DELIVERABLE DI PROGETTO DI RICERCA

DiGriFlex

Real-time Distribution Grid Control and Flexibility Provision under Uncertainties

RELAZIONE

SUB-ATTIVITA' OR2

Deliverable D2.3

Report tecnico sulla validazione dei sistemi di previsione, attraverso simulazioni numeriche di scenari e casi studio

(*Validation of the forecasting systems using appropriate test bench and scenarios in a numerical environment*)

Titolo del progetto: DiGriFlex: Real-time Distribution Grid Control and Flexibility Provision under Uncertainties

Pratica MIUR n.: ENSGPLUSREGSYS18_00016

Partners: Università di Napoli Federico II, Università di Napoli Parthenope

Relazione sulle attività di ricerca svolte nel periodo dal 01/09/2020 al 30/06/2022

1. INDEX

1. Index 2

2. Introduction 3

3. Objectives 4

4. Recalls on the theoretical background of the forecasting methodology 5

5. Results of activities for Tasks 2.4 11

5.1. Refinement and revisiting of the forecasting systems 11

5.2. Results of the validation stage: Background 12

5.3. Results of the validation stage: Day-Ahead Maximum PV Power Forecasting 16

5.4. Results of the validation stage: Real-Time Maximum PV Power Forecasting 18

5.5. Results of the validation stage: Day-Ahead Load Forecasting 19

5.6. Results of the validation stage: Real-Time Load Forecasting 21

6. Comparison with expected results 23

7. Deliverables 23

8. Profiles of human resources 23

9. Diffusion of the results 24

10. Conclusions 24

Bibliography 25

2. INTRODUCTION

This report is a Deliverable of the research project DiGriFlex (Real-Time Distribution Grid Control and Flexibility Provision under Uncertainties) and is related to the activities carried out in the framework of the work package WP2: "Development of appropriate day-ahead and real-time forecasting systems for renewable generation and loads". Particularly, the report refers to the activities developed as a result of Task 2.4 (Refinement and revisiting of the day ahead and real-time forecasting system with respect to the requirements of the optimization framework (WP2) and validation results (WP4))[1].

The activities included:

- the refinement and revisiting of the forecasting systems developed in Tasks 2.2 and 2.3. More specifically, the Bayesan Bootstrap Quantile Regression forecasting system has been adapted to suit the requirements related to the forecast lead time and time resolution, the limitations imposed on the computational time, to the limitations on the availability of input data.
- the integration of the forecasting systems, developed in MATLAB and R numerical environments, with the real-time acquisition measurement systems (working in LabVIEW environment) and with the two-level grid optimization model solver (developed in Python environment);
- the development of a forecasting system to predict also the maximum PV power that can be produced by a controlled PV system; note that this involves the prediction of the global solar irradiance, which is the weather variable that can be directly linked to the maximum producible PV power;
- the modification of the forecast lead times, in order to allow the real-time implementation of the forecast generation maintaining a sufficient margin to run the optimization model solver. In particular, the day-ahead forecasting systems were rearranged in order to be run at 18:00 of the day D-1, to generate the forecasts of the target variable for all the 144 ten-minute sub-intervals of the day D;
- the study for the selection of a single type of forecasting system to predict maximum PV power and load, in order to reduce the overall computational burden of the two-level grid optimization model solver;

[1] Attività AR 2.4 del Capitolato tecnico: *Rifinitura e rivisitazione dei sistemi di previsione in riferimento ai feedback forniti dal sistema di ottimizzazione a due livelli ed ai risultati della fase di validazione*

- the application of Bayesian bootstrap methodologies on underlying linear quantile regression models to obtain predictive quantiles of the maximum PV power and of the load active and reactive power for the day ahead predictions and a deterministic value (i.e., the conditional mean of the predictive distribution) of the maximum PV power and of the load active and reactive power in the case of the real-time forecast.
- the application of the developed forecasting systems to actual time series for day-ahead and real-time forecasting of the maximum PV power and the load (active and reactive power).

In the following Sections, the data used in the experiments, the benchmarks and the error indices for the performance evaluation are briefly presented, first. Then, the results of day-ahead and real-time maximum PV power forecasting are reported, based on the solar irradiance measurements taken by a dedicated weather station installed at the same location of the test distribution grid of the ReIne laboratory. Finally, the results of day-ahead and real-time load forecasting, in terms of active and reactive power are shown, based on the load data (active and reactive powers) collected at a college building in the same campus at which the ReIne laboratory is located.

3. OBJECTIVES

As described in the DiGriFlex project, the objective of the WP2 "Development of appropriate day-ahead and real-time forecasting systems for renewable generation and loads" is the development of an appropriate forecasting system for the prediction of uncertain parameters that suits two different scenarios: the day-ahead and the real-time forecasting. In the considered period the forecasting systems of renewable generation and loads were improved and refined, for the considered time horizons, to address the different needs in both scenarios. In order to cope with these different needs, the selected forecasting systems were based on non-parametric quantile regression methods, which proved to be flexible and versatile tools for PV power forecasting and load forecasting. The proposed methodologies were implemented in numerical environments for their application, including real measured data from the new fully configurable test distribution network at the HEIG-VD ReIne laboratory.

The activities of WP2 developed by the University of Naples Parthenope and the University of Naples Federico II, in the last period of the project, allowed reaching the objectives expected in Task 2.4 of the DiGriFlex project. The activities are related also to the definition of optimization requirements and the revisiting and refinement of the optimization tools.

In the work package WP2, the development of appropriate day-ahead and real-time forecasting systems for renewable generation and loads is carried out. In particular, it is focused on the forecasting systems of renewable generation and loads developed for different time horizons.

-

4. RECALLS ON THE THEORETICAL BACKGROUND OF THE FORECASTING METHODOLOGY

The proposed forecasting system is based on Bayesian Bootstrap Quantile Regression (BBQR). It is presented in this deliverable to predict the generic target variable y (e.g., the maximum PV power that can be produced by a controlled PV system or the power required by a load).

The structure of the proposed forecasting system is illustrated in Figure 4.1. The inputs of the forecasting system are the historical values $\boldsymbol{Y}$ of the target variable and historical values $\boldsymbol{EX}$ of exogenous variables.

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 Quantile Regression (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 $\boldsymbol{X}$ (i.e., independent variables in the QR model) which are informative for the target variable (i.e., the dependent variable in the QR model). Data are then partitioned into training datasets $\boldsymbol{Y}_{tra}$ and $\boldsymbol{X}_{tra}$ (a $1 \times T$ vector and a $T \times M$ matrix, respectively), and into validation datasets $\boldsymbol{Y}_{val}$ and $\boldsymbol{X}_{val}$ (a $1 \times V$ vector and a $V \times M$ matrix, respectively).

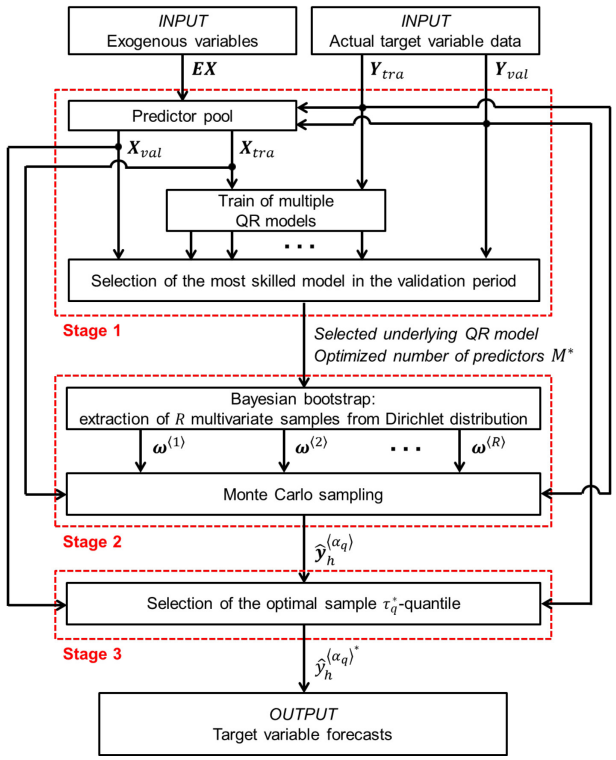

Fig. 4.1. Schematic procedure of the BBQR forecasting system.

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) 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 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 ($\hat{\boldsymbol{y}}_h^{\langle \alpha_q \rangle}$) of predictive α_q-quantiles of the target variable for the target horizon h.

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

The models and the stages of the forecasting system are recalled below.

1) Quantile regression modeling

A Quantile Regression (QR) model links the target variable (i.e., the predictive α_q-quantile $\hat{P}_h^{\langle\alpha_q\rangle}$ at the target time horizon h) to predictors $\boldsymbol{x}_h = \{x_{1_h}, \dots, x_{M_h}\}$ related to the time horizon h but available at the forecast origin $h-k$. The forecast lead time is indicated by k. This Deliverable focuses on a ten-minute time resolution, although the proposal can also be applied to other short-term frameworks. In order to lighten the notation, no reference to the forecast lead time k is reported in the symbols below, although the mathematical formulation is related to a specific lead time k.

