ORIGINAL RESEARCH   **Open Access**

Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting

Mokhtar Bozorg[1], Antonio Bracale[2], Pierluigi Caramia[2], Guido Carpinelli[3], Mauro Carpita[1] and Pasquale De Falco[2*]

Abstract

Photovoltaic (PV) systems are widely spread across MV and LV distribution systems and the penetration of PV generation is solidly growing. Because of the uncertain nature of the solar energy resource, PV power forecasting models are crucial in any energy management system for smart distribution networks. Although point forecasts can suit many scopes, probabilistic forecasts add further flexibility to an energy management system and are recommended to enable a wider range of decision making and optimization strategies. This paper proposes methodology towards probabilistic PV power forecasting based on a Bayesian bootstrap quantile regression model, in which a Bayesian bootstrap is applied to estimate the parameters of a quantile regression model. A novel procedure is presented to optimize the extraction of the predictive quantiles from the bootstrapped estimation of the related coefficients, raising the predictive ability of the final forecasts. Numerical experiments based on actual data quantify an enhancement of the performance of up to 2.2% when compared to relevant benchmarks.

Keywords: Bayesian bootstrap, Photovoltaic systems, Probabilistic forecasting, Renewable generation; smart grids

1 Introduction

The recent growth in distributed Photovoltaic (PV) power generation systems fosters the exploitation of renewable energy and adds further flexibility to electrical distribution systems, which now experience a significant amount of generation in close proximity to load, with obvious advantages in terms of reduced line congestion and losses. PV power is however nondeterministic, as it strictly depends on weather and environmental conditions [1]. This brings significant challenges for transmission and distribution system operators and for market operators. They continually have to deal with local production that may not respond coherently to day-ahead prior generation programs [2]. The impact of such uncertainty on the network operation is therefore severe. For example, large margins of energy reserve are necessary to be able to deal with the deviations from prior generation programs, and the distributed energy resources of smart grids must be managed under such uncertainty [3]. In order to optimally accomplish such tasks, forecasts of PV power for a specific time horizon must be available up to few days before the actual energy generation.

The relevant literature on PV power forecasting is quite heterogeneous [4]. Most researchers and practitioners use deterministic PV power forecasting, i.e., the disposal of a single value that predicts the actual PV power at a specific time horizon [5]. However, PV power forecasting is best addressed within probabilistic frameworks, because of the intrinsic randomness of the physical phenomena. Literature reviews and competition surveys [5–8] indicate a varied state-of-the-art, although the number of contributions devoted to probabilistic PV power forecasting is much smaller than that in the deterministic framework. Considering that a single spot value can always be extracted from probabilistic

* Correspondence: pasquale.defalco@uniparthenope.it
[2]Department of Engineering, University of Naples Parthenope, Naples, Italy
Full list of author information is available at the end of the article

forecasts while the opposite is clearly unfeasible, research and contributions in the latter are highly recommended.

Probabilistic PV power forecasting systems range from pure statistical models to hybrid physical-statistical models. High-performance solutions are based on Quantile Regression (QR) models [9–11], machine learning approaches (such as gradient boosting [12], quantile regression forests [10, 13, 14] and k-nearest neighbors [15]). It is worth noting that, although the analytic formulation of QR models is much simpler than machine learning approaches, QR predictions still are somehow competitive in most cases.

Generally speaking, the integration of Numeric Weather Predictions (NWPs) into PV power forecasting models is quite mandatory to improve predictions, particularly for large time horizons [6]. Usually several weather variables are available to forecasters, and thus model selection (i.e., the selection of the most informative predictors for the final model) is typically applied to discard uninformative inputs. To further maximize the exploitation of the available input data, ensemble approaches [7] have been applied with success in probabilistic energy forecasting [10, 16–19]. These approaches are commonly divided into boosting, stacking, and bagging. Boosting consists of building a "strong" model by combining several weaker ones, which are trained iteratively. Stacking consists of combining the outcomes of several models which are solved in parallel or in cascade, in order to form the final prediction. Bagging consists of building the final prediction as a combination of the output of the same model trained multiple times, using input data resampled with replacement (i.e., bootstrap aggregated) [7].

This paper provides a contribution to probabilistic ensemble PV power forecasting within the bagging framework, based on the interaction between a QR model and a Bayesian bootstrap. A Bayesian bootstrap is the Bayesian analogue for bootstrapping, originally presented in [20]. Bayesian models have been applied with success in probabilistic energy forecasting [16, 21–23], although applications of Bayesian bootstrap are still very rare. Thus further contributions are worthy of attention.

Like other bootstrapping techniques, the Bayesian bootstrap can improve the probabilistic forecasts by using resampled data with replacement, which allows for differentiating the output predictions. In the forecasting system presented in this paper, the Bayesian bootstrap works analytically to find the posterior distributions of the QR model parameters, thus differentiating itself from the traditional bootstrap which instead relies on random picks among the available input data. In this paper, a procedure to extract an optimal point from the posterior distributions of the QR model parameters is specifically developed, and this procedure is added to the forecasting system, in order to generate the final PV power forecasts for the target forecast horizon.

The proposal is validated using actual data collected at a real PV installation in Switzerland and actual data published in the framework of the Global Energy Forecasting Competition 2014 [8]. Extensive numerical experiments, based on these data and NWPs provided by an external source [24], are presented in this paper. Several related benchmarks are also presented to validate the accuracy of the forecasts under comparative analyses.

This paper is organized as follows. Section 2 presents the forecasting system and the analytic models used in each stage of the system. Section 3 presents the benchmarks and the error indices used to validate the proposal. Section 4 present the numerical experiments, and the paper is concluded in Section 5.

2 Methods

The proposed PV power forecasting system based on Bayesian Bootstrap Quantile Regression (BBQR) is illustrated in Fig. 1. The inputs of the system are NWPs NW and historical measured PV power data P. The proposed forecasting system consists of three stages.

The first stage is model selection, i.e., the selection of the most informative predictors among the available pool of predictors. This is performed by evaluating the performance of multiple QR models having different combinations of predictors, and by picking the model which returns the smallest error.

Inputs are therefore pooled in order to form predictor data X (i.e., independent variables in the QR model) which is informative for the PV power (i.e., the dependent variable in the QR model). Data are then partitioned into training datasets P_{tra} and X_{tra} (a $1 \times T$ vector and a $T \times M$ matrix, respectively), and into validation datasets P_{val} and X_{val} (a $1 \times V$ vector and a $V \times M$ matrix, respectively).

Training data thus contains T occurrences which are used only to train models, whereas validation data contains V occurrences for model selection to develop and refine the forecasting system. M is the number of predictors contemplated in the generic QR model, which therefore has $M + 1$ parameters.

