Advances in Applied Energy 9 (2023) 100120

Contents lists available at ScienceDirect

Advances in Applied Energy

journal homepage: www.elsevier.com/locate/adapen

Day-ahead probabilistic forecasting at a co-located wind and solar power park in Sweden: Trading and forecast verification

O. Lindberg [a,*], D. Lingfors [a], J. Arnqvist [b], D. van der Meer [c], J. Munkhammar [a]

[a] Department of Civil and Industrial Engineering, Uppsala University, Lägerhyddsvägen 1, Uppsala 752 37, Sweden
[b] Department of Earth Sciences, Uppsala University, Villavägen Uppsala 16 752 36, Sweden
[c] Mines Paris, PSL University, Centre for processes, renewable energy and energy systems (PERSEE), Sophia Antipolis 06904, France

ARTICLE INFO

Keywords:
Forecast value
Quantile forecasts
PV power
Wind power
Hybrid power park
Probabilistic forecasting

ABSTRACT

This paper presents a first step in the field of probabilistic forecasting of co-located wind and photovoltaic (PV) parks. The effect of aggregation is analyzed with respect to forecast accuracy and value at a co-located park in Sweden using roughly three years of data. We use a fixed modelling framework where we post-process numerical weather predictions to calibrated probabilistic production forecasts, which is a prerequisite when placing optimal bids in the day-ahead market. The results show that aggregation improves forecast accuracy in terms of continuous ranked probability score, interval score and quantile score when compared to wind or PV power forecasts alone. The optimal aggregation ratio is found to be 50%–60% wind power and the remainder PV power. This is explained by the aggregated time series being smoother, which improves the calibration and produces sharper predictive distributions, especially during periods of high variability in both resources, i.e., most prominently in the summer, spring and fall. Furthermore, the daily variability of wind and PV power generation was found to be anti-correlated which proved to be beneficial when forecasting the aggregated time series. Finally, we show that probabilistic forecasts of co-located production improve trading in the day-ahead market, where the more accurate and sharper forecasts reduce balancing costs. In conclusion, the study indicates that co-locating wind and PV power parks can improve probabilistic forecasts which, furthermore, carry over to electricity market trading. The results from the study should be generally applicable to other co-located parks in similar climates.

1. Introduction

The power output of wind and solar photovoltaic (PV) power is intermittent and non-dispatchable, hence accurate forecasts are crucial for balancing purposes and providing technical support when trading [1]. Conventional forecasts, which give the expected generation output as a single value at each lead time, are referred to as deterministic forecasts. However, every lead time has an inherent and irreducible uncertainty which is why forecasted uncertainty information at these time steps, often referred to as probabilistic forecasts, have gained attention in renewable energy forecasting over the last years [1,2]. Short-term forecasts, having a lead time from a few hours to a few days ahead, are used for day-ahead market trading where electricity generated from wind and solar are exchanged [1]. Such forecasts are typically produced by numerical weather predition (NWP) models in a first step and post-processed using statistical models in order to reduce forecast errors in a second step. Although more commonly used when forecasting wind power [1], using NWP forecasts to improve the accuracy of day-ahead PV and solar irradiance forecasts have become increasingly popular [3].

The variability in wind and PV power production can be reduced through spatial dispersion of the systems, commonly known as the *smoothing effect* [4]. For example, Katzenstein et al. [5] studied 20 interconnected wind power parks across Texas and found that the reduction in variability occurred on time scales shorter than 24 hours. In Lave et al. [6], the aggregation of six irradiance measurement stations within 3 km in San Diego was studied. The authors found that ramp rates became uncorrelated on time scales below 5 minutes which reduced the aggregated variability. In both cases, the smoothing occurs due to averaging independent variations in the signals. However, if the correlation is close to zero or even negative, which is often the case for wind and PV power [7], the smoothing effect can be achieved without physically separating the systems. This motivates co-locating wind and PV parks to mitigate the average variability in the combined output.

A co-located park is owned by a single power producer where the power sources share a common interconnection point to the electricity grid. This means that a co-located park is considered as a single power generating facility according to the network codes Requirements for

* Corresponding author.
E-mail address: oskar.lindberg@angstrom.uu.se (O. Lindberg).

https://doi.org/10.1016/j.adapen.2022.100120
Received 26 September 2022; Received in revised form 14 December 2022; Accepted 30 December 2022
Available online 3 January 2023
2666-7924/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Generators (RfG) [8,9]. In fact, individual co-located wind and PV parks are motivated as potential candidates to be part of electricity markets related to capacity services [10], ancillary services [11] and time-varying pricing [12] from which profitability of future individual power parks will likely depend on [13]. Unlike separated parks that are connected to the same grid, individual co-located parks can be oversized (i.e., the total nameplate capacity exceeds the maximum power that can be fed into the grid) to increase the production rate and also be complemented with battery storages. This means that co-located parks can, e.g., share control strategies [14] or provide flexibility to the electricity grid by coordinating energy management strategies within the park [15]. As recent as March 30, 2022, the IEA TCP Wind Task 50 was also initiated to coordinate and develop the field of co-located wind and PV parks [16]. Furthermore, in a recent review paper, Lindberg et al. [9] found that the scientific literature on co-located wind and PV parks is scarce, although the parks have been studied and motivated in various aspects, e.g., more effective usage of land [17], increased utilization of the electrical infrastructure [18] as well as shared operation and maintenance work [10]. However, the review paper highlights that the field of forecasting co-located parks is particularly limited, although, e.g., Long et al. [19] highlight the importance of accurate day-ahead forecasts from aggregated co-located wind and PV generation when trading. A natural first step in the field of co-located wind and PV forecasting is therefore to analyze the effect of aggregation of power sources in co-located parks in terms of forecast accuracy relative to forecast accuracy of single power source production. This guides to quantify the potential benefits of forecasting aggregated production for co-located parks relative to combining forecasts from each power source.

There are a few studies that have focused on forecasting co-located wind and PV. Alessandrini and McCandless [20] forecasted the output of a utility-scale co-located wind and PV park in Kuwait using analog ensemble combined with Schaake shuffle technique to recover the correlation and autocorrelation between the power sources. The corresponding paired solar and wind power members were thereafter summed to build an ensemble of combined generation that was statistically consistent. At the same co-located park, Haupt et al. [21] describes the forecasting system used to generate forecasts. In Pombo et al. [22], different physics-informed machine learning methods to issue deterministic forecasts for the individual wind and PV generation units in a small-scale co-located system in Denmark was assessed. The results showed that shallow machine learning methods perform better for 5 minute resolved data, while deep learning methods perform better for hourly resolved data for wind and solar, respectively. Furthermore, the authors found that PV power was easier than wind power to predict due to its deterministic diurnal variability.

Although Alessandrini and McCandless [20], Haupt et al. [21] and Pombo et al. [22] are the only studies known to the authors that explicitly forecast wind and PV production at a co-located park, there are other studies that forecast aggregated production. In Camal et al. [23], a multivariate copula was used to generate correlated scenarios from day-ahead probabilistic forecasts of dispersed wind, PV and small hydro power plants. The authors found that scenarios generated from aggregated production directly (as compared to separate forecasts from each power source) reproduced less variability due to the smoothing effect. Aggregated production also resulted in less forecast errors, which was explained by the efficient learning capacity in high dimension of the forecast model quantile regression forest (QRF, a statistical machine learning method introduced in detailed in Section 2.2.1). In the same paper, aggregated forecasts increased profit and reduced costs when used for reserve bidding and unit commitment. This was explained by such applications being sensible to the amplitude of the scenarios which was more present in scenarios generated from aggregated production. Yin et al. [24] proposed a coordinated day-ahead scheduling of power systems consisting of thermal, hydro, wind and PV power sources. For this, scenarios of wind and PV power production were generated using a Monte Carlo method and a bi-variate copula. The authors found that stronger negative correlations between wind and PV power output resulted in lower system operation cost.

In two studies by van der Meer et al. [25,26], it was observed that the aggregated net load (*Load* minus *PV power production*) probabilistic forecasts are better calibrated than electricity demand forecasts. Two plausible explanations to this are brought up which relate to the combined distributions being more Gaussian (the forecasts where generated using Gaussian Processes which produces a Gaussian probability density function (PDF)) [26] and that the combined time series are smoother and consequently more straightforward to predict [25]. In van der Meer et al. [25], it was also found that for time periods of high PV power variability, i.e., during spring, the accuracy of the PV power forecasts tended to be lower but the accuracy of the net load forecast was higher compared to electricity demand forecasts alone.

1.1. Aim of the study

The previous section detailed that the research on co-located wind and PV park generation forecasting is limited and many questions are still unanswered. Therefore, we clearly delimit the scope of this study. Specifically, the purpose of this study is not to produce the most accurate forecasts. Instead, we investigate whether co-locating wind and PV parks benefits forecast accuracy and value under a fixed modelling framework. To that end, we analyze what impact the smoothing effect – caused by local wind power and PV power time series aggregation – has on probabilistic forecast accuracy and electricity market participation. In this way, we cover the three types of "goodness" of a forecast, namely (i) consistency, (ii) quality and (iii) value [27]. Furthermore, we study the effect of aggregation as a function of season, since the variability of wind and PV power generation varies over time. The contributions of this paper can be summarized as follows:

1. We analyze the effect of aggregating varying shares of installed wind power and PV power in a co-located park using calibrated probabilistic forecasts generated using a fixed modelling framework.
2. We simulate day-ahead electricity trading to study the impact of aggregation on the value of the forecasts.
3. We investigate the influence of seasonal variation in power output variability on forecast accuracy.

The remainder of the paper is structured as follows. In Section 2 the data as well as the forecasting methods, verification metrics and algorithm for participation in the day-ahead market using probabilistic forecasts is presented. In Section 3, results from the study are presented, which are discussed in Section 4 including ideas for future studies. Finally, in Section 5, the main conclusions are put forward.

2. Data and methodology

This section presents the data and methodologies used in the study. In Section 2.1, the data used in the study is presented. Section 2.2 presents the forecasting algorithms used to generate probabilistic forecasts under a fixed modelling framework, whereas Section 2.3 presents visual verification tools and numerical scores. An overview of the electricity market and optimal trading strategies based on probabilistic forecasts are presented in Section 2.4. Finally, Fig. 1 provides an overview of the study's framework as a flowchart with regards to the data and methodologies presented in this section. More specifically, the figure shows how scenarios of different wind and PV capacity ratios in the co-located park are used as inputs to the forecast procedure and evaluated in terms of forecast performance metrics and profitability on the day-ahead spot market.

2.1. Data

This section presents the data used in the study. In Section 2.1.1, the co-located power park and corresponding production data are presented. Section 2.1.2 describes the NWP model used when forecasting

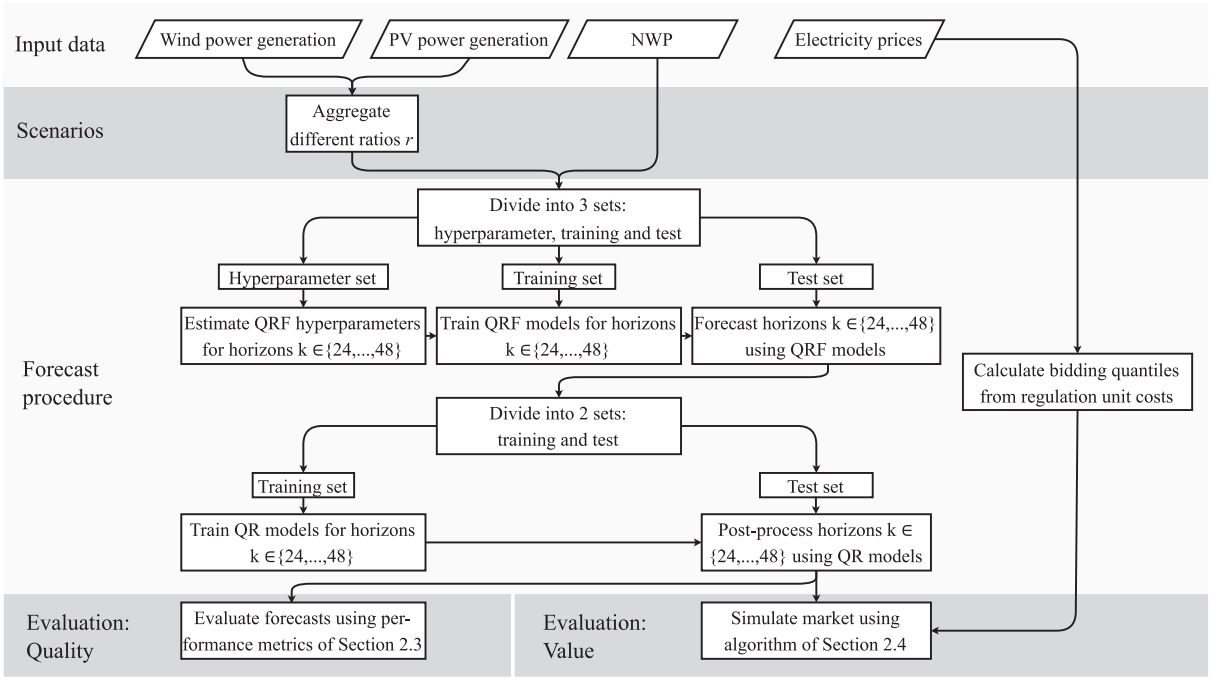

Fig. 1. Flowchart illustrating the input data, method steps and analysis in this study. The input data are illustrated as parallelograms.

the power output and Section 2.1.3 presents the electricity prices used in simulating the electricity market.