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

$$\hat{y}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \sum_{m=1}^{M} \hat{\beta}_m^{\langle\alpha_q\rangle} \cdot x_{m_h}, \tag{4.1}$$

where $\widehat{\boldsymbol{\beta}}^{\langle\alpha_q\rangle} = \left\{\hat{\beta}_0^{\langle\alpha_q\rangle}, \dots, \hat{\beta}_M^{\langle\alpha_q\rangle}\right\}$ are the $M+1$ estimated values of model parameters $\boldsymbol{\beta}^{\langle\alpha_q\rangle} = \left\{\beta_0^{\langle\alpha_q\rangle}, \dots, \beta_M^{\langle\alpha_q\rangle}\right\}$. Note that (4.1) 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).

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 to evaluate the accuracy of probabilistic forecasts. The minimization problem is:

$$\widehat{\boldsymbol{\beta}}^{\langle\alpha_q\rangle} = \underset{\boldsymbol{\beta}^{\langle\alpha_q\rangle}}{\operatorname{argmin}}\, PS\left(\boldsymbol{Y}_{tra}, \widehat{\boldsymbol{Y}}_{tra}^{\langle\alpha_q\rangle}\right), \tag{4.2}$$

where $PS\left(\boldsymbol{Y}_{tra}, \widehat{\boldsymbol{Y}}_{tra}^{\langle\alpha_q\rangle}\right)$ is the PS of the T forecasts $\widehat{\boldsymbol{Y}}_{tra}^{\langle\alpha_q\rangle}$ issued for the training period of length T, calculated with respect to the actual occurrences of the target variable in the training set $\boldsymbol{Y}_{tra} = \{y_{t_1}, \dots, y_{t_T}\}$.

Although it is not directly explained in (4.2), the forecasts $\widehat{\boldsymbol{Y}}_{tra}^{\langle \alpha_q \rangle}$ are obtained from (4.1), 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:

$$\widehat{\boldsymbol{Y}}_{tra}^{\langle \alpha_q \rangle} = \boldsymbol{f}_{QR}\left(\boldsymbol{\beta}^{\langle \alpha_q \rangle} | \boldsymbol{X}_{tra}\right), \tag{4.3}$$

and therefore (4.2) can be rewritten in compact form as:

$$\widehat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle} = \boldsymbol{G}(\boldsymbol{Y}_{tra}, \boldsymbol{X}_{tra}), \tag{4.4}$$

where $\boldsymbol{G}(\cdot)$ is a function obtained by combining (4.2) and (4.3).

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. Exogenous data in $\boldsymbol{EX}$ and lagged target variable manipulated from $\boldsymbol{Y}$, together with their coupled interactions, form the pool of candidate predictors. A hypothesis is added to reduce the search dimension for the optimal model: if a coupled interaction is a predictor of the model, the two individual variables are forced to occur in the model.

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

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 to 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 $\widehat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle}$ from (4.2)-(4.4) on the bootstrapped data. Therefore, the posterior distribution of the QR model

parameters is given by R samples $\widehat{\boldsymbol{\beta}}_1^{\langle\alpha_q\rangle}, \dots, \widehat{\boldsymbol{\beta}}_{M^*+1}^{\langle\alpha_q\rangle}$ for each parameter, and from these samples it is eventually obtains the bagged samples $\widehat{\boldsymbol{Y}}_h^{\langle\alpha_q\rangle}$ of the target variable.

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

The $T \times (M^* + 1)$ occurrence matrix $\boldsymbol{Z}_{tra} = [\boldsymbol{Y}'_{tra} \ \ \boldsymbol{X}_{tra}]$ is initially obtained from the transposed training set $\boldsymbol{Y}'_{tra}$ and the corresponding $T \times M^*$ matrix $\boldsymbol{X}_{tra}$. The n^{th} row vector $\boldsymbol{z}_{t_n} = \{y_{t_n}, x_{1_{t_n}}, \dots, x_{M^*_{t_n}}\}$, taken from the occurrence matrix $\boldsymbol{Z}_{tra}$, 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 $F(\boldsymbol{z})$, with T available realizations (i.e., past occurrences) $\boldsymbol{z}_{t_1}, \dots, \boldsymbol{z}_{t_T}$.

As shown earlier, the estimated parameters $\widehat{\boldsymbol{\beta}}^{\langle\alpha_q\rangle}$ of the QR model come from (4.4), and therefore they can be viewed as a function of $\boldsymbol{G}[F(\boldsymbol{z})]$:

$$\widehat{\boldsymbol{\beta}}^{\langle\alpha_q\rangle} = \boldsymbol{G}[F(\boldsymbol{z})]. \tag{4.5}$$

In bootstrap (either traditional or Bayesian), the unknown distribution $F(\boldsymbol{z})$ is searched for among distributions of the type $F_T(\boldsymbol{z})$:

$$F_T(\boldsymbol{z}) = \sum_{n=1}^{T} \omega_{t_n} \cdot \delta_{\boldsymbol{z}_{t_n}}, \tag{4.6}$$

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

$$\sum_{n=1}^{T} \omega_{t_n} = 1, \ \omega_{t_n} \geq 0 \ \forall n = 1, \dots, T. \tag{4.7}$$

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

$$f_{Mul}(T; 1/T, 1/T, \dots, 1/T). \tag{4.8}$$

and normalizing by T.

The Bayesian bootstrap differs from the traditional bootstrap since the bootstrapped weights $\omega_{t_1}^{\langle r \rangle}, \dots, \omega_{t_T}^{\langle r \rangle}$ are not obtained by random extraction from distribution (4.8). Instead, the vector $\boldsymbol{\omega} = \{\omega_{t_1}, \dots, \omega_{t_T}\}$ is the object of Bayesian analysis, which aims at evaluating a posterior distribution $p(\boldsymbol{\omega}|\boldsymbol{Z}_{tra})$ of this vector of weights, given occurrence data $\boldsymbol{Z}_{tra}$. A prior distribution $p(\boldsymbol{\omega})$ should be imposed upon the parameters $\boldsymbol{\omega}$ to start the Bayesian inference. A convenient choice is to select a Dirichlet distribution, which is a conjugate prior for the multinomial distribution of $\boldsymbol{z}$. In such a case, the posterior distribution $p(\boldsymbol{\omega}|\boldsymbol{Z}_{tra})$ is itself a Dirichlet distribution, $f_{Dir}(1, \dots, 1; 1, \dots, 1)$. This allows applying a Monte Carlo sampling method to get the Bayesian bootstrapped samples $\widehat{\boldsymbol{\beta}}_1^{\langle \alpha_q \rangle}, \dots, \widehat{\boldsymbol{\beta}}_{M^*+1}^{\langle \alpha_q \rangle}$ of the $M^* + 1$ estimated parameters of the QR model. The steps are:

i) R multivariate samples $\boldsymbol{\omega}^{\langle 1 \rangle}, \dots, \boldsymbol{\omega}^{\langle R \rangle}$ are independently extracted from the Dirichlet distribution $f_{Dir}(1, \dots, 1; 1, \dots, 1)$;

ii) $\boldsymbol{G}\left[F_T^{\langle 1 \rangle}(\boldsymbol{z})\right], \dots, \boldsymbol{G}\left[F_T^{\langle R \rangle}(\boldsymbol{z})\right]$ are calculated applying (4.2)-(4.4);

iii) R Bayesian bootstrapped samples $\widehat{\boldsymbol{\beta}}_1^{\langle \alpha_q \rangle}, \dots, \widehat{\boldsymbol{\beta}}_{M^*+1}^{\langle \alpha_q \rangle}$ for each of the $M^* + 1$ parameters of the QR model are obtained using (4.5). From these samples, it obtains R samples of the predictive α_q-quantile $\hat{y}_h^{\langle \alpha_q \rangle}$ of the target variable by applying (4.1). The set of these samples is indicated with $\widehat{\boldsymbol{y}}_h^{\langle \alpha_q \rangle}$.

4) Third stage: Extraction of a single value from the Bayesian bootstrapped predictive samples

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

Probabilistic 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 research activity to reduce the redundancy of the forecasts by extracting a single value from the samples of the predictive α_q-quantile. This single value is treated as the final predictive α_q-quantile $\hat{y}_h^{\langle \alpha_q \rangle^*}$ of the target variable 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 $\hat{y}_h^{\langle \alpha_q \rangle^*}$ is the sample quantile extracted from $\widehat{\boldsymbol{y}}_h^{\langle \alpha_q \rangle}$ as:

$$\hat{y}_h^{\langle \alpha_q \rangle^*} = \hat{y}_h^{\langle \alpha_q \rangle \langle \tau_q^* \rangle}, \tag{4.9}$$

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 $\boldsymbol{Y}_{val}$, i.e.:

$$\tau_q^* = \underset{\tau_q}{\operatorname{argmin}}\, PS\left(\boldsymbol{Y}_{val}, \hat{\boldsymbol{Y}}_{val}^{\langle \alpha_q \rangle \langle \tau_q \rangle}\right). \tag{4.10}$$

Note that this procedure is made independent for each quantile coverage $\alpha_1, \alpha_2, \alpha_3, \ldots$, for simplicity. Therefore, possible quantile crossing in the final forecasts is corrected by post-processing the results with simple sorting in ascending order across the nominal coverages.

