Energy 263 (2023) 125893

Contents lists available at ScienceDirect

Energy

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

Intelligent estimation of wind farm performance with direct and indirect 'point' forecasting approaches integrating several NWP models

Ghali Yakoub *, Sathyajith Mathew, Joao Leal

University of Agder, Jon Lilletunsvei 9, 4879 Grimstad, Norway

ARTICLE INFO

Keywords:
Wind power forecasting
NWP
Direct forecast
Indirect forecast
Machine learning

ABSTRACT

Reliable wind power forecasting is essential for profitably trading wind energy in the electricity market and efficiently integrating wind-generated electricity into the power grids. In this paper, we propose short- and medium-term wind power forecasting systems targeted to the Nordic energy market, which integrate inputs on the wind flow conditions from three numerical weather prediction sources. A point forecasting scheme is adopted, which forecasts the power at the individual turbine level. Both direct and indirect forecasting approaches are considered and compared. An automated machine-learning pipeline, built and optimized using genetic programming, is implemented for developing the proposed forecasting models. The turbine level power forecasts using different approaches are then combined into a single forecast using a weighting method based on recent forecast errors. These are then aggregated for the wind farm level power estimates. The proposed forecasting schemes are implemented with data from a Norwegian wind farm. We found that in both the direct and indirect forecasting approaches, the forecasting errors could be reduced between 8% and 22%, while inputs from several NWP sources are used together. The wind downscaling model, which is used in the indirect forecasting approach, could significantly contribute to the model's accuracy. The performance of both the direct and indirect forecasting schemes is comparable for the studied wind farm.

1. Introduction

Electric power industries and related markets face new challenges due to the increased wind power penetration into electric grids. One of the significant challenges in this large-scale integration of wind-generated electricity is the stochastic nature of wind. The speed and direction of the wind at a given site may significantly vary even within short intervals of time. As the power contained in the wind stream is proportional to the cube of wind velocity, these fluctuations in the velocity will be reflected in the wind power production on a magnified scale. This results in uncertainties in the generation expected from wind power plants. This is critical for System Operators (SOs) as they have to deal with these variabilities and uncertainties when scheduling and dispatching decisions. Moreover, these uncertainties pose challenges to Generating Companies (GENCOs) in operating their wind farms efficiently by providing stable and dispatchable energy, avoiding possible power quality issues and economic losses [1].

In many countries, wind power producers must participate in electricity markets in the same way as conventional power generators [2]. A good example is the Nordic power market (Nord Pool), which operates in 16 European countries, including the Nordic region. This market has three main alternatives: the day-ahead 'Elspot', the continuous intraday 'Elbas', and the balancing market.

In 'Elspot', the producers and consumers submit bids and offers to cover every hour of the following day, which is cleared before noon. Whereas in 'Elbas', the trade is settled individually between two parties which closes 1 h before the delivery hour in Norway. The balancing market regulates production or consumption up or down depending on keeping the grid's instantaneous balance [3].

Therefore, wind power forecasting (WPF) is becoming essential for realizing energy balance and scheduling decisions, particularly in regions with a significant share of wind electricity generation. Short- and medium-term forecasts are essential for successful participation in the electricity markets, particularly the day-ahead market, i.e., 36 h ahead, which traditionally has been the leading market mechanism for power trading. However, due to the increased penetration of wind electricity generation, the trading market moved to shorter horizons than the day-ahead market. The participants in the short-term markets use intra-day forecasts to adjust their day-ahead bids avoiding the associated imbalance costs, which are also used by system operators who must balance supply and demand in real-time, guaranteeing the system's security [4].

* Corresponding author. /
E-mail addresses: Ghali.yakoub@uia.no, Ghaliyakoub_89@hotmail.com (G. Yakoub).

https://doi.org/10.1016/j.energy.2022.125893
Received 9 February 2022; Received in revised form 23 September 2022; Accepted 24 October 2022
Available online 2 November 2022

Nomenclature

AI	Artificial Intelligence
ANN	Artificial Neural Network
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Auto-Regressive Moving Average
ECMWF-ifs	European Centre for Medium-Range Weather Forecasts - integrated forecast system
GA	Genetic Algorithm
GENCOs	Generating Companies
GP	Grey Prediction model
IoA	Index of Agreement
MAE	Mean Absolute Error
MEPS	MetCoOp Ensemble Prediction System
MET	The Norwegian Meteorological Institute
ML	Machine Learning
NMAE	Normalized Mean Absolute Error
NRMSE	Normalized Root Mean Squared Error
NWP	Numerical Weather Predictions
RMSE	Root Mean Squared Error
SCADA	Supervisory Control And Data Acquisition
SO	System Operator
TSO	Transmission System Operator
WPF	Wind Power Forecasting
R^2	Coefficient of Determination