2.1.1. Co-located wind and PV park

Lindberg et al. [9] found that the number of utility-scale co-located wind and PV power parks are limited on a global scale (27 in total in 2021). For this study, we used data from a co-located park on the west coast of Sweden (Latitude: 57.033577, Longitude: 12.381662). The wind power park consists of three Vestas V90 2 MW_p wind turbines with a combined nominal capacity of 6 MW_p. The PV power park, which was installed after the wind power park, consists of 9300 monocrystalline modules with a combined nominal capacity of 2.7 MW_p. This means that the ratio of PV to the whole park is roughly 31% in terms of nominal capacity. The dataset spans from 2016-07-02 (the date of commissioning of the PV park) to 2020-05-21 at 15 minute resolution. The NWP dataset used, which is explained in Section 2.1.2, was available from 2016-11-09. Prior to modeling, the data were preprocessed such that; (i) night-time values were set to zero for the PV production; (ii) negative production was set to zero for both generation units, e.g., indicating stand-by power from inverters at zero production; (iii) production above the nominal capacity was set to the nominal capacity for both generators; (iv) if the power production from the wind turbines were zero and the wind speed from the NWP was above 7 m/s, the whole day was removed to avoid corrupting the forecasts errors with down-time of the turbines due to, e.g., maintenance; (v) if the power production from the PV park was zero and the solar irradiance from the NWP was above 100 W/m^2, the whole day was removed (corresponding reason as in (iv)); (vi) data resolution was summed to hourly resolved values since they were given in watt-hours (Wh) and (vii) normalized by the nominal capacity of the total park output. The pre-processing resulted in 23 days being removed from the complete data set (30,936 hours).

Figure 2 shows the observed diurnal variability of PV and wind power output including 60% nominal coverage for the different seasons. In this paper, the seasons are defined as 3-month periods according to meteorological conventions; summer corresponds to June-August, fall corresponds to September-November, winter corresponds to December-February and spring corresponds to March-May. This means that the seasons are offset with respect to the solstices (or equinoxes), which means that the summer and spring receives almost the same irradiance in clear sky conditions. The PV power output has an obvious diurnal cycle where the magnitude and variability is a function of the seasons. The wind power output has a relatively stable diurnal profile, where the diurnal variability during the summer and spring is mainly due to temperature gradients due to the increase in solar irradiance. The highest wind power output and variability is during the winter and the lowest is during the summer and vice versa for PV. The capacity factor, i.e., the ratio of produced electric energy to the amount of energy if operated at rated power for a time period, is 14.1% and 35.6% for the PV and wind power park, respectively.

In order to study the different shares of wind and PV in the combined park, the first step is to combine different ratios (r) of wind (P_{WP}) and solar power (P_{PV}) to obtain the combined power output of the co-located park:

$$P_{tot} = rP_{WP} + (1-r)P_{PV}, \tag{1}$$

where $r \in \{0, 0.1, \dots, 1\} \times 100\%$ and the nominal capacities of P_{WP}, P_{PV} and P_{tot} are the same. This facilitates a fair comparison when verifying the accuracy of the forecasts since many of the diagnostic tools for probabilistic forecasts are directly dependent on the unit of the variables, which are explained in Section 2.3. Furthermore, this means that a park with a ratio of 0% consists of solely a PV park whereas a park with a ratio of 100% consists of solely a wind park, both with the same nominal capacity. A park with a ratio of 50% consists of wind and PV with equal nominal capacity.

2.1.2. NWP data

This study uses data from the mesoscale NWP Meterological Cooperation on Operational Numeric Weather Prediction (MetCoOp) with a horizontal resolution of approximately 2.5×2.5 km^2 [28,29]. The NWP data are freely available from the Norwegian Meteorological Institute [30]. From the NWP data, hourly time series of day-ahead forecasts (00 UTC cycle runs, lead times $t+k$, $k \in \{24, \dots, 48\}$) of the control ensemble member for selected meteorological variables were downloaded.

We post-process global horizontal irradiance (GHI) and wind speed at 80 m from MetCoOp to power output. MetCoOp provides the hourly cumulative surface downwelling shortwave radiation (J/m^2), which we

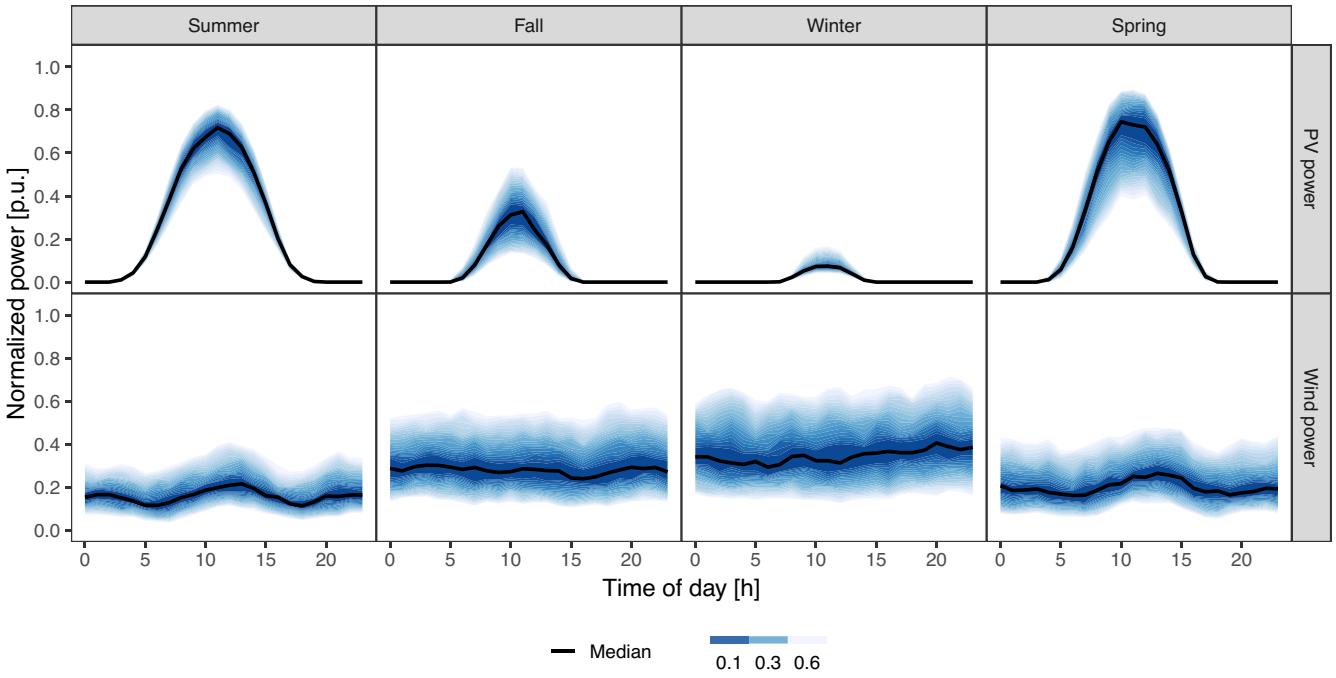

Fig. 2. Measured median of diurnal variability of PV power (upper subplot) and wind power (lower subplot) for the different seasons at the co-located wind and PV power park. The blue envelopes show the observed 20–80% quantile and the black solid lines show the median. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

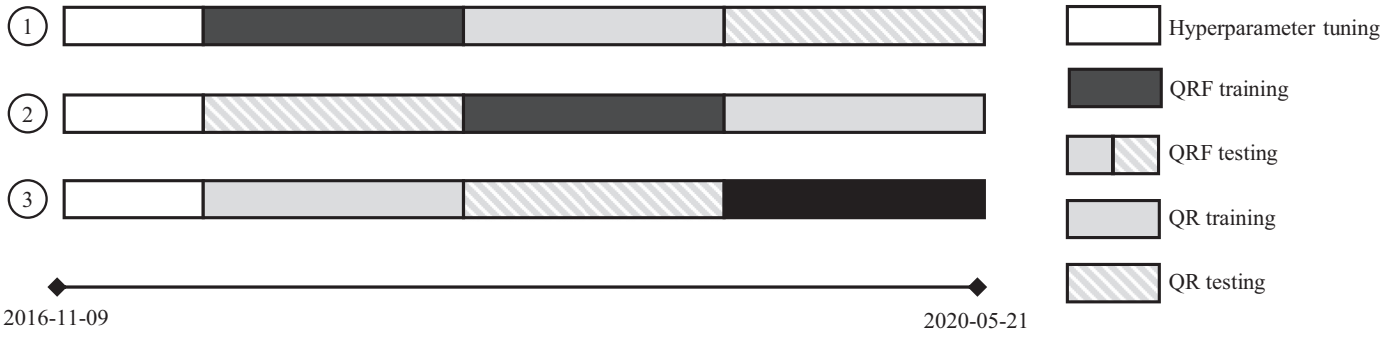

Fig. 3. The division of the data sets for the hyperparameter tuning (white block), quantile regression forest (QRF) training (black block), quantile regression (QR) training (gray block) and QR testing (gray striped block) for the three different model runs. The QRF testing is the combination of the QR training and testing data sets. The numbered circles corresponds to the three different shuffles.

convert to GHI (W/m^2) by taking the difference between the time steps and dividing it by 3600. The 80 m wind speeds were extracted by transformation of the model levels to pressure coordinates and finally height coordinates through application of the Hypsometric equation. Thereafter, the 80 m wind speeds were calculated using linear interpolation of the height coordinates. Forecast horizons corresponding to zenith angles below 85° (solar elevation angles larger than 5°) are used when verifying accuracy the forecasts. This means that we only include daytime values since there is no available PV power production during the night time. When simulating the value of the forecasts, all data are included. For the NWP data, no pre-processing was done to the data.

2.1.3. Electricity price data

The electricity pool for Scandinavia, in which the park operates, is commonly known as Nord Pool. Day-ahead as well as up and down regulating prices covering 26,280 hours of the complete dataset (2016-11-09 to 2020-05-21) were downloaded from Nord Pool [31] in order to simulate the day-ahead market. The reason for choosing three years of recent data is to shuffle the dataset to test the robustness of the results. The shuffling is done according to Fig. 3. We assume perfect forecasts of the electricity prices, which is warranted because we want to isolate the effect of time series aggregation on the forecast value. The co-located power park, described in Section 2.1.1, is located in the electricity market bidding zone 3 (SE3), for which prices were downloaded.

2.2. Forecasting

This section presents the forecast models in our fixed modeling framework. Section 2.2.1 presents the QRF model that we use to model the nonlinear relationships between the independent and dependent variables, while Section 2.2.2 describes QR, which is used to improve the calibration of the probabilistic forecasts. In Figure 1, the different steps for generating forecasts are presented.

2.2.1. Quantile regression forests

QRF is a commonly used probabilistic forecasting method in the fields of PV [2] and wind power [1] and therefore provides a baseline for analyzing how the aggregation of wind and PV affects the accuracy

of probabilistic forecasts. Furthermore, it has been shown that the forecast distribution can rarely be expressed by a common parametric distribution [32]. To that end, QRF is a nonparametric forecasting method where no assumptions about the predictive density is assumed. QRF is an extension of random forest (RF) [33] and works as follow [34]. First, k trees are grown where each branch of tree T has a random parameter vector θ_t that include the variables under consideration at each split point in each tree such that $T(\theta_t)$. After the trees are grown, the next step is to calculate weights $w_i(x,\theta_t)$ and $w_i(x)$ of observations $i \in \{1,\dots,n\}$ for every tree and every observation, for a certain realization x of predictor X. The weights are defined as [33]:

$$w_i(x,\theta_t) = \frac{\mathbb{1}\{X_i \in R_{\ell_{(x,\theta_t)}}\}}{\#\{j : X_j \in R_{\ell_{(x,\theta_t)}}\}}, \tag{2}$$

$$w_i(x) = k^{-1}\sum_{t=1}^{k} w_i(x,\theta_t), \tag{3}$$

where $R_{\ell_{(x,\theta_t)}}$ is a rectangular subset of space S in which X lives in for every leaf $\ell = 1,\dots,L$. This means that the leaf of which the unknown sample falls into in each tree is found. For each observation in the training set, a weight is given to each observation in each tree. If the observations are in the same leaf as the new sample, the weight is the fraction of samples in the same leaf and otherwise zero.

In RF, the conditional mean of the observed response variable is stored in each leaf. Instead of storing the conditional mean, QRF stores all information after which a conditional distribution can be constructed from the observed response variables. For every $x \in S$, there is only one leaf such that $x \in R_\ell$ which can be denoted as leaf $\ell(x,\theta)$ for tree $T(\theta)$. Finally, the estimated distribution function can be calculated as [34]:

$$\hat{F}(y|X=x) = \sum_{i=1}^{n} w_i(x)\mathbb{1}\{Y_i \le y\}, \tag{4}$$

where Y is a random response variable and y is a realization. From Equation (4), quantiles $\hat{q}_\tau$ can be obtained. A comprehensive description of QRF can found in [34].