5. RESULTS OF ACTIVITIES FOR TASK 2.4

The Task 2.4 of the DiGriFlex project focuses on the refinement and the revisiting of the forecasting systems developed in the Tasks 2.2 and 2.3, based on the feedbacks provided by the developers of the optimization system that constitutes the core of the project, and on the report of the results of the validation stage.

5.1. REFINEMENT AND REVISITING OF THE FORECASTING SYSTEMS

With reference to the refinement and the revisiting of the forecasting systems developed in the Tasks 2.2 and 2.3, the BBQR forecasting system has been adapted to suit the requirements related to the forecast lead time and time resolution, to the limitations imposed on the computational time, to the limitations on the availability of input data.

In particular, a ten-minute time resolution is considered in the optimization stage, so all the forecasts must refer to the value of the target variable during a ten-minute time interval. In the remainder of the text, the term "interval" will refer to a ten-minute interval. In the day-ahead forecasting system to be implemented for the scheduling management procedure, the forecasts must be issued at 18:00 of a day for all the 144 intervals of the next day, with lead times k ranging from 37 to 180. In the real-time forecasting system to be implemented for the

real-time management procedure, the forecasts must be issued for the next interval but one (e.g., a forecast is issued at 16:30 for the interval 16:40÷16:50), thus with lead time $k = 2$.

Some limitations are imposed on the computational time in order to ensure that the forecasts are available and ready at the forecast origin. This problem is particularly important in the real-time forecasting problem, as there is only a ten-minute window to run and complete the codes that give the prediction. To comply with this problem, the number of bootstrap samples R is kept low (i.e., searched in the 10÷500 interval) in the real-time forecasting framework.

Eventually, another limitation is imposed by the unavailability of some data. In particular, it is not possible to dynamically retrieve Numeric Weather Predictions (NWPs) from external providers. Therefore, NWPs are not included as candidate predictors in the considered forecasting systems.

5.2. RESULTS OF THE VALIDATION STAGE: BACKGROUND

With reference to the report of the results of the validation stage, this Deliverable presents some numerical experiments for day-ahead and real-time forecasting of the maximum PV power and of the load (active and reactive power). All the forecasts are developed in the R environment, exploiting the packages *quantreg* and *bayesboot*.

The data used in the experiments, the benchmarks and the error indices for the performance evaluation are briefly presented in this sub-Section.

1) Data

Three datasets are used for the activities presented in this Deliverable. These datasets are related to the target variables (maximum PV power, load active power, load reactive power) and to the external variables (i.e., weather data) that are used as additional inputs of the forecasting systems.

Dataset_SI1: this dataset includes solar irradiance measurements taken by a dedicated weather station installed at the same location of the test distribution grid of the ReIne laboratory. The data has been continuously collected since August 24, 2019, with a one-minute time resolution. Data are eventually averaged to get a ten-minute time resolution. The solar irradiance data taken from this dataset are used as target data for the validation of the maximum PV power forecasting system.

Dataset_college_load: this dataset includes load data (active and reactive powers) collected at a college building in the same campus at which the ReIne laboratory is located. The building is open to the public (academic staff, students) during weekdays (Monday-Friday), and it is closed on Saturdays, Sundays and holidays. Energy meters collect the aggregate electrical data at fifteen-minute time resolution; the data have eventually been rescaled to ten-minute time resolution. These data are used for the validation of the load forecasting system.

Dataset_weath_PV: a dedicated weather station is installed at the location of the test distribution grid of the ReIne laboratory. This station collected weather data at the same time resolution and for the same time periods of the data in Dataset_SI1, allowing their usage as exogenous variables for the maximum PV power forecasting system.

2) Benchmark

Two naïve benchmarks are considered as unbiased references for comparative purposes.

The Persistence Model (PM) is based on the invariance hypothesis of the target variable throughout the forecast lead time k. In practice, the PM prediction of the target variable y at the horizon h is equal to the last observed value; the PM deterministic prediction is:

$$\hat{y}_h = y_{h-k}, \tag{5.2.1}$$

whereas the PM probabilistic prediction in terms of α_q-quantile is:

$$\hat{y}_h^{\langle \alpha_q \rangle} = y_{h-k}. \tag{5.2.2}$$

The structure of the PM benchmark is inherently more indicated to be used for comparison in the real-time forecasting framework, as its error grows as the lead time increases.

The Seasonal Persistence Model (SPM) is based on the underlying periodicity that some target variables can show. For example, a PV power pattern is driven by the rotation of the Earth around its own axis and therefore shows a daily periodicity. Since this deliverable focuses on the validation of a maximum PV power forecasting system and of a load forecasting system, the considered periodicity is one day in both cases. In practice, the SPM prediction of the target variable y at the horizon h is equal to the last observed value in a homologous interval. Given the ten-minute resolution and the lead times of the presented forecasting framework, the deterministic SPM prediction is:

$$\hat{y}_h = \begin{cases} P_{h-144} & \text{if } k \leq 144 \\ P_{h-288} & \text{if } k > 144 \end{cases} \quad (5.2.3)$$

whereas the SPM probabilistic prediction in terms of α_q-quantile is:

$$\hat{y}_h^{\langle \alpha_q \rangle} = \begin{cases} P_{h-144} & \text{if } k \leq 144 \\ P_{h-288} & \text{if } k > 144 \end{cases} \quad (5.2.4)$$

The structure of the SPM benchmark is indicated to be used for comparison both in the day-ahead and in the real-time forecasting frameworks.

3) Error indices

Several error indices are used to quantify the accuracy of the forecasts.

The first two indices are the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE), used to assess deterministic forecasts. Indicating with y_h the actual value of the target variable at horizon h and with H the total number of horizons for which forecasts are issued, the MAE and the RMSE are, respectively:

$$MAE = \frac{1}{H}\sum_{h=1}^{H}|y_h - \hat{y}_h|, \quad (5.2.5)$$

$$RMSE = \sqrt{\frac{1}{H}\sum_{h=1}^{H}(y_h - \hat{y}_h)^2}. \quad (5.2.6)$$

The MAE and the RMSE are negatively oriented, so smaller values indicate more accurate forecasts.

The Normalized values of the MAE (i.e., NMAE) and of the RMSE (i.e., NRMSE) are, respectively:

$$NMAE = \frac{MAE}{\bar{y}}, \quad (5.2.7)$$

$$NRMSE = \frac{NRMSE}{\bar{y}}, \quad (5.2.8)$$

where the normalization term $\bar{y}$ is a fixed value (e.g., the rated power of the installation). The NMAE and the NRMSE are negatively oriented, too.

The third index is the abovementioned PS, which is a strictly proper score that simultaneously addresses the reliability and the sharpness of forecasts. Its formulation is:

$$PS_h\left(y_h, \hat{y}_h^{\langle\alpha_q\rangle}\right) = \left[\alpha_q - \mathrm{I}\left(y_h, \hat{y}_h^{\langle\alpha_q\rangle}\right)\right] \cdot \left(y_h - \hat{y}_h^{\langle\alpha_q\rangle}\right), \tag{5.2.9}$$

where the indicator function $\mathrm{I}\left(y_h, \hat{y}_h^{\langle\alpha_q\rangle}\right)$ is:

$$\mathrm{I}\left(y_h, \hat{y}_h^{\langle\alpha_q\rangle}\right) = \begin{cases} 1 & \text{if } y_h \leq \hat{y}_h^{\langle\alpha_q\rangle} \\ 0 & \text{if } y_h > \hat{y}_h^{\langle\alpha_q\rangle} \end{cases} \tag{5.2.10}$$

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

The Normalized value of the PS (i.e., NPS) is:

$$NPS_h = \frac{PS_h}{\bar{y}}, \tag{5.2.11}$$

where the normalization term $\bar{y}$ is a fixed value (e.g., the rated power of the installation). The NPS is negatively oriented, too.

The fourth 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. Because of its intrinsic properties, it can only be formulated for multiple forecast issues. It is calculated from the estimated α_q-coverage $\hat{\alpha}_q$, that is:

$$\hat{\alpha}_q = \frac{1}{H} \cdot \sum_{h=1}^{H} \mathrm{I}\left(y_h, \hat{y}_h^{\langle\alpha_q\rangle}\right), \tag{5.2.12}$$

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

$$ACE^{\langle\alpha_q\rangle} = \left|\alpha_q - \hat{\alpha}_q\right|. \tag{5.2.13}$$

The AACE across N_q coverages can be obtained as a percentage value ($AACE_{\%}$) as:

$$AACE_{\%} = \frac{100}{N_q} \cdot \sum_{q=1}^{N_q} ACE^{\langle \alpha_q \rangle}. \quad (5.2.14)$$

The $AACE_{\%}$ is negatively oriented, so a smaller AACE indicates more reliable forecasts.

