Downloaded from orbit.dtu.dk on: Nov 07, 2026

DTU Library

# Very Short-term Nonparametric Probabilistic Forecasting of Renewable Energy Generation - with Application to Solar Energy

**Golestaneh, Faranak; Pinson, Pierre; Gooi , Hoay Beng**

*Published in:*
IEEE Transactions on Power Systems

*Link to article, DOI:*
10.1109/TPWRS.2015.2502423

*Publication date:*
2016

*Document Version*
Peer reviewed version

Link back to DTU Orbit

*Citation (APA):*
Golestaneh, F., Pinson, P., & Gooi , H. B. (2016). Very Short-term Nonparametric Probabilistic Forecasting of Renewable Energy Generation - with Application to Solar Energy. *IEEE Transactions on Power Systems*, *31*(5), 3850-3863. https://doi.org/10.1109/TPWRS.2015.2502423

**General rights**
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Very Short-term Nonparametric Probabilistic Forecasting of Renewable Energy Generation - with Application to Solar Energy

Faranak Golestaneh, *Student Member, IEEE,* Pierre Pinson, *Senior Member, IEEE,* and Hoay Beng Gooi, *Senior Member, IEEE*

***Abstract*—Due to the inherent uncertainty involved in renewable energy forecasting, uncertainty quantification is a key input to maintain acceptable levels of reliability and profitability in power system operation. A proposal is formulated and evaluated here for the case of solar power generation, when only power and meteorological measurements are available, without sky-imaging and information about cloud passages. Our empirical investigation reveals that the distribution of forecast errors do not follow any of the common parametric densities. This therefore motivates the proposal of a nonparametric approach to generate very short-term predictive densities, i.e., for lead times between a few minutes to one hour ahead, with fast frequency updates. We rely on an Extreme Learning Machine (ELM) as a fast regression model, trained in varied ways to obtain both point and quantile forecasts of solar power generation. Four probabilistic methods are implemented as benchmarks. Rival approaches are evaluated based on a number of test cases for two solar power generation sites in different climatic regions, allowing us to show that our approach results in generation of skilful and reliable probabilistic forecasts in a computationally efficient manner.**

***Index Terms*—Solar power, forecasting, uncertainty quantification, quantile regression, Extreme Machine Learning.**

## I. Introduction

FORECASTING is of utmost importance in both integration and trading of intermittent renewable energy generation. The variability and limited predictability of intermittent generation pose new challenges in power system operations and electricity markets, such as the need for additional ancillary services [1], as well as an increase in voltage and frequency fluctuations [2], among others.

Improving photovoltaic power (abbreviated PV) forecasting methods, as an essential component to mitigating PV power intermittency, has attracted increased attention over the last few years. The online short-term PV forecasting using a clear sky model described in [3] comprises a notable example of a relevant approach.

Perez *et. al.* [4] have investigated the correlation between irradiance variability in neighboring sites as a function of their distances. The same authors have looked into the predictability of sub-hourly variability of solar resources using hourly satellite-derived insolation data in [5]. In [6], it has been shown that by considering multiple PV fleets comparing to a single one, the total relative power output variability drops quasi-exponentially as a function of dispersion factor of the PV fleets and can be further decreased if they are optimally-spaced.

Intra-hour solar irradiance forecasting is performed by predicting cloud motions and locations in [7]. Cloud locations are predicted for 5 minutes ahead and simulations showed that for coastal areas, cloud forecast error increases as the forecast horizon increases. A neural network-based reforecasting method is developed in [8] to improve forecast accuracy of (1) a physical model based on cloud tracking, (2) auto regressive moving average and (3) $k$-th nearest neighbor methods for forecast horizons of 15 minutes and less. A recent review of the state-of-the-art can be found in [9].

Nearly all previous works focused on single-valued (or point) forecasts, similar to the case of wind power forecasting [10]. However, point forecasts can only be helpful in cases where the loss function of decision-makers is truly quadratic, or when no significant uncertainty is involved, since it fails to dispense a full picture of all potential future outcomes [11]. On the other hand, this is the aim of probabilistic predictions to provide decision-making under uncertainty with the full information.

Indeed in recent years, there has been a surge of interest in stochastic optimization approaches to cover different uncertainties in power systems and electricity markets. Various problems have been analysed, e.g., stochastic security-constrained unit commitment [12], probabilistic optimal power flow [13] and market participation for renewables [14]. For a recent overview on uncertainty quantification and its integration in various decision-making problems related to electricity markets, the reader is referred to [15]. It is often assumed that the random variables involved have known parametric distributions. However, even in cases where observations form a known and well-behaved marginal distribution, there is no guarantee that conditional predictive densities (or distributions of forecast errors) follow that same distribution. Wrong distributional assumptions may directly yield biases in analyses and results. Stochastic optimization therefore calls for a thorough design and evaluation of probabilistic forecasting approaches.

As a relevant parallel, there has been comprehensive work carried out for the case of probabilistic forecasting of wind power generation. In [16], a hybrid intelligent algorithm (named HIA) is proposed to find nonparametric prediction intervals for wind generation. However, the evaluation criteria and objective function proposed by this paper verify only the coverage and the width of the prediction intervals. The quality of quantiles which is essential for any reliable probabilistic forecast is ignored. An Extreme Learning Machine (ELM)-Bootstrap (named BELM) method is developed in [17] to estimate the statistical properties of the wind generation and predict intervals for three forecast horizons. This method stands on the normality assumption for distribution of forecast errors. In applications where the distribution of forecast errors does not have a normal shape, its accuracy should be verified.

In contrast to this vast amount of work on probabilistic wind power forecasting, one can hardly find proposals and

empirical investigations for the solar energy case. Lorenz *et. al.* [18] proposed a method to obtain 95% prediction interval for global horizontal irradiance (GHI) based on standard error as a function of solar zenith angle and clear sky index. Two assumptions are made in their work, standard error follows normal distributions and the prediction interval is central on point forecasts. In [19], a hybrid method is developed to provide prediction intervals for one-minute average direct normal irradiance for three intra-hour horizons. Prediction intervals with three coverage rates are obtained and compared with those given by bootstrap-ANN and persistence methods. A heuristic formulation is used to evaluate and compare prediction intervals. Moreover, prediction intervals are obtained independent of central quantiles and the quality of the quantiles is not checked and concerned. A proposal for a geostrophic wind dependent probabilistic solar irradiance forecasts can be found in [20]. By post-processing NWP and comparing the North American Mesoscale model with the California Irrigation Management Information System observations and SolarAnywhere data, the model generates intervals for global horizontal solar irradiance. In the mentioned paper, only 95% prediction interval is sought and is compared with few methods including the method proposed in [18] to check if it can provide narrower intervals. Probabilistic skill scores are not calculated and the evaluation of intervals includes width and reliability metrics separately where the width of the intervals is compared visually.

In spite of basic similarities in wind and PV power forecasts, there are substantial differences in the stochastic process dynamics, the influential variables, as well as in the relationship between meteorological variables and the resulting PV power generation. According to the empirical investigation conducted here, where cloud passages forecast is not available (a common situation), forecast errors do not obey any known parametric distributions. Hence, in order to avoid restrictive assumptions on the shape of predictive distributions, we propose a framework to generate nonparametric density forecasts. These are assembled by predicting their quantiles for various nominal levels, in a quantile regression framework. The ELM [21] is used as a regression solver while a gradient-free optimization algorithm serves for parameter estimation, based on a suitable objective function that accounts for feasibility, reliability and, sharpness of predictive densities.

ELM is a recently proposed learning algorithm and is selected due to its good generalization performance with very fast learning mechanism, which makes it thousands of times faster than the traditional learning algorithms such as back-propagation. We compared its performance with that of more classical linear regression models, AutoRegressive (AR) with the same explanatory variables. It was found that the ELM led to the best accuracy with satisfactory computational costs, therefore motivating us to focus on using it in PV generation forecasts. For more information about the performance of the ELM with respect to its competitors, the reader is referred to [21].

## II. Experimental framework

As a basis for our empirical investigation, the time series with solar power generation at two sites with different climatic regions are used. Their temporal resolution is of one minute. Site $A$ is the St Lucia PV site at the University of Queensland (433kWp), with latitude and longitude of 27.498°S and 153.013°E respectively and 28m elevation above the sea level. To have a consistent forecasting duration for all days throughout a year, daytime data from 5:00 am to 7:00 pm is considered. Site $B$ is a 1kWp PV system at Nanyang Technological University (NTU), Singapore, with latitude and longitude of 1.345°N and 103.681°E respectively and 41m elevation above the sea level. NTU is located in a tropical region with mostly cloudy days all year round. The diurnal time for this case is assumed to be 7:00 am to 6:50 pm. Therefore, each day contains 710 data points. For both sites, the data from 2012 is used for training, while that for 2013 is used for genuine forecast evaluation. Fig. 1 illustrates PV generation for a typical day at Site $A$, showing its high variability and consequent forecast uncertainty. At such small temporal resolution and lead times, cloud passages are the main reason for PV power fluctuations, commonly causing 60% drop in the installed capacity [22].

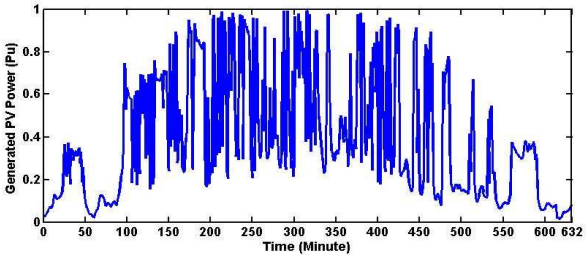


Fig. 1. PV power variability for a typical day at Site $A$ in March 2013

The lack of inertia in PV systems, high ramp rates due to cloud passages, and switching in power generation dynamics pose challenges at different time scales. Very short-term fluctuations in PV output have an impact on power quality (e.g., voltage flicker), generation-load imbalance, and regulation cost. Since increasing regulation reserve to compensate PV fluctuations is a costly option, using high-quality forecasts to preventively deal with potential imbalances is indispensable [2], [22]. For instance, a part of imbalances can be adjusted through re-dispatch rather than limited and expensive frequency regulation control. This motivates us to focus on very short-term and high frequency forecasting, in a probabilistic forecasting framework. We refer to "very short-term" as lead times up to one hour ahead. In general, a more accurate forecast can bring several benefits to the power systems. Better prediction of solar power generation will help system operators to better manage the dispatchable generators and the generation-load balance such that system production cost is minimized to the best extent possible. Moreover, accurate forecasts help firm up the frequency regulation and reserve markets by reducing the uncertainty associated with variable generation resources. This will avoid paying for extra generation in the frequency regulation market that will not be needed with more accurate forecast from solar units.