Prior to traning and forecasting using the QRF model, we perform hyperparameter tuning. The observations presented in Section 2.1.1 were used as target values and the NWP data described in Section 2.1.2 were used as feature variables for the QRF model. A hyperparameter tuning dataset (which consist of 4104 hours and the total number of data points is 30,384 hours) is used to optimize the performance of the QRF algoritm for each forecast horizon and aggregated ratio r (Equation (1)), respectively. The reason is that we use NWP issued at midnight and the production of wind and PV power behaves differently according to the time of the day, e.g., due to the diurnal variability of PV. The QRF model is mainly influenced by three hyperparameters [23]; the number of trees in the forest (*n_estimators*), the minimum number of observations to become a leaf (*min_samples_leaf*) and the number of features to consider when looking for the best split (*max_features*). *max_features* were chosen as 2 to give opportunity to choose among all available features at each tree. In methods with a large set of features, 1/3 of the available features is usually recommended for regression [35]. For the other two hyperparameters (*n_estimators* and *min_samples_leaf*), we perform an exhaustive grid search to find the set of parameters that minimize the continuous ranked probability score (CRPS) using 10-fold cross-validation. The hyperparameter *n_estimators* should generally be large to make sure that important local features in the data are captured in some of the trees. Therefore, the grid search was performed along n_estimators $\in \{400, 450, \dots, 2000\}$. For the minimum observations per leaf, *min_samples_leaf*, the minimum number of observations should not be too small since that might lead to overfitting the model. Considering this, we performed the grid search along min_samples_leaf $\in \{10, 11, \dots, 20\}$.

After the hyperparameter tuning, we use the training data (that consists of 8760 hours) to train one QRF model for each forecast horizon. After the QRF is built and trained, a vector of explanatory variables from the test data set (consisting of the remaining data, i.e., 17,520 hours) is dropped down the tree. The test data will be compared at each splitpoint and directed to the most similar branch and an output can be estimated from which we produce forecasts with evenly spaced nominal probabilities $\tau \in \{0.05,\dots,0.95\}$ (see Equation (4)). In order to quantify differences in the results within the time period of the study, the division into training and testing data sets are repeated for three different data intervals (in this study denoted as shuffles) according to Fig. 3.

2.2.2. Quantile regression

Although formally introduced in Section 2.3, one of the required properties of probabilistic forecasts relates to the calibration. Uncalibrated forecasts may introduce bias when trading. Initial verification of the QRF model showed that the forecasts were not sufficiently calibrated (see upper panel of Fig. 5), even though we tried to extend the hyperparameter and training datasets of the model. To post-process the QRF forecasts we use QR [36], as was initially suggested by Bremnes [37] for post-processing ensemble predictions [3,38]. QR produces a non-parametric cumulative distribution function (CDF) by assuming a linear relationship between the output and the independent variables. Koenker and Bassett [36] found that instead of minimizing the squared loss to find the conditional mean, the pinball loss function could be minimized to find the conditional quantile of nominal probability τ. QR is formulated as follows [36]: Similar to linear regression, we establish a linear relation between the dependent variable $\hat{F}_\tau$ (in our case the post-processed forecasts) and the regressors x (in our case the QRF predictive distributions) such that:

$$\hat{F}_\tau = \beta x + \epsilon, \tag{5}$$

where τ is the quantile probability and ϵ is an error term that describes everything but the input-output relationship that cannot be captured by the model. From this, $\hat{\beta}$ can be estimated from the following minimization problem:

$$\hat{\beta}_\tau = \arg\min_{\beta} \sum_{i=1}^{n} \rho_\tau(y_i - \beta x_i), \tag{6}$$

where ρ_τ is an asymmetric piece wise linear function, often denoted pinball loss function, given as:

$$\rho_\tau(u) = \begin{cases} (1-\tau)u, & \text{if} u < 0, \\ \tau u, & \text{if} u \ge 0. \end{cases} \tag{7}$$

To train the QR model, the QRF test set (17,520 hours) was divided into a new training (8760 hours) and testing data set (in which the new test data set consists of a year (8760 hours) of data). From the QR model, we produce the final forecast with 19 evenly spaced nominal probabilities such that $\tau \in \{0.05,\dots,0.95\}$. An example of two days of probabilistic forecasts, with different issue-time and weather conditions for wind, PV power and aggregation of equal ratio of wind and PV are shown in Fig. 4. The entire forecasting model consisting of QRF and QR is implemented using the Python library `scikit-learn` [35].

2.3. Verification

Diagnostic tools to evaluate probabilstic forecasts are to check the *calibration* (also called *reliability*), *sharpness* and *resolution* of the forecasts. A series of probabilistic forecasts is said to be reliable when statistically, they correspond to the observations. Sharpness relates to the concentration of the probabilistic forecasts. While reliability is a joint property of the forecast and the observations, sharpness is solely a property of the forecast. Gneiting and Raftery [39] argued that the sharpness of a probabilistic forecast should be maximized, subject to the reliability. Resolution, which should be maximized, relates to the capability of the forecast model to issue forecasts that are different from the average observation.

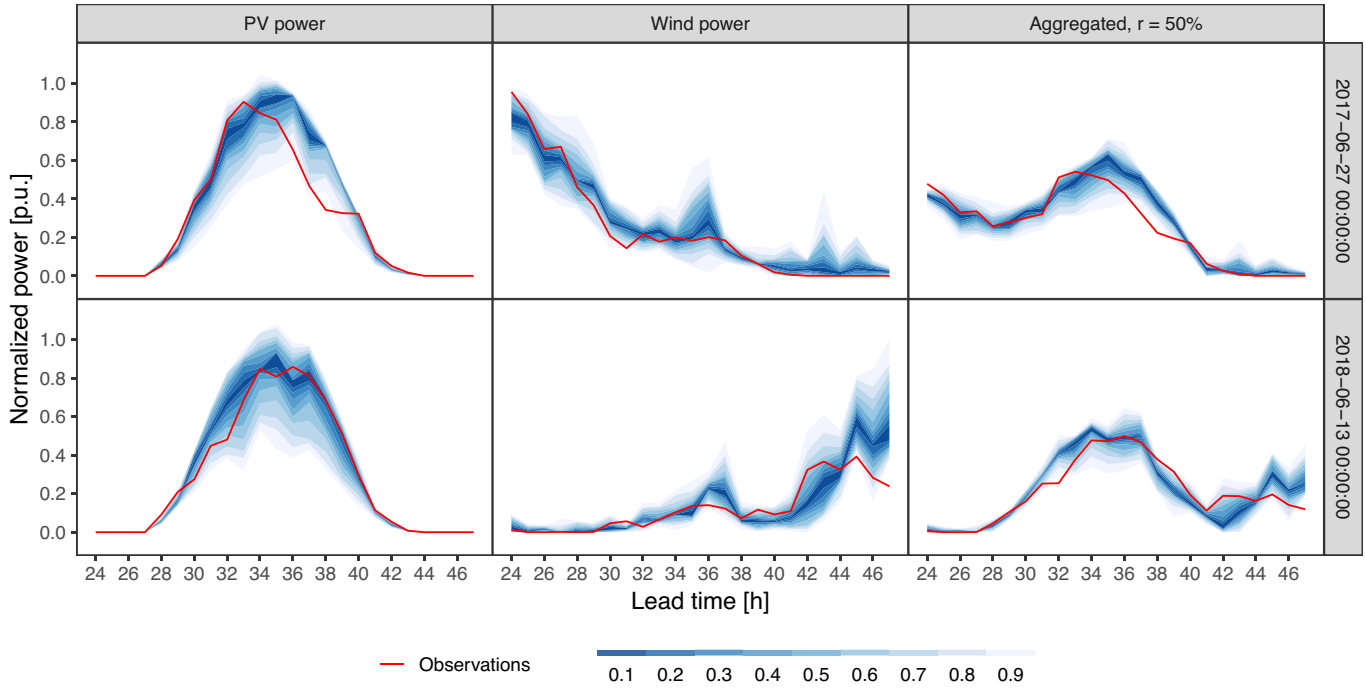

Fig. 4. Two example days of normalized PV, wind and aggregation of equal level of wind and PV power production forecasts at the co-located park. The columns represent the different power sources. The rows represent the different issue times in UTC. The prediction intervals are shown as envelopes with nominal coverage range from 10% to 90% and the observations as solid red lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

As pointed out by Lauret et al. [40]; it should be noted that the diagnostic tools used to evaluate probabilistic forecasts have different meanings from the point of view of a meteorologist and a statistician. From a statisticians viewpoint, sharpness refers to the concentration of the predictive distribution [41,42]. In meteorology, sharpness refers to the ability of a forecasting system to generate forecasts that deviates from the climatological value of the predictand [43]. From the meteorologists viewpoint, resolution measures the ability to produce predictive distributions based on the predictand (i.e., forecasts are case-dependent) while it refers to the ability to produce different density forecasts depending on the forecasts conditions (i.e., predictive distributions are not only conditioned by the values of the predictand) from a statistical point of view [42]. In this work, sharpness refers to the concentration of the prediction intervals and resolution quantifies the ability of the forecasting system to generate predictive distributions (i.e., the statisticians viewpoint). Also, as pointed out by Gneiting et al. [44], reliability can be improved by statistical post-processing (which was explained in Section 2.2.2), whereas this is not possible for resolution.

To understand potential benefits of aggregating wind and PV in terms of forecast quality, the quality appraisal will be made according to the verification framework proposed by Lauret et al. [40]. The framework consists of visual diagnostic tools and numerical scores to objectively rank the quality of the forecasts. In Section 2.3.1–Section 2.3.1, the visual verification tools are explained. In Section 2.3.3–Section 2.3.5, the numerical verification scores used in this study are presented.

2.3.1. Probability integral transform (PIT)

To assess the calibration of probabilistic forecasts, we use the probability integral transform (PIT). If F is the CDF of a calibrated probabilitic forecast for observations Y, it holds that $F(Y) \sim \mathcal{U}([0,1])$. For a larger number of forecasts and observation pairs, the calibration can be assessed empirically by plotting the histograms of the resulting PIT values. A probabilistic forecast is reliable if, statistically, the forecast probabilities agree with the observed frequencies. A flat PIT histogram is a necessary but not a sufficient condition for calibration [39]. However, due to limited amount of forecast and observation pairs in the testing set, a range of empirical probabilities could be observed despite being a reliable forecast. Considering this, Bröcker and Smith [45] proposed consistency intervals. The lower and upper bounds can be calculated by a the binomial distribution $\mathcal{B}(N, 1/(|Q|+1))$, where N is the number of time step in the testing set. $|Q|$ represents the total number of quantiles in a set of quantiles Q [46]. In this sense, a set of probabilistic forecasts are calibrated if the PIT histogram falls within the consistency interval.

2.3.2. Prediction interval average width (PIAW)

The sharpness is commonly calculated using prediction interval average width (PIAW). The sharpness of the predictive distribution is calculated by the mean size of the central prediction intervals for different nominal coverage rates $(1-\alpha)100\%$ according to:

$$PIAW(\alpha) = \frac{1}{N}\sum_{t=1}^{N}(\hat{q}_{t,\tau=1-\alpha/2} - \hat{q}_{t,\tau=\alpha/2}), \tag{8}$$

where $\hat{q}$ is the forecasted quantile at probability τ. Note that the sharpness is a property of the forecast alone and PIAW should not be normalized using observations. For this reason, the sharpness does not necessarily inform the forecaster about the quality of a forecasting method; rather, it can be used to rank calibrated forecasts.

2.3.3. Continuous ranked probability score (CRPS)

Perhaps the most common scoring rule for probabilistic forecasts is the CRPS, which evaluates sharpness and reliability simultaneously. The CRPS, is defined as [47]:

$$CRPS(F_t, y_t) = \int_{-\infty}^{\infty}\left(F_t(x) - \mathbb{1}\{y_t \le x\}\right)^2 dx, \tag{9}$$

where F_t is the predicted CDF, y_t is the observation and $\mathbb{1}$ is the Heaviside step function. Here, subscript t denotes the forecast and observation pair at time step t. The CRPS is expressed in the same unit as the forecasted variable and the formulation above is a negatively oriented score, i.e., a lower value of the CRPS indicates a better performance. The

CRPS generalizes to the mean absolute error when F_t is a deterministic forecasts which facilitates comparison between probabilistic and deterministic forecasts. The CRPS is commonly averaged over the testing set and can be decomposed into a reliability, uncertainty and resolution term as CRPS = *Reliability* + *Uncertainty* - *Resolution* [48]. We use the `crpsDecomposition` in R developed by NCAR [49] to calculate the decomposed CRPS. The uncertainty term, which only depends on the observations and furthermore a term that is used to cluster forecasts in this study, is calculated according to Lauret et al. [40] and is formulated as:

$$Uncert = \frac{\sum_{t=1}^{N}\sum_{i=1}^{t}|y_t - y_i|}{N^2}, \tag{10}$$

where N is the total number of observations. From Equation (10), it should be noted that the uncertainty term is not directly related to the difficulty in forecasting the power output. Instead, when considering the power output, the uncertainty term is a function of the variability and magnitude of the power output.

2.3.4. Interval score

Following Winkler [50] and Gneiting and Raftery [39], a proper score to assess the nominal coverage rates $(1-\alpha)100\%$ is the *interval score*. As for the CRPS, the interval score is negatively oriented, i.e., a lower score is preferable. The interval score is defined as:

$$IS(\alpha) = \frac{1}{N}\sum_{t=1}^{N}(U_t - L_t) + \frac{2}{\alpha}(L_t - y_t)\mathbb{1}\{y_t < L_t\} + \frac{2}{\alpha}(y_t - U_t)\mathbb{1}\{y_t > U_t\}, \tag{11}$$

where L_t and U_t represents the lower and upper quantile according to $\hat{q}_{t,\tau=\alpha/2}$ and $\hat{q}_{t,\tau=1-\alpha/2}$, respectively. From Equation (11), a forecast is rewarded for a narrow prediction interval while penalized if the observation, y_t is outside the interval. The penalty is dependent on α.