5.3. RESULTS OF THE VALIDATION STAGE: DAY-AHEAD MAXIMUM PV POWER FORECASTING

This sub-Section presents the results of day-ahead maximum PV power forecasting, based on the **Dataset_SI1** and on the **Dataset_weath_PV**. In particular, the forecasting system is set to provide predictive quantiles, at nominal coverages $\alpha_1, \alpha_2, \alpha_3 = 0.2, 0.5, 0.8$, of the maximum PV power P. As stated above, the forecasts are issued at 18:00 of a day for all the 144 intervals of the next day, and no NWP is included as candidate predictor.

The selected underlying model of the BBQR forecasting system is either:

$$\hat{P}_h^{\langle \alpha_q \rangle} = \hat{\beta}_0^{\langle \alpha_q \rangle} + \hat{\beta}_1^{\langle \alpha_q \rangle} \cdot at_{h-144} + \hat{\beta}_2^{\langle \alpha_q \rangle} \cdot pr_{h-144} + \hat{\beta}_3^{\langle \alpha_q \rangle} \cdot P_{h-144} + \\ + \hat{\beta}_4^{\langle \alpha_q \rangle} \cdot \overline{at}_{D1_h} + \hat{\beta}_5^{\langle \alpha_q \rangle} \cdot \overline{pr}_{D1_h} + \hat{\beta}_6^{\langle \alpha_q \rangle} \cdot \bar{P}_{D1_h} + \hat{\beta}_7^{\langle \alpha_q \rangle} \cdot P_{h-144} \cdot \bar{P}_{D1_h}, \quad (5.3.1)$$

or:

$$\hat{P}_h^{\langle \alpha_q \rangle} = \hat{\beta}_0^{\langle \alpha_q \rangle} + \hat{\beta}_1^{\langle \alpha_q \rangle} \cdot at_{h-288} + \hat{\beta}_2^{\langle \alpha_q \rangle} \cdot pr_{h-288} + \hat{\beta}_3^{\langle \alpha_q \rangle} \cdot P_{h-288} + \\ + \hat{\beta}_4^{\langle \alpha_q \rangle} \cdot \overline{at}_{D2_h} + \hat{\beta}_5^{\langle \alpha_q \rangle} \cdot \overline{pr}_{D2_h} + \hat{\beta}_6^{\langle \alpha_q \rangle} \cdot \bar{P}_{D2_h} + \hat{\beta}_7^{\langle \alpha_q \rangle} \cdot P_{h-288} \cdot \bar{P}_{D2_h}, \quad (5.3.2)$$

where:

- at_h and pr_h are the ambient temperature and the air pressure at time h;
- $\overline{at}_D$, $\overline{pr}_D$, and $\bar{P}_D$ respectively indicate the average ambient temperature, average pressure and average maximum PV power recorded during the day D;
- $D1_h$ and $D2_h$ respectively indicate the day before and two days before the day that includes the horizon interval h.

The model (5.3.1) is used to forecast the maximum PV power for the first 108 intervals of the target day, whereas the model (5.3.2) is used to forecast the maximum PV power for the last 36 intervals of the target day. In both cases, the number of predictors of the selected model is $M^* = 7$.

Table 5.3.1 shows the NPS and the $AACE_{\%}$ indices calculated for the predictive quantiles obtained through the BBQR and the SPM through a two-week test period (2016 forecast issues), characterized by variable and adverse

weather conditions. As seen, BBQR returns an NPS that is about 37% smaller than the benchmark. Figure 5.3.1 shows the maximum PV power pattern and the forecasted values, with the 60% prediction interval extracted from the predictive 0.2-quantile and from the predictive 0.8-quantile, through three days of the test period. Figure 5.3.2 shows the reliability diagram evaluated through the test period.

Table 5.3.1 Day-ahead maximum PV power forecasting results through two weeks (2016 forecast issues)

Method	NPS [-]	$AACE_{\%}$ [%]
BBQR	0.0220	4.113
SPM	0.0306	-

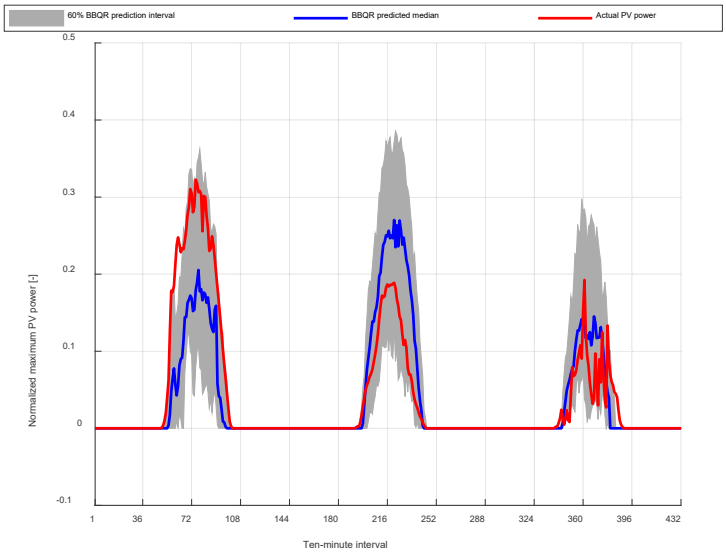

Fig. 5.3.1. BBQR day-ahead maximum PV power predictions through three days.

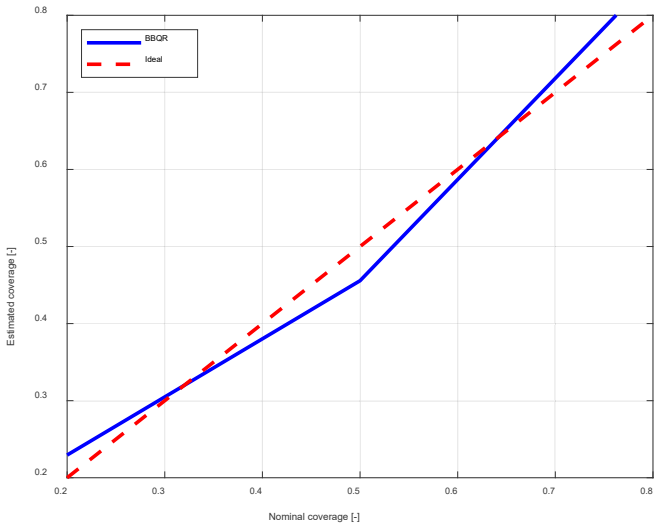

Fig. 5.3.2. Reliability diagram of the maximum PV power predictions through two weeks.

5.4. RESULTS OF THE VALIDATION STAGE: REAL-TIME MAXIMUM PV POWER FORECASTING

This sub-Section presents the results of real-time maximum PV power forecasting, based on the **Dataset_SI1** and on the **Dataset_weath_PV**. In particular, the forecasting system is set to provide a deterministic value (i.e., the conditional mean of the predictive distribution) of the maximum PV power P. As stated above, the forecasts are issued for the next interval but one, and no NWP is included as candidate predictor.

The selected underlying model of the BBQR forecasting system is:

$$\hat{P}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \hat{\beta}_1^{\langle\alpha_q\rangle} \cdot P_{h-144} + \hat{\beta}_2^{\langle\alpha_q\rangle} \cdot P_{h-2} + \hat{\beta}_3^{\langle\alpha_q\rangle} \cdot P_{h-3} + \hat{\beta}_4^{\langle\alpha_q\rangle} \cdot P_{h-4} + \\ + \hat{\beta}_5^{\langle\alpha_q\rangle} \cdot P_{h-5} + \hat{\beta}_6^{\langle\alpha_q\rangle} \cdot P_{h-6} + \hat{\beta}_7^{\langle\alpha_q\rangle} \cdot P_{h-144} \cdot P_{h-2} + \hat{\beta}_8^{\langle\alpha_q\rangle} \cdot P_{h-6} \cdot P_{h-2}. \quad (5.4.1)$$

The number of predictors of the selected model is $M^* = 8$.

Table 5.4.1 shows the NMAE and the NRMSE indices calculated for the forecasts obtained through the BBQR and the PM through a two-week test period (2016 forecast issues), characterized by variable and adverse weather conditions. As seen, BBQR returns a NMAE that is about 12% smaller than the benchmark and a NRMSE that is about 10% smaller than the benchmark. Figure 5.4.1 shows the maximum PV power pattern and the forecasted values through three days of the test period.

Table 5.4.1 Real-time maximum PV power forecasting results through two weeks (2016 forecast issues)

Method	NMAE [-]	NRMSE [-]
BBQR	0.009	0.026
PM	0.010	0.029

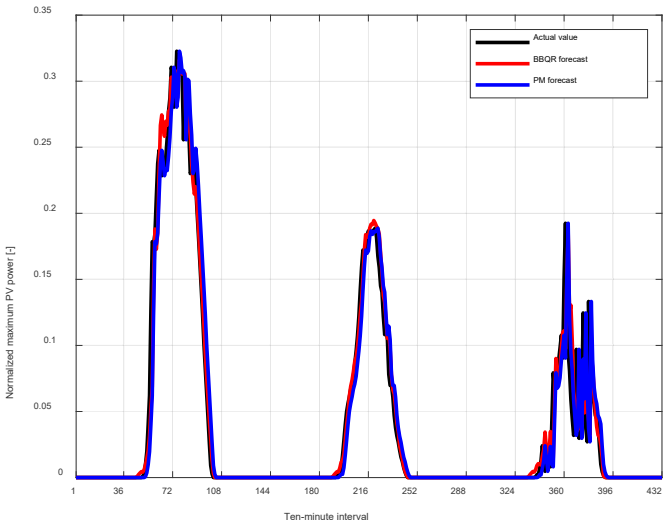

Fig. 5.4.1. BBQR real-time maximum PV power predictions through three days.