Multiple QR models are trained, and predictions for the validation period are issued with each model. Since predictions are given in terms of predictive quantiles, the Pinball Score (PS) [25] is considered in order to select the best model. In particular, the QR model returning the smallest PS for the validation period is considered as the most skilled, and is selected as the underlying
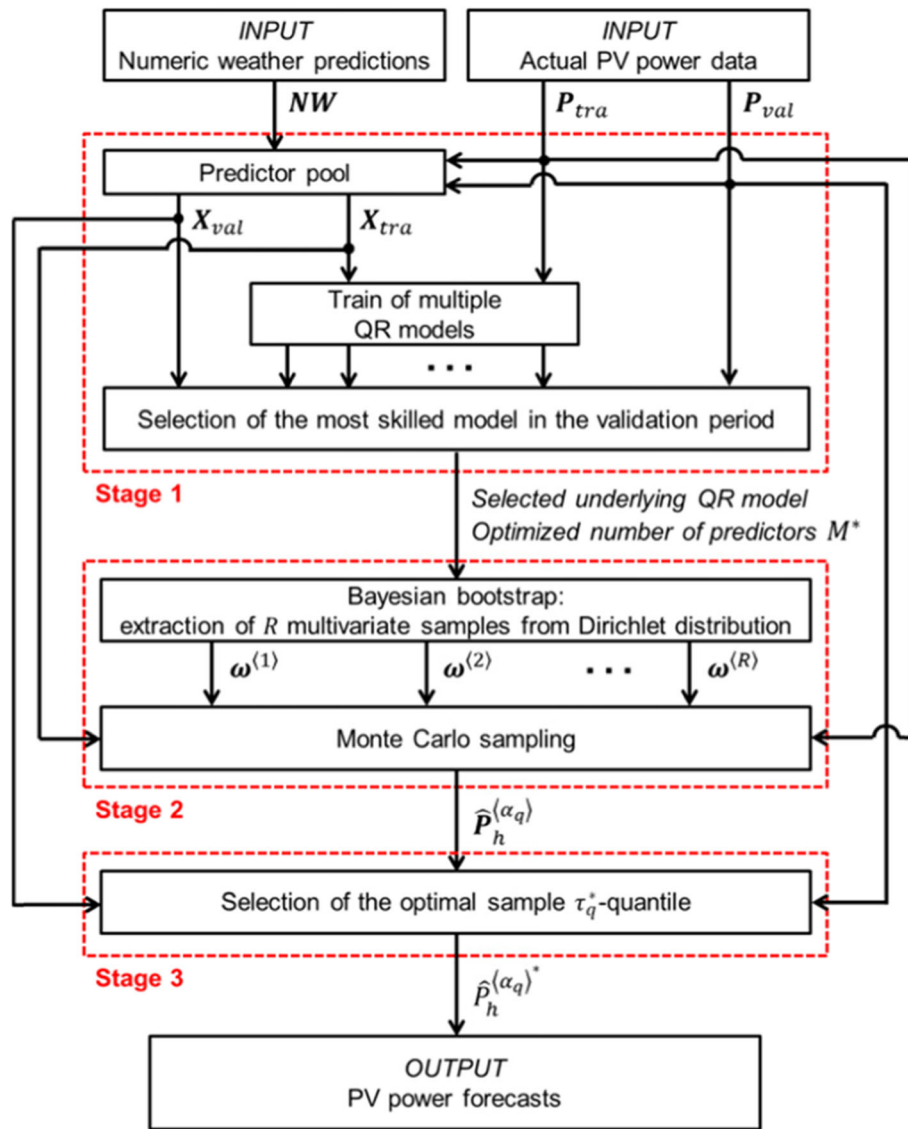
Fig. 1 Schematic procedure of the forecasting system based on Bayesian bootstrap quantile regression

QR model for the remainder of the system. For notation, the underlying QR model selected in this first stage of the system has M^* predictors and $M^* + 1$ parameters.

The second stage consists of applying Bayesian bootstrapping over the selected underlying QR model, in order to estimate the posterior distribution of the parameters of the QR model. Specifically, the Bayesian bootstrap returns R samples extracted from each of the $M^* + 1$ posterior distributions of the $M^* + 1$ parameters of the QR model. As will be shown later, these samples are extracted from a multivariate Dirichlet distribution. A Monte Carlo sampling method then extracts R samples ($\hat{P}_h^{\langle \alpha_q \rangle}$) of predictive α_q-quantiles of PV power for the target horizon h.

The third and last stage consists of extracting a single value $\hat{P}_h^{\langle \alpha_q \rangle^*}$ from the R samples of predictive quantiles of PV power for each coverage, in order to generate the prediction of PV power for the target horizon. A procedure dedicated to this purpose, based on the optimization of the sample τ_q-quantile of $\hat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$, is developed and presented in this paper. The entire predictive distribution of the final probabilistic PV power forecasts can be obtained by iteration for Q predictive quantiles.

The models and the stages of the forecasting system are discussed in the following sub-Sections.

2.1 Quantile regression modeling

A QR model links the target variable (i.e., the predictive α_q-quantile $\hat{P}_h^{\langle \alpha_q \rangle}$ of PV power at the target time horizon h) to predictors $\boldsymbol{x}_h = \{x_{1_h}, ..., x_{M_h}\}$ related to the time horizon h but available at the forecast origin $h - k$. The forecast lead time is indicated by k. This paper focuses on day-ahead forecasting with hourly time resolution, assuming that forecasts are issued at midnight of day D-1 for the entire day D (i.e., $k = 1$, ..., 24), although the proposal can also be applied to other short-term PV power forecasting frameworks. Note that theoretically 24 models should be developed (i.e., one for each hour of the day), but since the PV power production is deterministically zero during the night, only the 16 models corresponding to lead times $k = 5, ..., 20$ are considered. In order to lighten the notation in the entire paper, we will not make reference to the forecast lead time k in the symbols although the methodological section is related to a specific lead time k.

The link imposed by the generic QR model for the PV power α_q-quantile is:

$$\hat{P}_h^{\langle \alpha_q \rangle} = \hat{\beta}_0^{\langle \alpha_q \rangle} + \sum_{m=1}^{M} \hat{\beta}_m^{\langle \alpha_q \rangle} \cdot x_{m_h}, \tag{1}$$

where $\hat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle} = \{\hat{\beta}_0^{\langle \alpha_q \rangle}, ..., \hat{\beta}_M^{\langle \alpha_q \rangle}\}$ are the $M + 1$ estimated values of model parameters $\boldsymbol{\beta}^{\langle \alpha_q \rangle} = \{\beta_0^{\langle \alpha_q \rangle}, ..., \beta_M^{\langle \alpha_q \rangle}\}$. Note that (1) is linear with the parameters, although some predictors can be obtained as multiplicative terms between two or more variables (this allows the introduction of interaction effects among variables [26]).

Parameters are estimated in the training step by minimizing an error score on known data (i.e., supervised training). The PS fits this purpose well, since it can be applied directly on predictive quantiles and, for this reason, it is applied in this paper in order to evaluate the accuracy of PV power forecasts. The minimization problem is:

$$\hat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle} = \underset{\boldsymbol{\beta}^{\langle \alpha_q \rangle}}{\arg\min} \, PS\left(\boldsymbol{P}_{tra}, \hat{\boldsymbol{P}}_{tra}^{\langle \alpha_q \rangle}\right), \tag{2}$$

where $PS(\boldsymbol{P}_{tra}, \hat{\boldsymbol{P}}_{tra}^{\langle \alpha_q \rangle})$ is the PS of the T forecasts $\hat{\boldsymbol{P}}_{tra}^{\langle \alpha_q \rangle}$ issued for the training period of length T, calculated with respect to the actual occurrences of PV power in the training set $\boldsymbol{P}_{tra} = \{P_{t_1}, ..., P_{t_T}\}$.

