DELIVERABLE DI PROGETTO DI RICERCA

DiGriFlex

Real-time Distribution Grid Control and Flexibility Provision under Uncertainties

RELAZIONE

SUB-ATTIVITA' OR2

Deliverable D2.1

Sistemi di previsione e corrispondenti algoritmi risolutivi atti al pre-scheduling delle reti di distribuzione in presenza di incertezze

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 01/09/2019 al 31/08/2020

INDEX

1. 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 to carry out in the framework of the work package WP2: "Development of appropriate day-ahead and real-time forecasting systems for renewable generation and loads". In particular, this report includes the results of the research activities carried out within Task 2.1 ("Data collection, data pre-processing and exploratory data analysis") and Task 2.2 ("Development of day-ahead forecasting systems for renewable generation and loads").

a) With reference to Task 2.1 the activities included:
   - the collection of time series that include both target variables and predictor variables;
   - the description of methods to pre-process the data to eliminate outliers and bad values;
   - the exploratory data analysis in order to discard uninformative predictors.

b) With reference to Task 2.2 the activities included:
   - the development of methods and models based on multiple linear regression and random forests
   - the identification of the best combination type for the underlying models in the ensemble approach
   - the comparison between the results obtained applying the proposed approach and the ones given by the relevant state-of-the-art benchmarks

Activities in a) included all the data collection and data pre-processing tasks that were necessary to create a large robust database of variables which could be exploited to develop forecasting systems for renewable generation and loads. Accounting for the fact that the forecasting methodology will be integrated in the grid optimization models that are object of the WP3, the data collection and pre-processing targeted variables collected at the site of the installation of the test distribution grid of the ReIne laboratory. Other variables that were not available at the site of installation of the test distribution grid of the ReIne laboratory were taken from the literature and/or collected at different sites, in order to develop the forecasting systems.

Activities in a) were also devoted at making exploratory data analyses in order to individuate variables that are informative for the target variables (either renewable power generation or loads), as per building the input datasets for the forecasting systems developed in the Task 2.2. The effort required to individuate the informative variables significantly depend on the arrangement of the forecasting system, thus the activities in a) devoted to exploratory data analysis were developed in strict interaction with the activities in b) devoted to the development of forecasting systems.

Summarizing, the research activities completed in the Task 2.2 resulted in:

i. the collection of large, robust datasets of PV power generation and weather variables at the site of installation of the test distribution grid of the ReIne laboratory;
ii. the collection of large, robust datasets of wind power generation, industrial loads and weather variables available from the relevant literature and from public databases;
iii. the pre-processing of the collected data, aiming at individuating and correcting missing data, bad data and outliers;
iv. the exploratory data analysis to reduce the dimensionality of the input datasets, favoring the development of adequate forecasting systems for renewable generation and load.

Activities in b) included the development of methods and models able to obtain accurate day-ahead forecasting for renewable generation (wind and photovoltaic (PV) forecasting) and loads (active and reactive power). In addition, also the optimal operation of microgrids in presence of uncertainties is considered and the impact of inputs' forecasting error is analyzed in depth.

The widely spread of the generation systems based on wind and solar primary sources across MV and LV distribution systems lead to the need of methods able to predict the wind and photovoltaic generated power. In fact, since the uncertain nature of the solar and wind energy resource, PV and wind power forecasting models are crucial in any energy management system for smart distribution networks. On the other hand, industrial load takes a big portion of the total electricity demand. Skilled industrial load forecasts allow for optimally exploiting energy resources, managing the reserves, and market bidding, which are beneficial to distribution system operators and their industrial customers. Despite its importance, industrial load forecasting has never been a popular subject in the literature.

Eventually, for both generation and loads, in relevant literature point forecast and probabilistic forecast are proposed. 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. Thus, in our studies probabilistic forecast are considered.

Summarizing, the research activities completed in the Task 2.2 resulted in:

i. the development of a day-ahead probabilistic wind power forecasting based on ranking and combining Numeric Weather Predictions (NWPs);
ii. the development of a Bayesian bootstrap quantile regression model for probabilistic photovoltaic power forecasting;
iii. the development of a multivariate approach for probabilistic industrial load forecasting.

With reference to i), the volatility of the wind over large time horizons complicates the generation of skilled, reliable wind power forecasts. Exploiting NWPs is generally considered mandatory to increase the skill of probabilistic predictions, and forecasts may further be enhanced by adding several spatially-distributed predictions. However, feature selection becomes as more complicated and time-consuming as the number of NWPs increases. In our study, the power generated by a wind farm is predicted developing a new technique based on ranking and combining spatially-distributed NWPs, easing the feature selection and reducing the computational efforts, as well as maintaining high the skill of probabilistic forecasts. Several spatially-distributed NWPs, provided for the area surrounding the wind farm, are ranked for each individual generator, and the ranked NWPs are combined to form an ensemble set of predictors for the probabilistic forecasting model. This ensemble is obtained using three different weighted combination approaches. Gradient boosting regression tree models and quantile regression neural networks generate probabilistic wind power forecasts. The proposed methodology is applied for day-ahead wind power forecasting of individual generators and of the entire wind farm. Numerical experiments carried out on an actual wind farm in southern Italy.

With reference to ii), a probabilistic PV power forecasting based on a Bayesian bootstrap quantile regression model is developed. The Bayesian bootstrap is applied to estimate the parameters of a quantile regression model and 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.

With reference to iii), most existing methods for industrial load forecasting operate on the active power alone, partially or totally neglecting the reactive power. We developed a multivariate approach to probabilistic industrial load forecasting, which addresses active and reactive power simultaneously. The method is based on a two-level procedure, which consists of generating probabilistic forecasts individually for active and reactive power through univariate probabilistic models, and combining these forecasts in a multivariate approach based on a multivariate quantile regression model. The procedure to estimate the parameters of the multivariate quantile regression model is posed under a linear programming problem, to facilitate the convergence to the optimal solution. The proposed method is validated using actual load data collected at an Italian factory, under comparison with several probabilistic benchmarks.

This report is the final deliverable of activities carried out for Task 2.1 and 2.2 of DiGriFlex project and referred to the Milestone 1.

2. OBJECTIVES

Task 2.1 ("Data collection, data pre-processing and exploratory data analysis") and Task 2.2 ("Development of day-ahead forecasting systems for renewable generation and loads") are included in work package WP2: "Development of appropriate day-ahead and real-time forecasting systems for renewable generation and loads" that is focused on the forecasting systems of renewable generation and loads developed for different time horizons.

As reported in the final approved version of Technical Specifications of project, detailed tasks/subtasks of WP2 are Task 2.1: Data collection, data pre-processing and exploratory data analysis and Task 2.2: Development of day-ahead forecasting systems for renewable generation and loads.

The objective of the first task is to create large, robust datasets of variables which could be exploited to develop the forecasting systems for renewable generation and load. The dataset building creation requires i) to collect data from dedicated measurement systems and/or from accessing databases published in the relevant literature, ii) to pre-process these data in order to clean them from missing data, bad data and outliers, and iii) to make exploratory data analysis in order to reduce the dimensionality of the input datasets, removing variables that are uninformative for the target variables (either renewable power generation or load).

The objective of the Task 2.2 consists in identifying the most suitable techniques for the day-ahead forecast of the power produced by production plants from renewable sources (in particular from solar and wind sources) and the power required by the loads. In the Technical Specifications of project it is specified that probabilistic techniques (for example, Bayesian methods or ensembles of probabilistic methods) should be considered and compared with benchmarks. In particular, the methods must be based on the proper combination (ensemble) of multiple linear regression models and random forests due to their versatility and ease of implementation. An additional objective is the comparison between the results obtained applying the proposed approach and the ones given by the relevant state-of-the-art benchmarks.

3. STATE OF THE ART

Current practices in renewable energy forecasting and load forecasting are very heterogeneous. This Section reviews the state of the art on energy forecasting, starting from a classification of the methodologies and going into the details of the existing approaches.

3.1. CLASSIFICATION OF FORECASTING MODELS

The diversity in forecasting needs has a direct, intuitive consequence: no forecasting method is universally able to fit any purpose, but it has to be selected case by case on the basis of particular needs. The classifications of forecasting methods straightforwardly follows the diversity in terms of end user needs.

The first classification is made in terms of forecast lead time. Indeed, actions on power systems are performed on different time lines: e.g., improvement, replacement or realization of new infrastructures are planned several years before, while optimal management of distributed energy resources distribution grids is scheduled some minutes to some hours before [1,2].

Few papers [3] classify forecasting methods in 2 categories (short-term and long-term); however, the most complete practice is to individuate *Very Short-Term Forecasting* (VSTF), *Short-Term Forecasting* (STF), *Medium-Term Forecasting* (MTF), and *Long-Term Forecasting* (LTF) methods [4-8].

VSTF lead times range up to 24 hours; they are usually involved in power balancing and system optimal management and control. The influence of external variables (e.g., ambient temperature for load forecasting) is limited in this kind of applications, and therefore is often overlooked. VSTF partially covers the pre-scheduling of distribution grids, object of WP2.2, and it fully covers the real-time control of distribution grids, object of the WP2.3.

STF lead times range from 24 hours ahead to two weeks ahead; they are usually involved in power balancing for acquiring appropriate reserve, market participation, and system optimal management. STF partially covers the pre-scheduling of distribution grids, object of WP2.2.

MTF lead times range from 2 weeks to 3 year ahead; this wide interval of time makes MTF methods useful for market participation, system optimal management, and planning. Social and economic factors should be carefully investigated in MTF, especially for monthly and yearly scenarios.

LTF lead times start from 3 years and reach 20 (or more) years. These forecasts are involved in power system planning, and weather, social and economic long-term evaluations are mandatory in order to cope with evolutionary trends.

Table 4.1.1 associates forecasting methods, classified in terms of lead times, to corresponding needs [8].

A second classification involves the output of forecasting methods. This comes from the different risks linked to power system tasks that require forecasts to be completed.

Table 4.1.1 - Utility of forecasting methods in power system operation needs

Classification of forecasting methods	Power system operation need			
	Power balancing	Participation to electrical markets	Optimal management and control	Planning
VSTF	yes	no	yes	no
STF	yes	yes	yes	no
MTF	no	yes	yes	yes
LTF	no	no	no	yes

Let's think of a wind plant owner, who wants to sell energy on electrical markets [9]. He has to submit a selling offer, stating the (exact) amount of energy he will be able to produce; in several Countries, he is penalized if the resulting production is too far from the declared one. If he disposes of a forecasting method that provides only a single value of wind power as output, the plant owner has no other choice than submitting a selling offer of as much energy as the forecasted one. Instead, if he disposes of a forecasting tool that provides more values, or the probability distribution of wind powers, he can manage the forecasts and make the best choice for his needs.

In this context, *deterministic* forecasts provide as output only a single value of the variable of interest (point forecast). *Probabilistic* forecasts provide as output analytical distributions such as Probability Distribution Functions (PDFs), Cumulative Density Functions (CDFs), sampled distributions (discrete probabilities), quantiles, or moments of the predictive distribution (e.g., mean, variance and skewness) [10]. Note that the variable of interest is still treated as a random entity in both frameworks: the main difference is that a single value is given as forecast of the variable of interest in deterministic framework, while more values, or a function, are given as forecast of the variable of interest in probabilistic framework.

Probabilistic forecasts are generally preferable, since they provide also information about the uncertainty linked to the forecast itself. Therefore, they allow the risk assessment and the optimal selection of a single value, on the basis of different frameworks [11,12]. Indeed, it is always possible to extract a single, spot-value (e.g., the mean value of the predictive distribution) from probabilistic forecasts, while the reciprocal is obviously not valid. The main drawbacks of probabilistic forecasts are the increase of method complexity, and their greater computational burden. Then, if the forecast end user gains no benefit in having a probabilistic forecast, deterministic methods are still the best choice.

It is worth noting that probabilistic methods sometimes rely on an underlying deterministic method [13,14]; e.g., some parameters of the predictive probabilistic distribution could be set from the output value of a deterministic method. In this case, improving the performance of the underlying deterministic method is compulsory in order to increase the overall quality of the probabilistic forecasts. Thus, research efforts in the deterministic framework are always encouraged.

A third classification of forecasting methods is based on the characteristics of models involved in the forecasting method, and consequently on the solving procedure. The common classification is in terms of *parametric* and *non-parametric* methods.

Parametric methods are based on models that are univocally identified as several numerical parameters are known; e.g., a predictive analytical Gaussian distribution is univocally identified when its mean and variance are known. Therefore, solving a parametric forecasting method consists in finding estimations of unknown parameters, usually by minimizing or maximizing assigned objective functions (i.e., by minimizing an error index). In the particular case of parametric probabilistic methods, usually the problem of finding a prior probabilistic characterization of the variable of interest through a specific PDF has to be solved [13].

Non-parametric methods, instead, rely on the idea that forecasting future dynamics can be achieved by analogy with past dynamics. Indeed, the variable under study is not assessed through an analytic model, instead it is forecasted by means of a procedure that "learns" from the past. Note that the "non-parametric" definition could be misleading. It does not mean that no parameters are involved in non-parametric methods; indeed, some involved parameters could identify the order of the model, rather than the model itself.

In non-parametric methods, however, the complexity of the models (i.e., the number of parameters) grows with the dimension of the problem and, theoretically, is not constrained. The more the inputs (i.e., the elements of the training set) fed to the non-parametric method, the larger is the number of parameters to be estimated. Therefore, the structure of the model itself "grows" as the training set enlarges. On the other hand, the structure of models in parametric methods is fixed with the dimension of the problem; the same number of parameters has to be estimated, regardless of the size of the training set.

The fourth and last classification is based on the approach used to build and solve the forecasting problem. *Statistical* approaches rely on measurement data acquired in the past to produce forecasts for the future, starting from the assumption that past conditions are informative for the future. *Physical* approaches rely instead on a specific formulation of the problem under study in a rigorous fashion, by exploiting mathematical formulas based on physical principles. *Hybrid* approaches can be a combination of statistical approaches, physical approaches, or statistical and physical approaches together.

3.2. SURVEY ON FORECASTING SYSTEMS FOR RENEWABLE GENERATION AND LOADS

In the last two decades, hundreds of papers have dealt with energy forecasting and proposed a wide variety of forecasting systems. The literature review is conveniently conducted upon relevant surveys that have been published through the last decade.

Among the energy forecasting ones, load forecasting is the oldest topic tackled by researchers, especially considering LTF scenarios used for planning purposes [15]. Spatial load forecasting applied in this context allowed to get insight regarding where, when and how much load would grow during years. Although some research practices in this context belong to the last century, some of them are still operational practices in power industry [16]. However, it is only in the last 20 years that short-term load forecasting gained importance and attention due to the pushed pursue of the excellence in operating distribution and transmission systems. Early researches focused on the application of Artificial Neural Networks (ANNs) for short-term load forecasting, due to their easy implementation from a black-block perspective. Nevertheless, ANNs have been proved to suffer from flaws in theory and practice [17], thus other forecasting approaches such as time series models and regression-tree-based models have gained popularity as well. In these early approaches on load forecasting, the problem has been tackled within deterministic frameworks. Only in the last decade probabilistic load forecasting become consolidated [8], although far from being exhaustive due to the complexity of the probabilistic predictions and due to the increasing requests for accuracy and calibration even at low-aggregation load levels (e.g., for smart meters at residential or commercial building load levels [18]).

Renewable energy forecasting is definitely younger than load forecasting, as the widespread diffusion of renewable power generators occurred only in the last 15 years. Wind power forecasting initially gained the greatest attention due to the bigger installed power and the high volatility of the primary source (i.e., the wind speed). To keep the systems balanced, it was clear that wind power forecasts would be needed for lead times in the VSTF and in the STF horizons [19,20]. The volatility of wind speed also pushed get interactions among energy forecasters and meteorologists, and two important consequences arose from that:

i) the shift towards probabilistic wind power forecasting since weather forecasts are typically provided by probabilistic Numeric Weather Prediction (NWP) models [21,22];

ii) the development of parallel approaches based on physical and statistical models for wind power forecasting [23], which merged in the current state-of-the-art that incorporates both in wind power forecasting systems.

Solar and photovoltaic (PV) forecasting did not become a topic earlier than 2010. Following the teachings of wind power forecasting, the approaches directly considered external weather data, and particularly camera-based (for VSTF), satellite-based (for VSTF and STF) and NWP-based (for STF) inputs [24]. The current practices include these inputs in top-performing models, also in probabilistic contexts (although less maturity than wind power forecasting is evident).

From the literature review, forecasting combination and ensemble combination have recently arose as the frontiers in research. The two terms usually have the same meaning in energy forecasting, although the term "ensemble" takes its origin from ensemble weather modeling that returns NWPs.

The most common approach in ensemble forecasting is stacking. It refers to the combination of the outputs of different forecasting models using a custom approach. This combination can aim at maximizing or minimizing a particular loss function, or simply by using the outputs of individual forecasting models as inputs of another model. Forecast combination itself is not a new topic, as it was first discussed in late 1960s [25], but only recent works analyzed in details the pros and cons of stacking combinations [26-28]. Notable applications of stacking forecast combination and ensemble approaches are in load forecasting (e.g., combination of point forecasts [29] and of predictive quantiles [30]), with relevance also in disaggregated smart meter load forecasting [31], and in renewable energy forecasting (e.g., combination of predictive distributions [32] and predictive quantiles [33] for PV forecasting).

Other approaches to ensemble forecasting use the diversification in the data space and parameter space to quantify the uncertainty associated with a forecasting model, in a similar way as for the trajectories drawn for building ensemble NWPs by perturbing the initial conditions and making the system separately [34]. Notable examples of the ensemble space diversification include:

i) bootstrapping: it is a sampling technique consists of resampling the data used to train the models with replacement, in order to obtain datasets having the same size of the original data, that however differ from the original data, although mimicking the original data to retain their statistical properties;

ii) bagging (or bootstrap aggregating): it consists of training the same forecasting model several times, on different training datasets; these different training datasets are obtained by bootstrapping (i.e., resampling with replacement) the original training dataset, maintaining the same size. The outputs of these different models can be averaged in order to obtain point forecasts or treated as a predictive sampled distribution of the target variable;

iii) boosting: it consists of iteratively training weak models, adding them in order to build a final strong model. Each weak model fits the training data that however is re-weighted at each iteration; points of the training data that have been fitted with poor performances by the previous weak models are weighted for a greater value, whereas

points of the training data that have been fitted well receive a smaller weight. Therefore, the underlying idea of the boosting approach is that the iterative addition of a weak model compensates the shortcomings of the previous weak models, returning a strong forecasting model.

Eventually, a relevant outcome of the literature review determines the importance of the spatial-temporal analysis in the development of a forecasting system. This problem is prone to be tackled from a hierarchical perspective, proceeding by a series of successive merges of prediction made at a different level of aggregation [35,36]. In hierarchical forecasting, the information contained at low-level aggregation (for example, the load at the HV/MV nodes in a HV network, or the load of small regions) is used to predict at high-level aggregation (for example, the overall load of the entire HV network, or the overall load of a geographic region). Hierarchical methodologies to merge probabilistic forecasts have been recently applied with success [37-39]. They are however still challenging, mostly because the problem is usually formulated as a nonlinear and non-convex optimization problem, so that global optimality cannot be guaranteed and the combined results may be worse than individual forecasts [40].

In the following, a specific survey on relevant papers on these topics involved in PV power forecasting, wind power forecasting and industrial load forecasting is presented together with the open questions and challenges that arose from the literature review and that are tackled by the research carried out during this project.

3.2.1. SELECTED LITERATURE AND OPEN CHALLENGES ON PHOTOVOLTAIC POWER GENERATION

As stated above, literature reviews and competition surveys [40-43] 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 forecasts while the opposite is clearly unfeasible, research and contributions in the latter are highly recommended and only probabilistic PV power forecasting research is considered in this sub-Section.

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 [33,44,45], machine learning approaches (such as Gradient Boosting Regression Trees (GBRTs) [46], Quantile Regression Forests (QRFs) [33,47,48] and Quantile K-Nearest Neighbors (QKNN) [49]). 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. For example, these models proved their effectiveness also in recent energy forecasting competitions [40,50],

since the forecasting systems developed by the highest-ranking teams were based on these nonparametric probabilistic models.

Generally speaking, the integration of NWPs into PV power forecasting models is quite mandatory to improve predictions, particularly for large time horizons [42]. 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 (either boosting, stacking or bagging) [43] have been applied with success in probabilistic energy forecasting [32,33,43,51-53]. New trends in probabilistic PV power forecasting indeed individuate the probabilistic combination of individual forecasts as a suitable solution, in order to improve the accuracy of the results [32,33] Probabilistic forecast combination is not as straightforward as it seems to be at first inspection. Contrarily to combining point forecasts, for which the simple weighted averaging is often a plausible solution, combining probabilistic forecasts is a much more challenging task: the combined probabilistic forecasts indeed must retain adequate properties in terms of reliability and sharpness [32,54], and the main features of a probabilistic forecast (e.g., the ascending order of predictive quantiles) must be retained also by the combined forecasts [32]. Relevant literature has addressed these aspects under different points of view [43]. Individual probabilistic forecasts can be indeed merged: i) by a combination of the predictive cumulative distribution functions [32]; or ii) by a combination of the predictive quantiles [55]. Nevertheless, within these two approaches types, several strategies and architectures can be developed to combine forecasts; so there is room for further investigation and improvement.

3.2.2. SELECTED LITERATURE AND OPEN CHALLENGES ON WIND POWER GENERATION

Wind forecasters almost unanimously agree that exploiting NWPs enhances the skill of forecasts for short-term horizons (e.g., day ahead horizon), as stochastic-only forecasting systems usually perform well only for very-short lead times (e.g., 1-hour ahead), gradually losing relevance as the lead times increases. On the other hand, the availability of accurate NWPs has grown in the last years, thanks to the efforts in atmospheric modeling and in satellite acquisitions, as well as predictions of many weather variables are nowadays made available by weather institutions [56], and they can be used as input predictors of forecasting systems.

Getting the most from the available data to increase the skill of wind power forecasts is an object of energy data analysis. Exploiting spatially-distributed NWPs for an area surrounding the wind farm has been proved to be an efficient solution to further add skill [57-59]. However, adding too many predictors to the models could involve the risk of overfitting the data thus losing forecast skill. Standard techniques to reduce overfitting are based on dimensionality reduction [60,61] and feature selection (or feature engineering) [57,61-63], which are often

accomplished through validation, leave-one-out cross-validation or k-fold cross-validation frameworks. These latter procedures usually are computationally intensive, particularly when many NWPs are candidate predictors [64], as it can happen when handling spatially-distributed NWPs for an area surrounding the wind farm. This poses an arduous challenge to the development of forecasting systems in which models are re-trained as new observations become available.