## III. Non-Parametric Probabilistic Forecasts

Let $Y_t$ be the random variable for PV power generation at time $t$. $f_t$ and $F_t$ are its probability density function (PDF) and cumulative distribution function (CDF). The quantile $q_t^{\alpha_i}$ with level $\alpha_i$ can be defined as

$$q_t^{\alpha_i} = F_t^{-1}(\alpha_i) \tag{1}$$

A quantile forecast $\hat{q}^{\alpha_i}_{t+k|t}$ is consequently the estimate of $q^{\alpha_i}_{t+k}$ issued at time $t$ for lead time $t+k$ based on information available at that time. A single quantile forecast provides only limited information on forecast uncertainties. This motivates the joint consideration of $m$ number of quantile predictions, with increasing nominal levels $\alpha_i$, $i = 1, \ldots, m$, then yielding a nonparametric density forecast $\hat{f}_{t+k|t}$, i.e.,

$$\hat{f}_{t+k|t} = \{\hat{q}^{\alpha_i}_{t+k|t}; 0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_m \leq 1\} \quad (2)$$

with the full continuous density obtained by linear or nonlinear interpolation through these defined quantile forecasts.

Once quantile forecasts and predictive densities are generated, prediction intervals are a natural by-product allowing easing visualization of forecast uncertainties. This is because they represent a range of potential values in which the future observation $y_{t+k}$ may lie within a certain probability. Considering two quantiles forecasts whose nominal proportions are $\overline{\alpha}$ and $\underline{\alpha}$, central prediction intervals with nominal coverage $(1-\gamma)$ are defined by

$$\hat{I}^{1-\gamma}_{t+k|t} = \left[\hat{q}^{\overline{\alpha}}_{t+k|t}, \hat{q}^{\underline{\alpha}}_{t+k|t}\right] \quad (3)$$

where nominal levels of the quantile forecasts are centered around the median, i.e.

$$\overline{\alpha} - \underline{\alpha} = 1 - \gamma \quad (4)$$

The central prediction interval will serve as a basis for the illustration of our nonparametric probabilistic forecasts of solar power generation.

### A. Performance Evaluation

*1) Reliability:* High reliability is one of the major requirements for probabilistic forecasts, as lack of reliability will most likely result in a systematic bias in decision-making. Since $\hat{f}_{t+k|t}$ is composed of a set of quantile forecasts, the evaluation of forecast reliability also focuses on these quantiles. As a basis, let $\xi^{(.)}_{t,k}$ be an indicator variable

$$\xi^{\alpha_i}_{t,k} = \mathbf{1}\{y_{t+k|t} \leq \hat{q}^{\alpha_i}_{t+k|t}\} = \begin{cases} 1, & \text{if } y_{t+k|t} \leq \hat{q}^{\alpha_i}_{t+k|t} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

The empirical level $\hat{\alpha}_i$ of a quantile forecast with nominal proportion $\alpha_i$ is then

$$\hat{\alpha}_i = \frac{1}{N}\sum_{t=1}^{N} \xi^{\alpha_i}_{t,k} \quad (6)$$

Reliability is eventually assessed visually through the use of reliability diagrams, showing empirical vs. nominal levels of the defining quantile forecasts, $i = 1, \ldots, m$ [23]. Besides, the absolute deviation (denoted $Dev$) from the nominal proportion is a factor to assess overall reliability: quantile forecasts with a lower $Dev$ are to be seen as more reliable. The metric $Dev$ can be formulated as

$$Dev_i = |\alpha_i - \hat{\alpha}_i| \quad (7)$$

*2) Skill Score:* As for point prediction, forecast users are keen on using a unique score which is able to summarize the performance of rival probabilistic forecasting while accounting for their reliability and sharpness. Reliability evaluation alone does not allow us to do so, since probabilistic forecasts with very bad skill, can easily be made probabilistically reliable.

At time $t+k$, for a single forecast-observation pair, the score

$$Sc(\hat{f}_{t+k|t}, y_{t+k|t}) = -\sum_{i=1}^{m}\left(\xi^{\alpha_i}_{t,k} - \alpha_i\right)\left(y_{t+k|t} - \hat{q}^{\alpha_i}_{t+k|t}\right) \quad (8)$$

is to be seen as a proper skill score [24], hence allowing to objectively sort rival forecasting approaches. However, both reliability and sharpness should also be examined separately and visually via reliability diagrams and quantiles' plots to make sure there is a reasonable balance between these two attributes. Skill score values for a given lead time are obtained by averaging the score in the above equation over all forecast-observation pairs. Score values are positive, with lower score values meaning higher skill.

Probabilistic forecasts are a crucial input for decision-making problems which involve uncertainty. In such problems, traditionally point forecasts have been deployed. More recently in stochastic optimization, several assumptions are made to build a distribution for uncertain variables. For example, in interval optimization, intervals of the random variables are mostly assumed to be a certain percentages lower or upper the point predictions. In Monte Carlo or scenario based methods, the scenarios are extracted from an assumed distribution, usually normal. Therefore, the accuracy of the probabilistic approaches is highly dependent on how much those assumptions match the reality. In this proposal, instead of making assumptions to define the uncertainty of a random variable, a forecast density is predicted for each time point in the future. This can help the decision maker to get a better insight of the uncertainties involved in prediction and make more pragmatic decisions. In [25], as an example to demonstrate the benefits and necessities of the probabilistic forecasts, a bidding strategy in a pool based electricity market is investigated and mathematically it has been proved that the optimal bidding amount is equal to a special quantile produced by wind power probabilistic forecasts.

## IV. From Point Predictions to Nonparametric Probabilistic Forecasts

The ELM is employed as a regression model which can allow to generate both point and quantile forecasts, depending on the training criterion. For point predictions, the goal is to minimize the error between predictions and measurements. This can be translated to a least-square solution for parameters of the regression model. However, in probabilistic forecasting, the targets of the model are quantiles which should cover a certain proportion of the measurements over a defined time. In this case, a pure regression framework which tries to mimic the measurements based on least-square methods cannot be employed. Therefore, new criteria and formulation should be devised to train the model.

After introducing the overall ELM approach, the procedure for designing ELMs to directly generate predictive densities is explained in the following sections.

### A. Extreme Learning Machine (ELM)

For an ELM, parameters for the hidden layer are randomly selected while they do not need to be tuned through the training process. The output weights are analytically calculated at once without going through a lengthy iterative training [26].

Due to high scalability and very low computational burden, ELM not only unifies different widely used learning methods but also offers a unified approach for different practical applications.

Given a dataset of $N$ samples $\{(\mathrm{x}_j, \mathbf{g}_j)\}_{j=1}^{N}$ with $\mathrm{x}_j \in R_n$ as input and $\mathbf{g}_j \in R_m$ as output, for a single-hidden layer feed-forward neural network (SLFN) with $\tilde{N}$ hidden neurons and $\varphi(.)$ as an activation function, the relationship between inputs and outputs is defined by

$$\sum_{i=1}^{\tilde{N}} \beta_i \varphi\left(\mathrm{a}_i.\mathrm{x}_j + b_i\right) = \mathbf{g}_j, \quad j = 1, ..., N \tag{9}$$

where $\mathrm{a}_i = [a_{i1}, a_{i2}, ..., a_{in}]^\top$ is the weight vector which connects input neurons to the neurons in the hidden layer; $\mathrm{a}_i.\mathrm{x}_j$ represents the inner product of $\mathrm{a}_i$ and $\mathrm{x}_j$ and $\beta_i = [\beta_{i1}, \beta_{i2}, ..., \beta_{im}]^\top$ denotes the vector of weights of output nodes. The set of equations in (9) can be compacted as

$$\mathbf{H}\beta = \mathbf{G} \tag{10}$$

where

$$\mathbf{H}(\mathrm{a}_1, \ldots, \mathrm{a}_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}}, \mathrm{x}_1, \ldots, \mathrm{x}_N) = \begin{bmatrix} \varphi\left(\mathrm{a}_1.\mathrm{x}_1 + b_1\right) & \cdots & \varphi\left(\mathrm{a}_{\tilde{N}}.\mathrm{x}_1 + b_{\tilde{N}}\right) \\ \vdots & \ldots & \vdots \\ \varphi\left(\mathrm{a}_1.\mathrm{x}_N + b_1\right) & \cdots & \varphi\left(\mathrm{a}_{\tilde{N}}.\mathrm{x}_N + b_{\tilde{N}}\right) \end{bmatrix}_{N \times \tilde{N}} \tag{11}$$

$$\beta = \begin{bmatrix} \beta_1^\top \\ \vdots \\ \beta_{\tilde{N}}^\top \end{bmatrix}_{\tilde{N} \times m} \qquad \mathbf{G} = \begin{bmatrix} \mathbf{g}_1^\top \\ \vdots \\ \mathbf{g}_N^\top \end{bmatrix}_{N \times m} \tag{12}$$

If $N = \tilde{N}$ and the activation function is infinitely differentiable, ELM can train the SLFN with exactly zero error [27]. Once the parameters of the hidden layer are randomly assigned, the hidden layer matrix can be uniquely determined. Following (10), $\beta$ can be calculated by

$$\beta = \mathbf{H}^\dagger \mathbf{G} \tag{13}$$

where $\mathbf{H}^\dagger$ is the Moore-Penrose generalized inverse of $\mathbf{H}$ and it is generally derived by singular value decomposition [27].