5.5. RESULTS OF THE VALIDATION STAGE: DAY-AHEAD LOAD FORECASTING

This sub-Section presents the results of day-ahead load forecasting, in terms of active and reactive power, based on the **Dataset_college_load**. In particular, the forecasting system is set to provide predictive quantiles, at nominal coverages $\alpha_1, \alpha_2, \alpha_3 = 0.1, 0.5, 0.9$, of the load active power P and of the load reactive power Q. As stated above, the forecasts are issued at 18:00 of a day for all the 144 intervals of the next day.

The selected underlying model of the BBQR active power forecasting system is either:

$$\hat{P}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \hat{\beta}_1^{\langle\alpha_q\rangle} \cdot P_{h-144} + \hat{\beta}_2^{\langle\alpha_q\rangle} \cdot P_{h-145} + \hat{\beta}_3^{\langle\alpha_q\rangle} \cdot P_{h-1008} + \\ + \hat{\beta}_4^{\langle\alpha_q\rangle} \cdot P_{h-144} \cdot P_{h-145} + \hat{\beta}_5^{\langle\alpha_q\rangle} \cdot P_{h-144} \cdot P_{h-1008}, \quad (5.5.1)$$

or:

$$\hat{P}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \hat{\beta}_1^{\langle\alpha_q\rangle} \cdot P_{h-288} + \hat{\beta}_2^{\langle\alpha_q\rangle} \cdot P_{h-289} + \hat{\beta}_3^{\langle\alpha_q\rangle} \cdot P_{h-1008} + \\ + \hat{\beta}_4^{\langle\alpha_q\rangle} \cdot P_{h-288} \cdot P_{h-289} + \hat{\beta}_5^{\langle\alpha_q\rangle} \cdot P_{h-288} \cdot P_{h-1008}, \quad (5.5.2)$$

whereas the selected underlying model of the BBQR reactive power forecasting system is either:

$$\hat{Q}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \hat{\beta}_1^{\langle\alpha_q\rangle} \cdot Q_{h-144} + \hat{\beta}_2^{\langle\alpha_q\rangle} \cdot Q_{h-145} + \hat{\beta}_3^{\langle\alpha_q\rangle} \cdot Q_{h-1008} + \\ + \hat{\beta}_4^{\langle\alpha_q\rangle} \cdot Q_{h-144} \cdot Q_{h-145} + \hat{\beta}_5^{\langle\alpha_q\rangle} \cdot Q_{h-144} \cdot Q_{h-1008}, \quad (5.5.3)$$

or:

$$\hat{Q}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \hat{\beta}_1^{\langle\alpha_q\rangle} \cdot Q_{h-288} + \hat{\beta}_2^{\langle\alpha_q\rangle} \cdot Q_{h-289} + \hat{\beta}_3^{\langle\alpha_q\rangle} \cdot Q_{h-1008} + \\ + \hat{\beta}_4^{\langle\alpha_q\rangle} \cdot Q_{h-288} \cdot Q_{h-289} + \hat{\beta}_5^{\langle\alpha_q\rangle} \cdot Q_{h-288} \cdot Q_{h-1008}, \quad (5.5.4)$$

The models (5.5.1) and (5.5.3) are used to forecast the active and reactive power for the first 108 intervals of the target day, whereas the models (5.5.2) and (5.5.4) are used to forecast the active and reactive for the last 36 intervals of the target day. In all cases, the number of predictors of the selected model is $M^* = 5$.

Table 5.5.1 shows the PS and the $AACE_{\%}$ indices calculated for the predictive quantiles obtained through the BBQR and the SPM through a two-week test period (2016 forecast issues). As seen, BBQR returns an active power PS that is about 48% smaller than the benchmark, and a reactive power PS that is about 46% smaller than the benchmark. Figure 5.5.1 shows the active power pattern versus the forecasted values (Figure 5.5.1.a) and the reactive power pattern versus the forecasted values (Figure 5.5.1.b), with the 80% prediction interval extracted from the predictive 0.1-quantile and from the predictive 0.9-quantiles, through three days of the test period. Figure 5.5.2 shows the reliability diagram evaluated through the test period for the active power forecasts (Figure 5.5.2.a) and for the reactive power forecasts (Figure 5.5.2.b).

Table 5.5.1 Day-ahead load active and reactive power forecasting results through two weeks (2016 forecast issues)

Method	Active power		Reactive power	
	PS [kW]	$AACE_{\%}$ [%]	PS [kVAr]	$AACE_{\%}$ [%]
BBQR	4.211	3.902	0.586	5.030
SPM	8.516	-	1.100	-

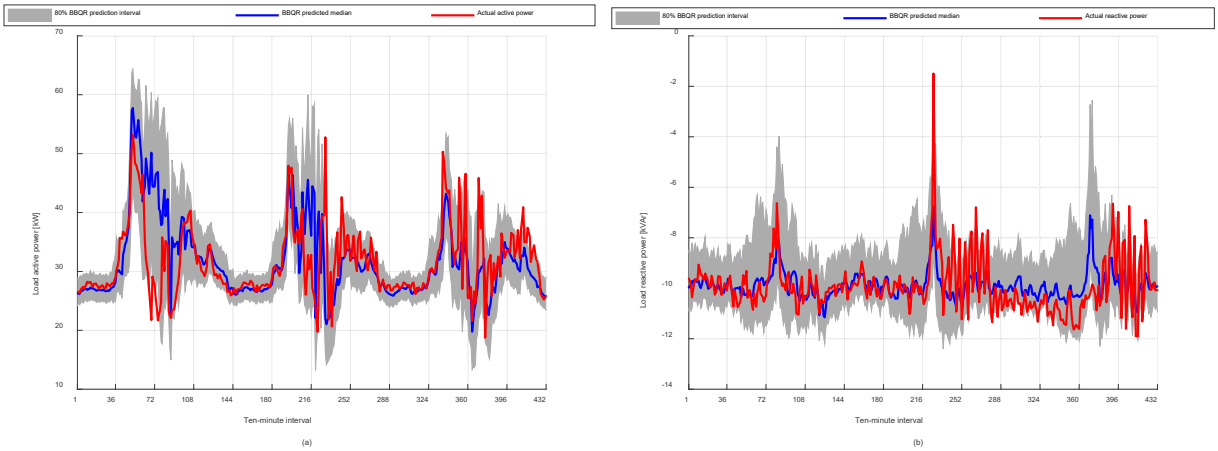

Fig. 5.5.1. BBQR day-ahead load active (a) and reactive (b) predictions through three days.

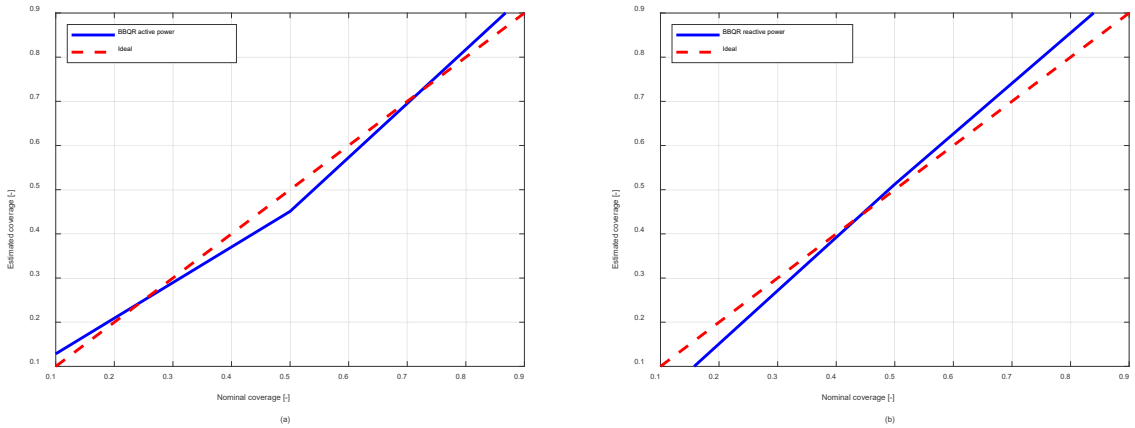

Fig. 5.5.2. Reliability diagram of the load active (a) and reactive (b) power predictions through two weeks.

5.6. RESULTS OF THE VALIDATION STAGE: REAL-TIME LOAD FORECASTING

This sub-Section presents the results of real-time load forecasting, in terms of active and reactive power, based on the **Dataset_college_load**. In particular, the forecasting system is set to provide a deterministic value (i.e., the conditional mean of the predictive distribution) of the load active power P and of the load reactive power Q. As stated above, the forecasts are issued for the next interval but one.