3.2.3. SELECTED LITERATURE AND OPEN CHALLENGES ON INDUSTRIAL LOAD FORECASTING

The literature on probabilistic load forecasting has expanded recently [8], and the focus has switched to nonparametric modeling [42]. Relevant probabilistic load forecasting systems are based on QR models [65] and QR averaging [66], QRFs [67] (also within optimal combination frameworks [68]), or GBRTs [69]. These models have been extensively applied either at sub-station or at smart-meter nodes; nevertheless, the performance of a forecasting system may drop when the target load is the demand of an industrial factory. This is due to the particular features which influence the industrial demand [67], due to the particular seasonal patterns of the industrial load profiles [70], and due to the smaller relative importance of the ambient temperature to predict industrial load [71]. The only recommended path to increase the accuracy of industrial load forecasts is to develop ad-hoc forecasting systems, accounting for all of these specific features [72].

Forecasters usually target the only active power, partially or totally disregarding reactive power. This choice may be considered obsolete nowadays due to i) the increased computer performance, which eases the generation of reactive power forecasts, and ii) the spread of smart grid tools, which manage and operate networks by the active and reactive power flows. For example, Volt/VAr optimization [73], harmonic compensation [74] and optimal energy dispatch [75,76] tools require prior estimations of the nodal reactive power. The multitude of power converters distributed across the grids can be controlled to compensate reactive power locally [77], increasing the total power transmission capacity of the grid and reducing losses [78]. It is also worth mentioning that reactive power support is an ancillary service, remunerated in several countries [79]. Improved frameworks for reactive power markets, either coupled with active power market [80] or developed within a probabilistic framework [81], have been proposed in relevant literature. Despite the rising interest of practitioners in reactive power management, the literature on reactive power forecasting slowly adapts to the practical needs. Reactive power forecasts are usually built by post-processing active power forecasts, on the basis of practitioners' experience or by average-power-factor corrections. Only a few papers directly deal with reactive power forecasting. The majority of literature is based on the application of artificial neural networks, either at domestic nodes [82], disaggregated levels [83] or at bulk supply nodes [84]. Other models applied to reactive power forecasting are fuzzy approaches [85], piecewise linear regression models

[70], time series [86], and support vector regression [70]. These research papers evidence that correctly modeling the mutual correlation between active and reactive power allows enhancing the performance of the forecasting system [70], suggesting for additional research in this sense.

4. RESULTS OF ACTIVITIES FOR TASK 2.1

The Task 2.1 focuses on data collection, data pre-processing and exploratory data analysis activities that are carried out in order to create large, robust databases of variables which could be exploited to test the performance of the forecasting systems that will eventually be used to predict energy at the site of the installation of the test distribution grid of the ReIne laboratory.

4.1. DATA COLLECTION

Several datasets are used for the activities related to the WPs 2.1 and 2.2 and presented in this Deliverable. These datasets are related to the target variables (loads and renewable generated power) and external variables (i.e., weather data) used as additional inputs of the forecasting systems.

Dataset_PVI1: this dataset includes PV power measurements taken at the 30-kWp PV installation (PVI1) equipped with 4 8.5-kWp inverters which is part of the test distribution grid of the ReIne laboratory. Due to its recent installation, the data collection started on August 24, 2019 with a 1-minute time resolution. The utilization of these data for validating and testing the PV power forecasting models is unfeasible at the progress stage of the first year of project activities due to the relatively short operation life of the PV system, therefore this dataset is not included in experimental frameworks.

Dataset_PVI2: this dataset includes PV power measurements taken at a second PV installation (PVI2), located close to the test distribution grid of the ReIne laboratory. PVI2 was monitored since January 1, 2016 until December 31, 2018 at a 1-minute time resolution. The related data are used for the initial validation of the PV power forecasting systems.

Dataset_PVI3: this dataset consists of zone-1 PV power data published in the framework of the Global Energy Forecasting Competition 2014 [40] at an unspecified location in Australia. Data span April 1, 2012 to June 30, 2014 with an hourly resolution. This dataset is included in the experimental frameworks in order to generalize the performance testing of the proposed forecasting systems.

Dataset_WG: this dataset includes wind power measurements collected at a wind farm located in southern Italy. The wind farm is constituted by ten 3-MVA generators. Wind power was measured at each generator throughout three years, with 10-minute acquisitions.

Dataset_indust_load: this dataset includes industrial load data (active and reactive powers) collected at an Italian factory that manufactures transformers. The factory operates on two work shifts during weekdays (Monday-Friday), on a single work shift during Saturdays, and it is closed on Sundays. Energy meters collect electrical data at fourteen single loads, at four LV power distribution feeders, and at the point of common coupling to the main MV network (i.e., the aggregate load of the factory). The time resolution of the data metering is 15 minutes. The layout of the industrial LV distribution system is illustrated in Fig. 5.1, where the buses of the monitored loads are in red font. The aggregate load is monitored at bus 3, and the four LV power distribution feeders are monitored at buses 5, 26, 43 and 66.

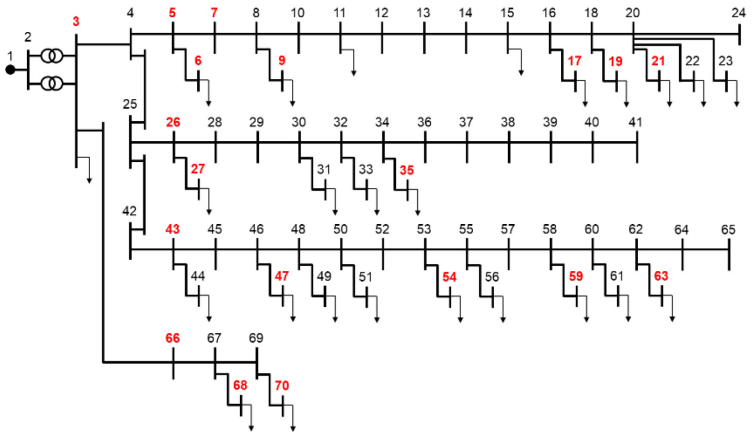

Figure 5.1. Layout of the industrial LV distribution system at which Dataset_indust_load was collected.

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 PV power data contained in Dataset_PVI1 and Dataset_PVI2, allowing their usage as exogenous variables for PV power forecasting models. Twenty-six variables are monitored in this way.

Dataset_weath_WG: this dataset includes absolute wind speeds and wind directions measured at each of the ten generators related to Dataset_WG. Wind data were collected throughout three years with 10-minute acquisitions, for the same time intervals of those related to Dataset_WG, allowing their usage as exogenous variables for wind power forecasting models.

Dataset_weath_ECMWF: weather forecast data are gathered from the European Centre for Medium-range Weather Forecasts (ECMWF) [56] for the locations and the time intervals corresponding to Dataset_PVI1, Dataset_PVI2, Dataset_PVI3 and Dataset_WG. Requests are prepared in Python3.7 and sent via the ECMWF Application Programming Interface (API). Forecasts for nine variables are obtained in this way. These data are related to the noon run (i.e., forecasts are issued at 12:00 A.M. of day D-1 for the entire day D) and to the midnight run (i.e., forecasts are issued at 12:00 P.M. of day D-1 for the entire day D). This differentiation in the weather forecast lead time allows developing models diversified for day-ahead control and real-time control of the distributed energy resources. In particular, the forecasts related to the noon run are used for the forecasting models aimed at the pre-scheduling control of the distribution grids, whereas the forecasts related to the midnight run are used for the forecasting models aimed at the real-time control of the distribution grids.

4.2. DATA PRE-PROCESSING: CLEANSING

The considered datasets are initially cleaned as an initial analysis revealed some potential outliers, missing and bad data. If not corrected, these data may significantly deteriorate the performance of the forecasting systems.

One of the principal objectives of the pre-processing activity is therefore to correct and remove this harmful effect by cleansing the data. Bad and missing data are easy to be individuated by visual inspection. Potential outliers instead are more subtle, since they cannot be immediately individuated by visual inspection. A slight modification of the Tukey's test [87] has been applied in order to individuate potential outliers. Tukey's test acts by examining and individuating data which lie beyond a specific band of tolerance, in which the null hypothesis can be rejected. For the generic variable y, the band of tolerance is individuated through its lower bound $y_{T,low}$ and upper bound $y_{T,up}$:

$$y_{T,low} = y^{\langle 0.25 \rangle} - 3 \cdot (y^{\langle 0.75 \rangle} - y^{\langle 0.25 \rangle})$$
$$y_{T,up} = y^{\langle 0.75 \rangle} + 3 \cdot (y^{\langle 0.75 \rangle} - y^{\langle 0.25 \rangle}) \quad (1)$$

where $y^{\langle 0.25 \rangle}$ and $y^{\langle 0.75 \rangle}$ are respectively the 0.25-quantile (25-percentile) and the 0.75-quantile (75-percentile) of the samples collected in the entire dataset. With reference to the PV and load datasets, due to the strong seasonality of the PV power data patterns and of the load patterns (which suggests heteroskedasticity), the Tukey's test has been differentiated for each hour of the day, accounting for different sample quantiles during the 24 hours of the day.

Potential outliers, bad data and missing data are treated in the same manner for each dataset, i.e., they are entirely discarded.

It is important to note that the only dataset not considered in the cleaning stage is the Dataset_weath_ECMWF, since the weather forecasts are already pre-processed by the ECMWF source.

4.3. DATA PRE-PROCESSING: AVERAGING

In all the considered experimental frameworks, the time resolution is one hour. An important objective of the data pre-processing activity is therefore to average values collected at different time resolution (for example, 1 minute for the Dataset_PVI2 of 10 minutes for the Dataset_WG) in order to obtain hourly data. This has been performed in R environment using the *lubridate* package [88]. In presence of too many removed data (i.e., beyond 30% of the total observations in the considered hour), the entire hourly value is set at the mean value from the two nearest hourly values. In presence of fewer removed data (i.e., less than 30% of the total observations in the considered hour), the entire hourly value is set at the mean value of the sub-hour observations within the considered hour.

It is important to note that ECMWF weather forecasts are already provided at hourly time resolution by the original source, therefore data averaging activity has not been developed on them.

4.4. DATA PRE-PROCESSING: NORMALIZATION

The last objective of the data pre-processing activity is to normalize hourly values in the range 0-1. This accommodation is usually necessary when the ranges in which the considered variables are included are very different. Although some forecasting models are insensitive to data normalization, other models may be significantly affected by the lack of normalization. All the data are normalized in order to be used in any case. The normalized value $\tilde{y}_h$ of the generic variable y occurred at hour h is:

$$\tilde{y}_h = \frac{y_h - y_{\min}}{y_{\max} - y_{\min}}, \tag{2}$$

where y_h is the value observed at hour h, and $y_{\min}$ and $y_{\max}$ are respectively the minimum and maximum values observed in the entire dataset.

4.5. EXPLORATORY DATA ANALYSIS

The database resulting from the data pre-processing activities consists of hourly observations of target variables (load, PV powers and wind powers) several exogenous weather variables. In order to reduce the dimensionality of the problem, an exploratory data analysis has been carried out to individuate exogenous variables which are informative for each target variable power, and those which are uninformative.

To avoid redundancy, in this sub-Section we present only the exploratory data analysis performed to individuate potential relationship between the PV power in Dataset_PVI2 and the ECMWF weather forecasts. Further information regarding specific exploratory data analysis are presented case-by-case in the experimental result Sections of this Deliverable.

The exploratory data analysis performed to individuate potential relationship between the PV power in Dataset_PVI2 and the ECMWF weather forecasts is performed via graphical inspection of relative scatter plots. As an example, scatter plots of the normalized PV power versus the normalized ECMWF clear-sky irradiance forecasts (Fig. 5.2a) and versus the normalized ECMWF solar irradiance forecasts (Fig. 5.2b) evidence clear relationship among these variables. Nevertheless, this relationship is not steady across the hour of the day, as patterns clearly differ considering, for example, 12 A.M., 9 A.M., and 6. P.M in the figures. From the graphical inspection of Fig. 5.2, it is suggested to add normalized ECMWF clear-sky irradiance forecasts and normalized ECMWF solar irradiance forecasts as candidate predictors of PV power forecasting models and to add a dummy variable to differentiate among the hours of the day.

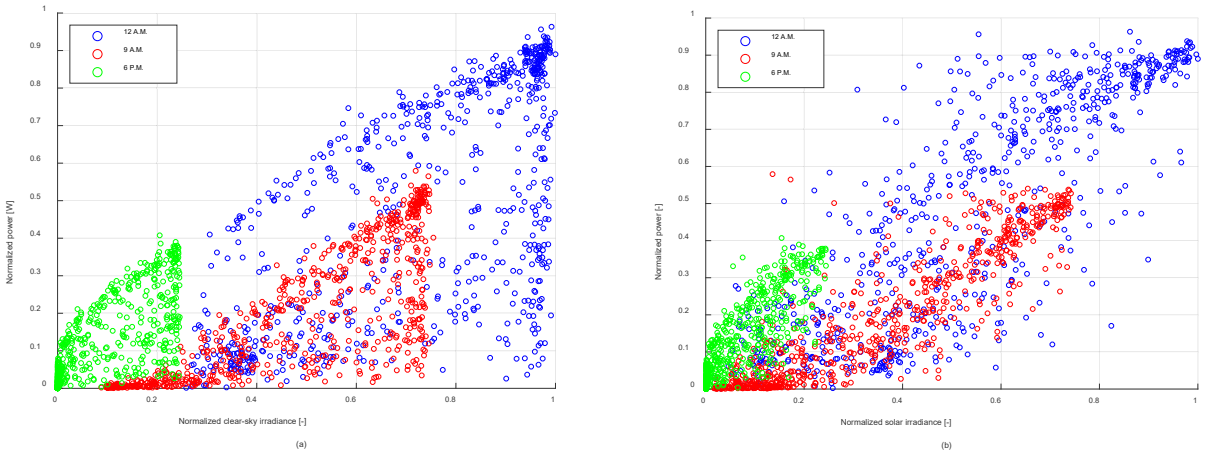

Figure 5.2. Scatter plots of the normalized PVI2 power versus the normalized ECMWF clear-sky irradiance forecasts (a) and versus the normalized ECMWF solar irradiance forecasts (b) for three different hours of the day.

The exploratory data analysis also allows discarding some variables which cannot be considered informative for predicting PV power. As significant example, Fig. 5.3 shows the scatter plots of the normalized PV power versus the normalized ECMWF forecasts of wind speed at 10 m. No clear relationship can be evidenced from this plot, as the cloud of points is very irregular. Also, there are no clear patterns differentiated among the hours of the day. For this reason, it can be considered safe to discard normalized ECMWF forecasts of wind speed at 10 m in predicting PVI2 power, thus reducing the dimensionality of the problem.

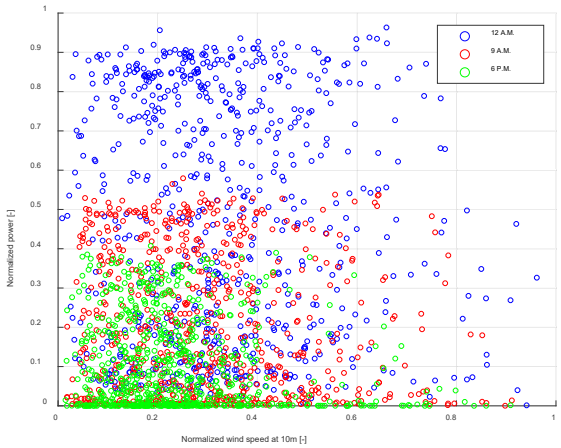

Figure 5.3. Scatter plots of the normalized PVI2 power versus the normalized ECMWF forecasts of wind speed at 10 m, for three different hours of the day.

5. RESULTS OF ACTIVITIES FOR TASK 2.2

The Task 2.2 focuses on the development of day-ahead forecasting systems for renewable generation and loads. With this reference, in the following Sections the following results are described:

i. the development of a Bayesian bootstrap quantile regression model for probabilistic photovoltaic power forecasting
ii. the development of a day-ahead probabilistic wind power forecasting based on ranking and combining Numerical Weather Predictions
iii. the development of a multivariate approach for probabilistic industrial load forecasting.

5.1. DEVELOPMENT OF DAY-AHEAD FORECASTING SYSTEMS FOR PHOTOVOLTAIC POWER GENERATION

Generally speaking, the integration of NWPs into PV power forecasting models is quite mandatory to improve predictions, particularly for large time horizons [42]. 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 [43] such as bagging, boosting and stacking have been applied with success in probabilistic energy forecasting [32,33,43,51-53]. The research activity presented in this Section provides a contribution to probabilistic ensemble PV power forecasting within the bagging framework, based on the interaction between a Quantile Regression (QR) model and a Bayesian bootstrap. A Bayesian bootstrap is the Bayesian analogue for bootstrapping, originally presented in [89]. Bayesian models have been applied with success in probabilistic energy forecasting [13,90,91], 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 below, 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. 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 **Dataset_PVI2** and **Dataset_PVI3**, together with the corresponding NWPs contained in **Dataset_weath_ECMWF** and **Dataset_weath_PV** [40]. Extensive numerical experiments, based on these data and NWPs provided by an external source [56], are presented. Several related benchmarks are also presented to validate the accuracy of the forecasts under comparative analyses.

To avoid verbosity in the analytic formulations, the symbols are consistent within the following sub-Sections 6.1.1, 6.1.2 and 6.1.3 only.

5.1.1. MODELS AND METHODS

The proposed PV power forecasting system based on Bayesian Bootstrap Quantile Regression (BBQR) is illustrated in Figure 6.1.1. The inputs of the system are NWPs $\boldsymbol{NW}$ and historical measured PV power data $\boldsymbol{P}$. 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 $\boldsymbol{X}$ (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 $\boldsymbol{P}_{tra}$ and $\boldsymbol{X}_{tra}$ (a $1 \times T$ vector and a $T \times M$ matrix, respectively), and into validation datasets $\boldsymbol{P}_{val}$ and $\boldsymbol{X}_{val}$ (a $1 \times V$ vector and a $V \times M$ matrix, respectively).

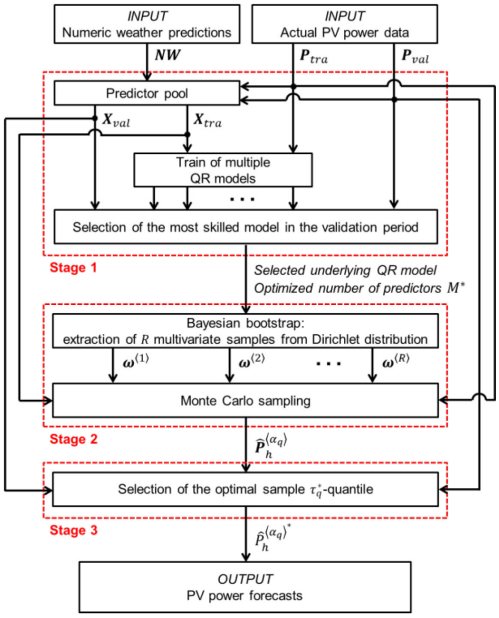

Fig. 6.1.1. Schematic procedure of the forecasting system based on Bayesian bootstrap quantile regression

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) [92] 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 ($\widehat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$) 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 $\widehat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$, is developed and presented here. 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 below.

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 $\boldsymbol{x}_h = \{x_{1_h}, \dots, x_{M_h}\}$ related to the time horizon h but available at the forecast origin $h - k$. The forecast lead time is indicated by k. This document 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., $k = 1, \dots, 24$), 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 $k = 5, \dots, 20$ are considered. In order to lighten the notation, 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:

$$\hat{P}_h^{\langle \alpha_q \rangle} = \hat{\beta}_0^{\langle \alpha_q \rangle} + \sum_{m=1}^{M} \hat{\beta}_m^{\langle \alpha_q \rangle} \cdot x_{m_h}, \qquad (6.1.1)$$

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 (6.1.1) 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 [14]).

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 PV power forecasts. The minimization problem is:

$$\widehat{\boldsymbol{\beta}}^{\langle\alpha_q\rangle} = \underset{\boldsymbol{\beta}^{\langle\alpha_q\rangle}}{\operatorname{argmin}}\, PS\left(\boldsymbol{P}_{tra}, \widehat{\boldsymbol{P}}_{tra}^{\langle\alpha_q\rangle}\right), \tag{6.1.2}$$

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

Although it is not directly explained in (6.1.2), the forecasts $\widehat{\boldsymbol{P}}_{tra}^{\langle\alpha_q\rangle}$ 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:

$$\widehat{\boldsymbol{P}}_{tra}^{\langle\alpha_q\rangle} = \boldsymbol{f}_{QR}\left(\boldsymbol{\beta}^{\langle\alpha_q\rangle} | \boldsymbol{X}_{tra}\right), \tag{6.1.3}$$

and therefore (6.1.2) can be rewritten in compact form as:

$$\widehat{\boldsymbol{\beta}}^{\langle\alpha_q\rangle} = \boldsymbol{G}(\boldsymbol{P}_{tra}, \boldsymbol{X}_{tra}), \tag{6.1.4}$$

where $\boldsymbol{G}(\cdot)$ is a function obtained by combining (6.1.2) and (6.1.3).

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. 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 [56]. 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 $\boldsymbol{P}_{val} = \{P_{v_1}, \dots, P_{v_V}\}$, 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 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 $\widehat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle}$ from (2)-(4) on the bootstrapped data. Therefore, the posterior distribution of the QR model parameters is given by R samples $\widehat{\boldsymbol{\beta}}_1^{\langle \alpha_q \rangle}, \dots, \widehat{\boldsymbol{\beta}}_{M^*+1}^{\langle \alpha_q \rangle}$ for each parameter, and from these samples it is eventually obtains the bagged samples $\widehat{\boldsymbol{P}}_h^{\langle \alpha_q \rangle}$ of PV power.

To facilitate the presentation of the BBQR formulation, a brief recap on traditional bootstrapping [93,94] 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 n^{th} row vector $\boldsymbol{y}_{t_n} = \{P_{t_n}, x_{1_{t_n}}, \dots, x_{M^*_{t_n}}\}$, 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) $\boldsymbol{y}_{t_1}, \dots, \boldsymbol{y}_{t_T}$.