### *B. Optimization Formulation*

The optimization methodology proposed in this paper aims at adjusting the ELM parameters in order to make them optimal in view of the skill score given in Section III-A, for the evaluation of predictive densities while satisfying the problem's constraints. The predicted quantiles should be non-crossing and lie in a feasible range of the PV generation. The proposed objective function is

$$\min F = \frac{-1}{N_T m} \sum_{t=1}^{N_T} \sum_{i=1}^{m} \Big( \gamma_1 \times \Big( (\xi_{t,k}^{\alpha_i} - \alpha_i)(y_{t+k|t} - \hat{q}_{t,k}^{\alpha_i}) \Big) + \gamma_2 \times \xi_{t,k}^{'\alpha_i} + \gamma_3 \times \xi_{t,k}^{''\alpha_i} \Big) \tag{14}$$

where $N_T$ is the number of samples over the training period and $m$ is the number of quantiles. $\gamma_1$ is a coefficient which represents the importance of the skill score in the objective function. $\gamma_2$ and $\gamma_3$ are penalty factors for violations from the problem constraints i.e. contravening the feasible range and crossing quantiles respectively. $\xi_{t,k}^{'\alpha_i}$ is an indicator variable which examines if a solution has violated from the feasible range $0 \leq [G]_{N \times m} \leq 1$ and is defined by

$$\xi_{t,k}^{'\alpha_i} = \begin{cases} 1, & \text{if } \left(\hat{q}_{t,k}^{\alpha_i} > 1\right) \text{ or } \left(\hat{q}_{t,k}^{\alpha_i} < 0\right) \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

To ensure that a quantile with a higher ratio encapsulates the one with a lower ratio, $\xi_{t,k}^{''\alpha_i}$ as defined in (16) is added to the objective function which checks if a violation has happened.

$$\xi_{t,k}^{''\alpha_i} = \begin{cases} 1, & \text{if } (\alpha_i > \alpha_m) \text{ and } \left(\hat{q}_{t,k}^{\alpha_i} < \hat{q}_{t,k}^{\alpha_m}\right) \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

In order to keep the values of three terms of the objective function in a roughly similar range, firstly $\gamma_1$, $\gamma_2$ and $\gamma_3$ are set to 1 and the problem is solved. Secondly, the calculated values for $c_2 = \frac{-1}{N_T m} \sum_{t=1}^{N_T} \sum_{i=1}^{m} \xi_{t,k}^{'\alpha_i}$ and $c_3 = \frac{-1}{N_T m} \sum_{t=1}^{N_T} \sum_{i=1}^{m} \xi_{t,k}^{''\alpha_i}$ are compared with those for $c_1 = \frac{-1}{N_T m} \sum_{t=1}^{N_T} \sum_{i=1}^{m} (\xi_{t,k}^{\alpha_i} - \alpha_i)(y_{t+k|t} - \hat{q}_{t,k}^{\alpha_i})$. Then, $\gamma_1$ is kept equal to 1 and $\gamma_j (j = 2, 3)$ is set to $1/a_j$ where $a_j = c_j/c_1 (j = 2, 3)$.

An optimization algorithm is developed and customized to optimize the proposed objective function. In the modelled optimization algorithm, the optimization procedure is performed through two updating methods. The first one is inspired by the Particle Swarm Optimization (PSO) and is defined by

$$\beta_\chi^1 = \beta_\chi^{'1} + W \times \Omega_\chi + C_1 \times (\Omega_\chi - P_\chi^b) + C_2 \times (\Omega_\chi - P^g) \tag{17}$$

where $\chi$ is the particles index; $P^g$ is the global best solution found by the population and $P_\chi^b$ is the best solution found by the particle $\chi$ so far. $\beta_\chi^{'1}$ is the solution found by particle $\chi$ in the previous iteration and $\beta_\chi^1$ is the new solution for particle $\chi$ following the updating method in (17). $\Omega$ is a repository which keeps a portion of the best solutions found by the whole population using two updating methods (17) and (18). Its size has been considered two times the population size. $W$, $C_1$ and $C_2$ are three random numbers between 0 and 1.

The second updating formula tries to increase diversity among solutions and is defined by

$$\beta_\chi^2 = \Omega_{n_1} + r_1 \times (\Omega_{n_2} - \Omega_{n_3}) + r_2 \times (\Omega_{n_4} - \Omega_{n_5}) \tag{18}$$

where $\beta_\chi^2$ is the updated solution using (18) and $r_1$ and $r_2$ are two random numbers between 0 and 1. $\Omega_i$ $(i = 1, ..., 5)$ are five randomly selected members from the repository.

For this specific optimization problem, if instead of searching within the repository, one searches the unconstrained search space; as the number of iterations increases, the number of infeasible solutions increases. This happens due to the fact that in this problem, the design variables are the output weight of the ELM ($[\beta]_{\tilde{N} \times m}$) while the constraints are on the output of the ELM ($[G]_{N \times m}$) as the quantiles of PV generations. There is no straightforward method to limit the search space of $[\beta]_{\tilde{N} \times m}$ based on the $[G]$ limitations.

We developed the above optimization algorithm and called it Repository-based PSO or RPSO, because several optimization functions in R programming and Matlab failed to provide satisfactory results for the problem formulated in this paper.

We tried "constrOptim", "optim" and "optimx" and "GenSA" in R and "fmincon", "fminunc" and "simulannealbnd" in Matlab. The mentioned methods resulted in long simulation time, trapping in local optima or infeasible solutions.

### C. Optimization Procedure

The optimization procedure can be described as follows:

*Step 1*) To initialize the optimal forecast density model, a set of training samples is defined as $\{(\mathrm{x}_{t,j}, \mathbf{g}_{t+k|t,j})\}_{j=1}^{N_T}$ where $\mathbf{g}_{t+k|t,j}$ is composed of a set of $m$ quantiles as

$$\mathbf{g}_{t+k|t,j} = \{\hat{q}^{\alpha_1}_{t+k|t,j}, \ldots, \hat{q}^{\alpha_i}_{t+k|t,j}, \ldots, \hat{q}^{\alpha_m}_{t+k|t,j}\} \tag{19}$$

for $j = 1, \ldots, N_T$, and

$$\hat{q}^{\alpha_i}_{t+k|t,j} = (\alpha_i + \mathrm{round}(\alpha_i))y_{t+k} \tag{20}$$

where $\mathrm{round}(\alpha_i)$ is a function which rounds $\alpha_i$ to the nearest integer (0 or 1). Other forms of initialization could be seen as relevant, like using the climatology method.

Initial values of $\beta$ are computed by following (13). The subsequent optimization phase then explores the search space to find better values for $\beta$, i.e., yielding a set of quantile forecasts with a better skill score value.

*Step 2*) Initialize the population by randomly generating $N_P$ particles around the initial $\beta$.

*Step 3*) Once the $N_P$ output weights $\beta$ are positioned using the proposed optimisation approach in subsection IV-B, their corresponding estimated quantiles and fitness values can be calculated according to (10) and (14), respectively.

*Step 4*) The iterative optimization procedure is carried out until the maximum number of iterations is reached or the convergence criterion is met. At each iteration, for each potential solution $\{\beta_p\}_{p=1}^{N_p}$, the outputs of ELM are calculated by (10) and the corresponding objective function is computed using (14).

*Step 5*) The final $\beta$ is retained, and then used to generate predictive densities out of the samples.

## V. Benchmark Methods

### A. HIA

HIA is formulated in [16] to directly find prediction intervals. Prediction intervals with associated confidence levels are estimated through an optimization procedure when the ELM serves as the regression model. The objective function of the optimization approach takes care of coverage and sharpness of intervals only, regardless of the quality of quantile forecasts. In HIA, calibration of intervals is evaluated by

$$A_t^{1-\gamma_i} = \hat{PIC}_t^{1-\gamma_i} - (1-\gamma_i) \tag{21}$$

where $\hat{PIC}_t^{1-\gamma_i}$ denotes the proportion of the observations lying within the interval with nominal coverage $1-\gamma_i$. The objective function is defined as

$$\min F = \sum_{i=1}^{m'} |A_t^{1-\gamma_i}| + |S_t^{1-\gamma_i}| \tag{22}$$

where $S_t^{1-\gamma_i}$ measures the width of the intervals and $m'$ is the number of intervals.

HIA is compared with the climatology, constant, persistence, exponential smoothing method (ESM), and quantile regression (QR) approaches and found to be superior for datasets used in [16].

### B. BELM

A bootstrap-based method is proposed in [17] to predict uncertainties involved in prediction. ELM serves as the regression model and a number of ELMs are trained with resampled training data. Two kinds of uncertainties are modelled, uncertainty related to the noise in data $\sigma_\epsilon^2$ and uncertainty in the ELM model $\sigma_{\hat{y}}^2$. These two uncertainties are assumed to be Gaussian, independent of each other and dependent on the inputs of the regression model. Total variance of prediction for time $t$ and $\mathrm{x}_j$ denoted by $\sigma_t^2(\mathrm{x}_j)$ is calculated by

$$\sigma_t^2(\mathrm{x}_j) = \sigma_{\hat{y}}^2(\mathrm{x}_j) + \sigma_{\hat{\varepsilon}}^2(\mathrm{x}_j) \tag{23}$$

$B_M$ training datasets are generated by resampling with replacement from the original training dataset $\{(\mathrm{x}_j, \mathbf{g}_j)\}_{j=1}^{N}$ and a separate ELM is trained for each of the bootstrapped datasets. With $\hat{y}_l(\mathrm{x}_j)$ as the output of the $l^{th}$ bootstrapped ELM, the mean of the distribution of the predictions is calculated by

$$\hat{y}(\mathrm{x}_j) = \frac{1}{B_M}\sum_{l=1}^{B_M} \hat{y}_l(\mathrm{x}_j) \tag{24}$$

$\sigma_{\hat{y}}^2$ is regarded as the variance of the outputs of $B_M$ bootstrapped ELMs $\{\hat{y}_l(\mathrm{x}_j)\}_{l=1}^{B_M}$.