Although it is not directly explained in (2), the forecasts $\hat{\boldsymbol{P}}_{tra}^{\langle \alpha_q \rangle}$ are obtained from (1), and thus they are functions of $\boldsymbol{\beta}^{\langle \alpha_q \rangle}$ and are dependent on the $T \times M$ matrix $\boldsymbol{X}_{tra}$ which contains the corresponding predictors for the training period. It is:

$$\hat{\boldsymbol{P}}_{tra}^{\langle \alpha_q \rangle} = \boldsymbol{f}_{QR}\left(\boldsymbol{\beta}^{\langle \alpha_q \rangle} | \boldsymbol{X}_{tra}\right), \tag{3}$$

and therefore (2) can be rewritten in compact form as:

$$\hat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle} = \boldsymbol{G}(\boldsymbol{P}_{tra}, \boldsymbol{X}_{tra}), \tag{4}$$

where $\boldsymbol{G}(\cdot)$ is a function obtained by combining (2) and (3).

2.2 First stage: model selection

In the first stage of the proposed forecasting system, the optimal model is selected among a pool of candidates, which differ in the predictors used to generate the predictions. In this paper, NWPs and 1-day lagged PV power, together with their coupled interactions, form the pool of candidate predictors. The considered NWPs are: total cloud coverage, clear-sky irradiance, total irradiance, total precipitation, pressure and air temperature [24]. Two hypotheses are added to reduce the search dimension for the optimal model: i) if a coupled interaction is a predictor of the model, the two individual variables are forced to occur in the model; ii) only models containing NWPs of total cloud coverage, clear-sky irradiance and total irradiance are considered because of their recognized importance in PV power forecasting.

The underlying QR model selected under these hypotheses is the one which minimizes the PS across the Q quantile coverages, i.e., the same optimal combination of M^* predictors is selected for the Q considered quantile coverages. To avoid overfitting, the minimum PS is evaluated on the validation dataset $\boldsymbol{P}_{val} = \{P_{v_1}, ..., P_{v_V}\}$, which is not used for training the model.

2.3 Second stage: Bayesian bootstrap quantile regression

Like traditional bootstrapping techniques, the Bayesian bootstrap can improve the probabilistic forecasts by using resampled data with replacement, which allows for differentiating the output predictions. The Bayesian bootstrap is specifically applied on the underlying QR model selected in the previous stage, in order to evaluate the posterior distribution of the $M^* + 1$ parameters. As shown in the remainder of this subsection, BBQR consists of extracting weights from a Dirichlet distribution R times (once for each bootstrap replicate), building R multinomial distributions using the occurrences and the weights, sampling with replacement from these R distributions and calculating $\hat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle}$ from (2)–(4) on the bootstrapped data. Therefore, the posterior distribution of the QR model parameters is given by R samples $\hat{\boldsymbol{\beta}}_1^{\langle \alpha_q \rangle}, ..., \hat{\boldsymbol{\beta}}_{M^*+1}^{\langle \alpha_q \rangle}$ for

each parameter, and from these samples it is eventually obtains the bagged samples $\hat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$ of PV power.

To facilitate the presentation of the BBQR formulation, a brief recap on traditional bootstrapping [27, 28] is provided.

The $T \times (M^* + 1)$ occurrence matrix $\boldsymbol{Y}_{tra} = [\boldsymbol{P}_{tra}' \ \boldsymbol{X}_{tra}]$ is initially obtained from the transposed training set $\boldsymbol{P}_{tra}'$ and the corresponding $T \times M^*$ matrix $\boldsymbol{X}_{tra}$. The nth row vector $\boldsymbol{y}_{t_n} = \{P_{t_n}, x_{1_{t_n}}, ..., x_{M_{t_n}^*}\}$, taken from the occurrence matrix $\boldsymbol{Y}_{tra}$, contains the target variable and the predictors at the time step t_n. It may be viewed as an item coming from some generic, unknown multinomial distribution $F(\boldsymbol{y})$, with T available realizations (i.e., past occurrences) $\boldsymbol{y}_{t_1}, ..., \boldsymbol{y}_{t_T}$.

As shown earlier, the estimated parameters $\hat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle}$ of the QR model come from (4), and therefore they can be viewed as a function of $\boldsymbol{G}[F(\boldsymbol{y})]$:

$$\hat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle} = \boldsymbol{G}[F(\boldsymbol{y})]. \tag{5}$$

In bootstrap (either traditional or Bayesian [20, 27, 28]), the unknown distribution $F(\boldsymbol{y})$ is searched for among distributions of the type $F_T(\boldsymbol{y})$:

$$F_T(\boldsymbol{y}) = \sum_{n=1}^{T} \omega_{t_n} \cdot \delta_{\boldsymbol{y}_{t_n}}, \tag{6}$$

where $\delta_{\boldsymbol{y}_{t_n}}$ is a degenerate probability measure for the nth vector $\boldsymbol{y}_{t_n}$ of occurrences, and ω_{t_n} is an assigned weight. For consistency, the weights must satisfy the following conditions:

$$\sum_{n=1}^{T} \omega_{t_n} = 1, \omega_{t_n} \geq 0, \forall n = 1, ..., T. \tag{7}$$

In a traditional bootstrap, the function $\boldsymbol{G}[F(\boldsymbol{y})]$ is estimated upon R distributions $F_T^{\langle 1 \rangle}(\boldsymbol{y}), ..., F_T^{\langle R \rangle}(\boldsymbol{y})$. With reference to the generic r^{th} replicate, the functional $\boldsymbol{G}[F_T^{\langle r \rangle}(\boldsymbol{y})]$ is calculated using the weights $\boldsymbol{\omega}^{\langle r \rangle} = \{\omega_{t_1}^{\langle r \rangle}, ..., \omega_{t_T}^{\langle r \rangle}\}$, that are obtained by a random extraction from the multinomial distribution:

$$f_{Mul}(T; 1/T, 1/T, ..., 1/T). \tag{8}$$

and normalizing by T.

The Bayesian bootstrap differs from the traditional bootstrap since the bootstrapped weights $\omega_{t_1}^{\langle r \rangle}, ..., \omega_{t_T}^{\langle r \rangle}$ are not obtained by random extraction from distribution (8). Instead, the vector $\boldsymbol{\omega} = \{\omega_{t_1}, ..., \omega_{t_T}\}$ is the object of Bayesian analysis, which aims at evaluating a posterior distribution $p(\boldsymbol{\omega} | \boldsymbol{Y}_{tra})$ of this vector of weights, given occurrence data $\boldsymbol{Y}_{tra}$. A prior distribution $p(\boldsymbol{\omega})$ should be imposed upon the parameters $\boldsymbol{\omega}$ to start the Bayesian inference [20, 27, 28]. A convenient choice is to select a Dirichlet distribution, which is a conjugate prior for the multinomial distribution of $\boldsymbol{y}$ [20, 27, 28]. In such a case, the posterior distribution $p(\boldsymbol{\omega} | \boldsymbol{Y}_{tra})$ is itself a Dirichlet distribution, $f_{Dir}(1, ..., 1; 1, ..., 1)$ [20, 27]. This allows applying a Monte Carlo sampling method to get the Bayesian bootstrapped samples $\hat{\boldsymbol{\beta}}_1^{\langle \alpha_q \rangle}, ..., \hat{\boldsymbol{\beta}}_{M^*+1}^{\langle \alpha_q \rangle}$ of the $M^* + 1$ estimated parameters of the QR model. The steps are:

i) R multivariate samples $\boldsymbol{\omega}^{\langle 1 \rangle}, ..., \boldsymbol{\omega}^{\langle R \rangle}$ are independently extracted from the Dirichlet distribution $f_{Dir}(1, ..., 1; 1, ..., 1)$;
ii) $\boldsymbol{G}[F_T^{\langle 1 \rangle}(\boldsymbol{y})], ..., \boldsymbol{G}[F_T^{\langle R \rangle}(\boldsymbol{y})]$ are calculated applying (2)–(4);
iii) R Bayesian bootstrapped samples $\hat{\boldsymbol{\beta}}_1^{\langle \alpha_q \rangle}, ..., \hat{\boldsymbol{\beta}}_{M^*+1}^{\langle \alpha_q \rangle}$ for each of the $M^* + 1$ parameters of the QR model are obtained using (5). From these samples, it obtains R samples of the predictive α_q-quantile $\hat{P}_h^{\langle \alpha_q \rangle}$ of PV power by applying (1). The set of these samples are indicated with $\hat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$.