As shown earlier, the estimated parameters $\widehat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle}$ of the QR model come from (6.1.4), and therefore they can be viewed as a function of $\boldsymbol{G}[F(\boldsymbol{y})]$:

$$\widehat{\boldsymbol{\beta}}^{\langle \alpha_q \rangle} = \boldsymbol{G}[F(\boldsymbol{y})]. \tag{6.1.5}$$

In bootstrap (either traditional or Bayesian [93,94]), the unknown distribution $F(\boldsymbol{y})$ is searched for among distributions of the type $F_T(\boldsymbol{y})$:

$$F_T(\boldsymbol{y}) = \sum_{n=1}^{T} \omega_{t_n} \cdot \delta_{\boldsymbol{y}_{t_n}}, \tag{6.1.6}$$

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

$$\sum_{n=1}^{T} \omega_{t_n} = 1, \ \omega_{t_n} \geq 0 \ \forall n = 1, \dots, T. \tag{6.1.7}$$

In a traditional bootstrap, the function $\boldsymbol{G}[F(\boldsymbol{y})]$ is estimated upon R distributions $F_T^{\langle 1 \rangle}(\boldsymbol{y}), \dots, F_T^{\langle R \rangle}(\boldsymbol{y})$. With reference to the generic r^{th} replicate, the functional $\boldsymbol{G}\left[F_T^{\langle r \rangle}(\boldsymbol{y})\right]$ 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:

$$f_{Mul}(T; 1/T, 1/T, \dots, 1/T). \tag{6.1.8}$$

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 (6.1.8). 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 $p(\boldsymbol{\omega}|\boldsymbol{Y}_{tra})$ 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 [93,94]. A convenient choice is to select a Dirichlet distribution, which is a conjugate prior for the multinomial distribution of $\boldsymbol{y}$ [93,94]. In such a case, the posterior distribution $p(\boldsymbol{\omega}|\boldsymbol{Y}_{tra})$ is itself a Dirichlet distribution, $f_{Dir}(1, \dots, 1; 1, \dots, 1)$ [93]. 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) $\boldsymbol{G}\left[F_T^{\langle 1 \rangle}(\boldsymbol{y})\right], \dots, \boldsymbol{G}\left[F_T^{\langle R \rangle}(\boldsymbol{y})\right]$ are calculated applying (6.1.2)-(6.1.4);

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 (6.1.5). 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 (6.1.1). The set of these samples are indicated with $\widehat{\boldsymbol{P}}_h^{\langle\alpha_q\rangle}$.

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

The R samples $\widehat{\boldsymbol{P}}_h^{\langle\alpha_q\rangle}$ 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 $\widehat{\boldsymbol{P}}_h^{\langle\alpha_q\rangle}$. Since it will be of use later, the generic sample τ_q-quantile estimated from $\widehat{\boldsymbol{P}}_h^{\langle\alpha_q\rangle}$ 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 research activity 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 $\widehat{\boldsymbol{P}}_h^{\langle\alpha_q\rangle}$ as:

$$\hat{P}_h^{\langle\alpha_q\rangle^*} = \hat{P}_h^{\langle\alpha_q\rangle\langle\tau_q^*\rangle}, \tag{6.1.9}$$

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.:

$$\tau_q^* = \underset{\tau_q}{\operatorname{argmin}}\, PS\left(\boldsymbol{P}_{val}, \widehat{\boldsymbol{P}}_{val}^{\langle\alpha_q\rangle\langle\tau_q\rangle}\right). \tag{6.1.10}$$

Note that this procedure is made independent for each quantile coverage $\alpha_1, \dots, \alpha_Q$, 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.

5.1.2. EXPERIMENTAL FRAMEWORK AND RESULTS

This section presents the experimental framework, in terms of benchmarks and the error indices used to assess the validity of the proposal and to compare the probabilistic forecasts, and the outcomes of the experimental results.

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 6.1.1 within the framework of the proposed forecasting system based on BBQR is also adopted for SQR.

Traditional Bootstrap QR (TBQR): the second benchmark [95] 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 in Section 6.1.1, are applied to TBQR.

Quantile Regression Neural Network (QRNN) and Gradient Boosting Regression Tree (GBRT): the third and fourth benchmarks of QRNN [96] and GBRT [97] 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 [96]. 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 [98].

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 [99].

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. In practice, each predictive quantile is the PV power observed the day before:

$$\hat{P}_h^{\langle \alpha_q \rangle} = P_{h-24}. \tag{6.1.11}$$

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

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 [92] that simultaneously addresses the reliability and the sharpness of forecasts. Its formulation is:

$$PS\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle}\right) = \left[\alpha_q - \mathrm{I}\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle}\right)\right] \cdot \left(P_h - \hat{P}_h^{\langle \alpha_q \rangle}\right), \tag{6.1.12}$$

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

$$\mathrm{I}\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle}\right) = \begin{cases} 1 & if\ P_h \leq \hat{P}_h^{\langle \alpha_q \rangle} \\ 0 & if\ P_h > \hat{P}_h^{\langle \alpha_q \rangle} \end{cases}. \tag{6.1.13}$$

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:

$$NPS = \frac{PS\left(P_h, \hat{P}_h^{\langle \alpha_q \rangle}\right)}{\bar{P}_{\text{rated}}}, \tag{6.1.14}$$

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 [23,42]. Because of its intrinsic properties, it can only be formulated for multiple forecast issues. For sake of clarity it is referred in (6.1.15) 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:

$$\hat{\alpha}_q = \frac{1}{V} \cdot \sum_{n=1}^{V} \mathrm{I}\left(P_{v_n}, \hat{P}_{v_n}^{\langle \alpha_q \rangle}\right), \tag{6.1.15}$$

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

$$ACE^{\langle \alpha_q \rangle} = \left|\alpha_q - \hat{\alpha}_q\right|. \tag{6.1.16}$$

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

$$AACE_{\%} = \frac{100}{Q} \cdot \sum_{q=1}^{Q} ACE^{\langle \alpha_q \rangle}. \tag{6.1.17}$$

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

3) Data and forecasting framework

The proposed forecasting system is assessed using **Dataset_PVI2** and **Dataset_PVI3**.

Data are normalized in the 0-1 range to accommodate for the very different intervals spanned by the variables, and they are partitioned in three subsets. **Dataset_PVI2** is split into a training set spanning February 1, 2016 to December 31, 2017 (i.e., $T = 16800$), a validation set covering first six months of 2018 (i.e., $V = 4344$), and a test set only used to assess the accuracy of forecasts, covering the remaining five months of 2018 (i.e., 3672 forecast issues). Similarly, **Dataset_PVI3** is split into a training set from April 1, 2012 to October 31, 2013 (i.e., $T = 13896$), a validation set covering the following five months (i.e., $V = 3624$), and a test set covering the remaining three months of 2014 (i.e., 2184 forecast issues).

The number R of bootstrapped is searched for in the range 1000-10000, 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 $\boldsymbol{P}_{val}$. 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 $\alpha_1, \ldots, \alpha_{19} = 0.05, \ldots, 0.95$. All forecasts are developed in the R environment, exploiting packages *quantreg* [100] and *bayesboot* [101], *qrnn* [98] and *gbm* [99]. The extensive results for **Dataset_PVI2** are presented below, whereas the results for **Dataset_PVI3** are presented in a more compact form, to avoid verbose presentation.

4) Results for Dataset_PVI2

The outcome of the first stage of the proposed forecasting system determines the model selected for the specific PV system. For **Dataset_PVI2**, the selected model is:

$$\hat{P}_h^{\langle\alpha_q\rangle} = \hat{\beta}_0^{\langle\alpha_q\rangle} + \hat{\beta}_1^{\langle\alpha_q\rangle} \cdot tcc_h + \hat{\beta}_2^{\langle\alpha_q\rangle} \cdot ti_h + \hat{\beta}_3^{\langle\alpha_q\rangle} \cdot csi_h + \hat{\beta}_4^{\langle\alpha_q\rangle} \cdot P_{h-24} + \hat{\beta}_5^{\langle\alpha_q\rangle} \cdot tcc_h \cdot ti_h + \\ + \hat{\beta}_6^{\langle\alpha_q\rangle} \cdot tcc_h \cdot csi_h + \hat{\beta}_7^{\langle\alpha_q\rangle} \cdot ti_h \cdot P_{h-24}, \quad (6.1.18)$$

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_PVI2** are shown in Table 6.1.1 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) outperform SQR, although the Bayesian-based procedure slightly outperforms the traditional bootstrapping procedure in terms of NPS.

Table 6.1.1 Forecast results for the test set of Dataset_PVI2

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	-

Further details on the skill of forecasts can be evaluated from the results of the experiments. Figure 6.1.2 shows the NPS, averaged through the test set of BBQR, SQR and TBQR forecasts for each nominal quantile level. Figure 6.1.3 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 patterns have small errors in proximity to dawn and dusk.

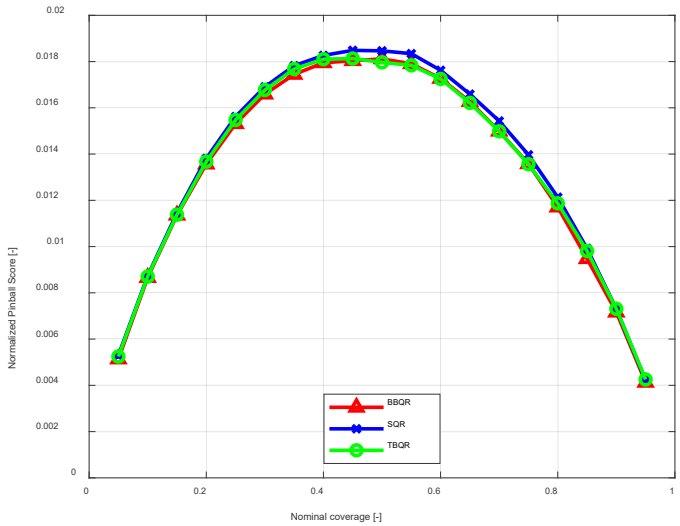

Fig. 6.1.2. NPS of probabilistic forecasts versus the nominal quantile coverage for the test set of Dataset_PVI2.

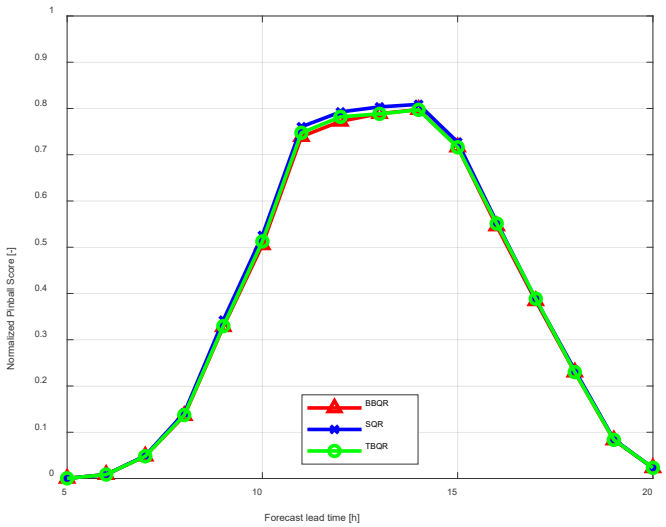

Fig. 6.1.3. NPS of probabilistic forecasts versus the forecast lead time for the test set of Dataset_PVI2.

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, Figure 6.1.4 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, Figure 6.1.5 shows the BBQR prediction intervals for one 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.

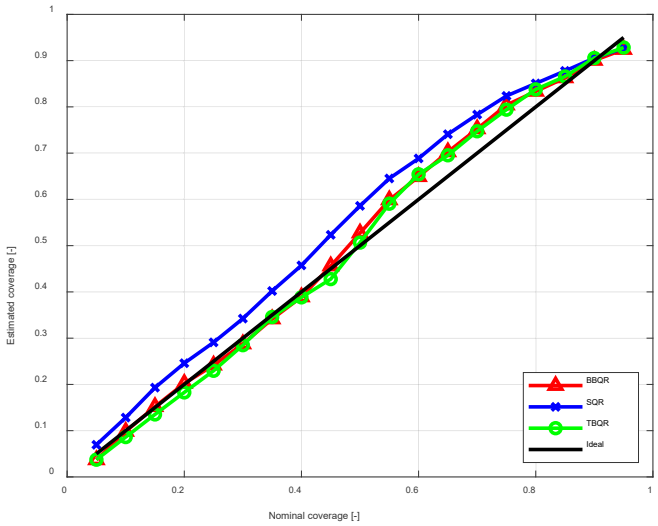

Fig. 6.1.4. Reliability diagrams of the QR-based probabilistic forecasts for the test set of Dataset_PVI2.

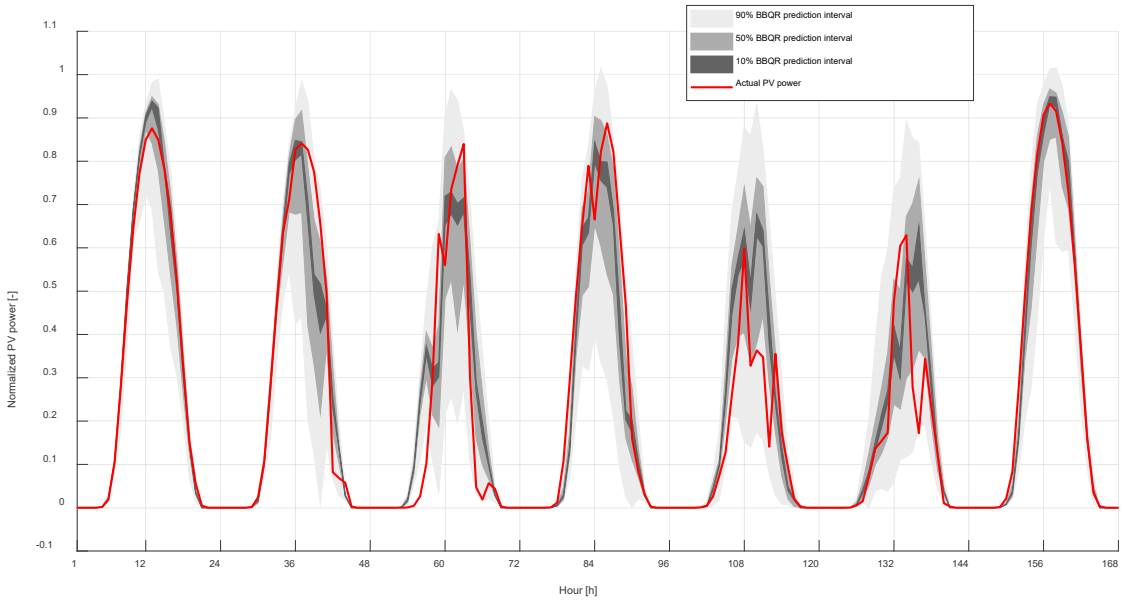

Fig. 6.1.5. BBQR prediction intervals during one week of the test set of Dataset_PVI2.

5) Results for Dataset_PVI3

The results for **Dataset_PVI3** are presented in a more compact form, to avoid duplication. Table 6.1.2 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 6.1.6 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.

Table 6.1.2 Forecast results for the test set of Dataset_PVI3

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	-

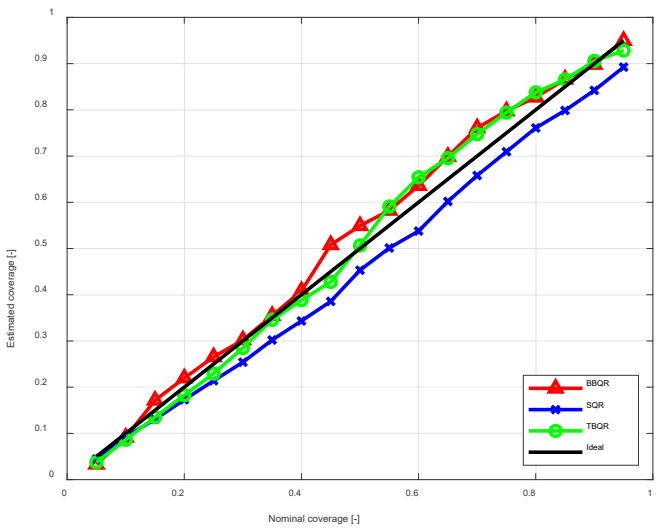

Fig. 6.1.6. Reliability diagrams of the QR-based probabilistic forecasts for the test set of Dataset_PVI3.

5.1.3. 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 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.

5.2. DEVELOPMENT OF DAY-AHEAD FORECASTING SYSTEMS FOR WIND POWER GENERATION

Forecasters almost unanimously agree that exploiting NWPs enhances the skill of wind power forecasts for short-term horizons (e.g., day ahead horizon), as stochastic-only forecasting systems usually perform well only for very-short lead times (e.g., 1-hour ahead), gradually losing relevance as the lead times increases. On the other hand, the availability of accurate NWPs has grown in the last years, thanks to the efforts in atmospheric modeling and in satellite acquisitions, as well as predictions of many weather variables are nowadays made available by weather institutions [56], and they can be used as input predictors of forecasting systems. Getting the most from the available data to increase the skill of wind power forecasts is an object of energy data analysis. Exploiting spatially-distributed NWPs for an area surrounding the wind farm has been proved to be an efficient solution to further add skill [57-59]. However, adding too many predictors to the models could involve the risk of overfitting the data thus losing forecast skill. Standard techniques to reduce overfitting are based on dimensionality reduction and feature selection (or feature engineering) [57,60-63], which are often accomplished through validation, leave-one-out cross-validation or k-fold cross-validation frameworks. These latter procedures usually are computationally intensive, particularly when many NWPs are candidate predictors [64], as it can happen when handling spatially-distributed NWPs for an area

surrounding the wind farm. This poses an arduous challenge to the development of forecasting systems in which models are re-trained as new observations become available.

In this research activity, we develop a simplified technique for feature selection in wind forecasting, based on ranking and combining spatially-distributed NWPs available for a given area surrounding the wind farm. We also develop a forecasting methodology to generate probabilistic forecasts of wind power on the basis of such a simplified technique.

Since wind farms usually cover a large area, in the proposed methodology NWPs are ranked in-sample and combined for each of the wind turbines, in order to provide more skilled forecasts for each individual generator. We expect that this approach may increase the exploitation of the information contained in all of the available NWPs and in the available individual-generator measured data, compared to the traditional case in which only the NWPs related to the specific location of the wind turbine are used.

The main advantage of the proposed procedure is that ranking NWPs by their in-sample accuracy reduces the risk of overfitting the training data, which would be a compelling problem if all of the available NWPs are naively added as predictors of the probabilistic model. This technique simplifies the feature selection for the underlying model which generates the forecasts. In addition, a single set of fictitious NWPs, gathered from an opportune combination of actual NWPs, is built to reduce the dimensionality of the problem.

In particular, two different NWP in-sample ranking procedures are proposed. One is based on the accuracy of absolute wind speed, and the other one is based on the combined accuracy of u- and v-components of the wind speed. The latter procedure is specifically tested in this research activity, due to the high influence of wind direction on wind generation.

Once NWPs are ranked in-sample, an ensemble (fictitious) set of NWP is obtained from the weighted combination of ranked NWPs. Three different approaches are developed to estimate the weights of the combination, in order to search for the most suitable solution and to provide diversity.

Two types of model are considered to predict wind generation. One is based on GBRT [97], due to the good performance shown during energy forecasting competitions [40]. The other is based on QRNN [96], due to the promising result obtained by neural networks in probabilistic energy forecasting. Selecting two theoretically dissimilar types of models allows testing the effectiveness of the NWP ranking and combination procedure in a fair, model-free manner. The proposed methodology is applied for day-ahead wind power forecasting of individual generators and of the entire wind farm.

In summary, the main contributions of this research activity are:

- the proposal of ranking and combining NWPs by their in-sample accuracy on individual wind generators, to form an ensemble set of predictors for the probabilistic forecasting model. Ranked predictors are put in an ensemble, which is a fictitious set of NWPs formed through three different weighted combination approaches;
- the comparison between two ranking procedures, the first based on in-sample accuracy of the absolute wind speed and the second based on in-sample combined accuracy of u- and v-components of the wind speed;
- the development of a new forecasting methodology based on this proposal, through the integration of the ranking procedure, of the combination approach, and either GBRT or QRNN models;
- the development of experimental results on an actual wind farm, composed of ten individual generators.

To avoid verbosity in the analytic formulations, the symbols are consistent within the following sub-Sections 6.2.1, 6.2.2 and 6.2.3 only.

5.2.1. Models and Methods

The probabilistic wind power forecasting methodology developed in this research activity is based on the proposal of ranking NWPs by their in-sample accuracy for each individual generator within the considered wind farm, and of combining them in a weighted combination approach. The scheme of the methodology is presented in Figure 6.2.1, and it includes five stages.

In the first stage of the methodology, the wind farm is characterized by identifying the location of individual generators and the data available from measurement stations installed at each generator. The exploratory data analysis allows to individuate the geographical matches among the N_s available sets $\boldsymbol{NWP}_1, \boldsymbol{NWP}_2, \dots, \boldsymbol{NWP}_{N_s}$ of NWPs and each of the N_g wind generators. NWPs are usually provided by weather forecasting services for elementary grid squares that have a fixed latitude-longitude resolution. It is here assumed that one NWP set is available for each elementary grid square of the globe, possibly partitioning the entire globe. Since each wind generator is located within a unique elementary grid square, it is possible to individuate the G^2 neighbor elementary grid squares surrounding each wind generator, assuming that these elementary grid squares form a composite grid square having a side of G. The corresponding G^2, related to the G^2 neighbor elementary grid squares surrounding each wind generator, are stored in N_g datasets $\boldsymbol{NWP}_1^*, \boldsymbol{NWP}_2^*, \dots, \boldsymbol{NWP}_{N_g}^*$, which are the outputs of the first stage of the methodology. Note that G is a hyper-parameter of the forecasting system, which is optimized in a cross-validation scheme. It is hereinafter assumed that NWPs include predictions of the absolute wind speed, and possibly of the u- and v-components of wind speed.

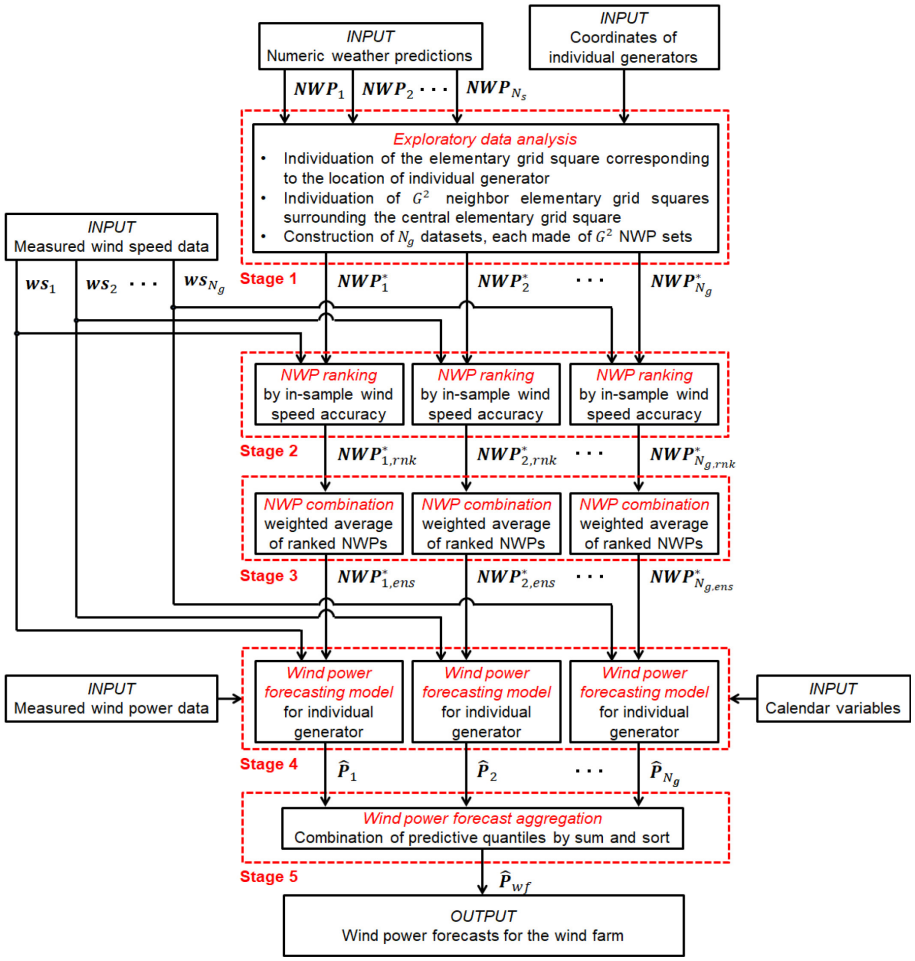

Fig. 6.2.1. Scheme of the proposed wind power forecasting methodology.

