lags *t* and *t-1* are also introduced in the models of Eq. 7 and 8.

The goal in this paper is to produce six lead-times ahead forecasts. Therefore, a different model (Eq. 7 and 8) is fitted for each lead-time.

For probabilistic forecasts, the component-wise GB described in section III is used with the quantile loss function [23] as the empirical risk function:

$$QL\left(p^{sub}, \hat{F}(x)\right) = \begin{cases} \tau \cdot \left(p^{sub} - \hat{F}(x)\right), & p^{sub} - \hat{F}(x) > 0 \\ -(1-\tau) \cdot \left(p^{sub} - \hat{F}(x)\right), & p^{sub} - \hat{F}(x) \le 0 \end{cases} \quad (9)$$

where τ is the quantile nominal proportion.

In this case, the negative gradient of step 2 in section III is given by:

$$-\frac{\partial QL\left(p^{sub}, \hat{F}(x)\right)}{\partial F} = \begin{cases} \tau, & p^{sub} - \hat{F}(x) > 0 \\ \tau - 1, & p^{sub} - \hat{F}(x) \le 0 \end{cases} \quad (10)$$

A model (Eq. 7 and 8) is fitted for each quantile τ ranging between 5% and 95% with 5% increments.

Although not explored in this paper, a spatial-temporal framework could be adopted to improve the forecasts for all substations from the distribution grid. For instance, a vector autoregressive with exogenous variables can be fitted using the extrapolated total solar power and the subset of EB [8].

V. Test Case Results

A. Description

The solar power dataset used as test case is from the city of Évora, with about 54,000 residents and an area of 1307 km^2. Concerning distributed generation, in August 2014, there were 225 micro-generation producers, mainly solar PV, with an installed capacity of 786.5 kW. During 2 years, more than 30,000 EB and 300 DTC (smart grid equipment in an MV/LV substation) were installed in Évora, including all customers and substations, in order to have the entire municipality covered.

In order to test the proposed forecasting framework, time-series from 44 EB were used, comprising domestic PV, with installed capacity ranging between 1.1 kWp and 3.7 kWp. These EB measurements are assumed to be connected to the same DTC, and the total solar power in the substation is the sum of the EB individual generation.

The original data was sampled in 15 minutes, but results with data resampled to hourly values are also presented (covering the electricity market time resolution). The period between 1 February 2011 and 31 January 2012 was used to fit the models, and the period between 1 February 2012 and 6 March 2013 was used to calculate the forecast errors.

The forgetting factor λ used in the RLS method was equal to 0.999.

The point forecast results are evaluated with the root mean square error (RMSE) calculated for the k^{th} lead-time:

$$\mathrm{RMSE}_k = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{P}^{sub}_{t+k|t} - P^{sub}_{t+k}\right)^2} \qquad (11)$$

The RMSE is normalized with the peak power.

The probabilistic forecast results are evaluated with the Continuous Ranking Probability Score (CRPS) modified to evaluate quantile forecasts [24]:

$$CRPS_k = 2 \cdot \int_0^1 QL\left(p^{sub}, \hat{F}(x); \tau\right) d\tau \qquad (12)$$

where *QL* is the average quantile loss function (see Eq. 9) calculated on the test period dataset. The integral is calculated through numerical integration with the Simpson's rule.

In the next section, results from the following three models are presented:

- "AR no extrapolation": model that assumes real-time observations from all the 44 EB, i.e., in Eq. 7 the real substation power is used instead of extrapolated power;
- "AR": model from Eq. 7 with extrapolated power at the substation level;
- "ARX": model from Eq. 8 with extrapolated power at the substation level and real-time observations from a subset of *s* EB.

The performance of the AR and ARX models is compared by computing the improvement over "AR no extrapolation" in terms of RMSE and CRPS. For the RMSE, the improvement is calculated as follows:

$$\mathrm{Imp}_k = \frac{RMSE_{k,AR\_no\_extra.} - RMSE_{k,AR}}{RMSE_{k,AR\_no\_extra}} \cdot 100\% \qquad (13)$$

The calculation is analogous for the CRPS.

B. Results

Based on the results depicted in Fig. 4 (section III), it was decided to use 12 coefficients, which, considering the constant term, means using 11 EB (25% of the total number of EB) with real-time observations. This results in an RMSE for the total solar power extrapolation of around 1%.