To estimate uncertainty related to noise of data, by replacing the targets $\{\mathbf{g}_j\}_{j=1}^{N}$ with $\{(\hat{y}(\mathrm{x}_j) - \mathbf{g}_j)\}_{j=1}^{N}$, a transformed dataset is designed as

$$D_\varepsilon = \{\mathrm{x}_j, (\hat{y}(\mathrm{x}_j) - \mathbf{g}_j)\}_{j=1}^{N} \tag{25}$$

Resampling with replacements is carried out to generate $B_N$ number of resampled datasets from $D_\varepsilon$. The output of $l^{th}$ bootstrapped ELM is denoted by $\hat{r}_l(\mathrm{x}_j)$. For the transformed training data, again a share of uncertainty is related to the ELM model. Estimated variance of noise $\hat{\sigma}_{\hat{\varepsilon}}^2(\mathrm{x}_j)$ and variance attached to the ELM model $\sigma_{\hat{r}}^2(\mathrm{x}_j)$ can be calculated by

$$\hat{\sigma}_{\hat{\varepsilon}}^2(\mathrm{x}_j) = \hat{r}(\mathrm{x}_j) = \frac{1}{B_N}\sum_{l=1}^{B_N} \hat{r}_l(\mathrm{x}_j) \tag{26}$$

$$\sigma_{\hat{r}}^2(\mathrm{x}_j) = \frac{1}{1-B_N}\sum_{l=1}^{B_N} (\hat{r}_l(\mathrm{x}_j) - \hat{r}(\mathrm{x}_j)) \tag{27}$$

The total variance of data noise can be obtained by

$$\sigma_{\hat{\varepsilon}}^2(\mathrm{x}_j) = \hat{\sigma}_{\hat{\varepsilon}}^2(\mathrm{x}_j) + \sigma_{\hat{r}}^2(\mathrm{x}_j) \tag{28}$$

Finally, the lower and upper bounds of interval predictions are obtained by

$$\hat{q}^{\overline{\alpha}}_{t+k|t} = \hat{y}(\mathrm{x}_j) + z_{1-\underline{\alpha}/2}\sqrt{\sigma_t^2(\mathrm{x}_j)} \tag{29}$$

$$\hat{q}^{\underline{\alpha}}_{t+k|t} = \hat{y}(\mathrm{x}_j) - z_{1-\underline{\alpha}/2}\sqrt{\sigma_t^2(\mathrm{x}_j)} \tag{30}$$

with $z_{1-\underline{\alpha}/2}$ as the critical value of standard Gaussian distribution dependent on the coverage level of prediction interval $1-\gamma$.

## VI. Empirical Investigation

Our empirical investigation aims at evaluating the feasibility of our approach for the solar power application, while evaluating it against a number of relevant benchmarks, namely the persistence [16], climatology [28], HIA [16] and BELM [17]. Persistence is a benchmark that is difficult to outperform when considering very short lead times, especially if issuing point forecasts only [29]. Climatology is the most widely used benchmark in probabilistic forecasting for weather and weather-related processes. It is not a dynamic forecast since it ignores recent power values. It is based on all available power measurements, though not making any distributional assumption (like Gaussian in the above) for predictive densities. HIA [16] and BELM [17] are also implemented as state-of-the-art methods to verify the comparative performance of the proposed method.

### A. ELM Tuning

For our ELM-based predictive densities, three case studies are considered:

- *Case* 1: Focus is given to Site $A$ and the lead time is of 10 minutes ahead. The predictors for this case are past solar power observations up to lag $n$ time intervals, clear sky solar irradiance [30], [31], lag values for air temperature, wind speed, humidity, and solar insolation. Therefore, in this case the number of predictors is $n+5$. Air temperature, wind speed, humidity are measured by Vaisala WXT520 Weather Transmitter and solar irradiance by Kipp & Zonen CMP11. The optimal value for $n$ is determined using 10-fold cross validation and forecasts frequency and resolution is of one minute.
- *Case* 2: Focus is given to Site $B$, while the model relies on past solar power observations up to lag $n$ time intervals, with the same setup as for *Case* 1.
- *Case* 3: Focus is given to Site $A$ except that the lead time of one hour is considered.

For all forecasting approaches, the relevant training window sizes $n$ for the proposed method, $n_p$ for the persistence and $n_c$ for the climatology are between 10 and 120 time steps (minutes), with 10-step increments. As an outcome of the cross-validation procedure, $n$ in *Cases* 1 and 2 is chosen to be 10 and for *Case* 3, $n = 30$. Besides, $n_c$ is chosen equal to 30, while the best value found for $n_p$ in *Cases* 1 and 2 is 60, and finally 120 for *Case* 3. The population size and maximum number of iterations for the optimization phase are set to 70 and 60 respectively and $\gamma_2$ and $\gamma_3$ are set to one-tenth of $\gamma_1$.

The pattern of PV generation varies significantly daily and seasonally. In order to assess the effectiveness and applicability of the proposed method in an computationally efficient manner and present results rationally, the dataset for each year is divided into quarters or so-called seasons. For *Case* 1 and *Case* 3, the definition of seasons in Southern hemisphere is used and for *Case* 2, the Northern hemisphere seasons are considered. For each season, a separate model is trained by using the archived data of the previous year for the same season. It should be noted that the diurnal hours are as described in Section II and the night-time hours are excluded in the study. For example, in *Case* 1 approximately $3$ (months) $\times$ $30$ (days) $\times$ $840$ (5 am to 7 pm) $=$ $75,600$ time instances are taken into account for each season. All PV generation data is normalized by the nominal capacity of their respective sites.

The simulation results for two forecast horizons, 10-minute (*Cases* 1 and 2) and 1-hour (*Case* 3) lead-times are analysed. The 10-minute look-ahead time is one of the most important forecast horizons due to its significant influence on imbalances in power systems. Moreover, the 1-hour lead time forecast has a notable role to play in several operation problems and PV power trading. We have performed one-minute forecast to ensure the applicability and effectiveness of the method for applications which require forecasts with high frequency and time resolution. System operators can deploy the proposed methodologies with their most preferred forecast frequency for their specific applications.

### B. Analysis of Forecast Error Distributions

To obtain the empirical PDF of the forecast errors, they are considered conditional on their corresponding predicted values. In principle, like for wind power, the conditional mean (i.e., the point forecasts) is expected to be the most important variable conditioning the prediction intervals or conditional PDFs [32].

The deterministic ELM trained with quadratic criterion is used to generate forecast time series. The data is paired as [forecast, error] where errors $\varepsilon_{t+k|t}$ for a given time and lead time are calculated as

$$\varepsilon_{t+k|t} = y_{t+k} - \hat{y}_{t+k|t} \tag{31}$$

with $\hat{y}_{t+k|t}$ the PV power forecast for lead time $t+k$ and $y_{t+k}$ the corresponding observation.

The pairs are then sorted by their predicted values and divided into 20 bins with 0.05 pu width each. Figs. 2 and 3 show kernel density estimates of forecast errors in all 20 bins for the predicted power values in *Case* 1 and *Case* 2 respectively. It is clear that PV forecast errors cannot be assumed to be *Gaussian*, *Beta* or following common parametric distributions. The same experiment was conducted for forecast errors for the persistence benchmark, yielding similar conclusions. Without using additional explanatory variables, e.g., from sky imaging, it is unlikely that these distributions could look much different.

Fig. 4 illustrates the overall kernel density of bins altogether versus fitted *Normal* distribution to the same data in *Case* 1 and *Case* 2 respectively. In the same figure, empirical CDF as well as fitted *Normal* CDF to the all forecasted errors are shown. From the figure, it is obvious that the overall distribution of the errors is not *Normal*. To support this argument, One-sample Kolmogorov-Smirnov (KS) test is conducted. Null hypothesis is defined as "the forecasted errors come from a *Normal* distribution" where mean and standard deviation of that distribution are calculated by fitting a normal distribution to the data. For both *Case* 1 and *Case* 2, KS test rejects the null hypothesis at the 5% significance level. The test repeated for *t-location-scale* and *Logistic* distributions and again KS test rejects the null hypothesis. *Rayleigh*, *Beta*, *Gamma* and *Weibull* distributions are ignored as they support only positive values. KS test also is conducted for each of the 40 bins illustrated in Figs. 2 and 3 with null hypotheses explained above. In all cases, the hypothesis is rejected.