In the second stage of the methodology, the spatially-distributed NWPs, available for the composite grid square surrounding individual generators, are ranked by the in-sample accuracy either of the absolute wind speed or of the u- and v-components of wind speed. The outputs of this stage are N_g ranked datasets $\boldsymbol{NWP}_{1,rnk}^*, \boldsymbol{NWP}_{2,rnk}^*, \dots, \boldsymbol{NWP}_{N_g,rnk}^*$.

In the third stage of the methodology, ranked NWPs are appropriately combined to form N_g ensemble (fictitious) datasets $\boldsymbol{NWP}_{1,ens}^*, \boldsymbol{NWP}_{2,ens}^*, \dots, \boldsymbol{NWP}_{N_g,ens}^*$, which are used as predictors of either GBRT or QRNN probabilistic model in the fourth stage of the methodology. In this research activity, three different combination approaches are considered to form the ensemble datasets. They are based on weighted average combinations of the NWPs.

In the fourth stage of the methodology, GBRT or QRNN probabilistic models forecast the power generated by individual turbines. These models exploit lagged wind power, lagged wind speed, the ensemble (fictitious) NWPs

returned by the third stage of the methodology, and calendar variables (hour of the day, day of the month and month of the year) to account for the possible seasonality of the wind power patterns. The outputs of the fourth stage of the methodology are the forecasts $\widehat{\boldsymbol{P}}_1, \widehat{\boldsymbol{P}}_2, \dots, \widehat{\boldsymbol{P}}_{N_g}$ of individual wind generators, given by predictive quantiles.

In the fifth stage of the methodology, the individual forecasts are aggregated to get the forecast $\widehat{\boldsymbol{P}}_{wf}$ of the power produced by the entire wind farm.

Details on the five stages are reported in the following subsections.

1) Exploratory data analysis

The exploratory data analysis aims at individuating correspondences among the available sets of NWPs and the data measured at each of the N_g wind generators. Due to its strict interaction with the proposal, it is presented here rather than in Section 5.5, although these activities were performed in the framework of the Task 2.1.

NWPs consist of predictions, provided by external weather forecasting institutions [56], of many weather variables. Only some of them are usually considered suitable candidate predictors for the wind power. To avoid unnecessary complexity, 7 variables [102] are considered in this proposal: pressure ap, precipitation tp, cloud cover cc, temperature at, u-component of wind speed us, v-component of wind speed vs, and absolute wind speed s.

NWPs are usually provided by weather forecasting services for elementary grid squares that have a fixed latitude-longitude resolution (e.g., 0.1x0.1 or 0.25x0.25 degrees of latitude-longitude). These elementary grid squares partition the entire globe, so that forecasts are theoretically available for the entire globe. Intuitively, spatially-distributed NWPs lose relevancy for the wind power of a specific wind generator as much as the related elementary grid is farther from the wind generator.

It is here assumed that N_s sets $\boldsymbol{NWP}_1, \boldsymbol{NWP}_2, \dots, \boldsymbol{NWP}_{N_s}$ of NWPs are available for a reasonably-wide area around the wind farm. Each set is a $N \times 7$ matrix, where N is the total number of historical NWP instances. For the generic k^{th} set, the n^{th} row has the form:

$$\boldsymbol{NWP}_{k_n} = \{NWP_{k_{n,1}}, \dots, NWP_{k_{n,7}}\} = \{\widehat{ap}_{k_n}, \widehat{tp}_{k_n}, \widehat{cc}_{k_n}, \widehat{at}_{k_n}, \widehat{us}_{k_n}, \widehat{vs}_{k_n}, \hat{s}_{k_n}\}. \quad (6.2.1)$$

Feature engineering may solve the problem of selecting only the most relevant sets of the available NWPs for the purpose of wind power forecasting, but it is very computationally challenging since it requires validation or cross-validation over a very large number of different combinations of predictors. This complicates the development of forecasting systems in which models are dynamically re-trained as new observations become available.

To cope with this problem by a simplified technique, in the first stage of the proposed methodology we propose to geographically match the available NWPs with individual wind generators, based on their spatial location. Each wind generator is located within a unique elementary grid square, which is individuated from the match between latitude-longitude coordinates of generators and latitude-longitude coordinates of elementary grid squares. It is straightforward to individuate G^2 elementary grid squares which surround each wind generator, assuming that these elementary grid squares form a composite grid square having side G and assuming that the central elementary grid square is the one corresponding to the location of the wind generator. Note that some of the elementary grid squares may be overlapped for different generators, on the basis of the spatial distribution of the generators within the wind farm. By applying this approach, feature engineering simply collapses into the search for an optimal value of the integer G, which is a hyper-parameter of the forecasting methodology. The spatial resolution of the available NWPs and the latitude at which the wind farm is located determine reasonable values among which G should be searched: if NWPs are too sparse and the wind farm is sufficiently far from the poles, it is redundant to search G among too large values. For example, if NWPs are provided for elementary grid squares of 0.25x0.25 degrees of latitude and longitude and latitude is around 45°N, a value $G = 10$ would approximately correspond to a composite grid square of about 250x200 km^2. Since in most applications G can be searched within small, positive integers, a simple grid search procedure is applied to optimize it in 10-fold cross-validation.

The exploratory data analysis is performed for each of the N_g generators of the wind farm, so that N_g NWP datasets related to the G^2 neighbor elementary grid squares surrounding each generator can be individuated. These N_g datasets $\boldsymbol{NWP}_1^*, \boldsymbol{NWP}_2^*, \dots, \boldsymbol{NWP}_{N_g}^*$, which are $N \times 7 \times G^2$ tensors, and they are the outcomes of the first stage of the proposed methodology. They are individually ranked in the second stage of the proposed methodology, as shown in the next Section.

2) Ranking NWPs for individual wind generators

Exploiting spatial information by ranking and combining NWPs has been successfully exploited in several energy forecasting applications [57-59,103]. In the second stage of the methodology, the N_g datasets of spatially-distributed NWPs are ranked by their in-sample accuracy. Practically, the data contained in each tensor $\boldsymbol{NWP}_1^*, \boldsymbol{NWP}_2^*, \dots, \boldsymbol{NWP}_{N_g}^*$ are ordered across the third dimension of the tensor, returning N_g ranked datasets $\boldsymbol{NWP}_{1,rnk}^*, \boldsymbol{NWP}_{2,rnk}^*, \dots, \boldsymbol{NWP}_{N_g,rnk}^*$, which are $N \times 7 \times G^2$ tensors too. The ranking is performed by analyzing the in-sample accuracy either of the absolute wind speed or of the u- and v-components of wind speed, to provide and assess diversification in the approaches. In particular, the second procedure is tested since it accounts also for wind direction, which particularly affects wind generation. The two procedures are presented below.

Ranking procedure by absolute wind speed: Ranking NWPs by the in-sample wind speed accuracy is quite straightforward. We refer hereinafter to the generic individual generator i.

Assuming that N_{tr} NWP instances are reserved to evaluate the in-sample accuracy, and that the corresponding wind speed $s_{i_1}, \dots, s_{i_{N_{tr}}}$ are measured at the i^{th} individual generator during the same time intervals, the in-sample wind speed Mean Squared Errors (MSEs) are calculated as:

$$MSE_{i_j}(s) = \frac{1}{N_{tr}} \cdot \sum_{n=1}^{N_{tr}} \left(s_{i_n} - \hat{s}_{i_{n,j}} \right)^2, \quad j = 1, \dots, G^2, \tag{6.2.2}$$

where $\hat{s}_{i_{n,j}}$ is the wind speed NWP at time n, contained in the tensor $\boldsymbol{NWP}_i^*$ and related to the j^{th} elementary grid square (i.e., the j^{th} slot on the third dimension of the tensor).

The ranked dataset $\boldsymbol{NWP}_{i,rnk}^*$ is obtained straightforwardly by ordering the G^2 MSEs in ascending order and mirroring the sort also over the third dimension of the tensor $\boldsymbol{NWP}_i^*$.

Ranking procedure by u- and v-components of wind speed: Ranking NWPs by the in-sample accuracy of u- and v-components of wind speed is less straightforward, since there are two variables to deal with.

Assuming that the u-components $us_{i_1}, \dots, us_{i_{N_{tr}}}$ and the v-components $vs_{i_1}, \dots, vs_{i_{N_{tr}}}$ of wind speed are measured during the same time intervals of the NWP instances, two in-sample MSEs are calculated:

$$MSE_{i_j}(us) = \frac{1}{N_{tr}} \cdot \sum_{n=1}^{N_{tr}} \left(us_{i_n} - \widehat{us}_{i_{n,j}} \right)^2, \quad j = 1, \dots, G^2, \tag{6.2.3}$$

$$MSE_{i_j}(vs) = \frac{1}{N_{tr}} \cdot \sum_{n=1}^{N_{tr}} \left(vs_{i_n} - \widehat{vs}_{i_{n,j}} \right)^2, \quad j = 1, \dots, G^2, \tag{6.2.4}$$

where $\widehat{us}_{i_{n,j}}$ and $\widehat{vs}_{i_{n,j}}$ are respectively the NWPs of the u-component and v-component of wind speed at time n, contained in the tensor $\boldsymbol{NWP}_i^*$ and related to the j^{th} elementary grid square (i.e., the j^{th} slot on the third dimension of the tensor).

In this case, the ranked dataset $\boldsymbol{NWP}_{i,rnk}^*$ is obtained straightforwardly by ordering the combined value $MSE_{i_j}(uvs)$ of these two MSEs:

$$MSE_{i_j}(uvs) = k_{us} \cdot MSE_{i_j}(us) + k_{vs} \cdot MSE_{i_j}(vs), \quad j = 1, \dots, G^2, \tag{6.2.5}$$

in ascending order and mirroring the sort also over the third dimension of the tensor $\boldsymbol{NWP}_i^*$. Theoretically, the weights k_{us} and k_{vs} assigned to the two MSEs could be different in situations in which there is a prevalent direction of the wind speed, in order to give more influence on the prevalent direction. Optimal values for these weights could be searched as well in the hyper-parameter optimization procedure. In this proposal, in order to keep the framework as much general and simple as it can be, we use same weights for the two MSEs (i.e., $k_{us} = k_{vs} = 1/2$), assuming that the MSEs of the u- and v-components of the wind speed bring the same information to generate forecasts.

3) Combination of the ranked NWPs for individual wind generators

The data contained in the ranked datasets are furthermore manipulated in the third stage of the methodology to create N_g ensemble datasets $\boldsymbol{NWP}_{1,ens}^*, \boldsymbol{NWP}_{2,ens}^*, \dots, \boldsymbol{NWP}_{N_g,ens}^*$, which are $N \times 7$ matrices used as predictors of the probabilistic model considered in the fourth stage.

The ensemble sets are obtained as a weighted average of the G^2 NWPs, where the weights depend on the rank (the highest ranks correspond to the greatest weights) [102]. In particular, the element $NWP_{i,ens_{r,c}}^*$ of $\boldsymbol{NWP}_{i,ens}^*$ at row r, column c is:

$$NWP_{i,ens_{r,c}}^* = \sum_{j=1}^{G^2} w_{i_j} \cdot NWP_{i,rnk_{r,c,j}}^*, \qquad r = 1, \dots, N \text{ and } c = 1, \dots, 7. \tag{6.2.6}$$

Note that weights w_{i_j} are the same for all of the considered weather variables contained in the NWP datasets. We develop three different combination approaches to estimate weights and to get the ranked NWPs. They are presented below.

Equally-Spaced Weighted Combination (ESWC): In this approach the weights are equally spaced and sum up to one. The greatest weight is assigned to the top-ranked NWP, and the smallest weight is assigned to the lowest-ranked NWP [102]. For the i^{th} wind generator, the j^{th} estimated weight $\widehat{w}_{i_j}$ is:

$$\widehat{w}_{i_j} = \frac{G^2 - r_{i_j} + 1}{\sum_{k=1}^{G^2} k}, \qquad j = 1, \dots, G^2, \tag{6.2.7}$$

where r_{i_j} is the rank of the NWP related to the j^{th} elementary grid square (i.e., the j^{th} slot on the third dimension of the tensor) surrounding the i^{th} wind generator.

Rank-Relevant Weighted Combination (RRWC): In this approach the weights are inversely-proportional to the values of the in-sample MSEs, which have led to the ranking order (thus, either $MSE_{i_j}(s)$ or $MSE_{i_j}(uvs)$). Weights sum up to one in this case too. For the i^{th} wind generator, the j^{th} estimated weight $\widehat{w}_{i_j}$ is:

$$\widehat{w}_{i_j} = \frac{1/MSE_{i_j}}{\sum_{k=1}^{G^2} 1/MSE_{i_k}}, \qquad j = 1, \dots, G^2. \tag{6.2.8}$$

In-sample MSE Weighted Combination (IMWC): In this approach the weights are estimated by optimizing the in-sample MSE, either of the ensemble absolute wind speed or of the ensemble u- and v-components of wind speed. In this case we do not constrain weights to sum up to one. For the i^{th} wind generator, considering respectively the ensemble absolute wind speed or the ensemble u- and v-components of wind speed, the estimated weights $\widehat{\boldsymbol{w}}_i$ are:

$$\widehat{\boldsymbol{w}}_i = \underset{\boldsymbol{w}_i}{\operatorname{argmin}} \frac{1}{N_{tr}} \cdot \sum_{n=1}^{N_{tr}} \left(s^*_{i,ens_n} - s_{i_n}\right)^2; \tag{6.2.9}$$

$$\widehat{\boldsymbol{w}}_i = \underset{\boldsymbol{w}_i}{\operatorname{argmin}} \frac{1}{N_{tr}} \cdot \sum_{n=1}^{N_{tr}} k_{us} \cdot \left(us^*_{i,ens_n} - us_{i_n}\right)^2 + k_{vs} \cdot \left(vs^*_{i,ens_n} - vs_{i_n}\right)^2, \tag{6.2.10}$$

where $s^*_{i,ens_n}, us^*_{i,ens_n}$ and vs^*_{i,ens_n} depend on weights $\boldsymbol{w}_i$, as it is shown in (6.2.6).

4) Gradient boosting regression tree

In the fourth stage of the methodology, the power generated by each individual wind turbine is probabilistically forecasted through a dedicated model. We validate the proposal using either GBRT or QRNN models to generate individual forecasts.

GBRT models are based on regression trees, which forecast the target variable (in this proposal, the power produced by an individual wind generator) grouping past observations by splitting on training data predictors [97]. At the root of the tree, all the training data are in the same group. A non-overlapping condition is posed upon a predictor, and branches grow from the split node. Children data subsets are individuated through the split, and they do not contain overlapping data. The local optimal condition for the split is searched by minimizing the MSE related to the children data subsets.

For the i^{th} generator, a regression tree iteratively fits the power residuals $P_i - \hat{P}_i$ of a weak learner $\hat{P}_i = \hat{f}_i(\boldsymbol{x}_i)$ until convergence over the minimization of a loss function $\psi_i\left(P_i, f_i(\boldsymbol{x}_i)\right)$ is reached (i.e., the relative improvement given by the last iteration is smaller than an assigned threshold). $\boldsymbol{x}_i$ is the vector containing the predictors of the

model (selected among lagged wind power, lagged wind speed, ensemble NWPs returned by the previous stage of the methodology, and calendar variables such as hour of the day, day of the month and month of the year, to account for the possible seasonality of the wind power patterns) for the i^{th} generator. At iteration $m = 0$, the function $\hat{f}_i^{(0)}(\boldsymbol{x}_i)$ is initialized at a constant value $\hat{\rho}_i^{(0)}$ given by:

$$\hat{\rho}_i^{(0)} = \underset{\rho}{\operatorname{argmin}} \sum_{n=1}^{N_{tr}} \psi_i(P_{i_n}, \rho) \,. \tag{6.2.11}$$

For subsequent iterations $m > 0$, the negative gradients of the loss function evaluated in $\hat{f}_i^{(m-1)}(\boldsymbol{x}_{i_n})$ are calculated for $n = 1, \dots, N_{tr}$:

$$g_{i_n} = -\left[\frac{\partial \psi_i\left(P_{i_n}, f_i^{(m-1)}(\boldsymbol{x}_{i_n})\right)}{\partial f_i^{(m-1)}(\boldsymbol{x}_{i_n})}\right]_{f_i^{(m-1)}(\boldsymbol{x}_{i_n})=\hat{f}_i^{(m-1)}(\boldsymbol{x}_{i_n})} . \tag{6.2.12}$$

A regression tree $T(\boldsymbol{x}_i)$ is fitted upon predictors $\boldsymbol{x}_{i_1}, \dots, \boldsymbol{x}_{i_{N_{tr}}}$ and corresponding outcomes $g_{i_1}, \dots, g_{i_{N_{tr}}}$, and the gradient descent step size $\hat{\rho}_i^{(m)}$ is estimated as:

$$\hat{\rho}_i^{(m)} = \underset{\rho}{\operatorname{argmin}} \sum_{n=1}^{N_{tr}} \psi_i\left(P_{i_n}, \hat{f}_i^{(m-1)}(\boldsymbol{x}_{i_n}) + \rho \cdot T(\boldsymbol{x}_i)\right) ; \tag{6.2.13}$$

the weak learner at the m^{th} iteration is then updated as:

$$\hat{f}_i^{(m)}(\boldsymbol{x}_{i_t}) = \hat{f}_i^{(m-1)}(\boldsymbol{x}_{i_t}) + \hat{\rho}_i^{(m)}, \tag{6.2.14}$$

and, given a threshold τ, the algorithm stops when $\hat{\rho}_i^{(m)} < \tau$, returning the prediction $\hat{P}_i = \hat{f}_i^{(m)}(\boldsymbol{x}_i)$. In order to generate probabilistic forecasts, the loss function $\psi_i\left(P_i, f_i(\boldsymbol{x}_i)\right)$ is the PS, defined in Section 6.1.2. PS is a function of the quantile coverage $0 \leq \alpha_q \leq 1$, so forecasts can be given by Q predictive quantiles $\hat{P}_{i_h}^{\langle \alpha_1 \rangle}, \dots, \hat{P}_{i_h}^{\langle \alpha_Q \rangle}$ for the target time horizon h.

5) Quantile regression neural network

QRNN is the other model used in the proposal to forecast the power generated by each individual wind turbine.

The QRNNs used for this purpose are extensions of Monotone Quantile Regression Neural Networks (MQRNNs), specifically adapted in order to simultaneously target multiple predictive conditional quantiles, avoiding quantile crossing [98]. The neural network architecture considered for MQRNN is the Multi-Layer Perceptron (MLP) with single hidden layer. For its formulation, the i^{th} generator is considered hereinafter. In traditional MQRNNs, the predictors $\boldsymbol{x}_i$ are separated into two groups: one group contains the A predictors $x_{a_1,i}, \ldots, x_{a_A,i}$ having a monotone increasing link with the target variable, and the other group contains the remaining B predictors $x_{b_1,i}, \ldots, x_{b_B,i}$. The predictive conditional α_q-quantile returned by the MQRNN for the target time horizon h is obtained as function of these predictors related to target time horizon h, as it follows:

$$\hat{P}_{i_h}^{\langle \alpha_q \rangle} = \gamma\left\{\sum_{j=1}^{J}\left[\eta\left(\sum_{n=1}^{A} x_{a_n,i_h}\cdot e^{\hat{\omega}_{a_n,j}^{\langle \alpha_q \rangle}} + \sum_{n=1}^{B} x_{b_n,i_h}\cdot \hat{\omega}_{b_n,j}^{\langle \alpha_q \rangle} + \hat{\sigma}_j^{\langle \alpha_q \rangle}\right)\cdot e^{\hat{\delta}_j^{\langle \alpha_q \rangle}}\right] + \hat{\beta}^{\langle \alpha_q \rangle}\right\}, \qquad (6.2.15)$$

where $\eta(\cdot)$ is the function operated at the level of the neurons of the hidden layer (in this proposal, it is the sigmoid hyperbolic tangent function illustrated in Figure 6.2.2), J is the number of neurons in the hidden layer, $\gamma(\cdot)$ is a non-decreasing function operated upon the outputs of the hidden layer (it is an inverse link function from generalized linear modeling, and in this proposal it is kept as the identity function), $\hat{\boldsymbol{\omega}}_j^{\langle \alpha_q \rangle}$ is the vector of $A+B$ parameters estimated at the level of the j^{th} neuron within the hidden layer for the α_q-quantile, $\hat{\boldsymbol{\delta}}^{\langle \alpha_q \rangle}$ and $\hat{\boldsymbol{\sigma}}^{\langle \alpha_q \rangle}$ are two vectors of J estimated parameters for the α_q-quantile, and $\hat{\beta}^{\langle \alpha_q \rangle}$ is a scalar estimated parameter for the α_q-quantile.

The parameters $\hat{\boldsymbol{\omega}}_j^{\langle \alpha_q \rangle}, \hat{\boldsymbol{\delta}}^{\langle \alpha_q \rangle}, \hat{\boldsymbol{\sigma}}^{\langle \alpha_q \rangle}$ and $\hat{\beta}^{\langle \alpha_q \rangle}$ are estimated by minimizing the PS loss function $\psi_i\left(P_i, f_i(\boldsymbol{x}_i)\right)$, defined in Section 6.1.2, for the given nominal quantile coverage α_q. In order to avoid quantile crossing in the training stage, the QRNNs used in this proposal simultaneously accounts for all the parameters across multiple quantile coverages, i.e., the parameters are estimated simultaneously for all of the considered quantile coverages [98]. This is performed by applying partial monotonicity constraints on the corresponding predictors (i.e., $\frac{\partial P_i^{\langle \alpha_q \rangle}}{\partial x_{a_n}} \geq 0$), as both $\eta(\cdot)$ and $\gamma(\cdot)$ are smooth non-decreasing functions [98].