Fig. 5 depicts the improvement results for the point forecast with hourly time resolution. Negative values mean no improvement (in fact, it means that the RMSE increased) over the model "AR no extrapolation". The AR model with extrapolated values only exhibits a small increase in forecast error (i.e., a maximum negative improvement of around -2%). This shows that it is possible to produce very short-term forecasts with acceptable accuracy for the total generation by substation even when only a subset of EB (or smart meters) have real-time communication capability.

The ARX model presents only a negative improvement in lead-time 2, while for the others it shows positive improvement. In this case, the additional information from exogenous variables (i.e., past observations from the EB subset) improves the forecast accuracy and, in some lead-times, it is even better than an AR model that uses non-extrapolated total generation values.

Fig. 6 depicts the results obtained with the same models, but for probabilistic forecasts. For the AR model, the conclusions are analogous to the ones obtained for point forecast. However, in this case, the results obtained with the ARX model are worse. For the first lead-time, the ARX presents a negative improvement of around -8.4%. This might be explained by the fact that the GB method used in this paper, contrary to the RLS algorithm, is not trained "on-line". Therefore, it is unable to capture structural changes in the time series.

Fig. 7 depicts the point forecast results for 15 minutes time resolution time series. Similar to the hourly resolution results, the ARX outperforms AR and presents positive improvement for lead-times 2, 5 and 6. Only for the first lead-time the models present a negative improvement greater than -6%.

The probabilistic forecast results for 15 minutes time resolution are depicted in Fig. 8. The conclusions are analogous to the point forecast and, in contrast to hourly resolution results, the ARX shows a good performance. An improvement of around 2% is achieved by the ARX model in lead-times 5 and 6.

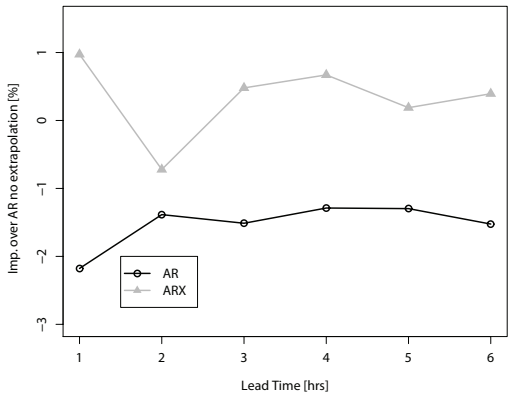

Figure 5. Point forecast and hourly time resolution: improvement of the AR and ARX models over the model “AR no extrapolation”.

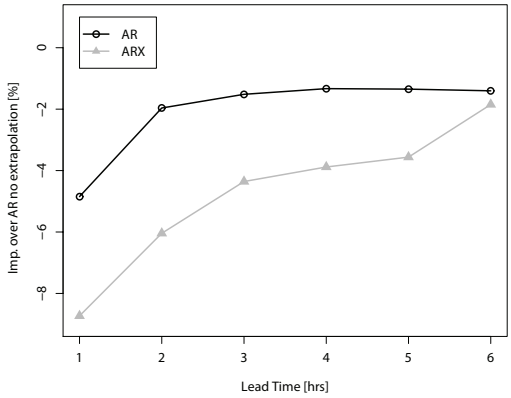

Figure 6. Probabilistic forecast and hourly time resolution: improvement of the AR and ARX models over the model “AR no extrapolation”.

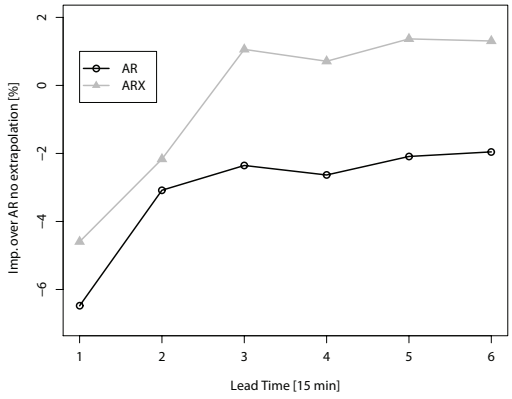

Figure 7. Point forecast and 15 min time resolution: improvement of the AR and ARX models over the model “AR no extrapolation”.

Fig. 9 presents the RMSE values for hourly and 15 minutes time resolution. For the point forecast and hourly values, the RMSE ranges between 6.8% and 12%, while for 15 minutes resolution it ranges between 5.1% and 11.5%. Note that the error obtained for the two time resolutions is not comparable, since the test sample size is different. The RMSE magnitude is consistent with the state of the art [7][22].

For probabilistic forecasts, the CRPS ranges between 3.2% and 5.4% for hourly resolution, and 2.5% and 5.5% for 15 minutes resolution.