The solar power generation process is nonlinear and bounded between zero and nominal capacity. Furthermore,

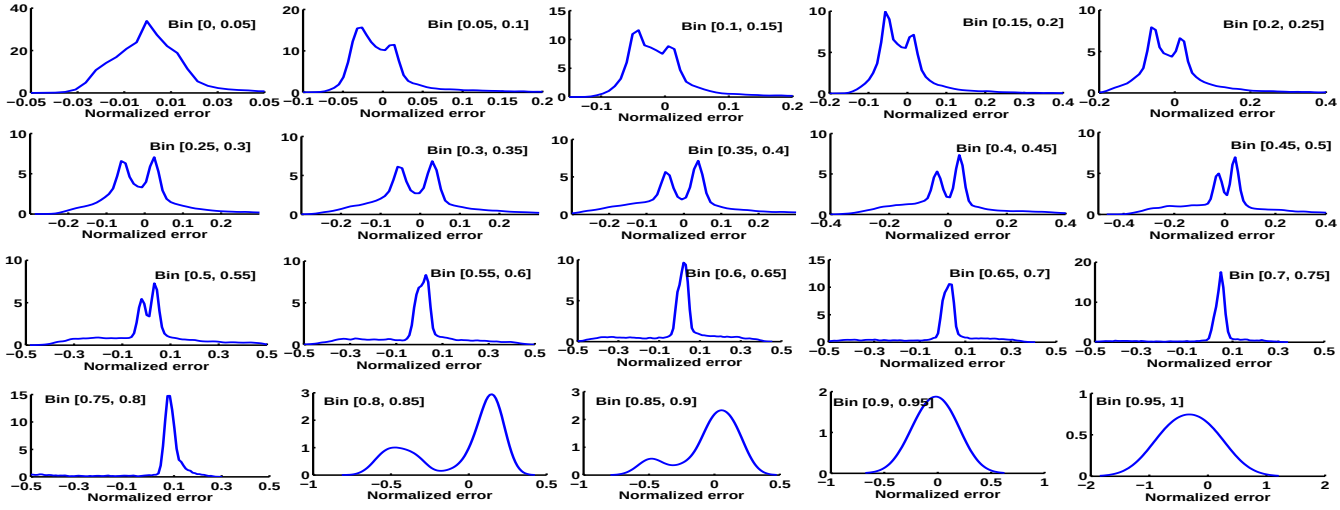


Fig. 2. Kernel smoothing density estimate of distribution of errors in forecasted values, each bin defines a range for point predictions of solar power values, for which all forecast errors are collected to illustrate their conditional empirical distributions (*Case* 1)

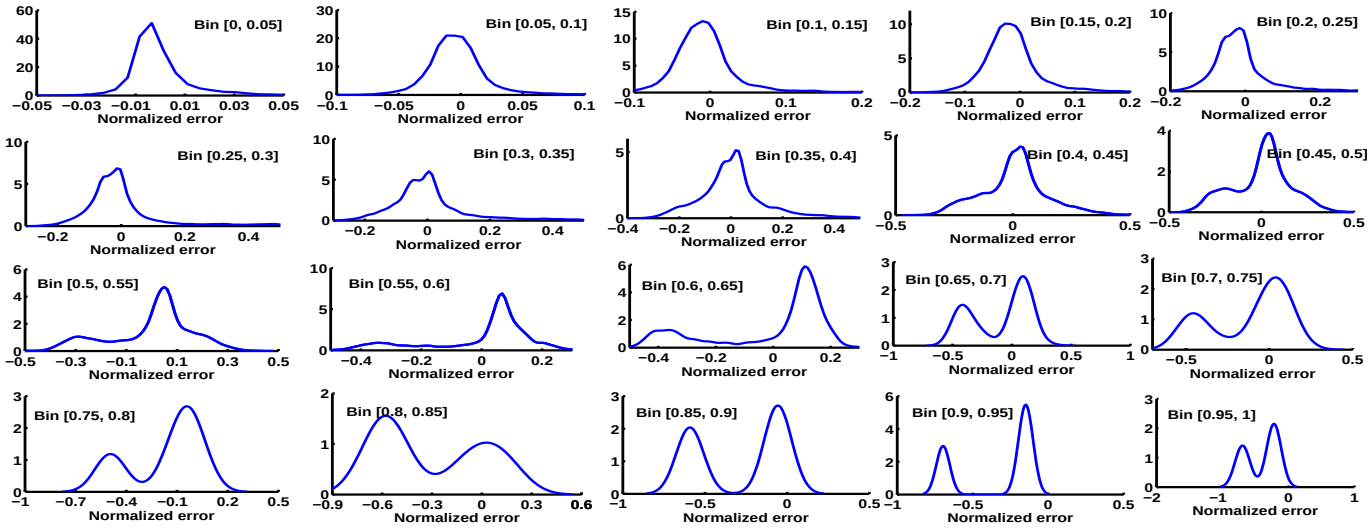


Fig. 3. Kernel smoothing density estimate of distribution of errors in forecasted values, each bin defines a range for point predictions of solar power values, for which all forecast errors are collected to illustrate their conditional empirical distributions (*Case* 2)

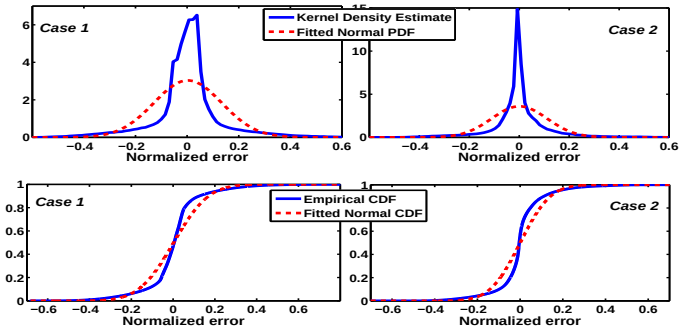


Fig. 4. Top: Kernel smoothing density estimate of distribution of errors in all bins and fitted *Normal* distribution for 220,000 and 200,000 forecasted values in *Case* 1 and *Case* 2 respectively. Bottom: Empirical CDF as well as fitted *Normal* CDF to the same data

its upper bound is actually given by the clear sky model for each and every time (for instance at night, the upper bound is zero). Similar to the case of wind power, this very nature of the process leads to the observed characteristics of the PDFs. When predicting power close to the upper bound, the potential realization can only get lower (hence a long lower tail), and when predicting close to zero, the power output can only get higher (hence a long upper tail). Besides, the bimodality is due to cloud passages, that can be regularly missed by the forecasting approach. When predicting, the method could inform that there are two "most probable states" for future power generation: one if no cloud is coming, and another one if cloud is coming. The presence of this bimodality depends on the local climate, and how much cloud passages may influence short-term variations in power production.

One could possibly envisage proposing mixtures of parametric distributions in the future, based on a hidden regime sequences. Here, emphasis is placed on obtaining highly-skilled nonparametric predictive distributions that accommodate these peculiar characteristics.

### C. Nonparametric Quantile Forecast Results

To describe predictive densities, 18 quantiles with nominal levels from 0.05 to 0.95, at increments of 0.05, except for the median (0.5), are issued. These quantiles form a set of 9 central prediction intervals, with nominal coverage rates from 10 to

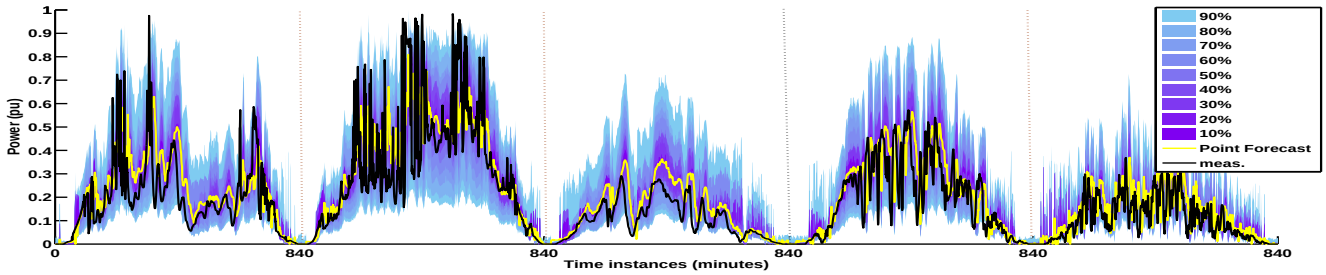


Fig. 5. Predictive densities as a set of prediction intervals for five successive days in January 2013 (*Case* 1)

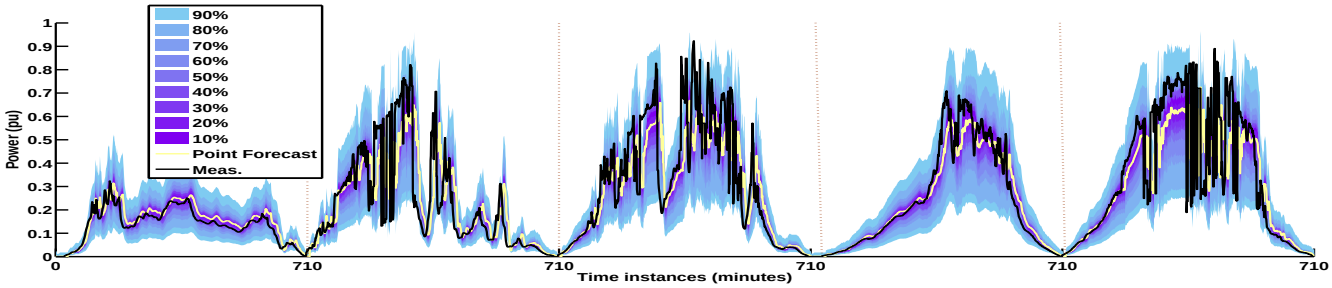


Fig. 6. Predictive densities as a set of prediction intervals for five successive days in August 2013 (*Case* 2)

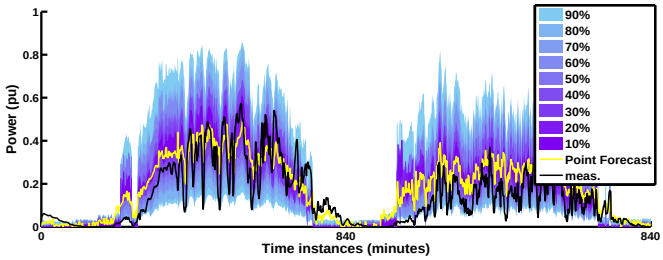


Fig. 7. Predictive densities as a set of prediction intervals for two successive days in January 2013 (*Case* 3)

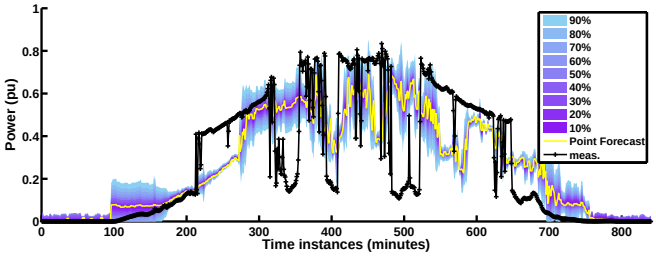


Fig. 8. BELM predictive densities as a set of prediction intervals for one day in March 2013 (*Case* 3)

90%. All quantiles in this study are censored such that PV generation is bounded between zero and the nominal capacity of PV at the respective sites.