2.3.5. Quantile score

By calculating the quantile score, information about the forecast quality at specific probability levels is obtained. As noted in Bentzien and Friederichs [51], CRPS is calculated over the complete CDF through integration of the Brier score. Therefore, deficiencies in different parts of the distributions might be hidden, e.g., the tails of the distribution. The average quantile score (QS), with N pairs of observations y_t and quantile forecast $\hat{q}_{t,\tau}$ at probability τ is given as:

$$QS(\tau) = \frac{1}{N}\sum_{t=1}^{N}\rho_\tau(y_t - \hat{q}_{t,\tau}), \tag{12}$$

where ρ_τ is the pinball loss function given by Equation (7).

2.4. Trading of stochastic production

As elaborated in Section 1, co-located parks could be seen as single entities that are connected to the electricity network at a point of common coupling. This means that the parks have to provide aggregated bids to electricity markets, where accurate forecasts of expected production is imperative. For the time periods analyzed in this study, the electricity pool is a dual-price market where electricity exchanges take place in two stages; the day-ahead and the balancing (also called regulating) market. In the first stage, the day-ahead market, power producers and consumers place bids for the next delivery period which is the next day. At gate closure, which occurs 12 hours before delivery, energy offers based on forecasts with horizons $t+k$, $k \in \{13,\dots,36\}$ should be proposed at hourly resolution. The market clearing yields a spot price π^c_{t+k} after matching production offers and consumption bids and a set of energy blocks y^c_{t+k} to be delivered by the power producer at every forecast horizon. Here, c defines a contract between y^c_{t+k} and π^c_{t+k} which means that the power producer is financially responsible for deviations from y^c_{t+k}. In the second stage, the balancing market, deviations from the contract are translated into financial penalties. For a positive deviation, the power producer has to sell on the balancing market for the cost of activating downward regulation services π^s_{t+k} and buy in the case of negative deviation for the cost of activating upward regulation services π^b_{t+k}. The revenue for a power producer operating in these markets can be formulated as:

$$R_{t+k} = \pi^c_{t+k}y_{t+k} + B_{t+k}(y_{t+k}, y^c_{t+k}), \tag{13}$$

where the imbalance term is defined as:

$$B_{t+k}(y_{t+k}, y^c_{t+k}) = \begin{cases} \pi^{\downarrow}_{t+k}(y_{t+k} - y^c_{t+k}), & y_{t+k} - y^c_{t+k} \le 0 \\ -\pi^{\uparrow}_{t+k}(y_{t+k} - y^c_{t+k}), & y_{t+k} - y^c_{t+k} > 0, \end{cases} \tag{14}$$

where y_{t+k} is the actual production, $\pi^{\downarrow}_{t+k}$ and $\pi^{\uparrow}_{t+k}$ are referred to as the regulation unit costs for downward and upward balancing costs, respectively. These are defined as follow:

$$\pi^{\downarrow}_{t+k} = \pi^c_{t+k} - \pi^s_{t+k} \tag{15}$$

$$\pi^{\uparrow}_{t+k} = \pi^b_{t+k} - \pi^c_{t+k} \tag{16}$$

This means that the imbalance term, $B_{t+k}(Y_{t+k}, y^c_{t+k})$, is formulated as a penalty (except for perhaps some unusual occasions). Pinson et al. [42] showed that the energy offer, i.e., the decision variable, only appears in the balancing market term and that the solution that maximize the revenue in the stochastic optimization problem is the optimal offer:

$$y^*_{t+k} = \hat{F}^{-1}_{t+k}\left(\frac{\pi^{\downarrow}_{t+k}}{\pi^{\uparrow}_{t+k} + \pi^{\downarrow}_{t+k}}\right), \tag{17}$$

where $\hat{F}_{t+k}$ is the predicted CDF, in this case the forecasted power production. This means that the optimal production offer corresponds to a specific quantile that is a direct function of the regulation unit costs. For derivation of the expression, the reader is referred to Bremnes [52]. We use perfect forecasts of spot prices and regulating unit costs, which means that we do not take into account any spatial or temporal dependencies affecting these prices. However, the reader should note that the regulating prices generally are difficult to estimate with hourly resolution [53]. An overview of the methodology and the data used in this study is presented in Fig. 1.

3. Results

Here results from various ratios of aggregated wind and PV are presented. The visual inspection of the forecasts are presented in Section 3.1 and the numerical verification scores are presented in Section 3.2. In Section 3.3, the results from the electricity market simulation are presented.

3.1. Visual verification tools

3.1.1. Calibration assessment

PIT histograms, averaged over all forecast horizons, for the respective power sources as well as aggregated output with equal nominal capacity are presented in Fig. 5 for the forecast and observation pairs in shuffle 2. Note that the other ratios and shuffles showed similar results which is why they are not included. Deviations from a uniform histogram (red lines in Fig. 5) occur due to limited amount of forecast and observation pairs in the testing set and show the range of empirical probabilities (dashed black lines in Fig. 5) that could be observed despite being a reliable forecast. In this case, the consistency intervals correspond to the nominal probabilities τ, i.e., the 5–95% probability interval.

Fig. 5 (upper row) shows that the forecasts generated from solely the QRF models are under-dispersive which is to say that the predictive distributions are too narrow on average. After post-processing the QRF forecasts with QR, Fig. 5 (lower row) shows that the probabilistic forecasts are indeed better calibrated, although not perfectly. After

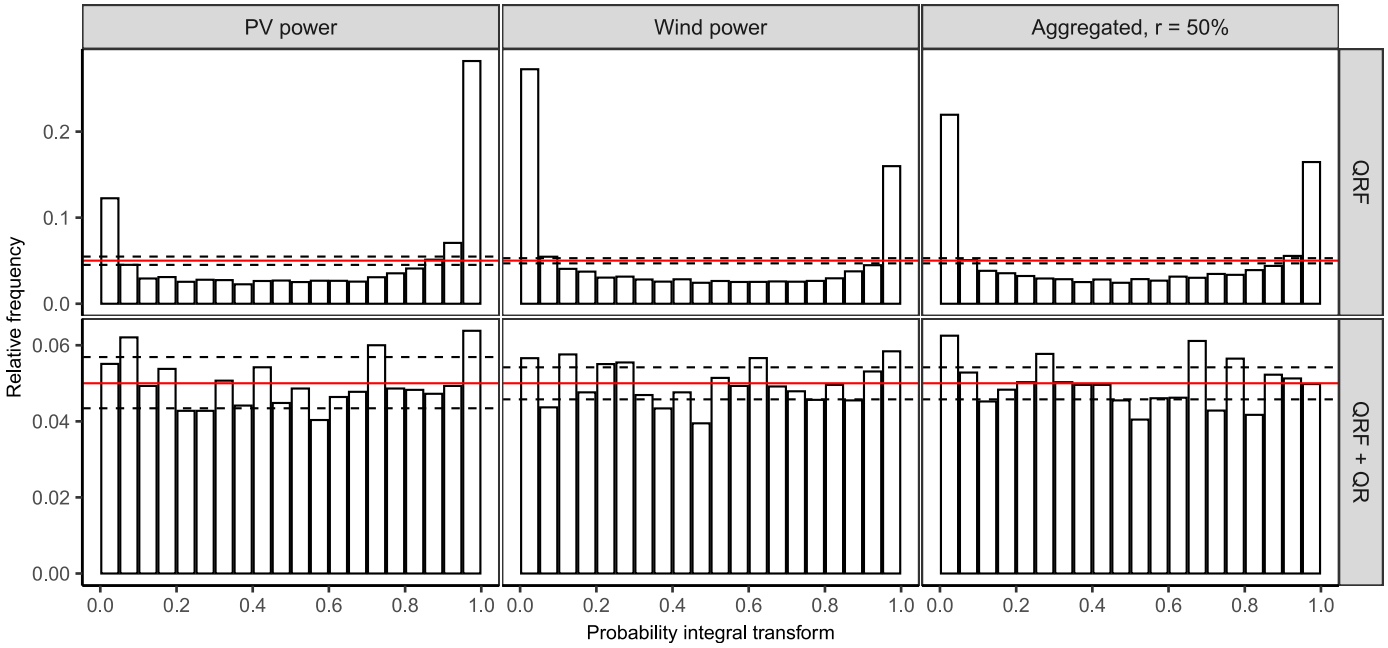

Fig. 5. PIT histograms for all look-ahead times for PV power, wind power and 50% aggregation level at the studied site for shuffle 2. Consistency intervals are denoted by the horizontal dashed lines and the red lines show a perfectly reliable forecast. Note the differences on the y-axes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

post-processing the QRF forecasts, the consistency band are wider because the size of the testing data set is halved. Also, the consistency intervals for PV forecasts are wider since these are limited to solar elevation angles of 5° or more. Most histogram bars remain within the margin that could be caused by randomness from which we conclude that the calibration is satisfactory for the purpose of the study since we cannot find any consistent miscalibration. To that end, we continue now only with forecasts generated from QRF and post-processed with QR.

3.1.2. Sharpness

Recall from Section 2.3 that sharpness is a property of the forecast only, i.e., observations are excluded. Therefore, a forecast system can produce sharp forecasts, despite being useless if they are not reliable. Fig. 6 shows the PIAW (which represents the sharpness) of probabilistic forecasts with $\alpha \in \{0.2, 0.3, \dots, 0.8\}$ generated for a park with ratios $r \in \{0, 0.2, \dots, 1\}$ for the different seasons. The reason for solely plotting the ratios in steps of 20% is to reduce clutter. For the complete test data set ("All year") and during the summer, any aggregated ratio of co-located production of more than 20% improve the sharpness of the forecasts compared to individual parks. However, during the fall in shuffle 3 the improvement is realized with a ratio of more than 60%. During the winter in shuffle 1 and 2, the improvement is realized with a ratio of more than 40% and solely 20%, respectively. No improvement over an individual PV park is realized in shuffle 2 during spring.

3.2. Numerical verification scores

3.2.1. Continuous ranked probability score (CRPS)

Fig. 7 shows the CRPS and its decomposition into reliability, resolution and uncertainty as well as the variance for different aggregated ratios of wind and PV in the co-located park calculated over the whole test set ("All year") and the different seasons for the shuffled data sets. 0% indicates a park with solely PV power and 100% indicates a park with solely wind power. For all seasons and shuffles, the CRPS is lowered when issuing probabilistic forecasts of aggregated production except for the spring in shuffle 2 where the lowest CRPS is for a wind park and in shuffle 1 during the winter when the lowest CRPS is for an individual PV park. In all cases, the CRPS is lowered when issuing probabilistic forecasts of aggregated production where the optimal mix is found at a 50% – 60% ratio considering all data as well as in the summer, fall and spring. This is explained by the aggregated power output from a co-located park being less variable, which is shown by the lowered variance.

The resolution terms, which should be maximized, are lowered when aggregating the data, indicating that the ability of the forecast model to produce different forecasts depending on the forecast conditions are becoming worse. At the same time, the uncertainty term, which is an artifact of the observation and should be minimized, is also decreased. The lowered uncertainty term and the reduction in variance is explained by the smoothing effect, where the average variability of the power output is lowered when the power output of the two sources are combined. Predicting a smoother signal means that the forecast model needs to capture less variability, which results in lower forecast resolution. Recall from Section 2.3 that the sharpness should be maximized subject the to reliability according to Gneiting and Raftery [39]. When the time series are aggregated, Fig. 6 and 7 show that the sharpness as well as the reliability are, in most cases, improved. This indicate that co-locating wind and solar has the effect of producing sharper and more reliable probabilistic forecasts as compared to generating forecasts from individual power parks.