2.4 Third stage: extraction of a single value from the Bayesian bootstrapped predictive PV power

The R samples $\hat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$ of the predictive α_q-quantile of PV power can be interpreted as probabilistic predictions for the predictive quantile. Sample quantiles and confidence intervals of the predictive α_q-quantile of PV power can therefore be estimated from $\hat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$. Since it will be of use later in the paper, the generic sample τ_q-quantile estimated from $\hat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$ is denoted by $\hat{P}_h^{\langle \alpha_q \rangle \langle \tau_q \rangle}$.

Probabilistic PV power forecasts are usually given in terms of predictive distribution or a set of predictive quantiles at different coverage levels, and the redundancy given by multiple samples for each quantile level can lead to misinterpretation of the results in practical utilization of forecasts. A dedicated procedure is developed in this paper to reduce the redundancy of the forecasts by extracting a single value from the samples of the predictive α_q-quantile of PV power. This single value is treated as the final predictive α_q-quantile $\hat{P}_h^{\langle \alpha_q \rangle *}$ of PV power returned by BBQR. The procedure effectively exploits the information contained in the available R samples, in order to further improve the final probabilistic forecasts.

The single value $\hat{P}_h^{\langle \alpha_q \rangle *}$ is the sample quantile extracted

from $\hat{P}_h^{\langle \alpha_q \rangle}$ as:

$$\hat{P}_h^{\langle \alpha_q \rangle *} = \hat{P}_h^{\langle \alpha_q \rangle \langle \tau_q^* \rangle}, \tag{9}$$

where the specific coverage τ_q^* of this sample quantile is the object of an optimization problem aimed at minimizing the PS of the final forecasts calculated on the validation set $\boldsymbol{P}_{val}$, i.e.:

$$\tau_q^* = \underset{\tau_q}{\operatorname{argmin}} PS\left(\boldsymbol{P}_{val}, \hat{\boldsymbol{P}}_{val}^{\langle \alpha_q \rangle \langle \tau_q \rangle} \right). \tag{10}$$

Note that this procedure is made independent for each quantile coverage α_1, ..., α_Q, for simplicity. Therefore, possible quantile crossing in the final PV power forecasts is corrected by post-processing the results with simple sorting in ascending order across the Q coverages.

3 Benchmarks and error indices
This section presents the benchmarks and the error indices used to assess the validity of the proposal and to compare the probabilistic forecasts.

3.1 Benchmarks
Five benchmarks are considered to provide a fair comparison of the results. They are listed below.

Simple QR (SQR): the first benchmark [10, 16] is introduced to be used as a reference in which each predictive quantile is directly provided as a single value, rather than by passing through the bootstrap. This allows assessing whether the bootstrap is effective or not in improving forecasts. To provide a fair comparison, the same model selection procedure presented in Section 2.2 within the framework of the proposed forecasting system based on BBQR is also adopted for SQR.
Traditional Bootstrap QR (TBQR): the second benchmark [29] is introduced in order to evaluate if the Bayesian bootstrap is more effective than the traditional bootstrap in increasing the skill of the final forecasts. For fair comparison, the same QR model selection procedure, the same Monte Carlo sampling method and the same procedure to extract a single value from the R samples, presented respectively in Sections 2.2, 2.3 and 2.4, are applied to TBQR.
Quantile Regression Neural Network (QRNN) and Gradient Boosting Regression Tree (GBRT): the third and fourth benchmarks of QRNN [30] and GBRT [31] are introduced to provide independent references that do not come from QR-based models.

QRNN is formulated to simultaneously predict several PV power quantiles to reduce the quantile crossing effect [30]. The basic neural network architecture used to develop QRNN is the multilayer perceptron with a single hidden layer. Hyperparameter optimization is performed through a validation procedure on the validation dataset $\boldsymbol{P}_{val}$, to maintain statistical fairness with the other models. QRNN is implemented using the *qrnn* package in R [32].

GBRT is developed individually for each considered quantile, and the predictions are post-processed in a sorting procedure to avoid quantile crossing. Also in this case, the hyperparameter optimization is performed through a validation procedure on the validation dataset $\boldsymbol{P}_{val}$, to maintain statistical fairness with the other models. GBRT is implemented using the *gbm* package in R [33].

Seasonal Persistence Model (SPM): the fifth benchmark is based on the underlying daily periodicity of the PV power pattern driven by the rotation of the Earth around its own axis [16]. In practice, each predictive quantile is the PV power observed the day before:

$$\hat{P}_h^{\langle \alpha_q \rangle} = P_{h-24}. \tag{11}$$

This benchmark is added in order to provide a naive, unbiased reference for comparison.

3.2 Error indices
Two error indices are used to quantify the accuracy of the forecasts. The first index is the abovementioned PS, which is a strictly proper score [25] that simultaneously addresses the reliability and the sharpness of forecasts [16, 25]. Its formulation is:

$$PS\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle} \right) = \left[\alpha_q - \mathrm{I}\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle} \right) \right] \cdot \left(P_h - \hat{P}_h^{\langle \alpha_q \rangle} \right), \tag{12}$$

where the indicator function $\mathrm{I}(P_h, \hat{P}_h^{\langle \alpha_q \rangle})$ is:

$$\mathrm{I}\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle} \right) = \begin{cases} 1 \ if \ P_h \leq \hat{P}_h^{\langle \alpha_q \rangle} \\ 0 \ if \ P_h > \hat{P}_h^{\langle \alpha_q \rangle} \end{cases} \tag{13}$$

A comprehensive PS can be obtained averaging across multiple forecast issues (e.g., the V issues in the validation set) and summing over the Q quantiles. PS is negatively oriented, so a smaller PS indicates better forecasts. The Normalized PS (NPS) is provided

in our numerical experiments to get scale-independent results. It is:

$$NPS = \frac{PS\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle}\right)}{\overline{P}_{\text{rated}}}, \tag{14}$$

where P_{rated} is the rated power of the PV system.