It is to be noted that the most important attributes of good probabilistic forecasts are reliability and sharpness. The reliability diagram gives information on how well the predicted probabilities of an event correspond to their observed frequencies. However, the probabilistic forecasts models cannot be judged only on their reliability. Considering both sharpness and reliability are required to benchmark methods. Those can be offered by the discussed skill score in subsection III-A2 which examines both mentioned attributes. Both reliability and sharpness should also be examined separately and visually via reliability and sharpness diagrams as well as the quantiles' plots. This helps to check if there is a reasonable balance between these two attributes. Some methods tend to provide too sharp quantiles that may result in good skill scores but if the reliability is too low, the efficiency is questionable.

Figs. 5, 6 and 7 depict example episodes with probabilistic forecasts generated for *Cases* 1, 2 and 3, respectively. They cover five successive days in *Cases* 1, 2 and two days (the same last two days shown in Fig. 5) in *Cases* 3 to provide a view of the high diurnal diversity in PV generation patterns. These plots illustrate how the proposed forecasting approach reacts to the high intermittency and variability in PV generation for different days. At first sight, the coverage of the prediction intervals appears reasonable. Besides, these intervals are not symmetric around the point forecasts. This is to be expected since point predictions correspond to the mean value of the distributions whereas prediction intervals are centered on the median.

Scores, maximum absolute deviation from nominal proportion, and the sum of the absolute deviations for all 18 quantiles are given in Tables I, II and III for *Cases* 1, 2 and 3 respectively. From the tables, the persistence and the BELM quantiles show acceptable performance in terms of skill score but they have very low reliability. The maximum deviations from perfect reliability using the BELM can be as high as 20%. This undermines its efficiency and reliability in practical applications. On the other hand, climatology provides the worst overall performance. This is of course reasonable since climatology is an unconditional forecast and ignores the heteroscedasticity of PV generation series. The HIA method also presents poor performance in terms of both reliability and sharpness. The poorly calibrated quantiles given by the HIA are the consequence of the evaluation criteria defined for this method in which only the calibration of intervals is taken into account and the quality of quantiles is neglected. The BELM method tends to provide very narrow and sharp quantiles which result in good skill scores. However, there should be

TABLE I
SCORES, MAXIMUM AND SUM OF COVERAGE DEVIATION OF ESTIMATED QUANTILES GIVEN BY DIFFERENT METHODS IN PERCENT FORM (%), $Case$ 1

| $Case$ 1 | Proposed Method | | | Climatology | | | Persistence | | | HIA | | | BELM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** |
| **Spring** | 2.92 | **4.52** | **28.71** | 4.63 | 10.47 | 75.87 | 2.96 | 6.77 | 57.07 | 4.51 | 10.23 | 82.41 | **2.34** | 19.88 | 127.55 |
| **Summer** | **3.30** | 5.99 | **35.81** | 5 | 4.82 | 36.8 | 3.63 | 8.8 | 53.71 | 4.21 | 8.43 | 61.21 | 3.31 | 6.21 | 61.51 |
| **Fall** | 3.12 | **3.74** | **26.94** | 3.79 | 11.08 | 93.34 | 3.67 | 9.92 | 79.48 | 4.35 | 9.54 | 67.54 | **2.78** | 13.77 | 111.46 |
| **Winter** | 2.10 | **3.44** | **30.98** | 2.67 | 9.26 | 95.01 | 2.23 | 8.38 | 49.81 | 2.91 | 11.21 | 75.43 | **1.53** | 19.81 | 182.82 |

TABLE II
SCORES, MAXIMUM AND SUM OF COVERAGE DEVIATION OF QUANTILES GIVEN BY DIFFERENT METHODS IN PERCENT FORM (%), $Case$ 2

| $Case$ 2 | Proposed Method | | | Climatology | | | Persistence | | | HIA | | | BELM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** | **Score** | **Max** | **Sum** |
| **Spring** | **2.39** | **3.71** | **14.22** | 4.7 | 4.62 | 36.49 | 2.66 | 10.05 | 85.79 | 2.97 | 9.18 | 82.67 | 2.98 | 13.81 | 152.23 |
| **Summer** | **2.41** | **3.80** | **22.10** | 4.6 | 6.60 | 50.93 | 2.6 | 10.23 | 66.98 | 3.45 | 7.52 | 63.45 | 2.62 | 8.93 | 79.93 |
| **Fall** | 2.78 | 8.00 | **56.93** | 4.9 | **4.77** | 57.08 | 3.13 | 10.17 | 66.13 | 3.21 | 6.33 | 69.52 | **2.53** | 9.71 | 88.44 |
| **Winter** | 2.51 | **5.36** | **41.03** | 4.68 | 13.76 | 167.6 | 3.23 | 9.87 | 82.67 | 3.15 | 8.42 | 78.43 | **2.48** | 13.03 | 113.68 |

a reasonable balance between sharpness and calibration. By looking at large deviations from the nominal proportions of the BELM quantiles, it can be concluded that the BELM gives more weight to sharpness which leads to unreliable quantiles. Overly narrow quantiles works in case of clear sky when no significant uncertainty is present and PV generation is predictable with good accuracy. However, they fail to cover most of the ramp ups and downs in cloudy days. This can be checked in Fig. 8 where BELM quantiles for the first day depicted in Fig. 7 are illustrated. The rational behind too narrow prediction intervals given by BELM is related to its formulation where the distribution of the uncertainty is assumed to be Gaussian. However, according to investigation discussed in subsection VI-B, for the case of PV generations, forecast errors are not distributed Gaussian. The variances among the output of different sets of bootstrapped ELMs are small and cannot represent the level of uncertainty in the measurements. Looking at Fig. 8, the centres of the prediction intervals are very close to the point predictions. As the prediction intervals according to (30) are centered on the average of the outputs of the bootstrapped ELMs, it can be concluded that the averages of forecasts by bootstrapped ELMs are very close to the point forecasts. It is to be noted that point predictions are generated using ELM [33] when the original dataset is treated as the input.

To get a better understanding of the sharpness of quantile forecasts generated by the various methods discussed, Fig. 10 is provided. Here, sharpness is examined by assessing the average width of central prediction intervals regardless of their similarity to the measurements. For each interval prediction, the distance between its two boundaries are calculated for each time and then the obtained values are averaged over the evaluation dataset. In Fig. 10, the nominal coverage rates of the intervals versus their average widths are depicted for four seasons of 2013, *Case* 2. According to this figure, BELM prediction intervals have the lowest width, followed by the persistence intervals. Although for the intervals with the lower coverage, the widths of the intervals given by the proposed method are close to those given by BELM, for the coverage rates of 70% and more, they are wider than those of BELM intervals. The reason for that is the inclusion of the calibration assessment in our proposed framework which leads to wider intervals to meet the calibration requirements.

It should be noted that in all cases, the temporal resolution of forecasting is one minute and the instantaneous measurements, not the average values, are taken into account. Needless to say, temporal averaging has smoothing effect on variability and if pursued it is expected to result in sharper quantiles.

Reliability diagrams in Fig. 9, for four seasons of 2013 in $Case$ 1, allow for a more comprehensive evaluation of the reliability of predictive densities. The red solid line represents perfect reliability and reliability curves for the predictive densities should be as close as possible to that of the ideal case. Here, the proposed method yields satisfactory reliability, with observed coverage close to the nominal proportions for most nominal levels. The BELM tends to underestimate quantiles with nominal levels below 50% while overestimating those greater the median. Among all methods mentioned in Fig. 9, the BELM shows the worst reliability and this can be a consequence of the normality assumption in this method as well as high intermittent datasets used in this study. This shows that the BELM is incapable of covering sharp ramp rates in PV generations.

As can be in Fig. 9, the reliability performance for the summer season is better than that of the other seasons. This also can be noticed by comparing the sum of the deviations from perfect reliability in Table I. Table IV provides point forecasts accuracy by season for *Case* 1 in terms of root mean score error (RMSE) and mean absolute error (MAE). Because the same predictors are used for all seasons, the lower RMSE can be interpreted as more predictability for the corresponding season. Looking at the table, the summer season has the lowest predictability among others which of course is expected. Thunderstorms are a normal part of Brisbane weather condition during December to March (summer season) and can be accompanied by hail and strong wind gusts, making the weather very unstable. PV generation experiences very intermittent and sharp ramps in the summer season. This will force the climatology, BELM and persistence methods to provide wider prediction intervals. The reason is that all the mentioned methods work based on observed prediction errors or PV power variations. When the ramps are sharper, these

methods tend to widen the intervals to envelop the variations. The low sharpness of the quantiles given by the mentioned methods can also be concluded by comparing the scores in Table I for different seasons. For the summer season, the worst score is obtained. As skill scores represent both sharpness and reliability, a better reliability with a worse skill score is an evidence of lower sharpness or wider quantiles. For the winter, when PV generation is more predictable, the best sharpness and the lowest reliability are obtained.

Our approach makes a compromise between sharpness and reliability, permitting us to obtain the best or close to best score values while maintaining reliability in an acceptable level. Analysing Tables I-III for *Case* 1, improvements with respect to the persistence benchmark are within 4-14%. Compared to the climatology, improvements reach 53%. Similarly from Table II, improvements with respect to the persistence range between 7 and 22% depending on the season of the year. The best improvements obtained by comparing against the climatology are up to 40%. Comparing the proposed method with the BELM, it is outperformed by the BELM for a couple of seasons, however, the deviations of the BELM quantiles are too high which undermine their reliabilities. Poorly reliable quantiles can cause systematic biases in decision making process.

Complementing our previous discussion on reliability assessment, the results in the various tables confirm the substantially better calibration of our approach. As an example, in Spring and for *Case* 2, the sum of deviations from the nominal levels for our approach is one-tenth of that for the BELM benchmark. The ratio is one-fifth in Fall for *Case* 1 and one-third in Summer for *Case* 3.