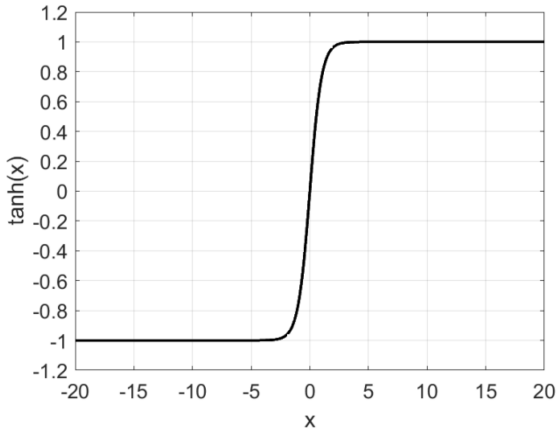

Fig. 6.2.2. Sigmoid hyperbolic tangent function used in the QRNN.

6) Aggregation of individual predictions

Forecasts provided for individual wind generators can be used per se or aggregated to form the probabilistic power forecast of the wind farm. Aggregating probabilistic predictions is not a trivial task [32]. The simple sum-and-sort of predictive quantiles is often effective to aggregate probabilistic predictions, although it sometimes may return slightly over-dispersed forecasts [32]. In this proposal we primarily use sum-and-sort to aggregate forecasts, as it appears to be a suitable solution on the basis of our experiments. This choice allows avoiding unnecessary complications and keeping the methodology as much simple as it can be.

The predictive α_q-quantiles of the power generated by the wind farm is:

$$\hat{P}_{wf_h}^{\langle \alpha_q \rangle} = \sum_{i=1}^{N_g} \hat{P}_{i_h}^{\langle \alpha_q \rangle}, \qquad q = 1, \dots, Q, \tag{6.2.16}$$

and they are eventually ordered to maintain statistical consistency (i.e., non-decreasing sorting) among predictive quantiles.

It is worth noting that some other aggregation solutions might be exploited as well, as they may return better aggregated forecasts for the wind farm. In this proposal a solution for the forecast aggregation, considering a Weighted Combination (WC) of individual predictive quantiles [68], is investigated and compared to the simple sum-and-sort. In particular, the WC predictive α_q-quantile of the power generated by the wind farm is:

$$\hat{P}_{wf_h}^{\langle \alpha_q \rangle} = \sum_{i=1}^{N_g} \hat{\varphi}_i^{\langle \alpha_q \rangle} \cdot \hat{P}_{i_h}^{\langle \alpha_q \rangle}, \qquad q = 1, \dots, Q, \tag{6.2.17}$$

where the weight parameters $\hat{\varphi}_i^{\langle \alpha_q \rangle}$ are optimized by minimizing the PS of the aggregated forecasts in a dedicated training window. The considered WC procedure does not account for any constraint upon the weight parameters, and it is assumed that potential crossing quantiles are sorted in a result post-process.

5.2.2. EXPERIMENTAL FRAMEWORK AND RESULTS

This section presents the experimental framework, in terms of benchmarks and the error indices used to assess the validity of the proposal and to compare the probabilistic forecasts, and the outcomes of the experimental results.

1) Benchmarks

We present several benchmarks in order to evaluate and compare the skill of the forecasts generated by the proposal. Each benchmark has a specific purpose, to evaluate single improvement brought by the proposal. Some of these benchmarks do not consider NWP ranking, in order to understand if and how much ranking NWPs adds skill to forecasts. Some of these benchmarks consider standard approaches for the feature selection and naïve approaches for unbiased comparison.

Single location (SL) NWP: in this benchmark, only the highest-ranked NWP set for the individual generator is considered within the vector $\boldsymbol{x}$ of predictors of the probabilistic forecasting model (either GBRT or QRNN). This benchmark aims at assessing the improvement brought by additional spatially-distributed NWPs in the surrounding area.

Average (AVG) NWP: in this benchmark, the entire pool of NWPs available for N_s elementary grid squares are simply averaged, in order to form fictitious NWPs for the considered area. These NWPs are considered within the vector $\boldsymbol{x}$ of predictors of the probabilistic forecasting model. This benchmark aims at assessing the improvement brought by differentiating the weights associated to the spatially-distributed NWPs in the surrounding area, rather than putting them all together in a simple average structure.

K-fold Cross-Validation (KCV) NWPs: in this benchmark, the entire pool of NWPs available the elementary grid squares surrounding the area are added as candidate predictors for either GBRT or QRNN models. Feature engineering is performed on this large set of candidate predictors through a 10-fold cross-validation procedure, in order to select only the most relevant ones to generate forecasts. This benchmark is based on a very standard technique which is used frequently in probabilistic forecasting, and thus it is expected to return skilled forecasts (probably even more skilled than the proposed technique). Nevertheless, it can be very time-consuming, particularly for many weather variables considered in the NWP sets.

Direct Wind Farm Forecast (DWFF): in this benchmark, the power generated by the wind farm is directly forecasted without transiting for forecasts of individual generators. In this case, we use the average NWPs as predictors of the probabilistic forecasting model. This benchmark has the purpose to investigate if and how much passing through individual predictions for individual generators increases the skill of forecasts.

Persistence Method (PM): this is a naïve benchmark added to provide unbiased results for a fair comparison. This benchmark assumes the wind power to be constant throughout the entire lead time so that each of the Q predictive quantiles is the same of the last observed value.

2) Error indices

Probabilistic forecasts are evaluated through a strictly proper score (the PS defined in Section 6.1.2) and by inspecting their reliability (i.e., the correspondence among estimated and nominal coverages) through reliability diagrams and through the AACE index defined in Section 6.1.2).

3) Data and forecasting framework

In the numerical experiments we assess the performance of the proposal as it was used to participate in day-ahead electrical markets. We use actual data collected at a wind farm located in southern Italy (**Dataset_WG** and the corresponding NWPs in **Dataset_weath_WG**). The wind farm is constituted by 10 3-MVA generators. Active power, absolute value of wind speed and wind direction have been measured at each generator throughout three years, with 10-minute acquisitions. The first two years are used for training the models, and the last year is completely reserved for testing. Data are post-processed to eliminate bad data and to average values to obtain data at hourly time resolution, since 1 hour is the bidding time slot in Italian day-ahead electrical market.

In summary, $H = 8760$ forecast instances are issued for each of the 365 days of the test year. Forecasts are generated using the *gbm* [99] and the *qrnn* [98] packages in R environment.

We start from a pool of $N_s = 121$ sets of NWPs provided for the area containing the wind farm, split into an 11x11 elementary-grid-square partition. For sake of clearness, these sets of NWPs are numbered from 1 (upper-left elementary grid square) to 121 (bottom-right elementary grid square), by rows (so that the set number 11 is at the upper-right elementary grid square).

The value for G, optimized in the cross-validation process, is $G = 5$.

In the framework of the Task 2.1, an exploratory data analysis is performed to check the validity of calendar variables, accounting for the possible seasonality of wind power patterns. As a significant example, Figure 6.2.3 shows the scatter plots of the power generated by the wind farm versus the hour of the day (a), the day of the month (b)

and the month of the year (c). The average wind power calculated for each class is shown in the same figures as a red line. From the scatter plots, we recognized only a slightly relevant pattern considering the hour of the day. Nevertheless, since the NWPs are developed 13 to 36 hours ahead, adding the class variable accounting for the hour of the day can be beneficial in order to allow the forecasting model to discern NWPs issued for different forecast lead times. Vice versa, class calendar variables accounting for the day of the month and for the month of the year are discarded, since they seemed not to add much information to predict wind power in the considered experiments.

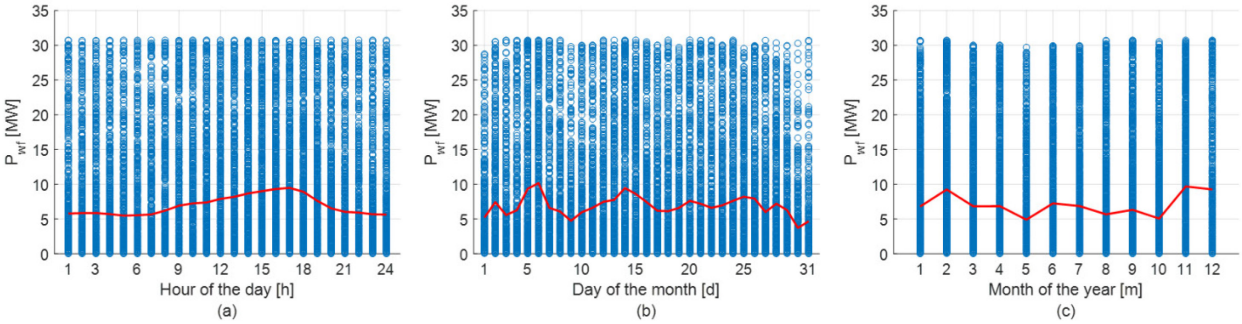

Fig. 6.2.3. Scatter plots of the power generated by the wind farm versus the hour of the day (a), the day of the month (b) and the month of the year (c).

4) Results of the ranking stage

To avoid unnecessary redundancy, we report the results of the ranking stage only for the first wind generator ($i = 1$). Table 6.2.1 shows the list of ranked NWPs, ranked by the absolute wind speed (Case 1) or by the u- and v-components of wind speed (Case 2). Note that the NWP for the location of the first generator is the set number 75.

The ranking stage results are compared to the outcomes of feature engineering by 10-fold cross validation for method KCV on GBRT and QRNN models. Actually, the KCV procedure allows selecting only the most informative among all of the seven variables (pressure, precipitation, cloud cover, temperature, u-component of wind speed, v-component of wind speed, and absolute wind speed) considered in the NWPs. In Table 6.2.1, however, only the identifying number of the wind speeds selected through KCV are reported, to avoid unnecessary verbosity and to provide a direct comparison with the results of the NWP ranking procedures. Apparently, there are no extraordinary dissimilarities between these approaches, suggesting that the proposal of ranking NWPs might be able to keep track of the relevant information contained in the available NWPs in generating forecasts.

Table 6.2.1. Results of the ranking stage and results of feature engineering for the first individual generator.

	Number of NWP sets	Required time
Ranked NWPs: Case 1	75-76-65-85-64-74-86-77-66-88-96-99-87-95-84-63-55-54-53-73-97-98-52-62-51	<1 s

Ranked NWPs: Case 2	75-65-76-64-66-74-85-86-77-99-63-54-96-53-95-87-84-88-52-73-97-62-55-98-51	<1 s
Feature engineering on GBRT (selected wind speed)	75-76-86-99-77-63-85	51.1 hours
Feature engineering on QRNN (selected wind speed)	75-76-96-77-84-65-64-86-74	72.8 hours

5) GBRT forecast results

PS and ACE values for the GBRT forecasts of the power generated by the 10 individual generators and for the aggregated wind farm are in Table 6.2.2. Bold values indicate the smallest PS for each generator and for the wind farm.

Table 6.2.2. Day-ahead GBRT forecast results for individual generators and for the aggregated wind farm. Bold values indicate the most skilled forecasts.

Method	Index	Generator										Wind farm
		1	2	3	4	5	6	7	8	9	10	
ESWC Case 1	PS [MW]	13.45	13.02	12.67	16.21	15.51	13.62	13.10	13.46	12.43	12.04	124.51
	AACE [%]	3.02	3.45	2.97	1.09	1.23	2.65	2.07	1.53	2.43	2.60	3.19
RRWC Case 1	PS [MW]	13.49	13.01	**12.63**	16.20	15.52	13.57	**13.00**	13.42	12.38	12.02	124.24
	AACE [%]	2.52	3.16	2.82	0.94	1.59	2.99	2.17	1.32	2.42	2.54	3.32
IMWC Case 1	PS [MW]	13.39	13.08	12.80	15.90	15.39	13.65	13.15	13.10	12.24	12.06	125.14
	AACE [%]	2.98	**2.27**	**2.40**	1.42	1.44	2.38	**1.70**	1.74	2.92	2.40	**2.28**
ESWC Case 2	PS [MW]	13.50	13.08	12.71	16.27	15.59	13.66	13.12	13.48	12.47	12.06	124.92
	AACE [%]	2.80	3.44	2.98	1.29	1.67	3.02	2.12	1.31	2.21	2.38	3.22
RRWC Case 2	PS [MW]	13.48	13.03	12.64	16.23	15.56	13.58	13.01	13.43	12.38	12.02	124.36
	AACE [%]	2.66	3.04	2.84	1.06	1.55	2.89	2.31	**1.10**	2.07	2.49	3.28
IMWC Case 2	PS [MW]	13.51	13.13	12.72	15.91	**15.30**	13.61	13.08	**12.92**	**12.19**	**12.01**	**124.16**
	AACE [%]	**2.45**	3.20	3.20	2.34	1.46	2.48	2.52	1.55	**2.03**	**1.97**	2.58
SL	PS [MW]	13.69	13.45	13.18	17.03	16.24	14.21	13.57	13.66	12.64	12.46	128.40
	AACE [%]	3.19	3.38	2.53	**0.68**	1.26	2.66	2.36	1.36	2.43	2.07	2.47
AVG	PS [MW]	13.99	13.52	13.23	16.72	16.02	14.07	13.69	13.93	12.87	12.53	127.29
	AACE [%]	2.59	3.08	2.44	1.16	1.47	2.83	1.88	1.40	2.36	2.65	3.24
KCV	PS [MW]	**13.29**	**12.96**	12.66	**15.88**	15.45	**13.56**	13.15	13.15	12.27	12.02	124.18
	AACE [%]	3.76	4.24	2.89	1.64	**1.15**	**1.78**	2.21	2.35	2.18	2.05	2.90
DWFF	PS [MW]	-	-	-	-	-	-	-	-	-	-	127.55
	AACE [%]	-	-	-	-	-	-	-	-	-	-	2.99
PM	PS [MW]	37.94	38.09	37.05	44.62	43.03	38.48	37.83	36.42	36.00	34.29	375.86
	AACE [%]	-	-	-	-	-	-	-	-	-	-	-

Regarding the individual generators, RRWC Case 1 and IMWC Case 2 generate the most skilled forecasts respectively for 2 and 4 generators, whereas KCV generate the most skilled forecasts for the other 4 generators. Even if the proposals are based on a simplified technique for feature selection, which requires a negligible amount of time

to complete, they still can sometimes outperform KCV, which is a standard and comprehensive technique for feature selection, which however is much more computationally intensive.

The relative improvement given by the best method with respect to the PM ranges from 64% to 66%, and it is quite stable versus the number of generators. When the proposal outperforms KCV, the PS is reduced up to 1.8%.

Compared to the methods in which only the highest-ranked NWP (SL) and in which the simple NWP average (AVG) are used as predictors, the relative improvements carried out by the proposal are respectively in the ranges 4.2 to 5.8% and 4.2 to 7.3%.

The reliability of RRWC Case 1 and IMWC Case 2 is checked and the ACE is always smaller than 3.20%. KCV is sometimes less reliable (see the ACE for generator 1 and 2). Figure 6.2.4 shows reliability diagrams obtained through RRWC Case 1, IMWC Case 2 and KCV for generator 1. Coverages are always slightly overestimated, and the greatest deviation from perfect reliability occurs at proximity to the highest quantile coverages.

Regarding the forecast for the entire wind farm, IMWC Case 2 generate the most skilled forecasts. The relative improvement with respect to KCV is quite small (less than 1%), but it comes also with the advantage of a lighter computational effort and with a slightly better reliability. Compared to the PM, the PS of the proposal is about 67% smaller, which is a slightly better improvement than for the individual generators. If the power generated by the wind farm is forecasted through the direct method (the DWFF), the PS would be about 2.7% greater. Compared to the methods in which only the highest-ranked NWP (SL) and in which the simple NWP average (AVG) are used as predictors, the relative improvements carried out by the IMWC Case 2 are respectively 4.3% and 2.5%. Figure 6.2.5 shows reliability diagrams obtained through RRWC Case 1, IMWC Case 2 and KCV for the wind farm. Also in this case, coverages are always slightly overestimated.

In Figure 6.2.6 and in Figure 6.2.7 we show the hourly wind farm power predicted respectively by the IMWC Case 2 and by the KCV, for an entire week within the test year, compared to the actual generated power. Forecasts are illustrated by the prediction intervals, centered around the median, which may be extracted from predictive quantiles (e.g., 80% prediction interval is the interval $\left[\hat{P}_{i_h}^{\langle 0.1 \rangle} \quad \hat{P}_{i_h}^{\langle 0.9 \rangle}\right]$).

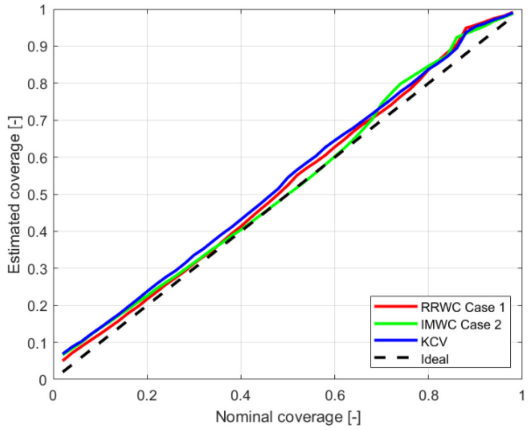

Fig. 6.2.4. GBRT reliability diagrams of RRWC Case 1, IMWC Case 2, and KVC for the first generator.

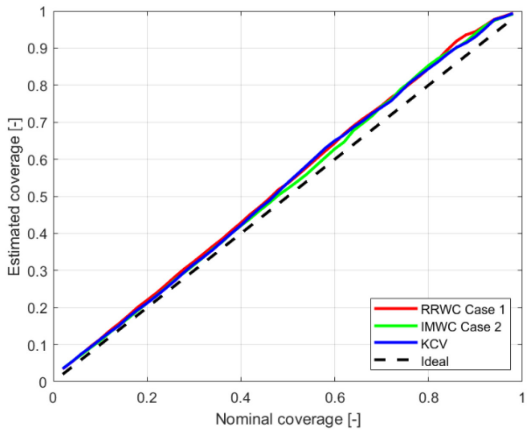

Fig. 6.2.5. GBRT reliability diagrams of RRWC Case 1, IMWC Case 2, and KVC for the wind farm.

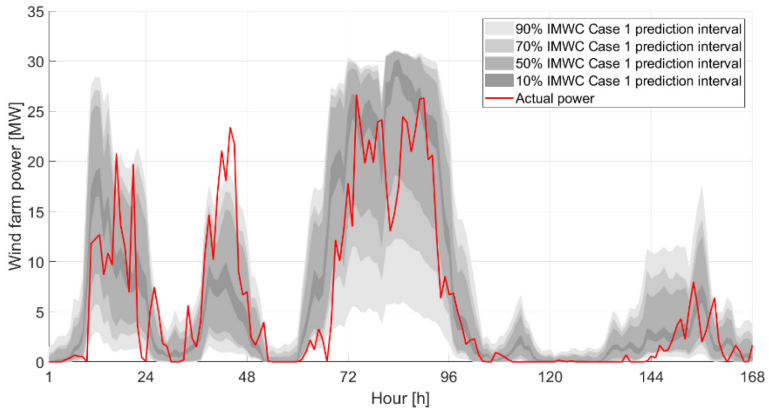

Fig. 6.2.6. IMWC Case 1 GBRT wind farm power forecasts and actual wind farm power during a week within the test year.

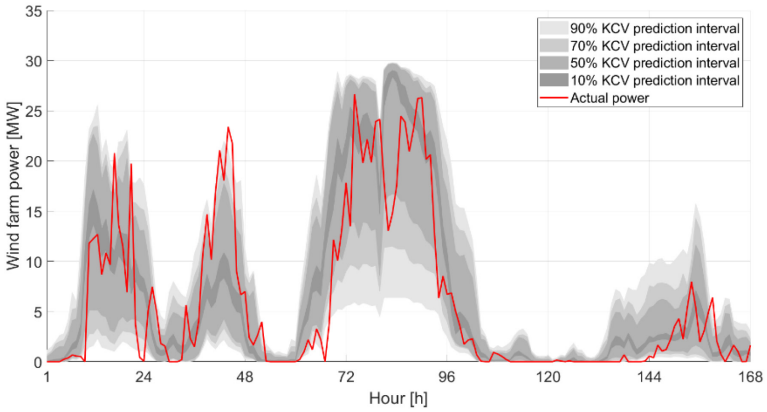

Fig. 6.2.6. KCV GBRT wind farm power forecasts and actual wind farm power during a week within the test year.

6) QRNN forecast results

PS and ACE values for the QRNN forecasts of the power generated by the 10 individual generators and by the aggregated wind farm are in Table 6.2.3. Bold values indicate the smallest PS for each generator and for the wind farm.

Table 6.2.3. Day-ahead QRNN forecast results for individual generators and for the aggregated wind farm. Bold values indicate the most skilled forecasts.

Method	Index	Generator										Wind farm
		1	2	3	4	5	6	7	8	9	10	
ESWC Case 1	PS [MW]	13.91	13.81	13.42	16.62	16.03	14.24	13.67	14.29	13.17	12.59	131.20
	ACE [%]	2.68	2.91	3.03	4.25	4.32	3.35	2.58	2.89	3.98	3.20	5.28
RRWC Case 1	PS [MW]	13.97	13.68	13.47	16.61	16.06	14.27	13.78	14.31	13.20	12.61	131.28
	ACE [%]	2.49	2.49	3.30	3.92	3.94	4.05	3.77	**2.35**	3.80	3.13	5.57
IMWC Case 1	PS [MW]	**13.80**	13.51	13.29	16.22	**15.75**	14.35	13.73	13.87	12.97	12.61	130.58
	ACE [%]	2.45	**2.38**	2.35	2.92	2.80	3.30	2.65	2.84	3.17	3.47	4.15
ESWC Case 2	PS [MW]	14.04	14.03	13.50	16.68	16.21	14.31	13.72	14.38	13.28	12.64	131.97
	ACE [%]	2.53	4.52	3.24	3.80	4.35	3.69	2.54	2.86	3.76	3.21	5.63
RRWC Case 2	PS [MW]	14.00	13.72	13.49	16.65	16.11	14.28	13.72	14.34	13.26	12.62	131.17
	ACE [%]	**2.32**	3.14	3.23	3.97	4.33	4.04	3.58	2.41	3.88	3.13	5.76
IMWC Case 2	PS [MW]	14.28	**13.50**	13.33	16.45	16.05	**14.23**	**13.55**	14.24	12.85	**12.24**	**130.08**
	ACE [%]	2.89	3.37	2.61	2.81	**2.69**	**2.93**	**2.25**	2.70	**3.01**	**2.25**	4.03
SL	PS [MW]	14.30	13.99	13.99	17.42	16.75	14.71	14.22	14.62	13.35	13.21	133.65
	ACE [%]	2.63	4.98	**2.31**	**2.13**	3.98	3.86	3.51	3.21	4.33	4.48	5.04
AVG	PS [MW]	14.50	13.96	13.49	16.68	16.10	14.27	13.77	14.33	13.19	12.63	132.43
	ACE [%]	2.46	4.36	3.19	3.93	4.35	3.93	3.25	2.40	3.24	3.08	5.93
KCV	PS [MW]	14.15	13.96	**13.17**	**16.13**	16.25	14.45	14.05	**13.54**	**12.84**	12.76	130.87
	ACE [%]	4.49	3.96	3.81	3.10	2.96	3.56	4.05	3.93	3.53	3.11	**3.75**
DWFF	PS [MW]	-	-	-	-	-	-	-	-	-	-	137.00

	ACE [%]	-	-	-	-	-	-	-	-	-	-	5.41
PM	PS [MW]	37.94	38.09	37.05	44.62	43.03	38.48	37.83	36.42	36.00	34.29	375.86
	ACE [%]	-	-	-	-	-	-	-	-	-	-	-

Regarding the individual generators, IMWC Case 1 and IMWC Case 2 generate the most skilled forecasts respectively for 2 and 4 generators, whereas KCV generate the most skilled forecasts for the other 4 generators. Also exploiting QRNN forecasts, even if the proposals are based on a simplified technique for feature selection, they still can sometimes outperform KCV, which is a standard and comprehensive technique for feature selection, but much more computationally intensive.