The second error index is the Average Absolute Coverage Error (AACE), and it addresses the reliability of the forecasts, i.e., the correspondence between the estimate and the nominal coverages of the predictive quantiles [34]. Because of its intrinsic properties, it can only be formulated for multiple forecast issues. For sake of clarity it is referred in (16) to the validation set (although it can be easily adapted to other data sets). In such a case, the estimated α_q-coverage $\hat{\alpha}_q$ is:

$$\hat{\alpha}_q = \frac{1}{V} \cdot \sum_{n=1}^{V} \mathrm{I}\left(P_{v_n}, \hat{P}_{v_n}^{\langle \alpha_q \rangle}\right), \tag{15}$$

and the Absolute Coverage Error (ACE) on the nominal α_q-quantile is:

$$ACE^{\langle \alpha_q \rangle} = \left| \alpha_q - \hat{\alpha}_q \right|. \tag{16}$$

The AACE across the Q coverages can be obtained as a percentage value as:

$$AACE_\% = \frac{100}{Q} \cdot \sum_{q=1}^{Q} ACE^{\langle \alpha_q \rangle}. \tag{17}$$

The AACE is negatively oriented, so a smaller AACE indicates more reliable forecasts.

4 Results and discussion

The proposed forecasting system is assessed using two actual PV power datasets.

Dataset #1 consists of data collected at a PV installation within the ReIne smart grid laboratory in Switzerland [35]. Data span February 1, 2016 to November 30, 2018, and are initially pre-processed to correct potential outliers, missing and bad data, and to average the measurements in order to obtain hourly PV power data.

Dataset #2 consists of zone-1 data published in the framework of the Global Energy Forecasting Competition 2014 [8] at an unspecified location in Australia, and data span April 1, 2012 to June 30, 2014. Since these data were already pre-processed by the competition organizers and arranged with an hourly resolution, no further processing is applied on dataset #2.

NWPs are provided in both cases by the European Centre for Medium-range Weather Forecast (ECMWF) [24] for the locations at which the PV systems are installed, and for the same time intervals related to datasets #1 and #2. All the weather forecasts refer to the midnight run, i.e., NWPs are available at 24:00 of day D-1 for the entire day D, in order to suit the desired forecasting framework [24].

Data are normalized in the 0–1 range to accommodate for the very different intervals spanned by the variables. The normalized value $\tilde{z}_h$ of the occurrence of the generic variable z at time h is:

$$\tilde{z}_h = \frac{z_h - z_{\min}}{z_{\max} - z_{\min}}, \tag{18}$$

where $z_{\min}$ and $z_{\max}$ are the minimum and maximum values registered for the variable, respectively.

The available datasets are partitioned in three subsets. Dataset #1 is split into a training set spanning February 1, 2016 to December 31, 2017 (i.e., $T = 16800$), a validation set covering first 6 months of 2018 (i.e., $V = 4344$), and a test set only used to assess the accuracy of forecasts, covering the remaining 5 months of 2018 (i.e., 3672 forecast issues). Similarly, dataset #2 is split into a training set from April 1, 2012 to October 31, 2013 (i.e., $T = 13896$), a validation set covering the following 5 months (i.e., $V = 3624$), and a test set covering the remaining 3 months of 2014 (i.e., 2184 forecast issues).

The number R of bootstrapped is searched for in the range 1000–10,000, considering the size of the two datasets. Four tests with $R = 1000, 2000, 5000, 10000$ were run, and the performances in these four cases were checked on the validation datasets P_{val}. A good compromise was found by selecting $R = 5000$ for both datasets. 5000 is the value used eventually to predict the PV power in the test periods.

Forecasts are issued for $Q = 19$ nominal quantile coverages $\alpha_1, ..., \alpha_{19} = 0.05, ..., 0.95$. All forecasts are developed in the R environment, exploiting packages *quantreg* [36] and *bayesboot* [37], *qrnn* [32] and *gbm* [33]. In the following subsections, the extensive results for dataset #1 are presented, whereas the results for dataset #2 are presented in a more compact form, to avoid verbose presentation.

4.1 Results for dataset #1

The outcome of the first stage of the proposed forecasting system determines the model selected for

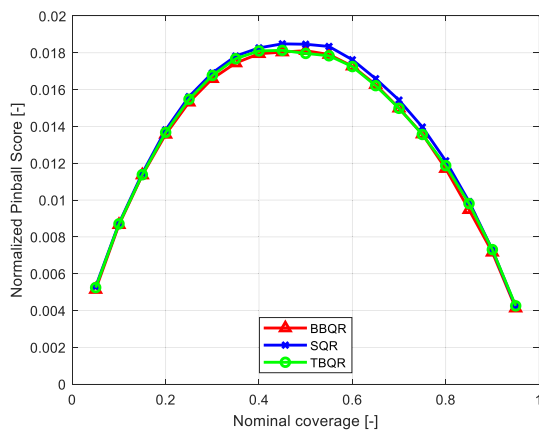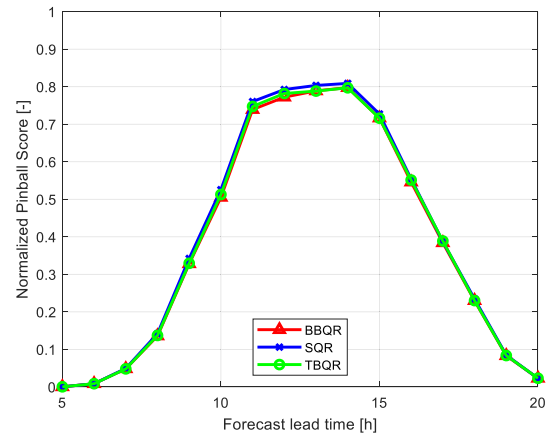**Table 1** Forecast results for the test set of dataset #1

Method	NPS [–]	AACE [%]
BBQR	0.2547	2.22
SQR	0.2604	5.41
TBQR	0.2562	2.38
QRNN	0.2692	5.08
GBRT	0.2583	5.72
SPM	0.5193	–

the specific PV system. For dataset #1, the selected model is:

$$
\begin{aligned}
\hat{P}_h^{\langle \alpha_q \rangle} = {} & \hat{\beta}_0^{\langle \alpha_q \rangle} + \hat{\beta}_1^{\langle \alpha_q \rangle} \cdot tcc_h \\
& + \hat{\beta}_2^{\langle \alpha_q \rangle} \cdot ti_h + \hat{\beta}_3^{\langle \alpha_q \rangle} \cdot csi_h \\
& + \hat{\beta}_4^{\langle \alpha_q \rangle} \cdot P_{h-24} \\
& + \hat{\beta}_5^{\langle \alpha_q \rangle} \cdot tcc_h \cdot ti_h \\
& + \hat{\beta}_6^{\langle \alpha_q \rangle} \cdot tcc_h \cdot csi_h \\
& + \hat{\beta}_7^{\langle \alpha_q \rangle} \cdot ti_h \cdot P_{h-24},
\end{aligned}
\tag{19}
$$

where tcc_h, csi_h and ti_h are NWPs of total cloud coverage, clear-sky irradiance and total irradiance, respectively. The number of predictors of the selected model is therefore $M^* = 7$.

The forecast results for the test set of dataset #1 are shown in Table 1 via NPS (summed across the $Q = 19$ quantiles and averaged through the test set) and AACE. BBQR returns a NPS smaller than SQR, TBQR, QRNN, GBRT and SPM benchmarks by 2.2%, 0.6%, 5.4%, 1.4% and 51.0%, respectively. Bootstrapping increases the accuracy of forecasts, since both the bootstrapped methods (BBQR and TBQR)


Fig. 3 NPS of probabilistic forecasts versus the forecast lead time for the test set of dataset #1

outperform SQR, although the Bayesian-based procedure slightly outperforms the traditional bootstrapping procedure in terms of NPS.