3.2.2. Variability in power output

The uncertainty term presented in Fig. 7 differ between seasons, e.g., the highest uncertainty term in PV power production is in the spring and summer whereas the highest uncertainty term for wind power production is during the winter. The uncertainty term is determined by the observed power output according to Equation (10), which means that the term depicts the variability in the power output. In order to understand the effect of aggregation at different levels of variability, we calculate the uncertainty term of wind power on a daily basis and cluster these values. The reason for not calculating the daily uncertainty terms of PV power is the strong seasonality of the insolation. Considering this, the highest values of the daily uncertainty terms (i.e., daily variability) in

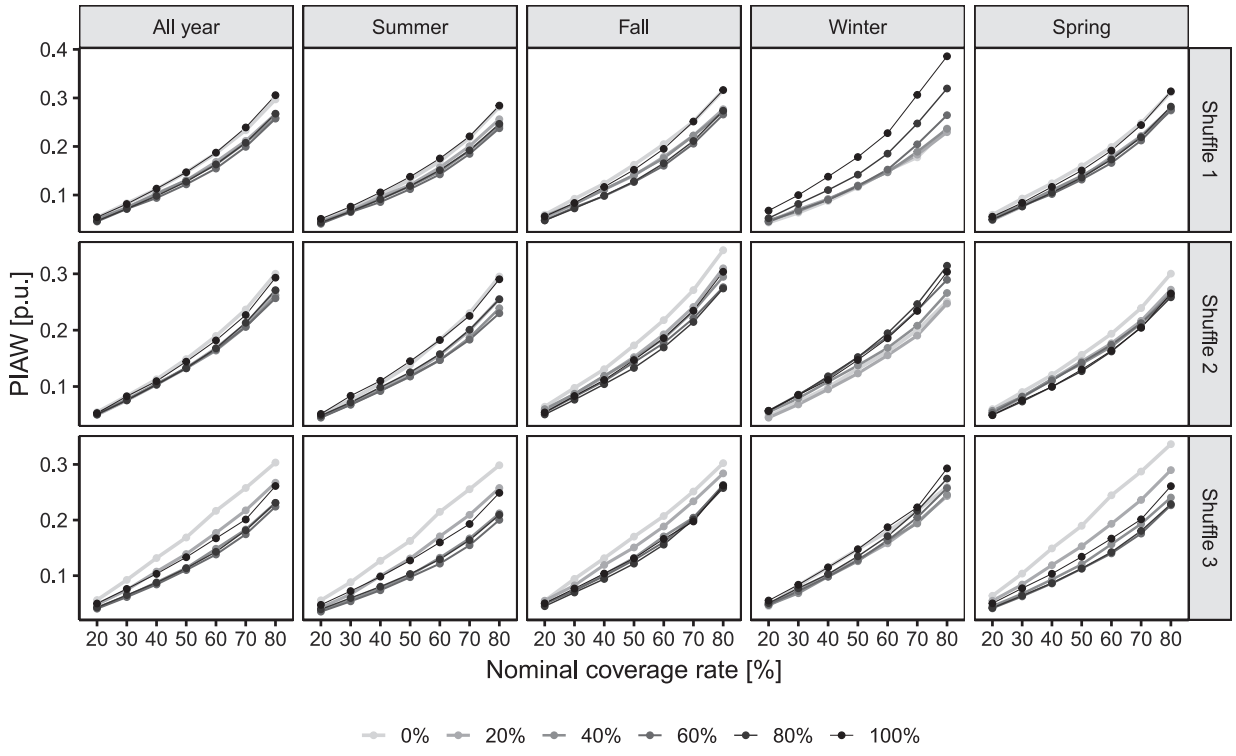

Fig. 6. PIAW as a function of the nominal coverage rate of the prediction intervals for different nominal capacity of wind power to the nominal capacity of the entire park. The rows corresponds to the shuffles and the columns to the seasons.

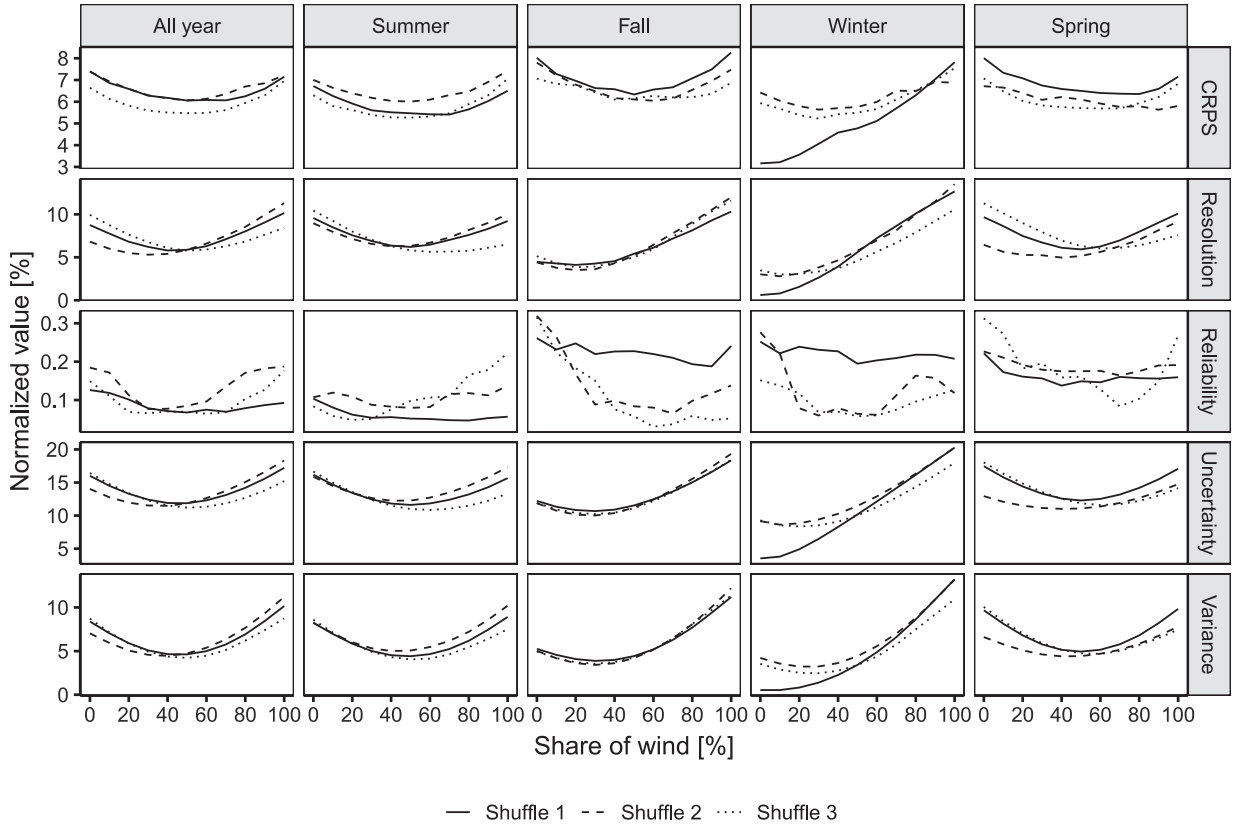

Fig. 7. The rows present the CRPS and corresponding decomposition scores as well as the variance for different shares of wind in the combined park for the different seasons and shuffled data sets, respectively. The columns corresponds to the seasons.

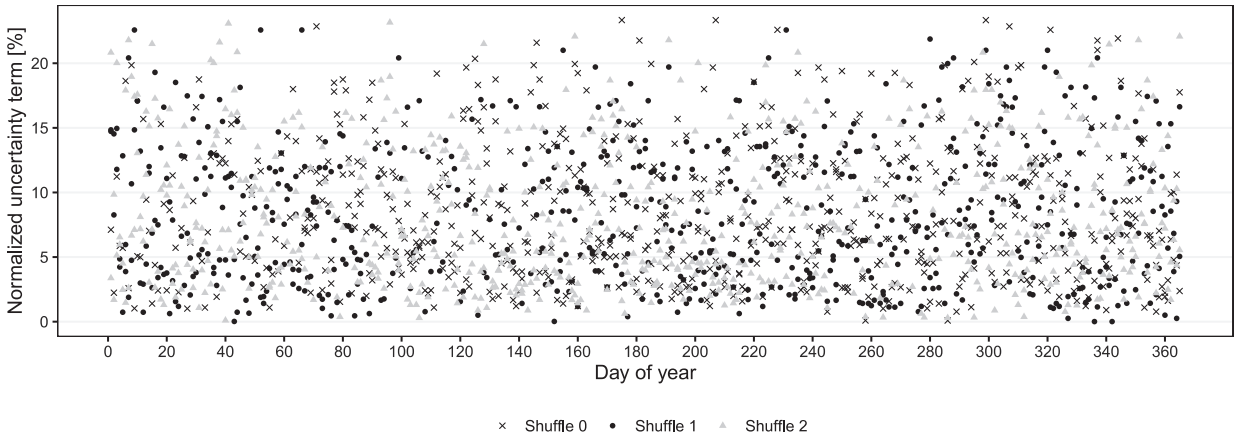

Fig. 8. Uncertainty terms calculated on a daily basis for wind power and normalized by the nominal capacity of the park for the shuffled datasets.

PV power is found in the spring, which was already assessed when dividing into seasons in Fig. 7. Fig. 8 shows the normalized daily wind power uncertainty terms for each day of the year for the shuffled data sets. The calculated daily uncertainty terms are relatively homogeneously spread throughout the year where values are in between 0–25% uncertainty level. The grid lines on the y-axis corresponds to the different uncertainty term clusters that are used for the analyis, i.e., 0–5%, 5–10%, 10–15%, 15–20% and 20–25%.

In Table 1, the decomposed CRPS for different uncertainty term clusters as well as different aggregated ratios of wind and PV is shown. During times when the daily wind power variability is the lowest (i.e., uncertainty term cluster 0–5%), the CRPS of a park consisting of 90% wind power is the lowest. As the daily variability in wind power is increasing, the CRPS is also increasing for a pure wind park, with the exception of the uncertainty term cluster 20–25%. From visual inspection (not shown in this paper), this uncertainty term cluster was restricted to days where the wind power production went from a small power output to rated power output within the course of the day. The non-increasing CRPS for this uncertainty term cluster is likely explained by the power curve, from which wind power output is approximately determined. Above the rated and below the cut-out wind speed electric power is constant, which is relatively straightforward to forecast and results in a lower CRPS score for this uncertainty term cluster. By assuming that the forecasts are statistically independent, one may calculate the standard error to gain information about the confidence in the result. Considering the 95% confidence limit of the standard errors for the calculated CRPS values (parenthesis values in Table 1), aggregation decrease the standard error. This means that not only are the forecasts accuracy improved from aggregation on average, but the variability in the forecast errors are also lowered. Furthermore, the differences between the calculated CRPS values for the different ratios are much larger than the calculated standard errors. We can thus conclude that the improved forecast accuracy is statistically significant.

Forecasting aggregated time series has the effect of dampening the combined uncertainty term, i.e., reducing the variability of the power production. Aggregation furthermore, similar to what is illustrated in Fig. 7 reduce the reliability term, which combined with the other decomposition terms result in a lower overall CRPS score. Another interesting aspect is that the CRPS for an individual PV power park is relatively stable (and slightly decreasing) as the wind power variability is increasing. This is explained by the fact that the variability in the power sources are anti-correlated, i.e., less solar power is produced during windy conditions while wind power production is lower during sunny conditions. This indicates that during periods when one of the power sources are better predicted (low CRPS), the other one is predicted slighly worse (higher CRPS).

3.2.3. Interval score

Similar to the CRPS score shown in Fig. 7, the interval score is a negatively oriented score. Table 2 shows the interval score for the 20% – 80% central prediction interval for different ratios of wind to the nominal capacity of the entire park. Considering the whole test data set as well as the summer, fall and spring; the lowest interval score of the 20%–80% central prediction interval is found at around 50% – 60% wind, except for shuffle 1 in the spring, where the sizing that lowers the 60% – 80% central prediction interval the most is a ratio of 70% wind power to the combined nominal capacity of the park. For the same central prediction intervals, a mix of 10% – 30% wind power to the combined park yields the lowest scores during the winter.

3.2.4. Quantile score

Fig. 9 shows the negatively oriented quantile score as a function of the probability levels $\tau \in \{0.1, 0.15, \dots, 0.9\}$ with $r \in \{0, 0.2, \dots, 1\}$ during the different seasons. Again, as in Fig. 6, the reason for solely plotting these ratios was to reduce clutter because results were found to follow the same trend. When forecasting aggregated production, the forecasts are improved by means of the quantile score compared to individual parks. However, in shuffle 3 during fall the improvement is found after the 40% aggregation level, all ratios except for 20% in shuffle 1 and 2 in the spring and all ratios except for 80% in shuffle 2 and 3 in the winter. In shuffle 1 in the winter, the quantile scores were always higher for an aggregated park as compared to a pure PV park. On top of an improved quantile score, aggregated production improves the symmetry across the range of quantiles compared to the forecasting of the individual power sources. This means that forecasts of the highest and lowest quantiles using aggregated time series are better estimated compared to forecasts of an individual wind or solar park.

3.3. Electricity trading

Based on the probabilistic forecasts of electricity production and the optimal trading strategy in the electricity market presented in Section 2.4, this section evaluates results from electricity market simulation. In Table 3, results with different ratios of wind and PV power are presented for the shuffled datasets. In order to make the result from trading comparable between the different ratios, the parks are scaled such that all ratios produce the same amount of energy when using perfect forecasts (20.0 GWh). The term *part of imbalance* is defined as the sum of deviations divided by the sum of the produced energy. The *performance ratio* is calculated by normalizing the total revenue from the forecasts by the revenue that would be obtained with perfect forecasts, which is a metric that can be used to evaluate rivaling bidding strategies in the electricity market [42]. Using perfect predictions, the total revenue and

Table 1
The CRPS (mean ± standard error with 95% confidence limits) and corresponding decomposition into reliability ("Rel"), resolution ("Res") and uncertainty ("Uncert") for different ratios of nominal wind power to the nominal power of the park and different binned uncertainty term levels for all shuffled datasets. The parenthesis in the column headers show the number of days with the given uncertainty term level. 0% indicates a park solely PV and 100% indicates a park with solely wind power. The values are normalized by the nominal power meaning that the unit is %. The lowest CRPS values and corresponding decomposition terms in each uncertainty term cluster are underlined and in bold font.