The selected underlying model of the BBQR active power forecasting system is:

$$\hat{P}_h^{\langle \alpha_q \rangle} = \hat{\beta}_0^{\langle \alpha_q \rangle} + \hat{\beta}_1^{\langle \alpha_q \rangle} \cdot P_{h-144} + \hat{\beta}_2^{\langle \alpha_q \rangle} \cdot P_{h-2} + \hat{\beta}_3^{\langle \alpha_q \rangle} \cdot P_{h-3} + \hat{\beta}_4^{\langle \alpha_q \rangle} \cdot P_{h-4} +$$

$$+\hat{\beta}_5^{\langle\alpha_q\rangle}\cdot P_{h-5}+\hat{\beta}_6^{\langle\alpha_q\rangle}\cdot P_{h-6}+\hat{\beta}_7^{\langle\alpha_q\rangle}\cdot P_{h-144}\cdot P_{h-2}+\hat{\beta}_8^{\langle\alpha_q\rangle}\cdot P_{h-6}\cdot P_{h-2}, \tag{5.6.1}$$

whereas the selected underlying model of the BBQR reactive power forecasting system is:

$$\hat{Q}_h^{\langle\alpha_q\rangle}=\hat{\beta}_0^{\langle\alpha_q\rangle}+\hat{\beta}_1^{\langle\alpha_q\rangle}\cdot Q_{h-144}+\hat{\beta}_2^{\langle\alpha_q\rangle}\cdot Q_{h-2}+\hat{\beta}_3^{\langle\alpha_q\rangle}\cdot Q_{h-3}+\hat{\beta}_4^{\langle\alpha_q\rangle}\cdot Q_{h-4}+$$
$$+\hat{\beta}_5^{\langle\alpha_q\rangle}\cdot Q_{h-5}+\hat{\beta}_6^{\langle\alpha_q\rangle}\cdot Q_{h-6}+\hat{\beta}_7^{\langle\alpha_q\rangle}\cdot Q_{h-144}\cdot Q_{h-2}+\hat{\beta}_8^{\langle\alpha_q\rangle}\cdot Q_{h-6}\cdot Q_{h-2}, \tag{5.6.2}$$

In both cases, the number of predictors of the selected models is $M^*=8$.

Table 5.6.1 shows the NMAE and the NRMSE indices calculated for the load active and reactive power forecasts obtained through the BBQR and the PM through a two-week test period (2016 forecast issues). As seen, BBQR returns an active power MAE that is about 4% smaller than the benchmark, an active power RMSE that is about 6% smaller than the benchmark, a reactive power MAE that is about 21% smaller than the benchmark, and a reactive power RMSE that is about 29% smaller than the benchmark. Figure 5.6.1 shows the active power pattern versus the forecasted values (Figure 5.6.1.a) and the reactive power pattern versus the forecasted values (Figure 5.6.1.b), through three days of the test period.

Table 5.6.1 Real-time load active and reactive power forecasting results through two weeks (2016 forecast issues)

Method	Active power		Reactive power	
	MAE [kW]	RMSE [kW]	MAE [kVAr]	RMSE [kVAr]
BBQR	2.017	3.232	0.450	0.646
SPM	2.103	3.427	0.567	0.908

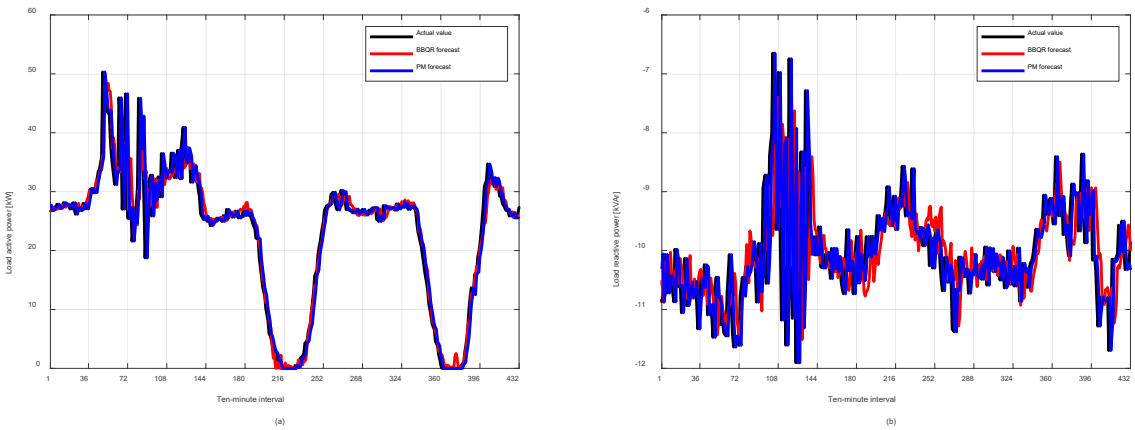

Fig. 5.6.1. BBQR real-time load active (a) and reactive (b) predictions through three days.

6. COMPARISON WITH EXPECTED RESULTS

With reference to the Task 2.4 of the Digriflex project, the expected results at the end of the Milestone include: Refinement and revisiting of day ahead forecasting system and real time forecasting system with respect to the requirements of the optimization framework (WP2) and validation results (WP4).

The activities were carried out by the University of Naples Parthenope and the University of Naples Federico II in strict interaction, and with the partner HEIG-VD for the acquisition of measurements available at the test distribution network at the ReIne laboratory.

The activities related to these tasks allowed giving a further confirmation of the good performance of the proposed methods with respect to the feedback on the requirements of the optimization framework.

The activities developed, illustrated above, allowed to fully reach the objectives expected in Task 2.4 of the DiGriFlex project.

7. DELIVERABLES

The deliverables of the research activities related to Task 2.4 consist of :

- this technical document, as expected per the Technical Specifications of the DiGriFlex Project;
- a participation with a poster presented during the 20th International Conference on Harmonics and Quality of Power 2022 (ICHQP 2022), Naples (Italy), 29 May – 2 June.
- a conference paper titled *Real-Time Distribution Grid Control and Flexibility Provision Under Uncertainties: Laboratory Demonstration*, Melecon 2022, Palermo (Italy) 14-16 June 2022.

8. PROFILES OF HUMAN RESOURCES

The researchers of the University of Naples Parthenope who participated at the research activities related to Task 2.4 are:

- Pierluigi Caramia, Full Professor of Power Systems
- Antonio Bracale, Associate Professor of Power Systems
- Pasquale De Falco, Assistant Professor of Power Systems

The researchers of the University of Naples Federico II who participated at the research activities related to Task 2.4 are:

- Daniela Proto, Associate Professor of Power Systems
- Fabio Mottola, Assistant Professor of Power Systems.

9. DIFFUSION OF THE RESULTS

The results of the research activities related to Task 2.4 were propaedeutic to the preparation of the following scientific paper:

- Mohammad Rayati, Mokhtar Bozorg, Mauro Carpita, Pasquale De Falco, Pierluigi Caramia, Antonio Bracale, Daniela Proto, Fabio Mottola: Real-Time Distribution Grid Control and Flexibility Provision Under Uncertainties: Laboratory Demonstration, Melecon 2022, Palermo (Italy) 14-16 June 2022.

Also, a poster including the activities of this and of other tasks of the project has been presented during the 20th International Conference on Harmonics and Quality of Power - ICHQP 2022, May 29th- June 1st 2022, Naples (Italy).

10. CONCLUSIONS

The research activities discussed in this document focused on the validation of day-ahead and real-time forecasting systems for renewable generation and loads. The forecasting systems have been developed by exploiting ensemble techniques within probabilistic frameworks.

The major contributions of the research activities related to Task 2.4 can be summarized as follows:

- Refinement and revisiting of the forecasting systems
- Results of the validation stage, including:
    - Day-Ahead Maximum PV Power Forecasting
    - Real-Time Maximum PV Power Forecasting
    - Day-Ahead Load Forecasting
    - Real-Time Load Forecasting

The above activities allowed fully reaching the objectives expected in Task 2.4 of the DiGriFlex project and providing the refined tools for the activities of WP4.