Further details on the skill of forecasts can be evaluated from the results of the experiments. Figure 2 shows the NPS, averaged through the test set of BBQR, SQR and TBQR forecasts for each nominal quantile level. Figure 3 shows the NPS (summed across the $Q = 19$ quantiles and averaged through the test set) of BBQR, SQR and TBQR forecasts versus the forecast lead time. The similar patterns illustrated in these two figures are determined by the same underlying QR model used in all three forecasting methods. It can be determined that peak NPS occurs for middle coverage levels, whereas the NPS changes with the lead time for two reasons: i) forecasts inevitably tend to lose accuracy as lead time increases, and ii) the "bell-shaped" PV power

Fig. 2 NPS of probabilistic forecasts versus the nominal quantile coverage for the test set of dataset #1
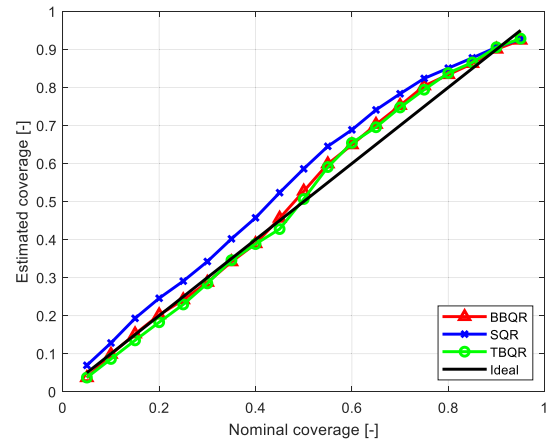
Fig. 4 Reliability diagrams of the QR-based probabilistic forecasts for the test set of dataset #1
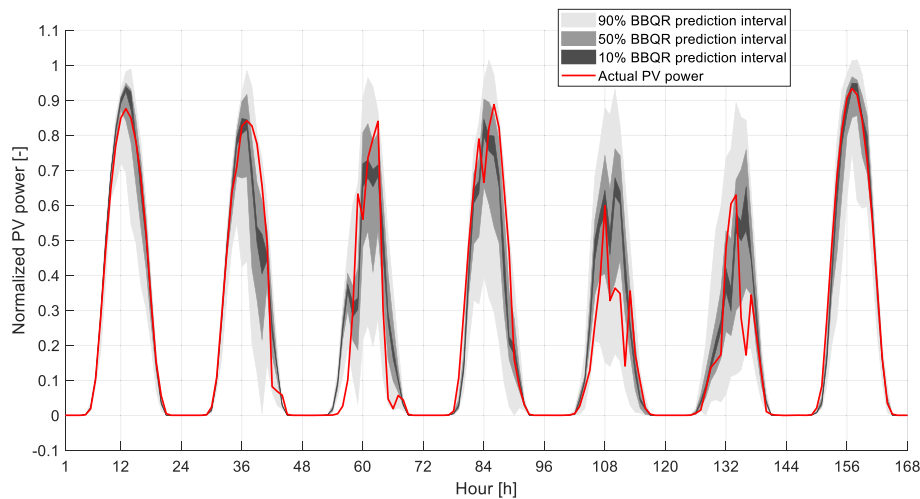
Fig. 5 BBQR prediction intervals during 1 week of the test set of dataset #1

patterns have small errors in proximity to dawn and dusk.

BBQR forecasts are also the most reliable, as the AACE is reduced by 59%, 6.7%, 56.3% and 61.2%, with respect to SQR, TBQR, QRNN and GBRT, respectively. SPM AACE is not presented, since SPM forecasts are the same for each quantile coverage. To compare the probabilistic QR-based forecasts in detail, Fig. 4 shows the reliability diagrams, outlining the estimated coverages versus nominal, of BBQR, SQR and TBQR. Both BBQR and TBQR show similar patterns, with slightly overestimated coverages in the range 0.5 to 0.8. However, the SQR coverages are overestimate for all nominal levels.

In order to provide a graphical interpretation of the PV power forecasts versus time, Fig. 5 shows the BBQR prediction intervals for 1 week of the test period. Prediction intervals are given for rates 90%, 50% and 10%, and they are plotted together with the actual PV power.

4.2 Results for dataset #2
The results for dataset #2 are presented in a more compact form, to avoid duplication. Table 2 shows the NPS and AACE for BBQR and the benchmarks. As seen, BBQR returns an NPS that is smaller than the benchmarks for this dataset, too. The respective reductions are around 4.7%, 2.4%, 6.3%, 2.0% and 53.4%, compared to SQR, TBQR, QRNN, GBRT and SPM. Bootstrapping is again proved to increase the accuracy of forecasts since both bootstrapped methods (BBQR and TBQR) outperform the SQR. The Bayesian-based procedure slightly outperforms the traditional bootstrapping procedure in terms of NPS.

The AACE of BBQR forecasts is again the smallest, accounting for reductions of 41.3%, 7.7%, 30.4% and 34.9% with respect to SQR, TBQR, QRNN and GBRT, respectively. Figure 6 shows the reliability diagrams of the probabilistic QR-based forecasts. BBQR and TBQR show similar patterns, with slightly overestimated coverages in the range 0.5 to 0.8, whereas the SQR coverages are underestimated for all nominal levels.

5 Discussion
The results obtained in the experiments denote the ability of BBQR to slightly increase the performance of the probabilistic predictions. Compared to traditional bootstrap approaches, the NPS is reduced in the range from 0.6% to 2.4%, and the overall reliability is slightly increased. The proposed method also performs well when compared with the state-of-the-art consolidated probabilistic models in energy forecasting, such as QRNN and GBRT.

These promising results indicate the applicability of Bayesian bootstrap techniques for estimating the parameters of different models, thus not limiting the

Table 2 Forecast results for the test set of dataset #2

Method	NPS [—]	AACE [%]
BBQR	0.2364	2.50
SQR	0.2480	4.26
TBQR	0.2422	2.71
QRNN	0.2522	3.59
GBRT	0.2412	3.84
SPM	0.5078	–
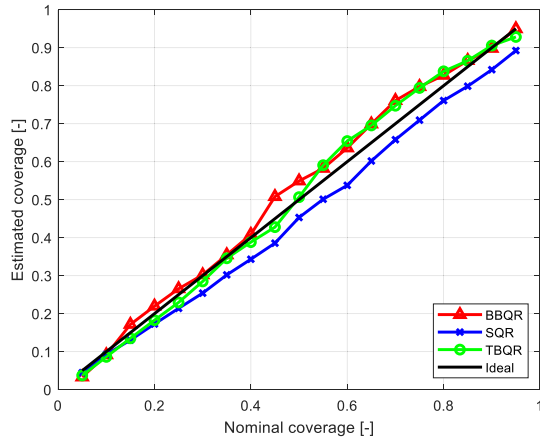
Fig. 6 Reliability diagrams of the QR-based probabilistic forecasts for the test set of dataset #2

analysis to QR-based models, in order to consolidate the technique in probabilistic energy forecasting.

Some limitations apply to the type of prior and posterior distributions of the parameter. As shown in the methodology section, although conjugate priors ease the process of Bayesian bootstrap sampling, numerical methods (e.g., Metropolis-Hastings or Gibbs sampling) can be applied to draw samples from the posterior distributions of parameters even if the prior is not conjugate, thus allowing for generalizing the approach under different assumptions.