	0–5% (359 days)				5–10% (343 days)				10–15% (247 days)				15–20% (117 days)				20–25% (29 days)			
	CRPS	Res	Rel	Uncert	CRPS	Res	Rel	Uncert	CRPS	Res	Rel	Uncert	CRPS	Res	Rel	Uncert	CRPS	Res	Rel	Uncert
0%	6.49 (±0.099)	9.93	0.07	16.35	6.45 (±0.103)	9.38	0.08	15.76	6.43 (±0.123)	7.9s3	0.12	14.24	6.42 (±0.206)	6.13	0.23	12.32	5.12 (±0.276)	8.39	0.12	13.39
10%	6.06 (±0.091)	8.76	0.07	14.75	5.90 (±0.094)	8.40	0.07	14.24	6.06 (±0.110)	6.97	0.09	12.94	6.30 (±0.192)	5.16	0.20	11.26	**5.01 ± (0.226)**	**7.69**	**0.08**	**12.62**
20%	5.65 (±0.083)	7.85	0.06	13.44	5.65 (±0.085)	7.38	0.06	12.97	**6.01±0.096)**	**6.07**	**0.08**	**12.00**	6.34 (±0.173)	4.48	0.18	10.64	5.17 (±0.234)	7.23	0.08	12.32
30%	5.20 (±0.072)	7.23	0.05	12.38	5.42 (±0.075)	6.55	0.05	11.92	6.13 (±0.100)	5.36	0.10	11.39	**6.26±0.152)**	**4.26**	**0.14**	**10.38**	5.60 (±0.246)	6.89	0.13	12.36
40%	4.88 (±0.065)	6.81	0.05	11.63	5.28 (±0.069)	5.90	0.05	11.13	6.48 (±0.101)	4.76	0.13	11.11	6.49 (±0.155)	4.17	0.16	10.50	6.14 (±0.291)	6.75	0.18	12.71
50%	4.52 (±0.056)	6.79	0.04	11.27	**5.25±0.066)**	**5.47**	**0.05**	**10.67**	6.83 (±0.107)	4.53	0.15	11.21	6.76 (±0.158)	4.43	0.18	11.00	6.38 (±0.305)	7.09	0.12	13.34
60%	4.22 (±0.051)	7.02	0.05	11.19	5.32 (±0.066)	5.33	0.05	10.60	7.32 (±0.117)	4.60	0.19	11.72	7.19 (±0.168)	4.88	0.24	11.83	7.00 (±0.349)	7.48	0.17	14.30
70%	4.01 (±0.048)	7.27	0.06	11.23	5.48 (±0.070)	5.46	0.05	10.89	7.87 (±0.130)	4.99	0.24	12.62	7.63 (±0.190)	5.58	0.29	12.92	7.26 (±0.382)	8.42	0.18	15.51
80%	3.83 (±0.048)	7.66	0.08	11.40	5.84 (±0.079)	5.74	0.07	11.51	8.66 (±0.146)	5.47	0.31	13.82	8.45 (±0.221)	6.12	0.35	14.22	7.92 (±0.438)	9.23	0.24	16.91
90%	**3.76±0.052)**	**8.19**	**0.11**	**11.83**	6.16 (±0.084)	6.38	0.08	12.46	9.28 (±0.159)	6.30	0.32	15.26	9.34 (±0.248)	6.81	0.45	15.70	8.28 (±0.450)	10.40	0.22	18.46
100%	3.95 (±0.057)	8.88	0.15	12.67	6.70 (±0.092)	7.10	0.09	13.71	10.16 (±0.175)	7.09	0.36	16.90	10.46 (±0.278)	7.42	0.54	17.34	8.98 (±0.512)	11.45	0.25	20.17

average energy price indicate that the spot prices when the PV park produces its energy are higher than during the times that the wind park produces energy, meaning that the potential revenue is higher for a PV park than a wind power park. This is explained by the penetration levels of PV and wind power in the Swedish power systems. In 2017, 2018, 2019 and 2020, PV produced 0.1%, 0.3%, 0.4% and 0.6% of net electricity production and wind power produced 11.0%, 10.4%, 12.0% and 17.1% of net electricity production [54]. Due to the low amount of PV to the total net production, the correlation of PV power production, as compared to wind power production, to the existing power production across the power system is lower which results in a higher value factor of PV [55]. For Shuffle 1 and 2, PV power forecasts performs worse than the wind power forecasts, yielding a lower performance ratio, while the opposite is true for Shuffle 3. However, when forecasting aggregated time series, the performance ratio is higher than for a park consisting of solely wind or PV. This indicate that aggregating forecasts at a co-located wind and PV power park translates into improved performance when trading electricity. When the aggregated time series are forecasted, deviations from the contracted energy production, i.e., surplus and shortage, is reduced causing the regulation costs to be lowered, which improves the performance ratio and lowers the imbalance. The optimal ratio is, similar to what was found in Section 3.1 and Section 3.2 found at around 40% – 60% wind power to the combined nominal capacity of the park. This is explained by the forecasts of aggregated time series being improved by means of accuracy, reliability and sharpness compared to forecasts of individual power sources which translates into improved trading in the day-ahead market.

4. Discussion

The scientific literature has motivated co-location of wind and PV park in several aspects, e.g., reduced project development costs, synergies in operation and maintenance, more effective land usage and better use of the electrical infrastructure [9]. Despite these foreseen benefits, the scientific literature on co-located wind and PV power parks is limited, in particular related to forecasting. Therefore, this study presents a first study on the effect of aggregating the power sources on probabilistic forecast quality and value and what possible synergies might be found when forecasting different ratios of co-located wind and PV under a fixed modelling framework.

An important and challenging aspect in probabilistic forecasting relates to calibration, which means that the forecasts should, statistically, correspond to the observations. Calibration is furthermore important in decision making processes, such as electricity market trading, in order to produce accurate approximation of the underlying forecast uncertainty. Trading using unreliable probabilistic forecasts will likely lead to biases. To produce calibrated probabilistic forecasts of wind and PV power, we use QRF post-processed with QR using inputs from the regional NWP model MetCoOp. Although the main contribution of this paper is not to produce the most accurate probabilistic forecasts, but rather analyze the effect of aggregation within the chosen model framework itself, possible improvements would be to test and compare additional statistical and physical forecasting models. Furthermore, future research should also include more explanatory variables or include ensemble members. Ensemble members were excluded because the configuration of the ensemble was changed several times within the studied time period. In this study, the forecast models are trained for each lead time which corresponds to the specific phase in the diurnal cycle that the NWP forecast pertains to. However, antecedent time steps (due to non-zero autocovariance) are interesting to consider as explanatory variables in future studies especially if the temporal dependence structure is of importance, e.g., if the co-located park includes storage. For this, the marginal probabilities could be transformed to scenarios that could be used in stochastic optimization.

When studying the effect of aggregating co-located wind and PV power, the results indicated that a substantial gain could be realized dur-

Table 2
The interval score ("IS") with varying prediction intervals for different ratios of nominal wind power to the nominal power of the complete park. 0% indicates a PV park and 100% indicates wind power park. The values are normalized by the nominal power meaning that the unit is %. The lowest interval score for each central prediction interval and time period are underlined and in bold font.

Shuffle 1.

	All year				Summer				Fall				Winter				Spring			
	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$
0%	10.38	11.17	12.89	16.55	9.27	10.01	11.78	15.54	11.43	12.33	14.40	18.05	7.47	8.51	10.47	**14.10**	11.03	11.98	14.03	17.98
10%	9.63	10.36	11.86	15.22	8.65	9.37	10.86	14.14	10.85	11.69	13.64	17.53	**7.40**	**8.32**	**10.35**	14.19	10.18	11.11	13.03	16.91
20%	9.25	10.03	11.57	14.85	8.28	8.86	10.26	13.40	10.08	10.97	12.76	16.45	7.61	8.73	10.52	14.60	9.60	10.51	12.24	15.64
30%	8.90	9.68	11.22	14.42	7.85	8.50	9.92	13.30	9.88	10.65	12.46	16.41	7.72	8.64	10.63	14.80	9.49	10.12	11.61	15.21
40%	8.76	9.52	10.98	14.23	7.64	8.26	9.62	12.84	9.46	10.33	12.05	15.52	8.05	8.87	10.93	15.28	8.92	9.76	11.46	14.90
50%	**8.54**	9.31	10.91	**14.09**	**7.45**	**8.13**	**9.55**	**12.62**	**9.42**	**10.17**	**11.82**	15.26	8.15	9.38	11.65	15.56	8.72	9.61	11.35	14.69
60%	8.55	**9.21**	**10.63**	14.18	7.62	8.35	9.70	12.91	9.54	10.21	**11.82**	**15.02**	8.45	9.53	11.89	16.16	**8.65**	**9.57**	**11.24**	14.96
70%	8.49	9.22	10.83	14.21	7.65	8.37	9.94	13.03	9.76	10.41	11.86	14.88	9.00	9.89	11.82	16.31	8.76	9.60	**11.24**	**14.73**
80%	8.69	9.48	11.09	14.84	7.88	8.62	10.27	13.52	9.92	10.62	12.29	15.71	9.61	10.58	12.96	17.39	9.03	9.74	11.49	15.17
90%	9.19	9.95	11.63	15.85	8.34	9.17	10.76	14.21	10.43	11.08	12.76	16.40	10.40	11.78	14.30	19.37	9.25	10.07	11.96	15.90
100%	9.94	10.86	12.72	16.92	8.85	9.68	11.40	14.67	11.07	11.83	13.78	17.56	11.92	13.53	16.53	21.40	10.27	11.23	13.23	16.81

Shuffle 2.

	All year				Summer				Fall				Winter				Spring			
	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$
0%	10.52	11.30	12.94	16.47	10.02	10.71	12.35	15.95	11.25	12.20	14.15	18.32	8.96	9.78	11.23	14.14	13.11	13.77	14.97	17.75
10%	9.79	10.58	12.07	15.52	9.44	10.11	11.54	15.02	10.30	11.44	13.25	17.19	8.41	9.05	**10.37**	**13.52**	12.08	12.71	14.01	16.73
20%	9.29	10.06	11.55	14.94	9.05	9.72	11.00	14.14	9.65	10.63	12.58	16.66	8.09	8.90	10.43	13.69	11.20	11.82	12.99	15.94
30%	8.96	9.74	11.24	14.52	8.83	9.45	10.69	13.66	9.36	10.32	12.12	15.96	**8.04**	**8.81**	10.53	13.86	9.98	10.87	12.39	15.45
40%	8.75	9.55	10.99	14.39	8.66	9.26	**10.42**	13.41	8.89	9.90	11.61	15.73	8.06	9.02	10.81	14.45	9.81	10.64	11.97	14.96
50%	**8.58**	**9.32**	**10.91**	**14.07**	**8.60**	**9.17**	**10.42**	**13.11**	8.75	9.63	11.50	15.09	8.16	9.24	11.20	14.80	8.75	**9.29**	**10.94**	14.25
60%	8.59	9.45	11.06	14.27	8.69	9.39	10.63	13.27	**8.55**	**9.48**	**11.33**	**15.05**	8.46	9.53	11.68	15.57	**8.50**	9.44	11.08	**14.21**
70%	8.87	9.77	11.46	14.82	9.00	9.70	11.04	14.04	8.61	9.49	11.37	15.06	9.15	10.50	12.76	16.72	8.52	9.52	11.18	14.28
80%	9.29	10.09	11.67	15.26	9.26	9.98	11.29	14.58	9.04	9.87	11.59	15.38	10.18	11.21	13.26	17.50	8.60	9.31	10.82	14.09
90%	9.43	10.28	12.13	15.85	9.75	10.55	12.23	15.88	9.55	10.47	12.44	16.15	9.51	10.43	12.57	16.67	7.88	8.63	10.38	13.78
100%	9.95	10.80	12.65	16.50	10.44	11.27	13.05	16.65	10.24	11.13	13.04	17.10	9.53	10.44	12.48	16.79	8.15	8.86	10.54	14.22

Shuffle 3.

	All year				Summer				Fall				Winter				Spring			
	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$	$IS_{0.2}$	$IS_{0.4}$	$IS_{0.6}$	$IS_{0.8}$
0%	9.35	10.74	13.23	16.36	8.86	10.31	12.89	15.96	11.24	12.37	14.3	17.09	8.51	9.47	11.14	14.07	9.93	11.55	14.58	18.04
10%	8.74	9.84	12.07	15.36	8.19	9.34	11.73	14.76	10.47	11.39	12.97	16.2	8.33	9.1	10.56	13.83	9.2	10.5	13.14	16.83
20%	8.23	9.21	11.15	14.41	7.84	8.84	10.83	13.9	9.57	10.63	12.32	15.61	7.84	8.49	**9.95**	**13.07**	8.59	9.69	11.93	15.53
30%	7.9	8.67	10.28	13.68	7.51	8.23	9.72	13.01	9.23	9.98	11.45	14.63	**7.55**	**8.35**	10.13	13.41	8.22	9.06	10.8	14.51
40%	7.75	8.4	9.93	12.73	7.37	7.9	9.25	11.81	8.51	9.18	10.89	13.94	7.63	8.53	10.35	13.36	8.15	8.8	10.37	13.29
50%	7.67	8.27	**9.53**	**12.43**	**7.3**	**7.75**	**8.76**	**11.42**	**8.09**	**8.88**	**10.4**	**13.62**	7.77	8.63	10.25	13.48	8.02	8.61	9.93	12.9
60%	**7.65**	**8.25**	9.62	12.53	7.33	7.83	8.97	11.5	8.19	9.06	10.68	14.55	7.9	8.78	10.6	14.07	8	**8.51**	9.85	12.69
70%	7.79	8.42	9.74	12.66	7.48	8.05	9.16	11.8	8.22	8.97	10.69	14.33	8.51	9.27	11.03	14.45	**7.94**	8.54	**9.78**	**12.58**
80%	8.21	8.87	10.17	13.16	8.04	8.66	9.73	12.27	8.28	9.02	10.68	14.69	9.05	9.83	11.55	15.39	8.26	8.88	10.16	12.96
90%	8.65	9.33	10.76	13.78	8.59	9.16	10.45	13.15	8.32	9.25	10.91	14.48	9.57	10.35	12.1	15.78	8.6	9.31	10.67	13.6
100%	9.56	10.36	11.87	15.09	9.59	10.34	11.72	14.73	9.02	9.8	11.39	14.86	10.39	11.32	13.12	16.8	9.44	10.25	11.74	14.9