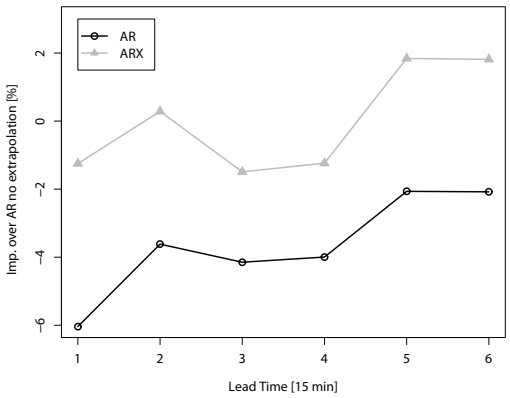

Figure 8. Probabilistic forecast and 15 min time resolution: improvement of the AR and ARX models over the model “AR no extrapolation”.

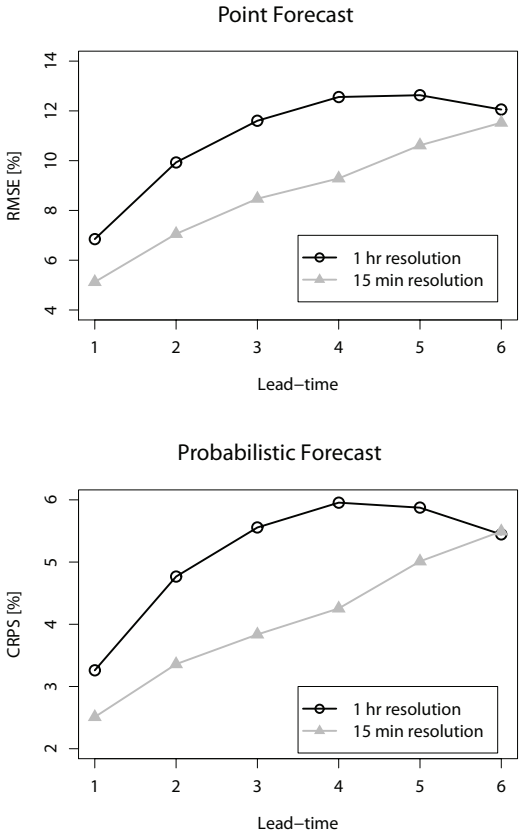

Figure 9. Normalized RMSE and CRPS (% of the rated power) of the solar power forecasts for hourly and 15 min time resolution.

VI. CONCLUSIONS

A new forecasting approach for very short-term solar power forecast, which overcomes the current communication constraints in the Smart Grid infrastructure, is proposed in this paper. It selects a subset of smart meters (EDP Box) with real-time communication and extrapolates the total generation by MV/LV substation level. Autoregressive models are constructed using the extrapolated time series.

The results for data from a Smart Grid pilot, in the city of Évora, Portugal, indicate that, using only 25% of the smart meters with real-time communication, it is possible to estimate the total power by MV/LV substation with an RMSE of around 1%. Moreover, forecasts produced with the extrapolated data only exhibit an accuracy decrease of

around 8% and, in some lead-times, present an improvement over a model that uses non-extrapolated values (i.e., assuming that all smart meters have real-time communication). Moreover, past observations from the smart meters with real-time information can help to improve the model. Therefore, the framework proposed in this paper can be used in a transitional phase with communication constraints between smart meters and DSO central systems. Nevertheless, the ideal scenario would be real-time data flow between smart meters and central systems, which would create opportunities to develop advanced data mining and machine learning algorithms.

The following research lines were identified for future work: development of a time-adaptive component-wise gradient boosting for probabilistic forecasting; inclusion of data from weather stations.

APPENDIX – RECURSIVE LEAST SQUARES

This appendix section describes the recursive least squares (RLS) method with a forgetting factor λ [25]. The update of the parameters of a linear model (such as the AR and ARX models described in section IV) is performed with the RLS method as follows for time step *t*:

$$\beta_t = \beta_{t-1} + K_t \cdot \left[p_t^{sub} - (\beta_{t-1} \cdot x) \right] \tag{14}$$

where K_t is given by

$$K_t = Q_t \cdot x \tag{15}$$

and Q_t by

$$Q_t = \frac{1}{\lambda} \cdot \left[Q_{t-1} - \frac{Q_{t-1} \cdot x \cdot x^T \cdot Q_{t-1}}{\lambda + x^T \cdot Q_{t-1} \cdot x} \right] \tag{16}$$

β_{t-1} is the coefficient matrix from time step *t-1*, β_t is the matrix with updated coefficients (i.e., after receiving the last observation), x is the vector of covariates.