6 Conclusions

This paper provides a new insight on bootstrapping in energy forecasting. A novel forecasting system based on Bayesian bootstrap is applied to day-ahead prediction of PV power. A Bayesian bootstrap is used to provide information regarding the parameters of an underlying QR model, and a Monte Carlo sampling method is specifically developed in order to generate the final forecasts.

The proposed forecasting system is compared to several benchmarks, to assess the efficacy of the Bayesian bootstrap with respect to traditional bootstrap and others not using a bootstrapping technique. Numerical experiments based on actual PV power data and on NWPs confirm that BBQR is a valid approach with slightly increased accuracy and reliability of PV power forecasts.

This paper opens new research perspectives in further applications of Bayesian approaches and the Bayesian bootstrap in probabilistic energy forecasting, and in the development of techniques to extract final forecasts from bootstrapped results.

Abbreviations
AACE: Average Absolute Coverage Error; ACE: Absolute Coverage Error; BBQR: Bayesian Bootstrap Quantile Regression; NPS: Normalized Pinball Score; NWP: Numeric Weather Prediction; PS: Pinball Score; PV: Photovoltaic; QR: Quantile Regression; SPM: Seasonal Persistence Model; SQR: Simple Quantile Regression; TBQR: Traditional Bootstrap Quantile Regression; $AACE_{\%}$: Percentage AACE value; $ACE^{\langle a_q \rangle}$: ACE value on the nominal a_q-quantile; csi_h: NWP of the clear-sky irradiance at the target horizon h; $F(\mathbf{y})$: Unknown multinomial distribution from which occurrences are extracted; $F_T(\mathbf{y})$: Family of multinomial distributions; $F_T^{\langle r \rangle}(\mathbf{y})$: Generic r[th] distribution evaluated from the rth bootstrap replicate; $f_{Dir}(\cdot)$: Dirichlet distribution; $f_{Mul}(\cdot)$: Multinomial distribution; $f_{QR}(\cdot | \cdot)$: Vector-valued function returning QR forecasts; $G(\cdot)$: Functional returning the estimated parameters of the QR model for the a_q-quantile; h : Target time horizon; $I(\cdot, \cdot)$: Indicator function; k : Forecast lead time; M: Number of predictors in the generic QR model; M^*: Number of predictors in the selected underlying QR model; NPS: NPS value; $\mathbf{NW}$: NWP data; $p(\boldsymbol{\omega})$: Prior distribution of weights $\boldsymbol{\omega}$; $p(\boldsymbol{\omega} | \mathbf{Y}_{tra})$: Posterior distribution of weights $\boldsymbol{\omega}$, given occurrence data $\mathbf{Y}_{tra}$; $\mathbf{P}$: PV power data; P_h : PV power at the target horizon h; $\overline{P}_{rated}$: Rated power of the PV system; $\mathbf{P}_{tra}$: PV power data of the training period; P_{t_n} : Generic n[th] element of the PV power data of the training period; $\mathbf{P}_{val}$: PV power of the validation period; P_{v_n} : Generic n[th] element of the PV power data of the validation period; $\hat{P}_{tra}^{\langle a_q \rangle}$: Predictive a_q-quantiles of PV power issued for the training period; $\hat{P}_{val}^{\langle a_q \rangle}$: Predictive a_q-quantiles of PV power issued for the validation period; $\hat{P}_h^{\langle a_q \rangle}$: Bagged predictive a_q-quantiles of PV power for the target horizon h; $\hat{P}_h^{\langle a_q \rangle}$: Predictive a_q-quantile of PV power for the target horizon h; $\hat{P}_h^{\langle a_q \rangle *}$: Final predictive a_q-quantile of PV power for the target horizon h; $\hat{P}_h^{\langle a_q \rangle \langle \tau_q \rangle}$: Generic sample τ_q-quantile estimated from the bagged $\hat{P}_h^{\langle a_q \rangle}$; $PS(\cdot, \cdot)$: PS value; Q: Number of quantiles forming the forecast; R: Number of bootstrap replicates; T: Length of the training period; tcc_h: NWP of the total cloud coverage at the target horizon h; ti_h: NWP of the total irradiance at the target horizon h; V: Length of the validation period; $\mathbf{X}$: Predictor data; $\mathbf{x}_h$: Vector of predictors related to the target horizon h; x_{m_h} : Generic m[th] element of the vector of predictors related to the target horizon h; $\mathbf{X}_{tra}$: Predictor data of the training period; $\mathbf{X}_{val}$: Predictor data of the validation period; $\mathbf{y}_{t_n}$: Generic n[th] row vector taken from the occurrence matrix $\mathbf{Y}_{tra}$; $\mathbf{Y}_{tra}$: Occurrence matrix, containing PV power data and predictor data of the training period; z_h: Variable z at the target horizon h; $\tilde{z}_h$: Normalized variable z at the target horizon h; z_{min}: Minimum value registered for variable z; z_{max}: Maximum value registered for variable z; a_q : Generic q[th] quantile coverage level; $\hat{a}_q$: Generic q[th] estimated coverage; $\boldsymbol{\beta}^{\langle a_q \rangle}$: Vector of parameters of the QR model for the a_q-quantile; $\beta_m^{\langle a_q \rangle}$: Generic m[th] parameter of the QR model for the a_q-quantile; $\hat{\boldsymbol{\beta}}^{\langle a_q \rangle}$: Vector of estimated parameters of the QR model for the a_q-quantile; $\hat{\boldsymbol{\beta}}_m^{\langle a_q \rangle}$: Bootstrapped samples of the generic m[th] parameter of the QR model for the a_q-quantile; $\hat{\beta}_m^{\langle a_q \rangle}$: Generic m[th] estimated parameter of the QR model for the a_q-quantile; $\delta_{\mathbf{y}_{t_n}}$: Degenerate probability measure for the generic n[th] vector $\mathbf{y}_{t_n}$ of occurrences; τ_q^*: Optimized quantile coverage level of the sample quantile estimated from the bagged $\hat{P}_h^{\langle a_q \rangle}$; $\boldsymbol{\omega}$: Vector of weights assigned to occurrences; ω_{t_n}: Generic n[th] weight assigned to the vector $\mathbf{y}_{t_n}$ of occurrences; $\boldsymbol{\omega}^{\langle r \rangle}$: Vector of weights assigned to the occurrences in the rth bootstrap replicate; $\omega_{t_n}^{\langle r \rangle}$: Generic n[th] weight assigned to the vector $\mathbf{y}_{t_n}$ of occurrences in the rth bootstrap replicate

Authors' contributions
The authors contributed equally to the development of this research. All authors read and approved the final manuscript.

Authors' information
M. Bozorg (1986-), male, PHD and Associate Professor at the University of Applied Sciences Western Switzerland (HES-SO), Major in power systems, smart grids and active distribution networks.
A. Bracale (1974-), male, PHD and Associate Professor at the University of Naples Parthenope, Major in power quality, energy forecasting and analysis and control of distribution systems.
P. Caramia (1963-), male, PHD and Professor at the University of Naples Parthenope, Major in power quality and analysis and control of distribution systems.
G. Carpinelli (1953-), male, Professor at the University of Naples Federico II, Major in power quality, probabilistic methods applied to power systems and analysis and control of distribution systems.
M. Carpita (1959-), male, PHD and Professor at the University of applied sciences Western Switzerland (HES-SO), Major in power electronics systems for energy, drives and smart grids applications.
P. De Falco (1989-), male, PHD and Research Fellow at the University of Naples Parthenope, Major in energy forecasting and probabilistic methods applied to power systems.