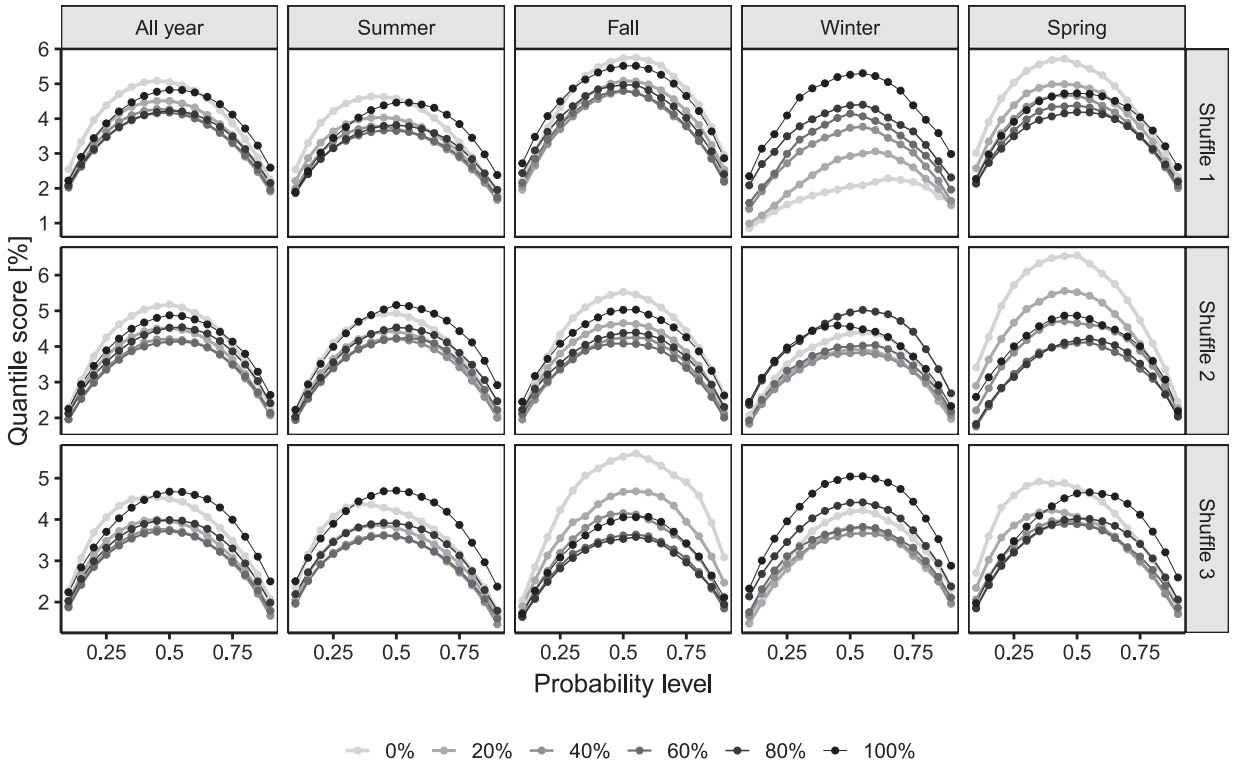

Fig. 9. Quantile score as a function of the probability level normalized by the nominal power. The columns show the different seasons and the rows show the shuffled data sets.

ing all seasons in terms of the CRPS, the interval score and the quantile score. The improvement was the lowest during the winter, which corresponds to a time period of low PV power output. On top of this, we found that the variability in the respective power source to be anti-correlated, which carry over to slightly anti-correlated CRPS values. These are interesting and useful results since the complementary characteristics in the power output variability carries over to complementary characteristics in terms of forecast performance. In fact, the output of wind and PV power production is generally anti-correlated [7].

In terms of the CRPS, we found that the variability in terms of the variance is lowered from aggregation, which translates into overall improved forecasts. As was shown by Lindberg et al. [56], the variability of a co-located wind and PV park is lowered both because of lack of correlation and negative correlation. Here it is shown that the lower variability indeed carries over to more accurate forecasts, at least for a majority of the time. To improve the confidence in the results, the datasets were shuffled, trained and evaluated over three separate years. Although the average production varied slightly between the different time periods, results showed that the lowered variability from co-location carries over to more accurate forecasts in all of the shuffled time periods. Apart from the general trend is that aggregation improves the forecast accuracy, the results also showed that winter sticks out, where a park with mostly PV power produce the lowest forecast errors. This is explained by the latitude location of the park and corresponding low amount in solar irradiation in the winter. On top of more accurate forecasts when aggregating the time series, results also showed that the standard errors are lowered, which means that the variations in the forecast errors is lowered. Although the reliability and sharpness should be improved subject to each other according to Gneiting and Raftery [39], the results showed that these characteristics both improved by means of aggregating co-located wind and PV. However, this comes at the cost of lowering the resolution of the forecasts. A lower resolution means that the ability to produce different density forecasts depending on the forecasts condition is reduced, which is explained by the variability of the aggregated time series being lowered.

In Petropoulos et al. [57], different types of uncertainties related to forecasting are brought up. According to the decomposition of the CRPS [48], the uncertainty term is a function of the variability in the observations. However, in a broader perspective, uncertainties might relate to, e.g., parametrizations or phase error in the NWP model, which carry over to power production. Considering that such uncertainties can be isolated, a future research direction would be to further extend the study, e.g., by estimating what uncertainties that relate to the physical NWP model and which uncertainties relate to the statistical model.

One aspect of the goodness of a forecast is whether or not it results in an increment economic value, according to Murphy [27]. Our results showed that aggregating co-located wind and PV power forecasts has the effect of improving the performance in the electricity market by reducing the deviations from the contracted energy sales. These deviations translate into a lower ratio between the regulation costs and the incomes. We assumed perfect forecasts of the spot price as well as the regulation costs. However, electricity prices are, like wind and PV power, generally seen as stochastic variables and requires forecast strategies which we left out of this study. On top of this, the electricity prices and power production are not necessarily independent variables, especially with large shares of wind power [58], which was also shown in this study. Although not specifically applied in this study, improved forecasts will likely be useful in other type of markets, such as capacity markets [10] and for ancillary services [11], for which co-located wind and PV have been proposed as future candidates. This study could also be extended and applied to those markets in order to study other aspects of potential forecast value from co-location of wind and PV power.

While the number of co-located wind and PV power parks are limited to a few on a global scale, it remains an interesting avenue for future work to validate the results in this study by forecasting aggregated wind and PV at other parks. On top of this, we have shown the effect

Table 3
Simulation results over market participation for different shares of wind and PV power in the co-located park. The columns represent the different shares and corresponding results in the rows. Values in parenthesis denote perfect forecasts.

	Shuffle 1.					
	0%	20%	40%	60%	80%	100%
Contracted [GWh]:	22.13 (20.0)	22.01 (20.0)	22.1 (20.0)	22.23 (20.0)	22.7 (20.0)	22.95 (20.0)
Surplus [GWh]	5.93 (0)	5.60 (0)	5.70 (0)	5.80 (0)	6.41 (0)	6.85 (0)
Shortage [GWh]	3.82 (0)	3.69 (0)	3.76 (0)	3.77 (0)	3.95 (0)	4.15 (0)
Down-regulation cost [10^3 Euro]	10.75 (0)	9.01 (0)	7.82 (0)	7.77 (0)	8.19 (0)	9.06 (0)
Up-regulation cost [10^3 Euro]	11.88 (0)	10.37 (0)	8.48 (0)	8.78 (0)	8.77 (0)	9.18 (0)
Total revenue [10^3 Euro]	722.2 (744.9)	707.3 (726.7)	697.9 (714.2)	688.6 (705.2)	681.3 (698.3)	674.6 (692.9)
Avg. energy price [Euro/MWh]	32.64 (37.21)	32.13 (36.15)	31.58 (35.43)	30.98 (34.91)	30.02 (34.52)	29.4 (34.21)
Avg. down-regulation unit cost [Euro/MWh]	4.48 (0)	4.28 (0)	4.32 (0)	4.36 (0)	4.38 (0)	4.43 (0)
Avg. up-regulation unit cost [Euro/MWh]	5.02 (0)	4.47 (0)	4.47 (0)	4.47 (0)	4.50 (0)	4.10 (0)
Part of imbalance [% of produced enery]	48.65 (0)	46.21 (0)	46.9 (0)	47.34 (0)	51.23 (0)	54.31 (0)
Performance ratio [%]	96.96 (100)	97.33 (100)	97.72 (100)	97.65 (100)	97.57 (100)	97.37 (100)
	Shuffle 2.					
	0%	20%	40%	60%	80%	100%
Contracted [GWh]:	24.85 (20.0)	23.04 (20.0)	21.82 (20.0)	21.32 (20.0)	21.28 (20.0)	21.21 (20.0)
Surplus [GWh]	9.02 (0)	6.90 (0)	5.70 (0)	5.30 (0)	5.30 (0)	5.32 (0)
Shortage [GWh]	4.17 (0)	3.87 (0)	3.87 (0)	3.97 (0)	4.00 (0)	4.11 (0)
Down-regulation cost [10^3 Euro]	6.16 (0)	5.06 (0)	5.64 (0)	6.12 (0)	6.41 (0)	7.14 (0)
Up-regulation cost [10^3 Euro]	17.65 (0)	13.97 (0)	10.18 (0)	9.12 (0)	10.13 (0)	9.42 (0)
Total revenue [10^3 Euro]	703.7 (727.5)	655.4 (674.4)	631.9 (647.7)	616.5 (631.7)	604.4 (621.0)	596.8 (613.3)
Avg. energy price [Euro/MWh]	28.32 (36.37)	28.45 (33.72)	28.96 (32.39)	28.91 (31.58)	28.4 (31.05)	28.13 (30.67)
Avg. down-regulation unit cost [Euro/MWh]	6.11 (0)	5.28 (0)	5.27 (0)	5.27 (0)	5.28 (0)	5.36 (0)
Avg. up-regulation unit cost [Euro/MWh]	6.01 (0)	4.58 (0)	4.60 (0)	4.59 (0)	4.54 (0)	4.43 (0)
Part of imbalance [% of produced enery]	65.96 (0)	53.85 (0)	47.86 (0)	46.33 (0)	46.43 (0)	47.14 (0)
Performance ratio [%]	96.73 (100)	97.18 (100)	97.56 (100)	97.59 (100)	97.34 (100)	97.3 (100)
	Shuffle 3.					
	0%	20%	40%	60%	80%	100%
Contracted [GWh]:	21.01 (20.0)	21.73 (20.0)	22.05 (20.0)	22.65 (20.0)	23.13 (20.0)	23.82 (20.0)
Surplus [GWh]	5.65 (0)	5.81 (0)	5.90 (0)	6.47 (0)	6.94 (0)	7.78 (0)
Shortage [GWh]	4.64 (0)	4.12 (0)	3.92 (0)	3.91 (0)	3.91 (0)	4.08 (0)
Down-regulation cost [10^3 Euro]	8.56 (0)	7.99 (0)	8.22 (0)	8.54 (0)	9.41 (0)	11.20 (0)
Up-regulation cost [10^3 Euro]	11.87 (0)	13.03 (0)	13.59 (0)	16.00 (0)	18.78 (0)	20.94 (0)
Total revenue [10^3 Euro]	910.2 (934.5)	904.3 (926.7)	899.3 (921.1)	892.4 (916.9)	885.5 (913.6)	878.9 (911.0)
Avg. energy price [Euro/MWh]	43.3 (46.72)	41.62 (46.24)	40.78 (45.9)	39.41 (45.64)	38.28 (45.43)	36.89 (45.27)
Avg. down-regulation unit cost [Euro/MWh]	5.176 (0)	5.075 (0)	5.066 (0)	5.08 (0)	5.066 (0)	5.044 (0)
Avg. up-regulation unit cost [Euro/MWh]	3.911 (0)	4.409 (0)	4.571 (0)	4.528 (0)	4.586 (0)	4.473 (0)
Part of imbalance [% of produced enery]	51.48 (0)	49.5 (0)	48.94 (0)	51.67 (0)	53.97 (0)	58.97 (0)
Performance ratio [%]	97.40 (100)	97.58 (100)	97.63 (100)	97.32 (100)	96.92 (100)	96.47 (100)

of temporal aggregation at a single park. Another interesting future research direction is to see how spatio-temporal aggregation, e.g., in a virtual power park with spatially sparse wind and PV parks, compares to temporal aggregation in terms of forecast accuracy. This could be done by analyzing if the local weather has a more pronounced effect on the uncertainty mitigation as compared to parks at different locations, not affected by the same type of weather.