The relative improvement given by the best method with respect to the PM ranges from 60% to 66%, and it is quite stable versus the number of generators. When the proposal outperforms KCV, the PS is reduced up to 4.1%.

Compared to the methods in which only the highest-ranked NWP (SL) and in which the simple NWP average (AVG) are used as predictors, the relative improvements carried out by the proposal are respectively in the ranges 0.1 to 7.3% and 0.3 to 3.3%.

The reliability of IMWC Case 1 and IMWC Case 2 is checked and the ACE is always smaller than 3.5% (slightly worse than GBRT forecasts). Figure 6.2.8 shows QRNN reliability diagrams obtained through IMWC Case 1, IMWC Case 2 and KCV for generator 1. For QRNN forecasts, low coverages are overestimated and high coverages are underestimated. The greatest deviation from perfect reliability occurs at proximity to the highest quantile coverages, which are severely underestimated (see particularly the KCV reliability diagram).

Regarding the forecast for the entire wind farm, IMWC Case 2 generates the most skilled forecasts also using QRNN. The relative improvement with respect to KCV is quite small (less than 1%), but it comes also with the advantage of a lighter computational effort and with a slightly better reliability. Compared to the PM, the PS of the proposal is about 65.4% smaller, which is a slightly better improvement than for most the individual generators. If the power generated by the wind farm is forecasted through the direct method (the DWFF), the PS would be about 5.3% greater. Compared to the methods in which only the highest-ranked NWP (SL) and in which the simple NWP average (AVG) are used as predictors, the relative improvements carried out by the IMWC Case 2 are respectively 2.7% and 1.8%. Figure 6.2.9 shows QRNN reliability diagrams obtained through IMWC Case 1, IMWC Case 2 and KCV for the wind farm. The reliability diagrams appear to be similar to those obtained for the individual generator.

In Figure 6.2.10 and in Figure 6.2.11 the hourly wind farm power predicted respectively by the IMWC Case 2 and by the KCV are shown for an entire week within the test year, and compared to the actual generated power.

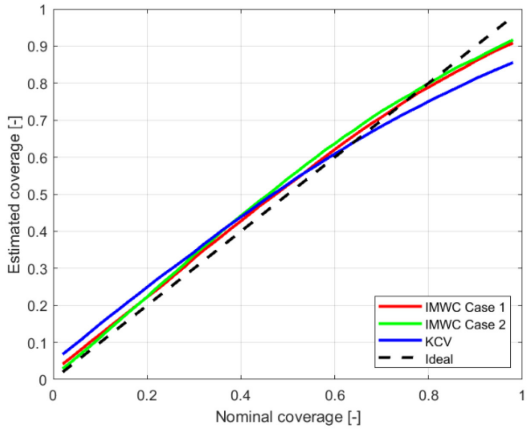

Fig. 6.2.8. QRNN reliability diagrams of RRWC Case 1, IMWC Case 2, and KVC for the first generator.

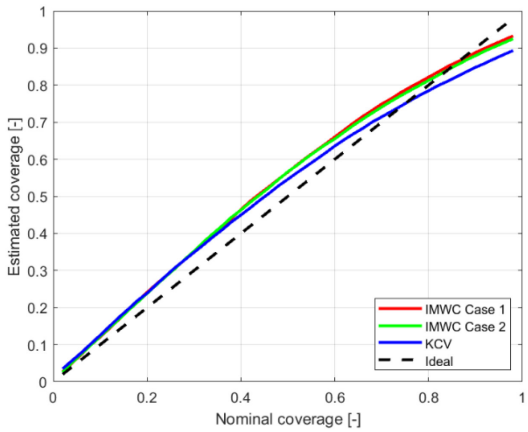

Fig. 6.2.9. QRNN reliability diagrams of RRWC Case 1, IMWC Case 2, and KVC for the wind farm.

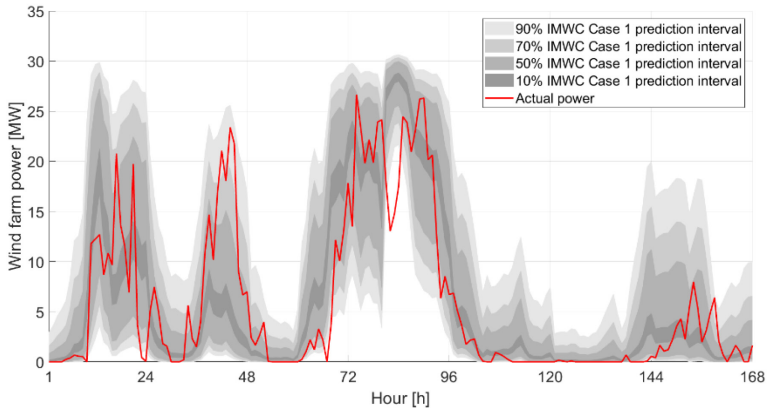

Fig. 6.2.10. IMWC Case 2 QRNN wind farm power forecasts and actual wind farm power during a week within the test year.

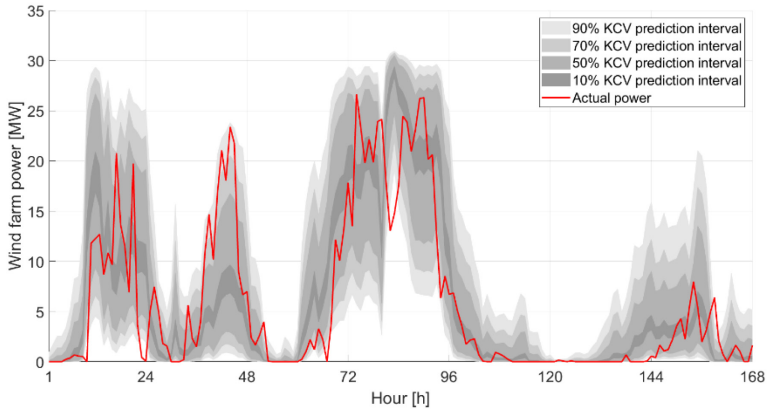

Fig. 6.2.11. KCV QRNN wind farm power forecasts and actual wind farm power during a week within the test year.

7) Analysis of different solutions for combining individual predictions

In this sub-Section the simple sum-and-sort solution for the aggregation of individual predictive quantiles is compared to the WC solution. Since the WC requires to optimize the weight parameters over a training sample, the first half (4380 instances) of the test year considered in previous sub-Section is reserved to optimize the weights, and the second half (4380 instances) of the test year is reserved to test the results.

The error indices calculated on wind farm power forecasts are reported in Table 6.2.4. To avoid unnecessary verbosity, only the most skilled methods (RRWC Case 1, IMWC Case 1, IMWC Case 2 and KCV) are considered. DWFF and PM results are provided for reference. The results indicate that the availability of individual probabilistic forecasts enables an optimized refinement of the forecasts for the aggregated wind farm.

Comparing the sum-and-sort and WC solutions, the PS slightly decreases using the latter solution (e.g., by about 1% for the IMWC Case 2 with GBRT model, which shows the smallest PS), but most importantly the reliability of forecasts dramatically increases (e.g., the ACE diminishes by about 41% for the IMWC Case 2), overcoming the well-known problem of the over-dispersion of aggregated quantiles. Also with reference to the WC results, the best proposal is slightly more skilled than the consolidated KCV using both GBRT or QRNN models.

Table 6.2.4. Day-ahead GBRT and QRNN forecasting results for the aggregated wind farm, considering sum-and-sort and WC solutions. Bold values indicate the most skilled forecasts.

Method	Index	GBRT sum-and-sort	GBRT WC	QRNN sum-and-sort	QRNN WC
RRWC Case 1	PS [MW]	114.82	114.26	122.88	122.40
	ACE [%]	3.07	1.49	5.21	2.55
IMWC Case 1	PS [MW]	114.70	114.90	122.62	120.96
	ACE [%]	2.16	1.78	4.53	2.86

IMWC Case 2	PS [MW]	113.56	**112.72**	121.56	119.94
	ACE [%]	3.18	1.87	4.24	3.08
KCV	PS [MW]	113.86	112.88	122.22	120.80
	ACE [%]	2.77	1.87	3.91	1.23
DWFF	PS [MW]	117.00		129.65	
	ACE [%]	2.68		4.51	
PM	PS [MW]	337.04			
	ACE [%]	-			

5.2.3. DISCUSSION

Reliable and skilled probabilistic wind power forecasting systems for day-ahead horizons exploit NWPs. In this research activity we exploit spatially-distributed NWPs provided for a given area surrounding the wind farm, to predict the generation from individual turbines and the aggregated generation of the wind farm. A new simplified technique for the feature selection, based on NWP ranking, is specifically developed in order to reduce the computational effort, although keeping the skill of forecast at high levels, compared to more standard techniques.

Two ranking procedures and three combination approaches are investigated. On average the IMWC Case 1 returned the most skilled forecasts, which are only slightly better (up to 1.8% PS for individual generators, and less than 1% PS for the entire wind farm) than a more standard benchmark based on cross-validation feature engineering. Nevertheless, the computational time dedicated to data analysis is significantly reduced, easing the exploitation of the technique in forecasting systems which dynamically re-train models as new observations become available.

This research also opens new research perspectives to increase the skill of forecasts. The ranking procedure may be adjusted in order to account for the predictive skill of the forecasts in sample. In this case, the procedure could not be performed offline and thus it would be more computationally intensive, but it may increase the accuracy of forecasts. The aggregation of forecasts for individual generators has been enhanced with solutions more sophisticated than the simple sum-and-sort of predictive quantiles, as much as some calibration techniques might allow to further reduce the width of prediction intervals. For these reasons, research on the aggregation of probabilistic forecasts is strongly encouraged. Eventually, the proposal might also be framed within a forecasting system in which models are re-trained as new observations become available, or within multi-objective optimization for the weights of the methodology.

5.3. DEVELOPMENT OF DAY-AHEAD FORECASTING SYSTEMS FOR INDUSTRIAL LOADS

Probabilistic forecasts are as more relevant to operate power systems as the size of the load is bigger, thus it is convenient to develop specific forecasting systems for the most impactful consumers. Industrial factories are perfect candidates: for example, more than 40% of the total electrical energy consumed in Italy in a year is delivered to industrial systems [104]. Due to the large consumption, owners of large enterprises and large industrial factories may have substantial benefits from directly participating to electrical markets and from improving their energy management. Industrial microgrids may be realized within large factories, as several manufacturing processes may be shed or shifted along time, allowing to optimally exploit distributed energy resources. Therefore, increasing the forecasting accuracy for industrial utilities may have positive effects for the operation and the management of the whole electrical system and for the owners of the installations.

Research on probabilistic load forecasting has expanded recently [1], but most methodologies have been extensively applied either at sub-station or at smart-meter nodes; nevertheless, the performance of a forecasting system may drop when the target load is the demand of an industrial factory. This is due to the particular features which influence the industrial demand [14], due to the particular seasonal patterns of the industrial load profiles [70], and due to the smaller relative importance of the ambient temperature to predict industrial load [71]. The only recommended path to increase the accuracy of industrial load forecasts is to develop ad-hoc forecasting systems, accounting for all of these specific features [72].

Forecasters usually target the only active power, partially or totally disregarding reactive power. This choice may be considered obsolete nowadays due to i) the increased computer performance, which eases the generation of reactive power forecasts, and ii) the spread of smart grid tools, which manage and operate networks by the active and reactive power flows. For example, Volt/VAr optimization [73], harmonic compensation [74] and optimal energy dispatch [75,76] tools require prior estimations of the nodal reactive power. The multitude of power converters distributed across the grids can be controlled to compensate reactive power locally [77], increasing the total power transmission capacity of the grid and reducing losses [78]. It is also worth mentioning that reactive power support is an ancillary service, remunerated in several countries [79]. Improved frameworks for reactive power markets, either coupled with active power market [80] or developed within a probabilistic framework [81], have been proposed in relevant literature. Despite the rising interest of practitioners in reactive power management, the literature on reactive power forecasting slowly adapts to the practical needs. Reactive power forecasts are usually built by post-

processing active power forecasts, on the basis of practitioners' experience or by average-power-factor corrections. Only a few papers directly deal with reactive power forecasting, although these research papers evidence that correctly modeling the mutual correlation between active and reactive power allows enhancing the performance of the forecasting system [70].

In this research activity, a multivariate approach is developed to forecast the active and reactive power of an industrial load. Individual probabilistic forecasting models for active and reactive power are initially developed, exploiting exogenous electrical predictors. In the proposed multivariate approach, the individual predictive quantiles of active power and reactive power are combined in a novel Multivariate Quantile Regression (MQR) model. The multivariate approach is expected to catch residual mutual information between the target variables, increasing the accuracy of the final active power and reactive power forecasts. To validate the proposed approach when it is applied to different underlying models, QRF and Univariate QR (UQR) models are separately investigated.

Estimating the MQR parameters consists of minimizing a loss function in the training period. Either a rolling window or a fixed-origin window may be used for this purpose. The type of window determines the size and the temporal distribution of the training data, and this may affect the accuracy of multivariate forecasts. In the proposal presented below, the MQR parameters are re-estimated as new observations become available and the performance of the two training schemes are compared by the out-of-sample verification of the final forecasts of active and reactive power. The loss-function minimization problem is formulated in this proposal under a linear programming form, adapting the trick presented in [105] for UQR to the MQR case. The solution of the loss-function minimization problem, therefore, may be provided by applying a linear programming solving algorithm (i.e., the dual simplex).

In summary, the main contributions of this research activity are:

1) the adaptation of QRF and UQR models to probabilistic industrial active and reactive power forecasting, exploring the validity of exogenous electrical predictors and of industry-related features;
2) the development of a novel multivariate approach to industrial active and reactive power forecasting, based on an MQR refinement of individual predictions;
3) the specific formulation of the MQR parameter estimation under a linear programming form;
4) a comparative analysis between the performances of two schemes for training the MQR models dynamically, as new observations become available.

To avoid verbosity in the analytic formulations, the symbols are consistent within the following sub-Sections 6.3.1, 6.3.2 and 6.3.3 only.

5.3.1. MODELS AND METHODS

The idea behind the multivariate approach was inspired from a previous study on industrial load data that includes the correlation analysis [70]. The probabilistic multivariate approach developed in this research activity for probabilistic industrial load forecasting consists of two main units: i) the generation of univariate predictive quantiles via underlying probabilistic models, and ii) the multivariate combination of active and reactive power predictive quantiles. Figure 6.3.1 illustrates the workflow of the proposed multivariate forecasting system.

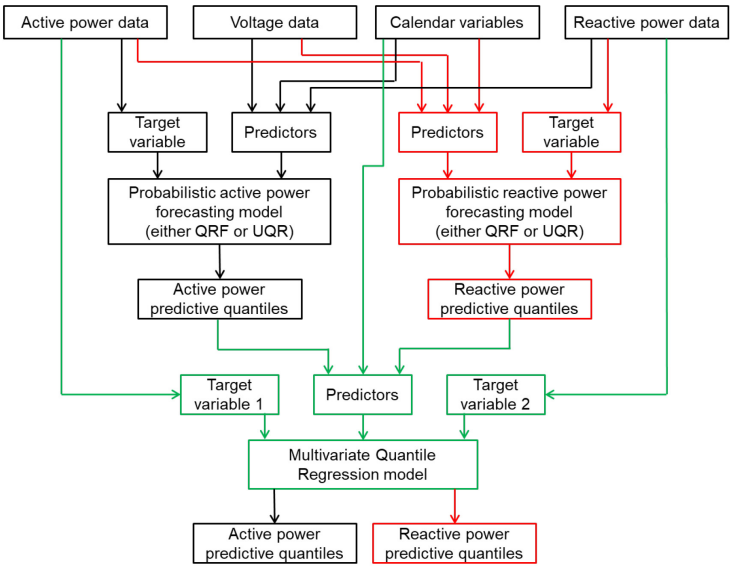

Fig. 6.3.1. Workflow of the proposed multivariate system for industrial active and reactive power forecasting.

QRF and UQR are alternatively used to generate individual predictive quantiles, in order to validate the proposal upon different underlying probabilistic models. QRF and UQR are developed following the outcomes presented in [70] for deterministic industrial load forecasting. Particularly, regarding the choice of input predictors, it has been noted that industrial load is usually less sensitive to ambient temperature, compared to load at bulk supply nodes in transmission systems or compared to the load of commercial and residential buildings. Adding temperatures to the models may result in a drop of the performance. Nevertheless, as demonstrated by the correlation analysis in [70], specific features of industrial load may be successfully exploited to generate active and reactive power predictions. Among them, electrical variables allow enhancing the modeling performance, by catching the mutual correlation between active and reactive power.

Motivated by above, the following predictors are included in the individual QRF and UQR models:

- Active power is selected as a predictor when targeting the reactive power, and vice versa.

- Average values of phase-to-ground voltage amplitudes are selected since they can be informative for active and reactive power, as shown in [70].
- Calendar variables are selected since they are frequently used predictors in load forecasting studies. Industrial load time series are often seasonal and heteroscedastic. Dummy calendar variables for hour of the day (***hod***), the type of day (***tod***), and the day of the month (***dom***) [70] are exploited to account for these characteristics, allowing to differentiate the prediction respectively based on daily seasonality, weekly seasonality and monthly seasonality. The variable ***dom***, which is not commonly found in load forecasting studies, is used here to model maintenance schedules, supply cycles and production cycles that may be arranged with monthly periodicity.

The interaction of calendar variables with other predictors allows differentiating the estimation of model parameters upon data that share calendar similarities. This allows for detailed characterization of load patterns. We are trying to avoid the situation that all load profiles are treated at the same manner even if they belong to different work regimes [72].

The individual predictive quantiles, returned either by QRF or by UQR models, are combined in the multivariate approach by applying a MQR model. The predictors of MQR model are the individual predictive quantiles returned either by QRF or by UQR models, and the calendar variables. The interactions among calendar variables and other predictors allow catching the seasonality in load patterns both at the underlying and at the multivariate level of the forecasting system.

In the proposed MQR model, the multivariate target variable is defined in $\mathbb{R}^{2L}$ (L predictive quantiles for active power and L predictive quantiles for reactive power), and thus it is not a simple bivariate target variable (active and reactive power). The multivariate approach is expected to enhance the accuracy of the final forecasts since: i) active and reactive power are mutually correlated [70], and this is modeled comprehensively by the multivariate approach; ii) MQR is suitable to be dynamically re-estimated as new observations become available, which may increase the accuracy of final forecasts.

In the proposed multivariate approach, the individual probabilistic models are trained only once upon a fixed training dataset, whereas the MQR parameters are dynamically re-estimated as new observations become available. In this proposal, 70% of the available data is dedicated to train the individual probabilistic models, 15% to train the MQR model, and 15% to test the performance of the proposed multivariate approach. Different percentage splits could be selected according to the specific situation. MQR parameter estimation is solved by a linear programming algorithm, which converges in a period much shorter than the forecast lead time. To investigate different data split procedures for determining MQR training data, two schemes for the MQR training schemes are compared: one is

based on a rolling window, and the other is based on a fixed-origin window. They are respectively illustrated in Figure 6.3.2a and 6.3.2b.

The next paragraphs briefly recall the underlying QRF and UQR models, and formulate the MQR model for active and reactive power forecasting.

Fig. 6.3.2. Timeline of the proposed multivariate forecasting system using: rolling window (a); fixed-origin window (b).

1) Quantile regression forests

Two QRFs (denoted by QRF$^{(P)}$ and QRF$^{(Q)}$) are used to forecast the active power P and the reactive power Q, respectively, for the target time horizon h. QRFs build predictive quantiles from an ensemble of T regression trees which form the forest [106]. To add diversity, each tree of the forest is grown independently from the others.

The data used to train the QRFs consist of N past observations of the target variable (i.e., the active power P for the QRF$^{(P)}$, and the reactive power Q for the QRF$^{(Q)}$), and of the corresponding N vectors of predictors (which form matrices $\boldsymbol{X}^{(P)}$ and $\boldsymbol{X}^{(Q)}$). The matrices $\boldsymbol{X}^{(P)}$ and $\boldsymbol{X}^{(Q)}$ of predictors respectively for the QRF$^{(P)}$ and QRF$^{(Q)}$ consist of lagged active power (such as $P_{h-k}, P_{h-k-1}, \dots, P_{h-24}, P_{h-168}$), lagged reactive power (such as $Q_{h-k}, Q_{h-k-1}, \dots, Q_{h-24}, Q_{h-168}$), lagged voltage (such as $V_{h-k}, V_{h-k-1}, \dots, V_{h-24}, V_{h-168}$) and calendar variables ($\boldsymbol{hod}_h$, $\boldsymbol{tod}_h$ *and* $\boldsymbol{dom}_h$). Note that lagged variables allow modeling temporal dependencies among variables. These data are split into T bagged subsets, and each subset is used to train one tree of the forest. The data cover the period corresponding to the red bars in Figure 6.3.2.

After training the QRFs, the predictors $\boldsymbol{x}_h^{(P)}$ and the predictors $\boldsymbol{x}_h^{(Q)}$, available at the forecast origin $h-k$ but related to the target time horizon h, individuate a unique path linking the roots of each tree to one leaf of each tree. Predictive quantiles are constructed as weighted averages of the outcomes contained these leaves (i.e., a subset of past values of the target variable). The detailed mathematical formulation of QRFs is in [106].

The hyperparameters of the QRFs and the best predictor configuration are optimized in 10-fold cross-validation, selecting the configuration with the smallest average PS.