As some of the operational problems such as state estimation requires quantiles with high coverages [34], additional tests were conducted to evaluate the capability of both persistence benchmark and our approach to generate quantiles with 99.5%, 99% and 97.5% proportions for *Case* 2. Similarly to less extreme quantiles, it was noticed that our approach yielded score values improvements of 67%, 53% and 47%, for nominal levels of 99.5%, 99% and 97.5% respectively.

In order to investigate the performance of the methodology for times close to sunset, for *Case* 2, we performed point prediction for data values between 6:05 pm to 7:05 pm using ELM methods. We used all data for 2012 as training data and all for 2013 as a test set. The yearly RMSE and MAE are found to be 2.6% and 1.6% respectively. Then, we forecasted quantiles for the same data using the proposed framework. Skill score, maximum deviation of the quantiles' coverage from perfect reliability and sum of the deviations for all 18 quantiles are calculated as 0.9%, 0.42% and 3.7% respectively and are much better than those calculated for the hours closer to midday (Table II).

It should be noted that the proposed probabilistic forecast is based on an offline training algorithm and it uses the historical data to train the model and only the values for predictors are updated in real-time. In this paper the forecasting model is updated seasonally as it results the best performance for the three case studies, however, the applicability of the idea is not limited to the updating period used here. Depending on statistical behaviour of the PV generation and methodological variables in a specific region, the updating periods as well as the selection of the data used as predictors may change. Once the model is trained, it is ready to be used for online or real-time prediction. To predict $m$ number of quantiles for each time instance in the future, only one simple matrix multiplication given in (10) is required. The multiplication can be executed in a small fraction of a second which makes the proposed method suitable for online and real-time applications.

### *D. Normalization with respect to Clear Sky Power*

In days with clear sky all day long, PV generations have a more predictable pattern. Therefore, one may expect that normalization of PV power with respect to the estimated clear sky power potentially can improve predictability PV generation. In [3], 85% quantile forecast of PV power is considered as the estimated clear sky power. Measurements are then normalized with respect to the clear sky power. The approach has improved the forecast accuracy of the point forecasts in terms of RMSE. In order to check the performance of the normalization with respect to the clear sky power, here, three definitions of clear sky power are provided. Those are 85% and 95% quantile forecasts and physical estimation of the clear sky power. For the later one, firstly using geographic coordinate data and elevation from the sea level, clear sky solar irradiance is calculated according to the model given in [31]. Then, clear sky solar irradiance is converted to the PV power according to the formulation given in [35]. We found physical estimation of the non-overcast power and 85% quantile closer to the measurements in days with clear sky all day. Therefore, investigations are carried out while considering these two as the clear sky power. In Fig. 11, normalized PV values with respect to the 85% quantile as well as physical clear sky power for 6 randomly selected days from the dataset for *Case* 2 are shown. As can be seen in the figure, some of the measurements in partially cloudy days exceed both the clear sky power and 85% quantile. The same problem is reported in [3]. Reflections from the clouds and varying level of water vapor in atmosphere are the main reasons for this phenomenon. Looking at the figure, the normalization can help to improve stationarity in days with clear sky and hours closer to noon time.

To verify the extent to which normalization of PV power measurements can improve the performance of probabilistic forecasts, quantile forecasts for normalized time series are generated by deploying the proposed method for *Case* 2. Based on the investigations carried out, the normalization approach results in no significant improvement in probabilistic forecasting according to the specific framework proposed in this paper. The following points can explain the results of the empirical investigations.

- The motivation behind the proposal of normalization of PV power with respect to the clear sky solar power in [3] is that in that paper classical time series methods are used for prediction. Classical time series methods require stationary time series as their inputs. Normalization can potentially create more stationary time series. However, there is no stationarity restriction for neural network based regression methods, such as ELM.
- Because normalization with respect to the low clear sky values for times close to sunrise and sunset leads to very high or infinity errors, according to [3], the daytimes with clear sky power lower than 20% of the maximum clear

TABLE III
SCORES, MAXIMUM AND SUM OF COVERAGE DEVIATION OF ESTIMATED QUANTILES IN PERCENT FORM (%), *Case* 3

| *Case* 3 | Proposed Method | | | Climatology | | | Persistence | | | HIA | | | BELM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Score | Max | Sum | Score | Max | Sum | Score | Max | Sum | Score | Max | Sum | Score | Max | Sum |
| **Spring** | 3.52 | **5.32** | **48.51** | 4.63 | 10.47 | 75.87 | 5.85 | 6.04 | 48.53 | 4.34 | 8.24 | 70.43 | 3.23 | 13.83 | 157.48 |
| **Summer** | **4.24** | 5.28 | 41.25 | 5.00 | **4.82** | **36.8** | 5.77 | 6.21 | 36.58 | 4.92 | 11.86 | 84.91 | 4.40 | 13.81 | 127.07 |
| **Fall** | **3.48** | **4.41** | **37.21** | 3.79 | 11.08 | 93.34 | 5.91 | 12.21 | 88.28 | 4.39 | 8.32 | 69.47 | 3.62 | 11.13 | 92.61 |
| **Winter** | 2.23 | **4.32** | **23.5** | 2.67 | 9.26 | 95.01 | 4.58 | 14.72 | 92.95 | 3.42 | 7.53 | 61.42 | **2.15** | 11.33 | 98.71 |

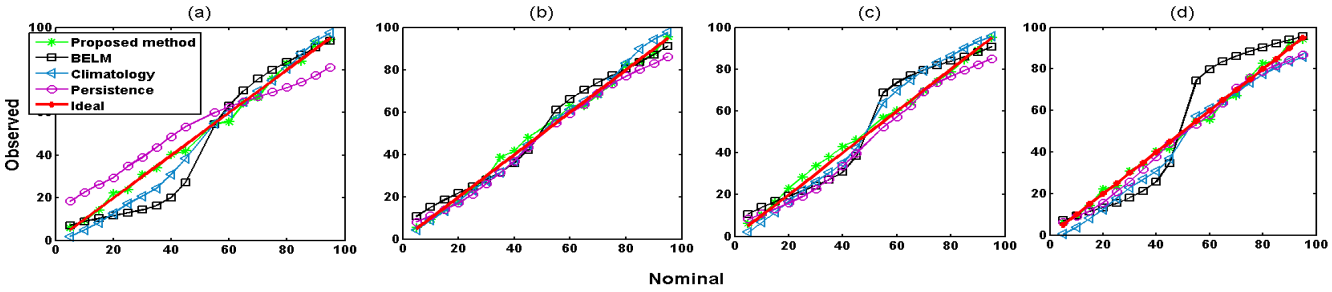


Fig. 9. Reliability diagrams for 18 quantiles with 0.05 increasing nominal proportion, *Case* 1 (a) Spring, (b) Summer, (c) Fall, (d) Winter

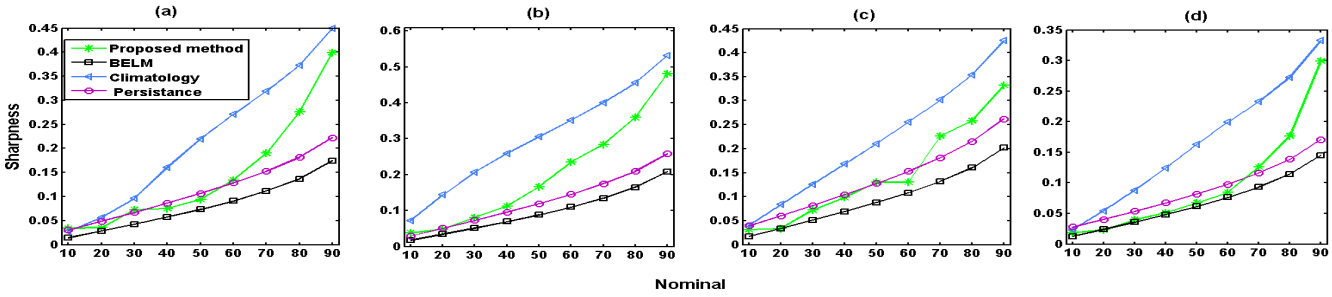


Fig. 10. Reliability diagrams for 18 quantiles with 0.05 increasing nominal proportion, *Case* 1 (a) Spring, (b) Summer, (c) Fall, (d) Winter

TABLE IV
POINT FORECASTS ACCURACY BY SEASON, *Case* 1

| Season\accuracy | RMSE | MAE |
|---|---|---|
| **Spring** | 13.26% | 8.08% |
| **Summer** | 16.43% | 10.73% |
| **Fall** | 13.85% | 8.48% |
| **Winter** | 7.67% | 4.16% |

sky power for the same day should be excluded from the dataset. This results in exclusion of 15% to 30% of the daytime period considered in this study (5 am to 7 pm). Therefore, normalization is not a pragmatic approach to improve forecasts for times close to sunrise and sunset.

- In [3], 15-min and 1-hour averaged values are used. In this paper, the resolution of data is one minute and both databases are from locations where most of days of year are cloudy or partially cloudy. With 1-minute instantaneous values, cloud passages are the main reason of intermittency. As the movement of clouds is random, for partially cloudy days, the clear sky power values can no longer be representative of real measurements.

## VII. CONCLUSION

Most of the previous works on PV forecasting focused on point prediction only, hence researching on conditional expectation of PV power output. We propose a systematic framework for generating probabilistic forecasts for PV power generation. The empirical investigations show that PV power forecast errors do not follow common distributions like *Gaussian*, *Beta*, etc. Therefore, to avoid restrictive assumptions on the shape of the forecast densities, we propose a nonparametric density forecasting method based on ELM as a regression tool, trained with an appropriate criterion. The proposed nonparametric method is successfully applied on two PV power datasets with one-minute resolution and highly fluctuating patterns. The results show that the proposed method is able to efficiently provide reliable and sharp predictive densities for the very short-term (10-minute and one-hour lead times). Part of our future works will focus on demonstrating the functionality, practicability, and value of such probabilistic solar power forecasts in operational problems. Besides, The approach could be extended to account for Numerical Weather Prediction (NWP) as input when looking at further lead times, say, more than two hours ahead.