Availability of data and materials
The PV power dataset #1 used in this study is available from the corresponding author on reasonable request. The corresponding NWP dataset is provided by ECMWF, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of ECMWF. The dataset #2 used in this study was published by the organizers of the Global Energy Forecasting Competition 2014, and it is available at https://doi.org/10.1016/j.ijforecast.2016.02.001.

Competing interests
The authors declare they have no competing interests.

Author details
[1]Institute of Energy and Electrical Systems, School of Management and Engineering Vaud, University of applied sciences Western Switzerland (HES-SO), Yverdon-les-Bains, Switzerland. [2]Department of Engineering, University of Naples Parthenope, Naples, Italy. [3]Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy.

References
1. Bessa, R. J., et al. (2017). Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry. *Energies, 10,* 1402.
2. Badal, F. R., Das, P., Sarker, S. K., & Das, S. K. (2019). A survey on control issues in renewable energy integration and microgrid. *Protection and Control of Modern Power Systems, 4*(8), 1–27.
3. Dobschinski, J., et al. (2017). Uncertainty forecasting in a nutshell: Prediction models designed to prevent significant errors. *IEEE Power and Energy Magazine, 15*(6), 40–49.
4. Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T. C., & Coimbra, C. F. M. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy, 168,* 60–101.
5. Sobri, S., Koohi-Kamali, S., & Rahim, N. A. (2018). Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management, 156,* 459–497.
6. van der Meer, D. W., Widén, J., & Munkhammar, J. (2018). Review on probabilistic forecasting of photovoltaic power production and electricity consumption. *Renewable and Sustainable Energy Reviews, 81,* 1484–1512.
7. Ren, Y., Suganthan, P. N., & Srikanth, N. (2015). Ensemble methods for wind and solar power forecasting—A state-of-the-art review. *Renewable and Sustainable Energy Reviews, 50,* 82–91.
8. Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., & Hyndman, R. J. (2016). Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting, 32*(3), 896–913.
9. Juban, R., Ohlsson, H., Maasoumy, M., Poirier, L., & Zico Kolter, J. (2016). A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014. *International Journal of Forecasting, 32*(3), 1094–1102.
10. Bracale, A., Carpinelli, G., & De Falco, P. (2019). Developing and comparing different strategies for combining probabilistic photovoltaic power forecasts in an ensemble method. *Energies, 12,* 1011.
11. Lauret, P., David, M., & Pedro, H. (2017). Probabilistic solar forecasting using quantile regression models. *Energies, 10*(10), 1591.
12. Pedro, H. T. C., Coimbra, C. F. M., David, M., & Lauret, P. (2018). Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. *Renewable Energy, 123,* 191–203.
13. Almeida, M. P., Perpiñán, O., & Narvarte, L. (2015). PV power forecast using a nonparametric PV model. *Solar Energy, 115,* 354–368.
14. Zhang, W., Quan, H., & Srinivasan, D. (2018). Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination. *Energy, 160,* 810–819.
15. Huang, J., & Perry, M. (2016). A semi-empirical approach using gradient boosting and k-nearest neighbors regression for GEFCom2014 probabilistic solar power forecasting. *International Journal of Forecasting, 32*(3), 1081–1086.
16. Bracale, A., Carpinelli, G., & De Falco, P. (2016). A probabilistic competitive ensemble method for short-term photovoltaic power forecasting. *IEEE Transactions on Sustainable Energy, 8*(2), 551–560.
17. Ni, Q., Zhuang, S., Sheng, H., Kang, G., & Xiao, J. (2017). An ensemble prediction intervals approach for short-term PV power forecasting. *Solar Energy, 155,* 1072–1083.
18. Mohammed, A. A., & Aung, Z. (2016). Ensemble learning approach for probabilistic forecasting of solar power generation. *Energies, 9,* 1017.
19. Mohammed, A. A., Yaqub, W., & Aung, Z. (2017). Probabilistic forecasting of solar power: An ensemble learning approach. In *International Conference on Intelligent Decision Technologies,* (pp. 449–458). Cham: Springer.
20. Rubin, D. B. (1981). The Bayesian bootstrap. *The Annals of Statistics, 9*(1), 130–134.
21. Bracale, A., Caramia, P., Carpinelli, G., Di Fazio, A., & Ferruzzi, G. (2013). A Bayesian method for short-term probabilistic forecasting of photovoltaic generation in smart grid operation and control. *Energies, 6*(2), 733–747.
22. Aryaputera, A. W., Verbois, H., & Walsh, W. M. (2016). Probabilistic accumulated irradiance forecast for Singapore using ensemble techniques. In *Proc. of IEEE 43rd Photovoltaic Specialists Conference (PVSC),* (pp. 1113–1118).
23. Bracale, A., Carpinelli, G., De Falco, P., Rizzo, R., & Russo, A. (2016). New advanced method and cost-based indices applied to probabilistic forecasting of photovoltaic generation. *Journal of Renewable and Sustainable Energy, 8*(2), 023505.
24. European Centre for Medium-range Weather Forecasts. https://www.ecmwf.int/ (Accessed on 30 Dec 2019).
25. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association, 102*(477), 359–378.
26. Liu, B., Nowotarski, J., Hong, T., & Weron, R. (2017). Probabilistic load forecasting via quantile regression averaging on sister forecasts. *IEEE Transactions on Smart Grid, 8*(2), 730–737.
27. Clyde, M. A., & Lee, H. K. H. (2001). Bagging and the Bayesian bootstrap. In T. Richardson, & T. Jaakkola (Eds.), *Artificial Intelligence and Statistics,* (pp. 169–174). New York: Elsevier.
28. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction,* (2nd ed.,) Springer Series in Statistics.
29. Koenker, R. (1994). Confidence intervals for regression quantiles. In P. Mandl, & M. Hušková (Eds.), *Asymptotic Statistics,* (pp. 349–359). Berlin Heidelberg: Springer-Verlag.

30. Cannon, A. J. (2018). Non-crossing nonlinear regression quantiles by monotone composite quantile regression neural network, with application to rainfall extremes. *Stochastic Environmental Research and Risk Assessment, 32*(11), 3207–3225.

31. Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis, 38*(4), 367–378.

32. R *qrnn* package: Quantile regression neural network. Available online: https://CRAN.R-project.org/package=qrnn (Accessed on 25 Mar 2020).

33. R *gbm* package: Generalized boosted regression models. Available online: https://CRAN.R-project.org/package=gbm (Accessed on 25 Mar 2020).

34. Alfieri, L., & De Falco, P. (2020). Wavelet-based decompositions in probabilistic load forecasting. *IEEE Transactions on Smart Grid, 11*(2), 1367–1376.

35. Carpita, M., Affolter, J., Bozorg, M., Houmard, D., & Wasterlain, S. (2019). Relne, a flexible laboratory for emulating and testing the distribution grid. In *Proc. of 21st European Conference on Power Electronics and Applications (EPE'19 ECCE Europe)*, (pp. 1–6).

36. R *quantreg* package: Quantile regression. Available online: https://CRAN.R-project.org/package=quantreg (Accessed on 30 Dec 2019).

37. R *bayesboot* package: An implementation of Rubin's (1981) Bayesian bootstrap. Available online: https://CRAN.R-project.org/package=bayesboot (Accessed on 30 Dec 2019).