A *forgetting factor* λ equal to 1 leads to a recursive estimation of the coefficients, while a smaller value discounts old data with an exponential decay.

This algorithm requires some initial values for B_0 and Q_0. A simple and robust approach is to initialize B_0 with zeros and Q_0 as a diagonal matrix with a large constant value.

REFERENCES

[1] 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.

[2] M. Coppo, P. Pelacchi, F. Pilo, G. Pisano, G.G. Soma, R. Turri, "The Italian smart grid pilot projects: Selection and assessment of the test beds for the regulation of smart electricity distribution," *Electric Power Systems Research*, in press, 2014. DOI: 10.1016/j.epsr.2014.06.018

[3] 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*, in press, 2014. DOI: 10.1016/j.epsr.2014.06.004.

[4] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, and H.G. Beyer, "Solar energy assessment using remote sensing technologies," *Rem. Sens. of Env.*, vol. 86, no. 3, pp. 423-432, 2003.

[5] L.A. Fernandez-Jimenez, A. Muñoz-Jimenez, A. Falces, M. Mendoza-Villena, E. Garcia-Garrido, P.M. Lara-Santillan, E. Zorzano-Alba, and P.J. Zorzano-Santamaria, "Short-term power forecasting system for photovoltaic plants," *Renew. Energ.*, vol. 44, pp. 311-317, Aug. 2012.

[6] J. Sumaili, R.J. Bessa, F. Rahman, R. Tomé, J.N. Sousa, "Electrical model parameter characterization for short-term solar power forecasting," *4th International Workshop on Integration of Solar Power into Power Systems*, Berlin, Germany, Nov. 2014.

[7] H. Pedro and C. Coimbra, "Assessment of Forecasting Techniques for Solar Power Production with no Exogenous inputs," *Sol. Energ.*, vol. 86, no. 7, pp. 2017-2028, Jul. 2012.

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

[9] R.J. Bessa, A. Madureira, J. Pereira, J. Peças Lopes, *et al.*, "Definition of overall system architecture," Deliverable D-2.3, EU Project SuSTAINABLE, Dec. 2013.

[10] R.J. Bessa, C.L. Moreira, B. Silva, M.A. Matos, "Handling renewable energy variability and uncertainty in power systems operation," *Wiley Inter. Rev.: Energ. and Env.*, vol. 3, no. 2, pp. 156-178, March/April 2014.

[11] A. Kaur, H. Pedro, C. Coimbra, "Impact of onsite solar generation on system load demand forecast," *Energy Conversion and Management*, vol. 75, pp. 701-709, 2013.

[12] A.G. Madureira and 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.

[13] H. Kanchev, D. Lu, F. Colas, V. Lazarov, and B. Francois, "Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications," *IEEE Trans. on Ind. Elect.*, vol. 58, no. 10, pp. 4583-4592, Oct. 2011.

[14] P. Bühlmann, "Boosting for high-dimensional linear models," *Ann. of Stat.*, vol. 34, no. 2, pp. 559-583, 2006.

[15] P. Godinho Matos, P. Daniel, A. Veiga, A. Messias, M. Oliveira, P. Monteiro, "InovGrid, a smart vision for a next generation distribution system," in *Proceedings of the 22nd International Conference on Electricity Distribution (CIRED)*, Stockholm, Sweden, 10-13 June 2013.

[16] 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.

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

[18] M. Zdrallek, C. Oerter, H. Brunner, R. Calone, *et al.*, "Smart grids on the distribution level – Hype or vision? CIRED's point of view," Technical Report, CIRED Working Group on Smart Grids, March 2013.

[19] S. Renner, M. Albu, H. Elburg, C. Heinemann, *et al.*, "European smart metering landscape report," Deliverable D-2.1, EU Project SmartRegions, Feb. 2011.

[20] T. Hothorn, P. Buehlmann, T. Kneib, M. Schmid, and B. Hofner, *mboost: Model-Based Boosting*, R package version 2.3-0, 2014, http://CRAN.R-project.org/package=mboost

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

[22] P. Bacher, H. Madsen, and H.A. Nielsen, "Online short-term solar power forecasting", *Sol. Energ.*, vol. 83, no. 10, pp. 1772-1783, 2009.

[23] R. Koenker and G. Bassett, "Regression quantiles", *Econometrica*, vol. 46, pp. 33-50, 1978.

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

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