## REFERENCES

[1] E. Ela, M. R. Milligan, and B. Kirby, *Operating Reserves and Variable Generation: A comprehensive review of current strategies, studies, and fundamental research on the impact that increased penetration of variable renewable generation has on power system operating reserves*. National Renewable Energy Laboratory, 2011.

[2] E. Ela, V. Diakov, E. Ibanez, and M. Heaney, "Impacts of variability and uncertainty in solar photovoltaic generation at multiple timescales," *Contract*, vol. 303, pp. 275–3000, 2013.

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

[4] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, and T. E. Hoff, "Short-term irradiance variability: Preliminary estimation of station pair correlation as a function of distance," *Solar Energy*, vol. 86, no. 8, pp. 2170–2176, 2012.

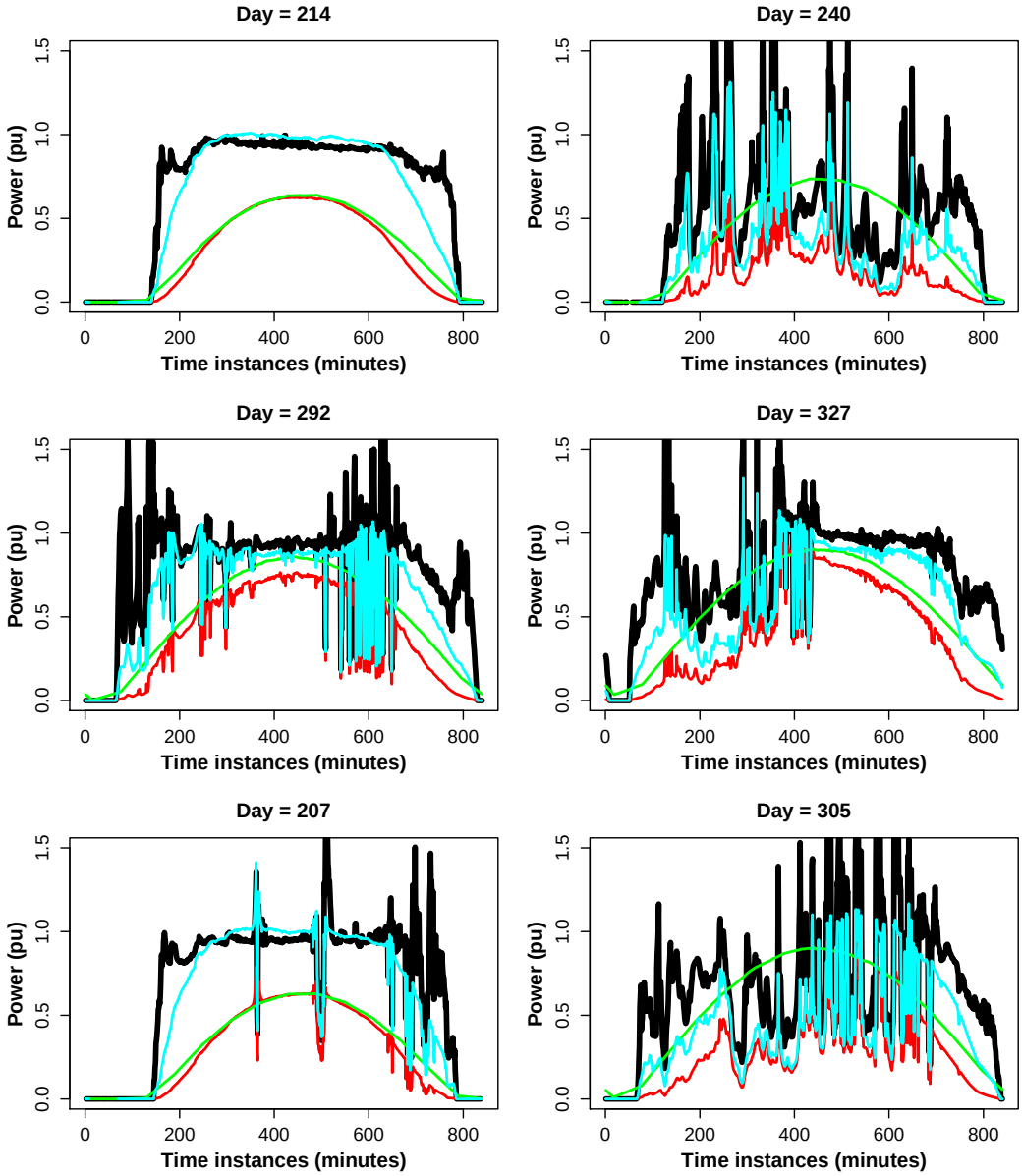


Fig. 11. PV measurements (red colour curve), physical clear sky power (green colour curve), normalized PV power with respect to the clear sky values (black colour curve) and normalized values with respect to 85% quantile (cyan colour curve) for 6 randomly selected days from the database for Case 2. The diagram covers values for 5 am to 7 pm. Counting minutes of day is started from 5 am. Power values more than 1.5 pu are not shown.


[5] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, and T. Hoff, "Parameterization of site-specific short-term irradiance variability," *Solar Energy*, vol. 85, no. 7, pp. 1343–1353, 2011.

[6] T. E. Hoff and R. Perez, "Quantifying pv power output variability," *Solar Energy*, vol. 84, no. 10, pp. 1782–1793, 2010.

[7] C. W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, and B. Washom, "Intra-hour forecasting with a total sky imager at the uc san diego solar energy testbed," *Solar Energy*, vol. 85, no. 11, pp. 2881–2893, 2011.

[8] Y. Chu, B. Urquhart, S. M. Gohari, H. T. Pedro, J. Kleissl, and C. F. Coimbra, "Short-term reforecasting of power output from a 48 mwe solar pv plant," *Solar Energy*, vol. 112, pp. 68–77, 2015.

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

[10] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 6, pp. 1725–1744, 2008.

[11] D. Sarewitz, R. A. Pielke, and R. Byerly, *Prediction: science, decision making, and the future of nature*. Island Press, 2000.

[12] L. Wu, M. Shahidehpour, and T. Li, "Stochastic security-constrained unit commitment," *Power Systems, IEEE Transactions on*, vol. 22, no. 2, pp. 800–811, 2007.

[13] J. M. Morales and J. Perez-Ruiz, "Point estimate schemes to solve the probabilistic power flow," *Power Systems, IEEE Transactions on*, vol. 22, no. 4, pp. 1594–1601, 2007.

[14] P. Pinson, C. Chevallier, and G. N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," *Power Systems, IEEE Transactions on*, vol. 22, no. 3, pp. 1148–1156, 2007.

[15] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*. Springer, 2014.

[16] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *Power Systems, IEEE Transactions on*, vol. 29, no. 3, pp. 1166–1174, 2014.

[17] C. Wan, Z. Xu, P. Pierre, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *Power Systems, IEEE Transactions on*, vol. 29, no. 3, pp. 1033–1044, 2014.

[18] E. Lorenz, J. Hurka, D. Heinemann, and H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 2, no. 1, pp. 2–10, 2009.

[19] Y. Chu, M. Li, H. T. Pedro, and C. F. Coimbra, "Real-time prediction intervals for intra-hour dni forecasts," *Renewable Energy*, vol. 83, pp. 234–244, 2015.

[20] P. Mathiesen, J. M. Brown, and J. Kleissl, "Geostrophic wind dependent

probabilistic irradiance forecasts for coastal california," *Sustainable Energy, IEEE Transactions on*, vol. 4, no. 2, pp. 510–518, 2013.

[21] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *Neural Networks, IEEE Transactions on*, vol. 17, no. 6, pp. 1411–1423, 2006.

[22] S. Sayeef, S. Heslop, D. Cornforth, T. Moore, S. Percy, J. K. Ward, A. Berry, and D. Rowe, "Solar intermittency: Australias clean energy challenge," 2012.

[23] P. Pinson, P. McSharry, and H. Madsen, "Reliability diagrams for non-parametric density forecasts of continuous variables: Accounting for serial correlation," *Quarterly Journal of the Royal Meteorological Society*, vol. 136, no. 646, pp. 77–90, 2010.

[24] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, 2007.

[25] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[26] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 2, pp. 513–529, 2012.

[27] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.

[28] P. Pinson and R. Hagedorn, "Verification of the ecmwf ensemble forecasts of wind speed against analyses and observations," *Meteorological Applications*, vol. 19, no. 4, pp. 484–500, 2012.

[29] J. Ma, Y. V. Makarov, C. Loutan, and Z. Xie, "Impact of wind and solar generation on the california iso's intra-hour balancing needs," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–6.

[30] M. J. Reno, C. W. Hansen, and J. S. Stein, "Global horizontal irradiance clear sky models: Implementation and analysis," *SANDIA report SAND2012-2389*, 2012.

[31] I. A. Walter, R. G. Allen, R. Elliott, M. Jensen, D. Itenfisu, B. Mecham, T. Howell, R. Snyder, P. Brown, S. Echings *et al.*, "Asces standardized reference evapotranspiration equation," in *Task Committee on Standardization of Reference Evapotranspiration of the Environmental and Water Resources Institute*, 2005.

[32] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *Power Systems, IEEE Transactions on*, vol. 25, no. 4, pp. 1845–1856, 2010.

[33] F. Golestaneh and H. B. Gooi, "Batch and sequential forecast models for photovoltaic generation," in *Power & Energy Society General Meeting, 2015 IEEE*. IEEE, 2015, pp. 1–5.

[34] V. Akhmatov, "Influence of wind direction on intense power fluctuations in large offshore windfarms in the north sea," *Wind Engineering*, vol. 31, no. 1, pp. 59–64, 2007.

[35] A. Yona, T. Senjyu, A. Y. Saber, T. Funabashi, H. Sekine, and C.-H. Kim, "Application of neural network to one-day-ahead 24 hours generating power forecasting for photovoltaic system," in *Intelligent Systems Applications to Power Systems, 2007. ISAP 2007. International Conference on*. IEEE, 2007, pp. 1–6.