2) Univariate quantile regression

Two UQRs (denoted by UQR$^{(P)}$ and UQR$^{(Q)}$) are used to forecast the active power P and the reactive power Q, respectively, for the target time horizon h. UQR models allow to estimate predictive quantiles as linear combinations of the predictors. The predictive α_l-quantile of the generic target variable z (i.e., active power P or reactive power Q) is:

$$\hat{z}_h^{(\alpha_l)} = \boldsymbol{x}_h^{(z)} \cdot \hat{\boldsymbol{\gamma}}^{(\alpha_l)} , \tag{6.3.1}$$

where $\hat{\boldsymbol{\gamma}}^{(\alpha_l)}$ is the vector of coefficients of the model, estimated by minimizing the PL in the training period corresponding to the red bars in Figure 6.3.2.

The PS minimization problem is formulated and solved in the linear programming form presented in [105]. The data used to train the UQR consist of N past observations of the target variable (i.e., the active power P for the UQR$^{(P)}$, and the reactive power Q for the UQR$^{(Q)}$), and of the corresponding N vectors of predictors (which form matrices $\boldsymbol{X}^{(P)}$ and $\boldsymbol{X}^{(Q)}$). The matrices $\boldsymbol{X}^{(P)}$ and $\boldsymbol{X}^{(Q)}$ of predictors respectively for the UQR$^{(P)}$ and UQR$^{(Q)}$ consist of lagged active power (such as $P_{h-k}, P_{h-k-1}, \dots, P_{h-24}, P_{h-168}$), lagged reactive power (such as $Q_{h-k}, Q_{h-k-1}, \dots, Q_{h-24}, Q_{h-168}$), lagged voltage (such as $V_{h-k}, V_{h-k-1}, \dots, V_{h-24}, V_{h-168}$) and calendar variables ($\boldsymbol{hod}_h$, $\boldsymbol{tod}_h$ *and* $\boldsymbol{dom}_h$).

The best predictor configuration is optimized in 10-fold cross-validation, selecting the configuration with the smallest average PS.

3) Multivariate quantile regression

The MQR is an adaptation of UQR to multiple target variables. It is specifically formulated in this proposal for the case of active and reactive power at the time horizon h. Extending (6.3.1) to the bivariate target variable $\boldsymbol{y}$, the vector $\boldsymbol{x}_h^{(z)}$ of predictors for the UQR in (6.3.1) becomes a matrix $\boldsymbol{X}_h^{(\alpha_l)}$ of predictors, whereas the vector $\hat{\boldsymbol{\gamma}}^{(\alpha_l)}$ of parameters for the UQR in (6.3.1) becomes a vector $\hat{\boldsymbol{\beta}}^{(\alpha_l)}$ of parameters having an increased size.

The vector $\hat{\boldsymbol{y}}_h^{(\alpha_l)}$ containing the predictive α_l-quantiles $\hat{P}_h^{(\alpha_l)}, \hat{Q}_h^{(\alpha_l)}$ of active and reactive power is:

$$\hat{\boldsymbol{y}}_h^{(\alpha_l)} = \begin{bmatrix} \hat{P}_h^{(\alpha_l)} \\ \hat{Q}_h^{(\alpha_l)} \end{bmatrix} = \boldsymbol{X}_h^{(\alpha_l)} \cdot \hat{\boldsymbol{\beta}}^{(\alpha_l)} , \tag{6.3.2}$$

where:

- $\boldsymbol{X}_h^{(\alpha_l)}$ is the following $2 \times 2J$ matrix:

$$\boldsymbol{X}_h^{(\alpha_l)} = \begin{bmatrix} \boldsymbol{x}_h^{(PQ)} & \boldsymbol{0}_{1\times J} \\ \boldsymbol{0}_{1\times J} & \boldsymbol{x}_h^{(PQ)} \end{bmatrix}, \tag{6.2.3}$$

containing the row vector $\boldsymbol{x}_h^{(PQ)}$ of J predictors for the bivariate target variable (univariate forecasts of active and reactive power, provided either by QRF or UQR, and calendar variables $\boldsymbol{hod}_h$, $\boldsymbol{tod}_h$, and $\boldsymbol{dom}_h$) and row vector $\boldsymbol{0}_{1\times J}$ of J zeros. The matrix $\boldsymbol{X}_h^{(\alpha_l)}$ is the same for all of the considered quantiles (i.e., $\forall l = 1, \dots, L$), to avoid unnecessary complications in the model selection procedure;

- $\widehat{\boldsymbol{\beta}}^{(\alpha_l)}$ is a $2J \times 1$ vector of the estimated coefficients of the MQR model for the α_l-quantile:

$$\widehat{\boldsymbol{\beta}}^{(\alpha_l)} = \begin{bmatrix} \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)} \\ \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_{1_P}^{(\alpha_l)} \\ \hat{\beta}_{2_P}^{(\alpha_l)} \\ \vdots \\ \hat{\beta}_{J_P}^{(\alpha_l)} \\ \hat{\beta}_{1_Q}^{(\alpha_l)} \\ \hat{\beta}_{2_Q}^{(\alpha_l)} \\ \vdots \\ \hat{\beta}_{J_Q}^{(\alpha_l)} \end{bmatrix}, \qquad l = 1, \dots, L. \tag{6.3.4}$$

Parameters $\widehat{\boldsymbol{\beta}}^{(\alpha_l)}$ are estimated by minimizing the PS in the training period corresponding to the green bars in Figure 6.3.2. The training data consist of M observations of active and reactive power until the forecast origin $h-k$. M is a constant number if the MQR model is trained under a rolling-window scheme, whereas M increases in time if the MQR model is trained under a fixed-origin-window scheme.

An index $o = h - k + 1$ is defined to ease the analytical formulation. Let $\boldsymbol{P} = \{P_{o-M}, P_{o-(M-1)}, \dots, P_{o-1}\}$ and $\boldsymbol{Q} = \{Q_{o-M}, Q_{o-(M-1)}, \dots, Q_{o-1}\}$ be the column vectors of the available past active and reactive power (notice that each vector contains M values measured until the forecast origin $h-k$). The parameters $\widehat{\boldsymbol{\beta}}^{(\alpha_l)}$ in (6.3.3) are estimated by solving the following optimization problem, which aims at minimizing the PS over the training data:

$$\widehat{\boldsymbol{\beta}}^{(\alpha_l)} = \underset{\widehat{\boldsymbol{\beta}}^{(\alpha_l)}}{\operatorname{argmin}} \sum_{m=1}^{M} \left[r_{P_m}^{(\alpha_l)} \cdot \left(\alpha_l - 1\left\{ r_{P_m}^{(\alpha_l)} < 0 \right\} \right) + r_{Q_m}^{(\alpha_l)} \cdot \left(\alpha_l - 1\left\{ r_{Q_m}^{(\alpha_l)} < 0 \right\} \right) \right], \tag{6.3.5}$$

where $\boldsymbol{\beta}^{(\alpha_l)} = \begin{bmatrix} \boldsymbol{\beta}_P^{(\alpha_l)} \\ \boldsymbol{\beta}_Q^{(\alpha_l)} \end{bmatrix}$, as for the estimated values in (6.3.4), and:

$$r_{P_m}^{(\alpha_l)} = P_{o-m} - \boldsymbol{x}_{o-m}^{(PQ)} \boldsymbol{\beta}_P^{(\alpha_l)} , \tag{6.3.6}$$

$$r_{Q_m}^{(\alpha_l)} = Q_{o-m} - \boldsymbol{x}_{o-m}^{(PQ)} \boldsymbol{\beta}_Q^{(\alpha_l)} , \tag{6.3.7}$$

are respectively the m^{th} residuals of the active and reactive power, i.e., the differences between the m^{th} measured value and the corresponding value given by the model.

In this proposal, MQR is extended to multiple quantiles handled simultaneously for the two target variables. In the MQR parameter estimation procedure, quantiles are constrained to be in non-diminishing order during the training period. This does not guarantee that the predictive quantiles during the test period are in non-diminishing order too, however it reduces the need to post-process the results by sorting the predictive quantiles.

The MQR parameter estimation is formulated hereinafter under a linear programming form, adapting the formulation presented in [105] to L quantile levels of active and reactive power:

$$\widehat{\boldsymbol{y}}_h = \begin{bmatrix} \hat{P}_h^{(\alpha_1)} \\ \vdots \\ \hat{P}_h^{(\alpha_L)} \\ \hat{Q}_h^{(\alpha_1)} \\ \vdots \\ \hat{Q}_h^{(\alpha_L)} \end{bmatrix} = \boldsymbol{X}_h \cdot \widehat{\boldsymbol{\beta}} , \tag{6.3.8}$$

where the $2L \times 2J \cdot L$ matrix $\boldsymbol{X}_h$ is:

$$\boldsymbol{X}_h = \begin{bmatrix} \boldsymbol{x}_h^{(PQ)} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} \\ \boldsymbol{0}_{1\times J} & \boldsymbol{x}_h^{(PQ)} & \cdots & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{x}_h^{(PQ)} & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} \\ \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} & \boldsymbol{x}_h^{(PQ)} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} \\ \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \boldsymbol{x}_h^{(PQ)} & \cdots & \boldsymbol{0}_{1\times J} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \boldsymbol{0}_{1\times J} & \cdots & \boldsymbol{x}_h^{(PQ)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{1,h} \\ \boldsymbol{X}_{2,h} \\ \vdots \\ \boldsymbol{X}_{L,h} \\ \boldsymbol{X}_{L+1,h} \\ \boldsymbol{X}_{L+2,h} \\ \vdots \\ \boldsymbol{X}_{2L,h} \end{bmatrix}, \tag{6.3.9}$$

$\widehat{\boldsymbol{\beta}}$ is a $2J \cdot L \times 1$ vector of coefficients:

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\boldsymbol{\beta}}_P^{(\alpha_1)} \\ \vdots \\ \widehat{\boldsymbol{\beta}}_P^{(\alpha_L)} \\ \widehat{\boldsymbol{\beta}}_Q^{(\alpha_1)} \\ \vdots \\ \widehat{\boldsymbol{\beta}}_Q^{(\alpha_L)} \end{bmatrix}, \tag{6.3.10}$$

and $\mathbf{0}_{n \times n}$ is a $n \times n$ matrix of zeros.

The same notation used before to present the single-quantile MQR is exploited hereinafter. $\boldsymbol{P}$, $\boldsymbol{Q}$ and o have the same meaning explained before, and $\boldsymbol{X}^{(PQ)}$ is the $M \times J$ matrix of past predictors (univariate forecasts of active and reactive power, provided by one of the underlying probabilistic models, and calendar variables):

$$\boldsymbol{X}^{(PQ)} = \begin{bmatrix} \boldsymbol{x}_{o-1}^{(PQ)} \\ \vdots \\ \boldsymbol{x}_{o-M}^{(PQ)} \end{bmatrix}. \tag{6.3.11}$$

The parameters in (6.3.10) are obtained as a subset of the solutions $\widehat{\boldsymbol{b}}$ of the following constrained linear programming problem:

$$\widehat{\boldsymbol{b}} = \underset{\boldsymbol{b}}{\operatorname{argmin}}\, \boldsymbol{a}' \cdot \boldsymbol{b} \tag{6.3.12}$$

$$\text{s.t.} \quad \boldsymbol{A} \cdot \boldsymbol{b} = \begin{bmatrix} [\boldsymbol{P}]_{L \times 1} \\ [\boldsymbol{Q}]_{L \times 1} \end{bmatrix},$$

$$b_{2J \cdot L + i} \geq 0 \qquad \forall i = 1, \dots, 4L \cdot M,$$

$$\boldsymbol{C} \cdot \boldsymbol{b} \leq \mathbf{0}_{2(L-1) \cdot M \times 1},$$

where $[\boldsymbol{P}]_{L \times 1}$ and $[\boldsymbol{Q}]_{L \times 1}$ are the vectors $\boldsymbol{P}$ and $\boldsymbol{Q}$ replicated L times,

$$\boldsymbol{a} = \begin{bmatrix} \mathbf{0}_{2J\cdot L\times 1} \\ \alpha_1 \cdot \mathbf{1}_{M\times 1} \\ (1-\alpha_1)\cdot \mathbf{1}_{M\times 1} \\ \vdots \\ \alpha_L \cdot \mathbf{1}_{M\times 1} \\ (1-\alpha_L)\cdot \mathbf{1}_{M\times 1} \\ \alpha_1 \cdot \mathbf{1}_{M\times 1} \\ (1-\alpha_1)\cdot \mathbf{1}_{M\times 1} \\ \vdots \\ \alpha_L \cdot \mathbf{1}_{M\times 1} \\ (1-\alpha_L)\cdot \mathbf{1}_{M\times 1} \end{bmatrix}, \tag{6.3.13}$$

$$\boldsymbol{b} = \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{r}_P^{(\alpha_1)} \\ \vdots \\ \boldsymbol{r}_P^{(\alpha_L)} \\ \boldsymbol{r}_Q^{(\alpha_1)} \\ \vdots \\ \boldsymbol{r}_Q^{(\alpha_L)} \end{bmatrix}; \tag{6.3.14}$$

$$\boldsymbol{r}_P^{(\alpha_l)} = \begin{bmatrix} 1\left\{\left[P_{o-1} - \boldsymbol{X}_{l,o-1}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)}\right] \geq 0\right\}\cdot \left[P_{o-1} - \boldsymbol{X}_{l,o-1}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)}\right] \\ \vdots \\ 1\left\{\left[P_{o-M} - \boldsymbol{X}_{l,o-M}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)}\right] \geq 0\right\}\cdot \left[P_{o-M} - \boldsymbol{X}_{l,o-M}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)}\right] \\ 1\left\{\left[P_{o-1} - \boldsymbol{X}_{l,o-1}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)}\right] < 0\right\}\cdot \left[\boldsymbol{X}_{l,o-1}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)} - P_{o-1}\right] \\ \vdots \\ 1\left\{\left[P_{o-M} - \boldsymbol{X}_{l,o-M}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)}\right] < 0\right\}\cdot \left[\boldsymbol{X}_{l,o-M}\cdot \widehat{\boldsymbol{\beta}}_P^{(\alpha_l)} - P_{o-M}\right] \end{bmatrix}; \tag{6.3.15}$$

$$\boldsymbol{R}_Q^{(\alpha_l)} = \begin{bmatrix} 1\left\{\left[Q_{o-1} - \boldsymbol{X}_{L+l,o-1}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)}\right] \geq 0\right\}\cdot \left[Q_{o-1} - \boldsymbol{X}_{L+l,o-1}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)}\right] \\ \vdots \\ 1\left\{\left[Q_{o-M} - \boldsymbol{X}_{L+l,o-M}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)}\right] \geq 0\right\}\cdot \left[Q_{o-M} - \boldsymbol{X}_{L+l,o-M}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)}\right] \\ 1\left\{\left[Q_{o-1} - \boldsymbol{X}_{L+l,o-1}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)}\right] < 0\right\}\cdot \left[\boldsymbol{X}_{L+l,o-1}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)} - Q_{o-1}\right] \\ \vdots \\ 1\left\{\left[Q_{o-M} - \boldsymbol{X}_{L+l,o-M}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)}\right] < 0\right\}\cdot \left[\boldsymbol{X}_{L+l,o-M}\cdot \widehat{\boldsymbol{\beta}}_Q^{(\alpha_l)} - Q_{o-M}\right] \end{bmatrix}; \tag{6.3.16}$$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{X}^{(PQ)} & \mathbf{0}_{M\times J} & \dots & \mathbf{0}_{M\times J} & [\boldsymbol{I}_{M\times M} & -\boldsymbol{I}_{M\times M}] & \mathbf{0}_{M\times 2M} & \dots & \mathbf{0}_{M\times 2M} \\ \mathbf{0}_{M\times J} & \boldsymbol{X}^{(PQ)} & \dots & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times 2M} & [\boldsymbol{I}_{M\times M} & -\boldsymbol{I}_{M\times M}] & \dots & \mathbf{0}_{M\times 2M} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \ddots & \vdots \\ \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} & \dots & \boldsymbol{X}^{(PQ)} & \mathbf{0}_{M\times 2M} & \mathbf{0}_{M\times 2M} & & \dots & [\boldsymbol{I}_{M\times M} & -\boldsymbol{I}_{M\times M}] \end{bmatrix}, \tag{6.3.17}$$

$$\boldsymbol{C} = \begin{bmatrix} \begin{bmatrix} \boldsymbol{X}^{(PQ)} & -\boldsymbol{X}^{(PQ)} & \mathbf{0}_{M\times J} & \dots & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} \\ \mathbf{0}_{M\times J} & \boldsymbol{X}^{(PQ)} & -\boldsymbol{X}^{(PQ)} & \dots & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} & \dots & \boldsymbol{X}^{(PQ)} & -\boldsymbol{X}^{(PQ)} \end{bmatrix} & \mathbf{0}_{(L-1)\cdot M\times J\cdot L} \\ \mathbf{0}_{(L-1)\cdot M\times J\cdot L} & \begin{bmatrix} \boldsymbol{X}^{(PQ)} & -\boldsymbol{X}^{(PQ)} & \mathbf{0}_{M\times J} & \dots & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} \\ \mathbf{0}_{M\times J} & \boldsymbol{X}^{(PQ)} & -\boldsymbol{X}^{(PQ)} & \dots & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} & \mathbf{0}_{M\times J} & \dots & \boldsymbol{X}^{(PQ)} & -\boldsymbol{X}^{(PQ)} \end{bmatrix} \end{bmatrix}, \tag{6.3.18}$$

In these equations, $\mathbf{1}_{n\times n}$ is a $n \times n$ matrix of ones and $\boldsymbol{I}_{n\times n}$ is the $n \times n$ identity matrix. The last line of constraints in (6.3.12) is added to avoid quantile crossing in the training period.

5.3.2. EXPERIMENTAL FRAMEWORK AND RESULTS

This section presents the experimental framework, in terms of benchmarks and the error indices used to assess the validity of the proposal and to compare the probabilistic forecasts, and the outcomes of the experimental results.

1) Benchmarks

The proposed approach is validated in this research activity by comparing its performance to several probabilistic benchmarks, briefly presented below.

Benchmark 1: Persistence method: The first benchmark is a simple method based on the persistence of the target variable. Regardless of the forecast lead time k, the predictive quantiles returned by the Persistence Method (PM) for the target time horizon h are the last observed values of active and reactive power:

$$\hat{P}_h^{(\alpha_l)} = P_{h-k}\,,\ \hat{Q}_h^{(\alpha_l)} = Q_{h-k} \qquad \forall l = 1, \dots, L \tag{6.3.19}$$

Benchmark 2: Quantile regression forest benchmark: The Quantile Regression Forest Benchmark (QRFB) returns univariate forecasts of active and reactive power through the QRF$^{(P)}$ and the QRF$^{(Q)}$, respectively. This benchmark is introduced to assess the refinement of individual QRF forecasts through an MQR model. In particular, this benchmark aims at checking if the score of MQR predictive quantiles is better than the score of univariate QRF predictive quantiles in the test period. This would confirm the validity of the multivariate approach applied to QRF underlying models.

The data used to train the $QRF^{(P)}$ and the $QRF^{(Q)}$ cover the period corresponding to the purple bars in Figure 6.3.2. The hyperparameters of the QRFB and the best predictor configuration are optimized in 10-fold cross-validation, selecting the configuration with the smallest average PS.

Benchmark 3: Univariate quantile regression benchmark: The Univariate Quantile Regression Benchmark (UQRB) returns univariate forecasts of active and reactive power through the $UQR^{(P)}$ and the $UQR^{(Q)}$, respectively. This benchmark is introduced to assess the refinement of UQR forecasts through an MQR model. In particular, this benchmark aims at checking if the score of MQR predictive quantiles is better than the score of UQR predictive quantiles in the test period. This would confirm the validity of the multivariate approach applied to UQR underlying models.

The data used to train the $UQR^{(P)}$ and the $UQR^{(Q)}$ cover the period corresponding to the purple bars in Figure 6.3.2. The best predictor configuration is optimized in 10-fold cross-validation, selecting the configuration with the smallest average PS.

2) Error indices

Two probabilistic error indices are used to assess the results. The first index is the PS, the strictly proper score defined in Section 5.1.2, which accounts for the reliability and the sharpness of forecasts [92]. The second considered index is the AACE defined in Section 5.1.2, which accounts for the reliability of the forecasts.

3) Data and forecasting framework

The data used for the experiments consists of **Dataset_indust_load**. Three loads with different characteristics are selected to validate the proposed multivariate approach: the factory aggregate load, the load of one of the LV power distribution feeders, and the load of a painting machine.

Available data are partitioned to meet the timeline in Figure 6.3.2. The first set, corresponding to the purple bars in Figure 6.3.2, is used to train models. It covers 20 months from April 2016 to November 2017: 16 months (from April 2016 to July 2017) are exclusively used to train the individual probabilistic models, whereas the remaining 4 months (from August 2017 to November 2017) are used to train the MQR model. The second set, corresponding to the pink bars in Figure 6.3.2, covers 3 months from December 2017 to February 2018 and it is used to test the forecasting system.

Hereinafter the multivariate system with QRF underlying models is denoted by "MQR on QRF", and the multivariate system with UQR underlying models is denoted by "MQR on UQR". In the numerical experiments, hour-ahead forecasts (lead time $k = 1$) and day-ahead forecasts (lead time $k = 24$) are generated by 99 predictive quantiles at levels 0.01, 0.02, …, 0.99. The two forecasting time horizons are selected in order to fulfill both the real-

time and the day-ahead purposes of the forecasting systems developed in the project, partially covering also activities related to Task 2.3. In this deliverable, hour-ahead forecasts of the aggregate load are thoroughly assessed, and the other results are presented with fewer details to avoid verbosity.

4) Results of hour-ahead aggregate load forecasts

Table 6.3.1 presents the outcomes of the optimized model selection procedures for the MQR on QRF, for the MQR on UQR, for the QRB, and for the QRFB. Pluses indicate selected predictors, minuses indicate discarded predictors, and NC indicates variables that have not been considered candidate predictors. Note that the forecasts obtained either through the underlying QRF or the underlying UQR are indicated by $\hat{P}_h^{(\alpha_l)}$, $\hat{Q}_h^{(\alpha_l)}$, to avoid repetitiveness. Obviously, univariate QRF forecasts are used in the MQR on QRF, and UQR forecasts are used in the MQR on UQR.