BIBLIOGRAPHY

1) A. Muñoz, E.F. Sánchez-Úbeda, A. Cruz, and J. Marín, "Short-term forecasting in power systems: a guided tour," in Handbook of power systems II, Springer, 2010.
2) H. Jiayi, J. Chuanwen, and X. Rong, "A review on distributed energy resources and MicroGrid," Renewable and Sustainable Energy Reviews, vol. 12, no. 9, pp. 2472-2483, 2008.
3) J. Xie, T. Hong, and J. Stroud, "Long-Term Retail Energy Forecasting With Consideration of Residential Customer Attrition," IEEE Transactions on Smart Grid, vol. 6, no. 5, pp. 2245-2252, 2015.
4) L. Hernández, C. Baladrón, J.M. Aguiar, B. Carro, A. Sánchez-Esguevillas, and J. Lloret, "Artificial neural networks for short-term load forecasting in microgrids environment," Energy, vol. 75, 2014.
5) A. Carpinone, M. Giorgio, R. Langella, and A. Testa, "Markov chain modeling for very-short-term wind power forecasting," Electric Power Systems Research, vol. 122, pp. 152-158, 2015.
6) T. Hong, and P. Wang, "Fuzzy interaction regression for short term load forecasting," Fuzzy optimization and decision making, vol. 13, no. 1, pp. 91-103, 2014.
7) F. Barbieri, S. Rajakaruna, and A. Ghosh, "Very short-term photovoltaic power forecasting with cloud modeling: A review," Renewable and Sustainable Energy Reviews, vol. 75, pp. 242-263, 2017.
8) T. Hong, and S. Fan, "Probabilistic electric load forecasting: A tutorial review," International Journal of Forecasting, vol. 32, no. 3, 2016.
9) T. Jónsson, P. Pinson, and H. Madsen, "On the market impact of wind energy forecasts," Energy Economics, vol. 32, no. 2, pp. 313-320, 2010.
10) Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," Renewable and Sustainable Energy Reviews, vol. 32, 2014.
11) M. Zugno, T. Jónsson, and P. Pinson, "Trading wind energy on the basis of probabilistic forecasts both of wind generation and of market quantities," Wind Energy, vol. 16, no. 6, pp. 909-926, 2013.
12) P. Pinson, C. Chevallier, and G.N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," IEEE Transactions on Power Systems, vol. 22, no. 3, pp. 1148-1156, 2007.
13) A. Bracale, P. Caramia, G. Carpinelli, A.R. Di Fazio, and G. Ferruzzi, "A Bayesian method for short-term probabilistic forecasting of photovoltaic generation in smart grid operation and control," Energies, vol. 6, no. 2, pp. 733-747, 2013.
14) B. Liu, J. Nowotarski, T. Hong, and R. Weron, "Probabilistic Load Forecasting via Quantile Regression Averaging on Sister Forecasts," IEEE Transactions on Smart Grid, vol. 8, no. 2, pp. 730-737, 2017.

15) T. Hong, "Energy forecasting: past, present and future," Foresight: The International Journal of Applied Forecasting, no. 32, pp. 43–48, 2014.

16) H. L. Willis and J.E. Northcote-Green, "Spatial electric load forecasting: a tutorial review," Proceedings of the IEEE, vol. 71, no. 2, pp.232–253, 1983.

17) H.S. Hippert, C.E. Pedreira, and R.C. Souza, "Neural networks forshort-term load forecasting: A review and evaluation," IEEE Transactions on Power Systems, vol. 16, no. 1, pp. 44–55, 2001.

18) H. Zhao and F. Magoules, "A review on the prediction of building energy consumption," Renewable and Sustainable Energy Reviews, vol. 16, no. 6, pp. 3586–3592, 2012.

19) S.S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," IEEE North American Power Symposium 2010, pp. 1–8, 2010.

20) 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,pp. 1725–1744, 2008.

21) 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.

22) P. Pinson, "Wind energy: Forecasting challenges for its operational management," Statistical Science, vol. 28, no. 4, pp. 564–585, 2013.

23) C. Sweeney, R. Bessa, J. Browell, and P. Pinson, "The future of forecasting for renewable energy," WIREs Energy and Environment, vol. 9, p. e365, 2020.

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

25) J. M. Bates and C. W. J. Granger, "The combination of forecasts," Journal of the Operational Research Society, vol. 20, no. 4, pp. 451–468, 1969.

26) R.T. Clemen, "Combining forecasts: A review and annotated bibliography," International Journal of Forecasting, vol. 5, no. 4, pp. 559–583,1989.

27) J.S. Armstrong, "Combining forecasts," in Principles of forecasting. Springer, 2001, pp. 417–439.

28) K.F. Wallis, "Combining forecasts – forty years later," Applied Financial Economics, vol. 21, no. 1–2, pp. 33–41, 2011.

29) J. Nowotarski, B. Liu, R. Weron, and T. Hong, "Improving short term load forecast accuracy via combining sister forecasts," Energy, vol. 98, pp. 40–49, 2016.

30) B. Liu, J. Nowotarski, T. Hong, and R. Weron, "Probabilistic load forecasting via Quantile Regression Averaging on sister forecasts," IEEE Trans. Smart Grid, vol. 8, pp. 730–737, 2017.

31) Y. Wang, Q. Chen, M. Sun, C. Kang, and Q. Xia, "An ensemble forecasting method for the aggregated load with subprofiles," IEEE Trans. Smart Grid, vol. 9, no. 4, pp. 3906–3908, 2018.

32) A. Bracale, G. Carpinelli, P. De Falco, "A probabilistic competitive ensemble method for short-term photovoltaic power forecasting," IEEE Transactions on Sustainable Energy, vol. 8, no. 2, pp. 551-560, 2017.

33) A. Bracale, G. Carpinelli, P. De Falco, "Developing and comparing different strategies for combining probabilistic photovoltaic power forecasts in an ensemble method," Energies, vol. 12, no. 6, 1011, 2019.

34) M. Roulston and L. Smith, "Combining dynamical and statistical ensembles," Tellus A: Dynamic Meteorology and Oceanography, vol. 55, no. 1, pp. 16–30, 2003.

35) R.J. Hyndman, R.A. Ahmed, G. Athanasopoulos, and H. L. Shang, "Optimal combination forecasts for hierarchical time series," Computational Statistics & Data Analysis, vol. 55, no 9, pp. 2579–2589, 2011.

36) R. J. Hyndman, A. J. Lee, and E. Wang, "Fast computation of reconciled forecasts for hierarchical and grouped time series," Computational Statistics & Data Analysis, vol. 97, pp. 16–32, 2016

37) D. Yang, "Reconciling solar forecasts: Probabilistic forecast reconciliation in a nonparametric framework," Solar Energy, in press, 2020.

38) D. Yang, H. Quan, V.R. Disfani, and L. Liu, "Reconciling solar forecasts: Geographical hierarchy," Solar Energy, vol. 146, pp. 276–286, 2017.

39) T. Hong, J. Xie, and J. Black, "Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting," International Journal of Forecasting, vol. 35, no. 4, pp. 1389–1399, 2019.

40) T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," International Journal of Forecasting, vol. 32, no. 3, pp. 896–913, 2016.

41) S. Sobri, S. Koohi-Kamali, and N.A. Rahim, "Solar photovoltaic generation forecasting methods: A review," Energy Conversion and Management, vol. 156, pp. 459-497, 2018.

42) D.W. van der Meer, J. Widén, and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," Renewable and Sustainable Energy Reviews, vol. 81, pp. 1484-1512, 2018.

43) Y. Ren, P.N. Suganthan, and N. Srikanth, "Ensemble methods for wind and solar power forecasting—A state-of-the-art review," Renewable and Sustainable Energy Reviews, vol. 50, pp. 82-91, 2015.

44) R. Juban, H. Ohlsson, M. Maasoumy, L. Poirier, J. Zico Kolter, "A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014," International Journal of Forecasting, 32(3), 1094-1102, 2016.

45) P. Lauret, M. David, and H. Pedro, "Probabilistic solar forecasting using quantile regression models," Energies, 10(10), 1591, 2017.

46) H.T.C. Pedro, C.F.M. Coimbra, M. David, P. Lauret, "Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts," Renewable Energy, 123, 191-203, 2018.

47) M.P. Almeida, O. Perpiñán, L. Narvarte, "PV power forecast using a nonparametric PV model," Solar Energy, 115, 354-368, 2015.

48) W. Zhang, H. Quan, D. Srinivasan, "Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination," Energy, 160, 810-819, 2018.

49) J. Huang, and M. Perry, "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, 2016.

50) T. Hong, P. Pinson, S. Fan, "Global Energy Forecasting Competition 2012," International Journal of Forecasting, 39(2), 357-363, 2014.

51) Q. Ni, et al. "An ensemble prediction intervals approach for short-term PV power forecasting." Solar Energy, 155, 1072-1083, 2017.

52) A.A. Mohammed, Z. Aung, "Ensemble learning approach for probabilistic forecasting of solar power generation," Energies, 9, 1017, 2016.

53) A.A. Mohammed, W. Yaqub, Z. Aung, "Probabilistic forecasting of solar power: An ensemble learning approach,". International Conference on Intelligent Decision Technologies, Springer, Cham, 449-458, 2017.

54) F. Golestaneh, P. Pinson, H.B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation— With application to solar energy." IEEE Transactions on Power Systems, 31, 3850-3863, 2016.

55) Y. Wang, et al. "Combining probabilistic load forecasts," IEEE Transactions on Smart Grid, 2018.

56) European Centre for Medium-range Weather Forecasts (ECMWF) website. Available online: https://www.ecmwf.int/.

57) C. Gilbert, J. Browell, and D. McMillan, "Leveraging turbine-level data for improved probabilistic wind power forecasting," IEEE Transactions on Sustainable Energy, 2019.

58) L. Cai, J. Gu, J. Ma, and Z. Jin, "Probabilistic wind power forecasting approach via instance-based transfer learning embedded gradient boosting decision trees," Energies, vol. 12, 159, 2019.

59) M. Khalid, and A.V. Savkin, "A method for short-term wind power prediction with multiple observation points," IEEE Transactions on Power Systems, vol. 27, pp. 579-586, 2012.

60) B.R. Prusty, and D. Jena, "A spatiotemporal probabilistic model-based temperature-augmented probabilistic load flow considering PV generations," International Transactions on Electrical Energy Systems, vol. 29, e2819, 2019.