5. Conclusions

In this work, we investigated the effect on probabilistic forecast quality and value when aggregating wind and PV power at a co-located park by validating the result using three years of observed power output. This was done by comparing different ratios of wind and PV to the total capacity of the park, where the forecasts are thoroughly evaluated using qualitative and quantitative methods and by simulating day-ahead market trading under a fixed modelling framework. The results presented in this study should be generally valid for other co-located parks in similar climates and show that the smoothing effect from aggregation carries over to more accurate forecasts quantified in terms of the CRPS, quantile score and interval score. The optimal ratio of wind and PV in the co-located park was found at around 50% – 60% wind power to the total nominal capacity of the park when studying the whole time period. Forecasting co-located production also has the effect of producing more reliable and sharper predictive distributions which is also most prominently the case during the summer, fall and spring and comes at the cost of lowered forecast resolution, which means that the capacity to issue case-dependent forecasts is lowered. The daily variability in the wind and PV power output were shown to be anti-correlated, which is beneficial in terms of forecasting aggregated time series. Although the power sources show slightly anti-correlated characteristics in terms of the CRPS and variability, it was also found that the PV forecasts being insensitive to the increasing variability of wind power production. Finally, the results indicate that a pure PV park yield the highest revenues on the day-ahead market, which is explained by the small amount of PV to the total net generation in Sweden. However, when forecasting aggregated time series in a co-located park, the imbalance costs are lowered as compared to the individual power sources. This is explained by better accuracy, reliability and sharpness from co-located forecasts as compared to forecasts from individual power sources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

O. Lindberg: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **D. Lingfors:** Funding acquisition, Writing – review & editing, Supervision. **J. Arnqvist:** Funding acquisition, Writing – review & editing, Supervision. **D. van der Meer:** Conceptualization, Writing – review & editing. **J. Munkhammar:** Conceptualization, Supervision, Writing – review & editing.

Data availability

The NWP forecasts from MetCoOp Ensemble Prediction System (MEPS) are available from [30]. The electricity prices for the day-ahead market are available from [31].

Acknowledgement

This study was conducted within the research project 'Increased utilisation of the grid with combined solar- and wind power parks' (49421-1), funded by the Swedish Energy Agency. The study has also been, in part, financed by StandUp for Energy. Dennis van der Meer was supported by the European Union Horizon 2020 Framework Program (Smart4RES, No. 864337).

References

[1] Zhang Y, Wang J, Wang X. Review on probabilistic forecasting of wind power generation. Renewable Sustainable Energy Rev 2014;32:255–70. doi:10.1016/j.rser.2014.01.033.
[2] van der Meer DW, Widén J, Munkhammar J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. Renewable Sustainable Energy Rev 2018;81:1484–512. doi:10.1016/j.rser.2017.05.212.
[3] Yang D, van der Meer D. Post-processing in solar forecasting: ten overarching thinking tools. Renewable Sustainable Energy Rev 2021;140:110735. doi:10.1016/j.rser.2021.110735.
[4] Widén J, Carpman N, Castellucci V, Lingfors D, Olauson J, Remouit F, et al. Variability assessment and forecasting of renewables: a review for solar, wind, wave and tidal resources. Renewable Sustainable Energy Rev 2015;44:356–75. doi:10.1016/j.rser.2014.12.019.
[5] Katzenstein W, Fertig E, Apt J. The variability of interconnected wind plants. Energy Policy 2010;38(8):4400–10. doi:10.1016/j.enpol.2010.03.069.
[6] Lave M, Kleissl J, Arias-Castro E. High-frequency irradiance fluctuations and geographic smoothing. Sol Energy 2012;86(8):2190–9. doi:10.1016/j.solener.2011.06.031. Progress in Solar Energy 3
[7] Jurasz J, Canales FA, Kies A, Guezgouz M, Beluco A. A review on the complementarity of renewable energy sources: concept, metrics, application and future research directions. Sol Energy 2020;195(April 2019):703–24. doi:10.1016/j.solener.2019.11.087.
[8] WindEurope. Renewable hybrid power plants: Exploring the benefits and market opportunities. 2020.
[9] Lindberg O, Arnqvist J, Munkhammar J, Lingfors D. Review on power-production modeling of hybrid wind and PV power parks. J Renewable Sustainable Energy 2021;13(4):042702. doi:10.1063/5.0056201.
[10] Dykes K, King J, DiOrio N. Research opportunities in the physical design optimization of hybrid power plants: Preprint. 4th International Hybrid Power Systems Workshop; 2019.
[11] Petersen L, Iov F, Tarnowski GC, Gevorgian V, Koralewicz P, Stroe DI. Validating performance models for hybrid power plant control assessment. Energies 2019;12(22). doi:10.3390/en12224330.
[12] Das K, Philippe Grapperon ALT, Sørensen PE, Hansen AD. Optimal battery operation for revenue maximization of wind-storage hybrid power plant. Electr Power Syst Res 2020;189:106631. doi:10.1016/j.epsr.2020.106631.
[13] Ahlstrom M, Ela E, Riesz J, O'Sullivan J, Hobbs BF, O'Malley M, et al. The evolution of the market: designing a market for high levels of variable generation. IEEE Power Energy Mag 2015;13(6):60–6. doi:10.1109/MPE.2015.2458755.
[14] Vázquez Pombo D, Iov F, Stroe D-I. A novel control architecture for hybrid power plants to provide coordinated frequency reserves. Energies 2019;12(5). doi:10.3390/en12050919. 1996-1073
[15] Zhu R., Das K., Sørensen P., Hansen A.. Energy management of hybrid power plants in balancing market. 6th Hybrid Power Systems Workshop; 2022. [https://doi.org/10.1049/icp.2021.2634].
[16] IEA. IEA wind TCP task 50: Hybrid power plants. 21st Wind and Solar Integration Workshop; 2022. [Accessed: 2022-10-26].
[17] Ludwig D, Breyer C, A Solomon A, Seguin R. Evaluation of an onsite integrated hybrid PV-wind power plant. AIMS Energy 2020;8(5):988–1006. doi:10.3934/energy.2020.5.988.
[18] Lindberg O., Lingfors D.. Oversizing co-located wind and solar parks to increase the capacity factor. 21st Wind and Solar Integration Workshop; 2022.
[19] Long Q., Zhu R., Das K., Sørensen P.. Interfacing energy management with supervisory control for hybrid power plants. 20th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants (WIW 2021); 2022. [https://doi.org/10.1049/icp.2021.2634].
[20] Alessandrini S, McCandless T. The Schaake Shuffle technique to combine solar and wind power probabilistic forecasting. Energies 2020;13(10). doi:10.3390/en13102503.
[21] Haupt SE, McCandless TC, Dettling S, Alessandrini S, Lee JA, Linden S, Petzke W, Brummet T, Nguyen N, Kosović B, Wiener G, Hussain T, Al-Rasheedi M. Combining artificial intelligence with physics-based methods for probabilistic renewable energy forecasting. Energies 2020;13(8). doi:10.3390/en13081979.
[22] Pombo DV, Rincón MJ, Bacher P, Bindner HW, Spataru SV, Sørensen PE. Assessing stacked physics-informed machine learning models for co-located wind-solar power forecasting. Sustainable Energy Grids Networks 2022;32:100943. doi:10.1016/j.segan.2022.100943.
[23] Camal S, Teng F, Michiorri A, Kariniotakis G, Badesa L. Scenario generation of aggregated wind, photovoltaics and small hydro production for power systems applications. Appl Energy 2019;242:1396–406. doi:10.1016/j.apenergy.2019.03.112.
[24] Yin Y, Liu T, He C. Day-ahead stochastic coordinated scheduling for thermal-hydro-wind-photovoltaic systems. Energy 2019;187:115944. doi:10.1016/j.energy.2019.115944.
[25] van der Meer DW, Munkhammar J, Widén J. Probabilistic forecasting of solar power, electricity consumption and net load: investigating the effect of seasons, aggregation and penetration on prediction intervals. Sol Energy 2018;171:397–413. doi:10.1016/j.solener.2018.06.103.
[26] van der Meer DW, Shepero M, Svensson A, Widén J, Munkhammar J. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using gaussian processes. Appl Energy 2018;213:195–207. doi:10.1016/j.apenergy.2017.12.104.
[27] Murphy AH. What is a good forecast? an essay on the nature of goodness in weather forecasting. Weather Forecasting 1993;8(2):281–93. doi:10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2.
[28] Bengtsson L, Andrae U, Aspelien T, Batrak Y, Calvo J, de Rooy W, et al. The HARMONIEAROME model configuration in the ALADINHIRLAM NWP system. Mon Weather Rev 2017;145(5):1919–35. doi:10.1175/MWR-D-16-0417.1.
[29] Frogner I-L, Andrae U, Bojarova J, Callado A, Escribá P, Feddersen H, et al. HarmonEPSthe HARMONIE ensemble prediction system. Weather Forecasting 2019;34(6):1909–37. doi:10.1175/WAF-D-19-0030.1.
[30] Institute N.M.. Met norway thredds service. 2022. https://thredds.met.no/.
[31] Nord-Pool. Market data - nordpool. https://www.nordpoolgroup.com/, [Accessed: 2022-11-09] 2022.
[32] Golestaneh F, Pinson P, Gooi HB. Very short-term nonparametric probabilistic forecasting of renewable energy generation with application to solar energy. IEEE Trans Power Syst 2016;31(5):3850–63. doi:10.1109/TPWRS.2015.2502423.
[33] Breiman L. Random forests. Mach Learn 2001;45(1):5–32. doi:10.1023/a:1010933404324.
[34] Meinshausen N. Quantile regression forests. Journal of Machine Learning Research 2006;7(35):983–99. http://jmlr.org/papers/v7/meinshausen06a.html
[35] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research 2011;12:2825–30.
[36] Koenker R, Bassett G. Regression quantiles. Econometrica 1978;46(1):33–50. http://www.jstor.org/stable/1913643
[37] Bremnes JB. Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. Mon Weather Rev 2004;132(1):338–47. doi:10.1175/1520-0493(2004)132<0338:PFOPIT>2.0.CO;2.
[38] Wilks DS. Chapter 3 - univariate ensemble postprocessing. In: Vannitsem S, Wilks DS, Messner JW, editors. Statistical Postprocessing of Ensemble Forecasts. Elsevier; 2018. p. 49–89. ISBN 978-0-12-812372-0. doi:10.1016/B978-0-12-812372-0.00003-0.
[39] Gneiting T, Raftery AE. Strictly proper scoring rules, prediction, and estimation. J Am Stat Assoc 2007;102(477):359–78. doi:10.1198/016214506000001437.
[40] Lauret P, David M, Pinson P. Verification of solar irradiance probabilistic forecasts. Sol Energy 2019;194:254–71. doi:10.1016/j.solener.2019.10.041.
[41] Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 2007;69(2):243–68. doi:10.1111/j.1467-9868.2007.00587.x.
[42] Pinson P, Nielsen HA, Møller JK, Madsen H, Kariniotakis GN. Non-parametric probabilistic forecasts of wind power: required properties and evaluation. Wind Energy 2007;10(6):497–516. doi:10.1002/we.230.
[43] Wilks DS. Chapter 2 - review of probability. In: Wilks DS, editor. Statistical Methods in the Atmospheric Sciences (Fourth Edition). Elsevier; 2019. p. 7–19. ISBN 978-0-12-815823-4. doi:10.1016/B978-0-12-815823-4.00002-X.
[44] Gneiting T, Raftery AE, Westveld AH, Goldman T. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. Mon Weather Rev 2005;133(5):1098–118. doi:10.1175/MWR2904.1.
[45] Bröcker J, Smith LA. Increasing the reliability of reliability diagrams. Weather Forecasting 2007;22(3):651–61. doi:10.1175/WAF993.1.
[46] Pinson P, McSharry P, Madsen H. Reliability diagrams for non-parametric density forecasts of continuous variables: accounting for serial correlation. Q J R Meteorolog Soc 2010;136(646):77–90. doi:10.1002/qj.559.
[47] Gneiting T, Katzfuss M. Probabilistic forecasting. Annu Rev Stat Appl 2014;1(1):125–51. doi:10.1146/annurev-statistics-062713-085831.

[48] Hersbach H. Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather Forecasting 2000;15(5):559–70. doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.
[49] NCAR, Laboratory R.A.. verification: Weather forecast verification utilities. 2015. https://cran.r-project.org/package=verification.
[50] Winkler RL. A decision-theoretic approach to interval estimation. J Am Stat Assoc 1972;67(337):187–91. doi:10.1080/01621459.1972.10481224.
[51] Bentzien S, Friederichs P. Decomposition and graphical portrayal of the quantile score. Q J R Meteorolog Soc 2014;140(683):1924–34. doi:10.1002/qj.2284.
[52] Bremnes JB. Probabilistic wind power forecasts using local quantile regression. Wind Energy 2004;7(1):47–54. doi:10.1002/we.107.
[53] Ilieva I, Bolkesjø TF. An econometric analysis of the regulation power market at the nordic power exchange. Energy Procedia 2014;58:58–64. doi:10.1016/j.egypro.2014.10.409. Renewable Energy Research Conference, RERC 2014
[54] Agency SE. Energy in sweden 2021 - an overview. Tech. Rep.. Swedish Energy Agency; 2021. https://energimyndigheten.a-w2m.se/Home.mvc?ResourceId=198022
[55] Hirth L. Market value of solar power: is photovoltaics cost-competitive? IET Renewable Power Gener 2015;9(8):37–45. https://digital-library.theiet.org/content/journals/10.1049/iet-rpg.2014.0101
[56] Lindberg O, Lingfors D, Arnqvist J. Analyzing the mechanisms behind temporal correlation between power sources using frequency separated time scales: a swedish case study on PV and wind. Energy 2022;259:124817. doi:10.1016/j.energy.2022.124817.
[57] Petropoulos F, Hyndman RJ, Bergmeir C. Exploring the sources of uncertainty: why does bagging for time series forecasting work? Eur J Oper Res 2018;268(2):545–54. doi:10.1016/j.ejor.2018.01.045.
[58] Unger EA, Ulfarsson GF, Gardarsson SM, Matthiasson T. The effect of wind energy production on cross-border electricity pricing: the case of western denmark in the nord pool market. Econ Anal Policy 2018;58:121–30. doi:10.1016/j.eap.2018.01.006.