Table 6.3.1. Outputs of the model selection procedures

Candidate predictor	Model						Candidate predictor	Model					
	MQR on QRF	MQR on UQR	UQRB		QRFB			MQR on QRF	MQR on UQR	UQRB		QRFB	
			$UQR^{(P)}$	$UQR^{(Q)}$	$QRF^{(P)}$	$QRF^{(Q)}$				$UQR^{(P)}$	$UQR^{(Q)}$	$QRF^{(P)}$	$QRF^{(Q)}$
P_{h-1}	NC	NC	+	−	+	+	V_{h-1}	NC	NC	−	−	−	+
P_{h-2}	NC	NC	+	−	+	+	V_{h-2}	NC	NC	−	−	−	+
P_{h-3}	NC	NC	+	−	+	+	V_{h-3}	NC	NC	−	−	−	+
P_{h-4}	NC	NC	+	−	+	−	V_{h-4}	NC	NC	−	−	−	−
P_{h-5}	NC	NC	−	−	−	−	V_{h-5}	NC	NC	−	−	−	−
P_{h-6}	NC	NC	−	−	−	−	V_{h-6}	NC	NC	−	−	−	−
P_{h-24}	NC	NC	+	−	+	+	V_{h-24}	NC	NC	−	−	−	+
P_{h-168}	NC	NC	+	−	+	+	V_{h-168}	NC	NC	−	−	−	+
Q_{h-1}	NC	NC	−	+	+	+	$\hat{P}_h^{(0.01)}, \hat{Q}_h^{(0.01)}$	+	+	NC	NC	NC	NC
Q_{h-2}	NC	NC	−	+	+	+	$\hat{P}_h^{(0.10)}, \hat{Q}_h^{(0.10)}$	+	+	NC	NC	NC	NC
Q_{h-3}	NC	NC	−	+	+	+	$\hat{P}_h^{(0.20)}, \hat{Q}_h^{(0.20)}$	+	+	NC	NC	NC	NC
Q_{h-4}	NC	NC	−	−	+	−	$\hat{P}_h^{(0.30)}, \hat{Q}_h^{(0.30)}$	+	+	NC	NC	NC	NC
Q_{h-5}	NC	NC	−	−	−	−	$\hat{P}_h^{(0.40)}, \hat{Q}_h^{(0.40)}$	+	+	NC	NC	NC	NC
Q_{h-6}	NC	NC	−	−	−	−	$\hat{P}_h^{(0.50)}, \hat{Q}_h^{(0.50)}$	+	+	NC	NC	NC	NC
Q_{h-24}	NC	NC	−	+	+	+	$\hat{P}_h^{(0.60)}, \hat{Q}_h^{(0.60)}$	+	+	NC	NC	NC	NC
Q_{h-168}	NC	NC	−	+	+	+	$\hat{P}_h^{(0.70)}, \hat{Q}_h^{(0.70)}$	+	+	NC	NC	NC	NC
$\boldsymbol{hod}_h$	+	+	+	+	+	+	$\hat{P}_h^{(0.80)}, \hat{Q}_h^{(0.80)}$	+	+	NC	NC	NC	NC
$\boldsymbol{dom}_h$	+	+	+	+	+	+	$\hat{P}_h^{(0.90)}, \hat{Q}_h^{(0.90)}$	+	+	NC	NC	NC	NC
$\boldsymbol{tod}_h$	+	+	+	+	+	+	$\hat{P}_h^{(0.99)}, \hat{Q}_h^{(0.99)}$	+	+	NC	NC	NC	NC

While day-of-the-month may not be important at all in conventional load forecasting studies that cover mostly residential loads or aggregate loads that have significant portion of residential usage, it can be a very effective variable to capture maintenance schedules, supply cycles and production cycles in industrial load forecasting. From the outcomes shown in Table 6.3.1, the calendar variable ***dom*** is picked by the model selection procedure.

To provide a quantitative estimation of the relative importance of the calendar variable ***dom***, Table 6.3.2 reports the relative predictor importance values, obtained from QRF, for the variables ***dom, hod*** and ***tod*** when considered

alone (i.e., without interactions with other variables). Values are relativized to the predictor importance of the ***hod*** variable, which is always the greatest among the three.

From Table 6.3.2, although the ***dom*** variable always has the smallest relative importance, still it is comparable to the relative importance of the ***tod*** variable, which is instead more frequently used in traditional load forecasting studies.

The error indices obtained for the test period are in Table 6.3.3. With reference to active power, the MQR on UQR with fixed-origin window returns the smallest PS. The PS is about 13% smaller than the best benchmark (i.e., the UQRB), about 12% smaller than the MQR on QRF with fixed-origin window, and less than 1% smaller than the MQR on UQR with rolling window. Forecasts are reasonably reliable as the AACE is smaller than 4%, except for the UQRB model which has a large AACE (about 60% greater than MQR-on-UQR forecasts). MQR-on-UQR forecasts are also more reliable than the MQR-on-QRF forecasts. The PS of the multivariate approaches on QRF (UQR) is smaller than the PM PS by about 15% (22.5%).

Table 6.3.2. Analysis of the relative importance of calendar variables

Variable	Load			
	Aggregate load		*Painting machine*	
	Active power P	Reactive power Q	Active power P	Reactive power Q
hod	1	1	1	1
dom	0.186	0.198	0.041	0.045
tod	0.674	0.260	0.103	0.090

Table 6.3.3. Error indices of hour-ahead aggregate load forecasts

Model	Active power		Reactive power	
	PS [kW]	AACE [%]	PS [kVAr]	AACE [%]
MQR on QRF, rolling window	512.57	3.86	352.92	3.99
MQR on QRF, fixed-origin window	511.68	3.91	347.91	2.98
MQR on UQR, rolling window	453.30	3.22	316.44	2.97
MQR on UQR, fixed-origin window	451.27	3.04	313.69	2.84
PM	606.08	-	423.81	-
UQRB	518.02	7.71	357.94	9.96
QRFB	522.76	2.77	371.85	2.12

With reference to reactive power, the MQR on UQR with fixed-origin window returns the smallest PS among the competitors. The PS is about 12.5% smaller than the best benchmark (i.e., the UQRB), about 10% smaller than the MQR on QRF with fixed-origin window, and less than 1% smaller than the MQR on UQR with rolling window. Also in this case, forecasts are reasonably reliable as the AACE is smaller than 4%, with the exception for the UQRB model which has a too large AACE (about 70% greater than MQR-on-UQR forecasts). Together with their smaller

PS, forecasts built by the MQR on UQR are also more reliable than the MQR-on-QRF forecasts. The PS of the multivariate approach on QRF (UQR) is smaller than the PM PS by about 18% (26%), respectively.

In order to give a graphical evaluation of the results, Figures 6.3.3 and 6.3.4 respectively show the prediction intervals obtained from MQR-on-UQR hour-ahead predictive quantiles of active and reactive power of the aggregate load. Values are plotted for a week of the test period, and compared to the actual load.

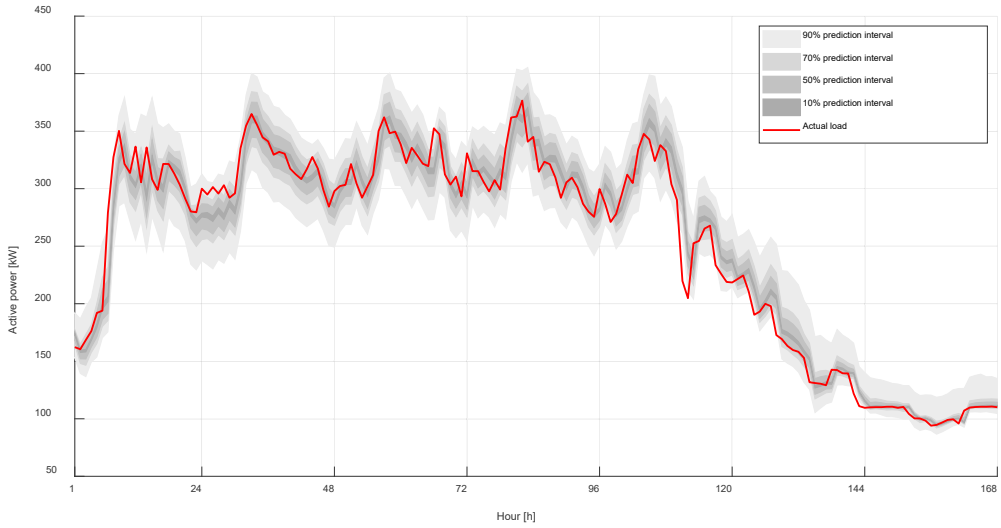

Fig. 6.3.3. MQR-on-UQR hour-ahead active power forecasts during a week of the test period.

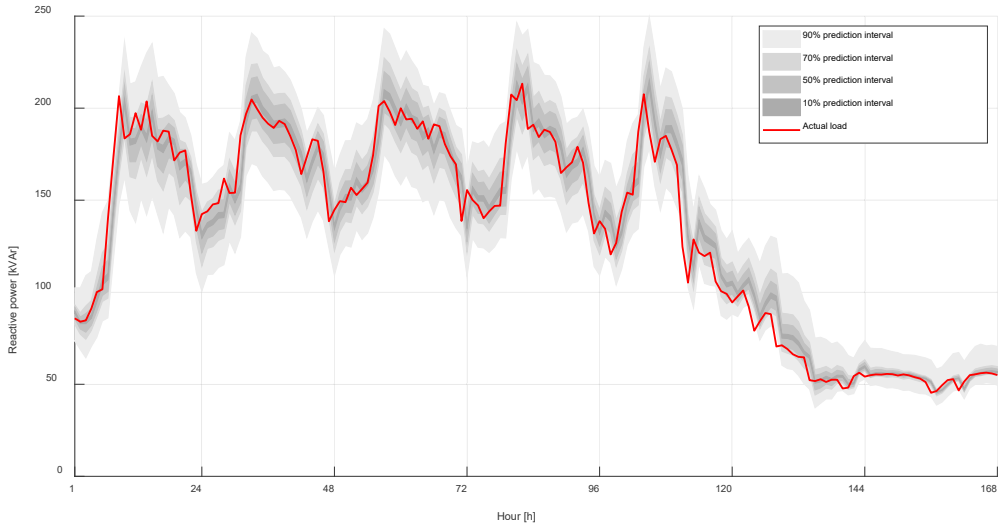

Fig. 6.3.4. MQR-on-UQR hour-ahead reactive power forecasts during a week of the test period.

5) Results of hour-ahead forecasts of the carpentry feeder and of the painting machine

Only the error indices for the forecasts of the carpentry feeder and of the painting machine are discussed, to avoid the verbose presentation of the outcomes of the model selection procedure. The error indices averaged during the test period are shown in Table 6.3.4.

With reference to carpentry feeder active power, the MQR on UQR with fixed-origin window returns the smallest PS among the competitors. The PS is about 10% smaller than the best benchmark (i.e., the UQRB), about 8% smaller than the MQR on QRF with fixed-origin window, and less than 1% smaller than the MQR on UQR with rolling window. The MQR with rolling window fails to outperform the UQRB. Forecasts generated by MQR-on-UQR models are less reliable than the forecasts generated by MQR-on-QRF models, although the AACE is smaller than 4.20% in all cases. Therefore, forecasts with the smaller PS should be picked over the competitors.

With reference to painting machine active power, the MQR on QRF with fixed-origin window returns the smallest PS among the competitors. The PS is about 6.5% smaller than the best benchmark (i.e., the QRFB), about 4.5% smaller than the MQR on UQR with fixed-origin window, and less than 1% smaller than the MQR on QRF with rolling window. MQR on QRF returns also the most reliable forecasts, thus it is the privileged pick over competitors.

With reference to carpentry feeder reactive power, MQR on UQR returns forecasts slightly more accurate than the MQR-on-QRF forecasts. The improvement is however smaller than 1% of the PS, and this is probably due to the worse performance of the UQR upon carpentry feeder reactive power. The improvement is mainly imputable to the better performance of the UQR upon carpentry feeder active power, which is fruitfully exploited in the multivariate approach. Since MQR-on-UQR forecasts are less reliable than MQR-on-QRF forecasts and the PS of the former is only slightly smaller than the PS of the latter, the latter could be picked over competitors if reliability matters more for the end-user applications.

With reference to painting machine reactive power, MQR on UQR returns forecasts about 5% more accurate than the MQR-on-QRF forecasts. The UQR has worse performance than the QRF both upon active and reactive power, and this has implications on the accuracy of the multivariate forecasts.

On average, the MQR on UQR has better performance than the competitors for the aggregate load and for the load of the carpentry feeder, whereas the MQR on QRF outperforms the competitors for the load of the painting machine. This can be explained by the nature of the latter load: the painting machine follows an irregular pattern, with many zero-load intervals. QRFs appear to better catch the uncertainties of this type of load, and this has implications also on the accuracy of the multivariate forecasts build upon QRF underlying models.

Table 6.3.4. Error indices of hour-ahead forecasts of the carpentry feeder and of the painting machine

Model	Carpentry feeder				Painting machine			
	Active power		Reactive power		Active power		Reactive power	
	PS [kW]	AACE [%]	PS [kVAr]	AACE [%]	PS [kW]	AACE [%]	PS [kVAr]	AACE [%]
MQR on QRF, rolling window	261.31	2.90	193.00	2.41	43.61	1.87	39.28	4.32
MQR on QRF, fixed-origin window	250.25	2.96	192.74	2.51	43.46	1.81	39.02	4.22
MQR on UQR, rolling window	231.23	4.18	191.58	4.22	45.75	4.53	41.26	3.88
MQR on UQR, fixed-origin window	230.97	4.05	191.05	4.80	45.55	3.89	41.10	5.22
PM	314.80	-	233.60	-	61.48	-	50.95	-
UQRB	257.42	12.40	215.11	10.14	50.10	5.36	46.00	8.19
QRFB	266.57	2.43	197.58	4.81	46.47	5.47	40.71	2.48

6) Results of day-ahead load forecasts

Day-ahead forecasts of aggregate load, of the carpentry feeder, and of the painting machine are assessed in this sub-Section. The error indices averaged during the test period are shown in Table 6.3.5. The MQR on UQR with fixed-origin window returns the most accurate forecasts among the competitors for the aggregate load and for the carpentry feeder. Compared to the MQR on QRF with fixed-origin window, the PS improvements are about 7% for the aggregate active power, 5% for the aggregate reactive power, 10% for the carpentry feeder active power, and 10.5% for the carpentry feeder reactive power. Compared to the best univariate benchmarks, the improvements are greater than 12.5%. The AACE of MQR on UQR with fixed-origin window is smaller than 4.50% in all of the considered cases.

The MQR on QRF with fixed-origin window returns the most accurate forecasts among the competitors for the load of the painting machine. Compared to the MQR on UQR with fixed-origin window, the PS improvements are about 2% and 1% for the active and reactive power, respectively. Compared to the best univariate benchmarks (i.e., the QRFB), the improvement is about 6% for both active and reactive power. In the case of reactive power forecasting, the AACE of MQR on QRF models is over 6.50%, so that MQR on UQR with fixed-origin window could be picked over it if reliability significantly matters for the end-user applications.

The PM PS of day-ahead forecasts is about 3÷4 times the PM PS of hour-ahead forecasts (see Tables 6.3.3 and 6.3.4). Nevertheless, the PSs of day-ahead forecasts built by the multivariate approaches instead are only 1.7÷2.5 times the MQR PS of hour-ahead forecasts. This suggests that the proposed multivariate approach is more skilled than a naïve approach in handling longer-term forecasting.

Table 6.3.4. Error indices of day-ahead forecasts

Model	Aggregate load				Carpentry feeder				Painting machine			
	Active power		Reactive power		Active power		Reactive power		Active power		Reactive power	
	PL [kW]	CE [%]	PL [kVAr]	CE [%]	PL [kW]	CE [%]	PL [kVAr]	CE [%]	PL [kW]	CE [%]	PL [kVAr]	CE [%]

MQR on QRF, rolling window	1281.46	4.85	873.09	3.52	485.62	6.49	427.31	5.13	74.45	2.32	79.47	8.41
MQR on QRF, fixed-origin window	1162.21	3.29	835.79	4.16	484.82	6.55	424.04	4.98	74.35	3.08	79.32	6.66
MQR on UQR, rolling window	1105.48	3.95	798.93	4.10	444.05	5.58	380.42	2.82	76.38	4.15	80.74	6.36
MQR on UQR, fixed-origin window	1081.23	3.34	793.00	4.39	437.19	4.46	379.12	3.01	75.92	4.02	80.27	5.00
PM	2458.38	-	1656.41	-	987.81	-	840.99	-	140.98	-	149.14	-
UQRB	1239.04	6.01	904.22	7.13	518.85	6.03	437.99	5.71	85.01	4.25	92.32	4.19
QRFB	1330.80	11.26	917.65	6.94	511.36	4.31	462.30	5.86	79.12	7.70	84.52	5.79

5.3.3. DISCUSSION

This research activity proposes a new contribution to probabilistic industrial load forecasting. A multivariate approach is proposed to model the physical, mutual correlation between active and reactive power.

Univariate predictive quantiles of active and reactive power are built either by QRF or by UQR. These forecasts are passed to an MQR model, which returns multivariate forecasts of active and reactive power in the form of predictive quantiles. The MQR model simultaneously addresses the two target variables, considering also different quantile levels. The MQR parameters are estimated by solving a linear programming problem in which inequality constraints avoid quantile crossing occurrences. Two MQR training schemes, either based on a fixed-origin window or on a rolling window, are compared in the experiments. The proposal is validated using electrical data of three industrial loads, monitored at an actual Italian factory.

Numerical experiments suggest that the MQR training scheme with fixed-origin window returns a smaller PS, and in most cases also a smaller AACE, than the corresponding rolling-origin window counterpart. Whether the QRF or the UQR are the underlying probabilistic models used to build univariate forecasts, the MQR enhances the skills of final forecasts. This enhancement ranges from 6% to 13.5%, compared to the univariate benchmarks. The MQR on QRF is the best pick for two loads (the factory aggregate and the load of a feeder), whereas the MQR on UQR is the best pick for the other load (a painting machine).

Future research on the applications of probabilistic industrial load forecasting are encouraged, mainly with reference to the development of decision-making tools for intelligent participation to electrical markets and to the development of management and control strategies in industrial microgrids.

6. COMPARISON WITH EXPECTED RESULTS

With reference to the Task 2.1 and the Task 2.2 of the Digriflex project, the expected results at the end of the Milestone 1 include:

a) Task 2.1:
   - collection of time series that include both target variables and predictor variables;
   - description of methods to pre-process the data to eliminate outliers and bad values;
   - exploratory data analysis in order to discard uninformative predictors.

b) Task 2.2:
   - development of methods and models based on multiple linear regression and random forests
   - identification of the best combination type for the underlying models in the ensemble approach
   - comparison between the results obtained applying the proposed approach and the ones given by the relevant state-of-the-art benchmarks.

The activities related to Task 2.1 allowed to create large, robust databases of variables which could be exploited to test the performance of the forecasting systems used to predict energy at the site of the installation of the test distribution grid of the ReIne laboratory.

The activities related to Task 2.2 allowed to develop forecasting systems of renewable generation and loads for the considered day-ahead time horizons. As expected, the developed forecasting systems were based on data-driven ensemble approaches, as they are flexible and versatile tools. Eventually, the proposed methodologies were validated and their effectiveness was demonstrated under realistic uncertainty sources, exploiting the databases built through the activities related to Task 2.2.

Part of these activities were carried out in strict interaction with the partner University of Naples Federico II for the data collection, data pre-processing and exploratory data analysis and for the theoretical development of the forecasting systems, and with the partner HEIG-VD for the acquisition of measurements available at the test distribution network at the ReIne laboratory.

In summary, the activities developed by University of Naples Parthenope in Milestone 1, illustrated above, allowed to fully reach the objectives expected in Task 2.1 and Task 2.2 of the DiGriFlex project. The activities will continue, in strict interaction with the partners, during the Task 2.4 in order to adapt the forecasting systems to the feedbacks provided by the two-level optimization strategy and to the validation results that are object of the WP3.

7. DELIVERABLES

The deliverables of the research activities related to Task 2.1 and Task 2.2 consist of this technical document, as expected per the Technical Specifications of the DiGriFlex Project, and of the following scientific papers published on international journals and proceedings of international conferences:

1) A. Bracale, P. Caramia, G. Carpinelli, P. De Falco, "Day-ahead probabilistic wind power forecasting based on ranking and combining NWPs," *International Transactions on Electrical Energy Systems*, vol. 30, no. 7, e12325, 2020.
2) A. Bracale, P. Caramia, P. De Falco, T. Hong, "A multivariate approach to probabilistic industrial load forecasting," *Electric Power Systems Research*, vol. 187, 106430, 2020.
3) M. Bozorg, A. Bracale, P. Caramia, G. Carpinelli, M. Carpita, P. De Falco, "Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting," *Journal of Protection and Control of Modern Power Systems*, vol.5, 21, pp. 1-12, 2020.

8. PROFILES OF HUMAN RESOURCES

The researchers of the University of Naples Parthenope who participated at the research activities related to Task 2.1 and Task 2.2 are:

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

9. DIFFUSION OF THE RESULTS

The results of the research activities related to Task 2.1 and Task 2.2 have been published in the following scientific papers in international journals and in proceedings of international conferences:

1) A. Bracale, P. Caramia, G. Carpinelli, P. De Falco, "Day-ahead probabilistic wind power forecasting based on ranking and combining NWPs," *International Transactions on Electrical Energy Systems*, vol. 30, no. 7, e12325, 2020.
2) A. Bracale, P. Caramia, P. De Falco, T. Hong, "A multivariate approach to probabilistic industrial load forecasting," *Electric Power Systems Research*, vol. 187, 106430, 2020.

3) M. Bozorg, A. Bracale, P. Caramia, G. Carpinelli, M. Carpita, P. De Falco, "Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting," *Journal of Protection and Control of Modern Power Systems*, vol.5, 21, pp. 1-12, 2020.

10. CONCLUSIONS

The research activities discussed in this document focused on the development of day-ahead forecasting systems for renewable generation and loads. A dedicated task aimed at building the input datasets by collecting and pre-processing available data. The forecasting systems have been developed by exploiting ensemble techniques within probabilistic frameworks.

The major contributions of the research activities related to Task 2.1 and Task 2.2 can be summarized as follows:

i. the collection of large, robust datasets of PV power generation, wind power generation and industrial loads;
ii. the pre-processing of the collected data, aiming at individuating and correcting missing data, bad data and outliers;
iii. the exploratory data analysis to reduce the dimensionality of the input datasets, favoring the development of adequate forecasting systems for renewable generation and load;
iv. the development of a Bayesian bootstrap quantile regression model for probabilistic PV power forecasting;
v. the development of a day-ahead probabilistic wind power forecasting based on ranking and combining NWPs;
vi. the development of a multivariate approach for probabilistic industrial load forecasting.

In summary, the activities developed by University of Naples Parthenope in Milestone 1, illustrated above, allowed to fully reach the objectives expected in Task 2.1 and Task 2.2 of the DiGriFlex project. The activities will continue, in strict interaction with the partners, during the Task 2.4 in order to adapt the forecasting systems to the feedbacks provided by the two-level optimization strategy and to the validation results that are object of the WP3.

BIBILIOGRAPHY

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.