61) J. Wang et al. "A novel framework of reservoir computing for deterministic and probabilistic wind power forecasting," IEEE Transactions on Sustainable Energy, 2019.

62) H. Zheng, and Y. Wu, "A XGBoost model with weather similarity analysis and feature engineering for short-term wind power forecasting," Applied Sciences, vol. 9, no. 15, 3019, 2019.

63) X.Y. Wu et al, "Data-driven wind speed forecasting using deep feature extraction and LSTM," IET Renewable Power Generation, 2019.

64) K. Zhou, C. Fu, and S. Yang, "Big data driven smart energy management: From big data to big insights," Renewable and Sustainable Energy Reviews, vol. 56, pp. 215-225, 2016.

65) P. Gaillard, Y. Goude, R. Nedellec, "Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting," Int. J. Forec., vol. 32, no. 3, 2016.

66) B. Liu, J. Nowotarski, T. Hong, R. Weron, "Probabilistic load forecasting via quantile regression averaging on sister forecasts," IEEE Trans. Smart Grid, vol. 8, no. 2, pp. 730-737, 2017.

67) A. Bracale, P. De Falco, G. Carpinelli, "Comparing univariate and multivariate methods for probabilistic industrial load forecasting," 2018 5th Int. Symp. Environment-Friendly Energies Appl. (EFEA), Rome, pp. 1-6, 2018.

68) Y. Wang, N. Zhang, Y. Tan, T. Hong, D.S. Kirschen, C. Kang, "Combining probabilistic load forecasts," IEEE Trans. Smart Grid, vol. 10, no. 4, pp. 3664-3674, 2018.

69) S. Ben Taieb, R. Huser, R.J. Hyndman, M.G. Genton, "Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression," IEEE Trans. Smart Grid, vol. 7, no. 5, pp. 2448-2455, 2016.

70) A. Bracale, G. Carpinelli, P. De Falco, T. Hong, "Short-term industrial reactive power forecasting," Int. J. Electr. Power Energy Syst., vol. 107, pp. 177-185, 2019.

71) P. Kou, F. Gao, "A sparse heteroscedastic model for the probabilistic load forecasting in energy-intensive enterprises," Int. J. Electr. Power Energy Syst., vol. 55, pp. 144-154, 2014.

72) K. Berk, A. Hoffmann, A. Müller, "Probabilistic forecasting of industrial electricity load with regime switching behavior," Int. J. Forec., vol. 34, no. 2, pp. 147-162, 2018.

73) Y. Xu, Z.Y. Dong, R. Zhang, D.J. Hill, "Multi-timescale coordinated voltage/var control of high renewable-penetrated distribution systems," IEEE Trans. Power Syst., vol. 32, no. 6, pp. 4398-4408, 2017.

74) L. Alfieri, G. Carpinelli, A. Bracale, P. Caramia, "On the optimal management of the reactive power in an industrial hybrid microgrid: A case study," 2018 Int. Symp. Power Electr., Electr. Drives, Autom. Motion (SPEEDAM), Amalfi, pp. 982-989, 2018.

75) S. Pirouzi, J. Aghaei, V. Vahidinasab, T. Niknam, A. Khodaei, "Robust linear architecture for active/reactive power scheduling of EV integrated smart distribution networks," Electr. Power Syst. Res., vol. 155, pp. 8-20, 2018.

76) M. Ghaljehei, Z. Soltani, J. Lin, G.B. Gharehpetian, M.A. Golkar, "Stochastic multi-objective optimal energy and reactive power dispatch considering cost, loading margin and coordinated reactive power reserve management," Electr. Power Syst. Res., vol. 166, pp. 163-177.

77) N.S. Coleman, K.N. Miu, "Distribution load capability with nodal power factor constraints," IEEE Trans. Power Syst., vol. 32, no. 4, pp. 3120-3126, 2017.

78) R.H.A. Zubo, G. Mokryani, R. Abd-Alhameed, "Optimal operation of distribution networks with high penetration of wind and solar power within a joint active and reactive distribution market environment," Appl. Energy, vol. 220, pp. 713-722, 2018.

79) H. Nezamabadi, M.S. Nazar, "Arbitrage strategy of virtual power plants in energy, spinning reserve and reactive power markets," IET Gen., Transm. Distr., vol. 10, no. 3, pp. 750-763, 2016.

80) A. Samimi, M. Nikzad, P. Siano, "Scenario-based stochastic framework for coupled active and reactive power market in smart distribution systems with demand response programs," Renew. Energy, vol. 109, pp. 22-40, 2017.

81) A. Kargarian, M. Raoofat, M. Mohammadi, "Probabilistic reactive power procurement in hybrid electricity markets with uncertain loads," Electr. Power Syst. Res., vol. 82, no.1, pp. 68-80, 2012.

82) A. Songpu, M.L. Kolhe, L. Jiao, Q. Zhang, "Domestic load forecasting using neural network and its use for missing data analysis," 9th Int. Symp. Adv. Topics Electr. Eng. (ATEE), Bucharest, pp. 535-538, 2015.

83) E.F. Arruda, R.B. Muller, M.P. Di Salvo, G.C.C. Rocha, M.T. Coelho, L.B.T. Baran, "Disaggregated active and reactive demand forecasting using first difference measured data and neural networks," Int. Conf. Electr. Distr., Lyon, 2015.

84) Y. Xu, J.V. Milanović, "Accuracy of ANN based methodology for load composition forecasting at bulk supply buses," 2014 Int. Conf. Prob. Meth. Appl. Power Syst. (PMAPS), Durham, pp. 1-6, 2014.

85) X.S. Han, L. Han, H.B. Gooi, Z.Y. Pan, "Ultra-short-term multi-node load forecasting – a composite approach," IET Gen., Transm. Distr., vol. 6, no. 5, pp. 436-444, 2012.

86) J.M. Cho, J.H. Kim, W.H. Park, Y. Lee, J. Kim, "Short-term reactive power load forecasting using multiple time-series model," IFAC Proc. Vol., vol. 36, no. 20, pp. 985-990, 2003.

87) I. Ben-Gal, "Outlier detection," in Data Mining and Knowledge Discovery Handbook, O. Maimo and L. Rokach, eds., Springer, Boston, MA, 2005.

88) R lubridate package. Available online: https://CRAN.R-project.org/package=lubridate.

89) D.B. Rubin, "The Bayesian bootstrap," The Annals of Statistics, vol. 9, no. 1, pp. 130-134, 1981.

90) A.W. ryaputera, H. Verbois, and W.M. Walsh, "Probabilistic accumulated irradiance forecast for Singapore using ensemble techniques," Proc. of IEEE 43rd Photovoltaic Specialists Conference (PVSC), 1113-1118, 2016.

91) A. Bracale, et al. "New advanced method and cost-based indices applied to probabilistic forecasting of photovoltaic generation," Journal of Renewable and Sustainable Energy, 8(2), 023505, 2016.

92) T. Gneiting, and A.E. Raftery, "Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association," 102(477), pp. 359-378, 2006.

93) Clyde, M.A., & Lee, H.K.H. (2001). Bagging and the Bayesian bootstrap. Artificial Intelligence and Statistics, eds. Richardson, T. & Jaakkola, T. New York: Elsevier, 169-174.

94) T. Hastie, R. Tibshirani, and J. Friedman. The elements of statistical learning: Data mining, inference, and prediction, 2nd edition. Springer Series in Statistics, 2009.

95) R. Koenker, Confidence intervals for regression quantiles. Asymptotic Statistics, eds. Mandl, P., & Hušková, M. Berlin Heidelberg: Springer-Verlag, pp. 349-359, 1994.

96) A.J. Cannon, "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), pp. 3207-3225, 2018.

97) J. Friedman, "Stochastic gradient boosting," Computational Statistics & Data Analysis, 38(4), pp. 367-378, 2002.

98) R qrnn package: Quantile regression neural network. Available online: https://CRAN.R-project.org/package=qrnn.

99) R gbm package: Generalized boosted regression models. Available online: https://CRAN.R-project.org/package=gbm.

100) R quantreg package: Quantile regression. Available online: https://CRAN.R-project.org/package=quantreg.

101) R bayesboot package: An implementation of Rubin's (1981) Bayesian bootstrap. Available online: https://CRAN.R-project.org/package=bayesboot.

102) T. Ouyang, et al. "A combined multivariate model for wind power prediction," Energy Conv Management, vol. 144, pp. 361-373, 2017.

103) M. Sobhani, et al., "Combining weather stations for electric load forecasting," Energies, vol. 12, 1510, 2019.

104) Terna (Italian Transmission System Operator) website. Statistics and forecast department. Available online: http://www.terna.it/SistemaElettrico/StatisticheePrevisioni/ConsumiEnergiaElettricaperSettoreMerceologico.aspx.

105) J.K. Møller, H.A. Nielsen, H. Madsen, "Time-adaptive quantile regression," Comp. Stat. Data Analysis, vol. 52, no. 3, pp. 1292-1303, 2008.

106) N. Meinshausen, "Quantile regression forests," Journal of Machine Learning Research, vol. 7, pp. 983-999, 2006.