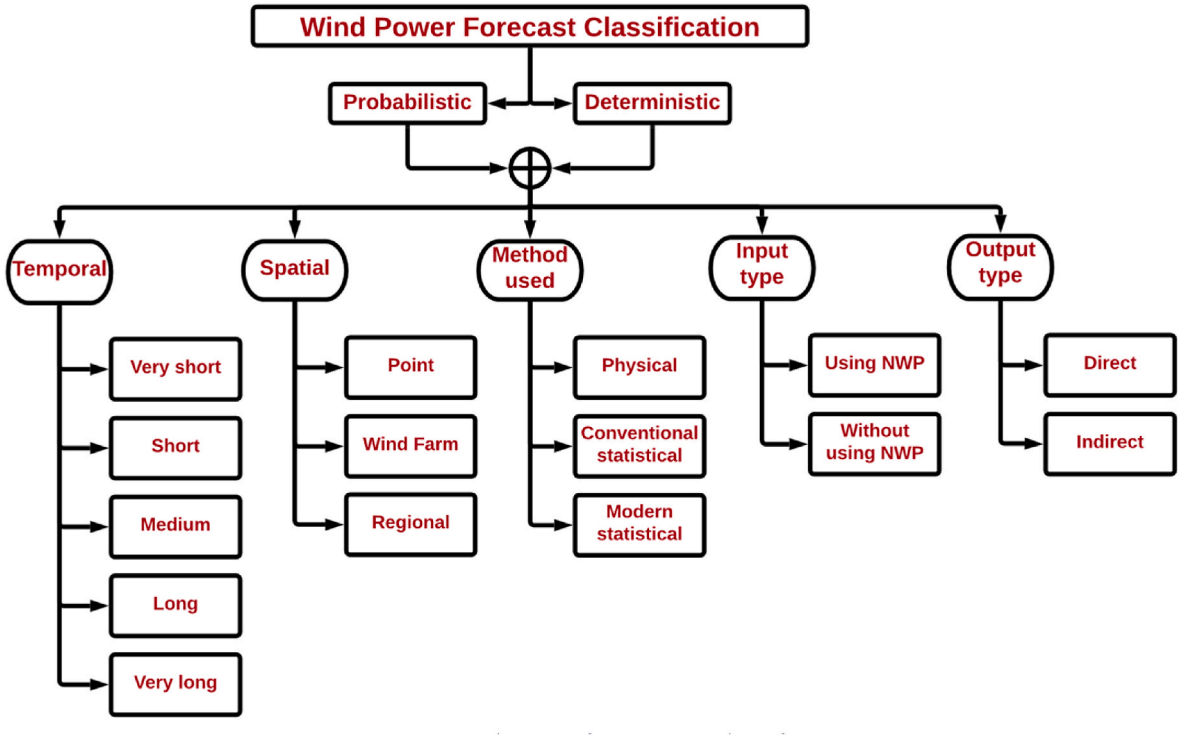

Fig. 1. Wind power forecasting classification.

Moreover, The power producers are accountable for the cost of real-time balancing of their deviations from their forward-contracted volumes [2]. Hence, it is essential to estimate the day-ahead and hour-ahead productions from the wind farms with an acceptable level of accuracy to avoid financial losses. Hour-ahead predictions are also crucial for formulating the dispatch schedules for the grid operation. Therefore, reliable short-to medium-term wind power forecasting (WPF) has become increasingly important in the efficient management of wind farms.

The WPF can be classified into several categories depending on the governing characteristics and the need for the forecast [5]. As shown in Fig. 1, in general, the forecast can be classified as either deterministic or probabilistic. Deterministic forecasts comprise a single value estimation for estimated future wind power generation. In contrast, under the probabilistic forecast, wind power is modelled as a stochastic process with confidence levels that can quantify the uncertainties in the predicted wind power generation [6–8]. Nevertheless, with the advances in the deterministic forecasting methods evolved over the years [9–11], many end-users still prefer the deterministic approach with a single forecasted value as interpreting the probabilistic forecasting results could be challenging [4,10,12].

Both the deterministic and probabilistic forecasts can further have several sub-categories, as shown in Fig. 1 and discussed below.

- Based on the temporal scale, WPF can be classified into five sub-categories [13] as seen in Fig. 1. The very short-term and short-term forecasts range from a few seconds to 30 min ahead and up to 6 h ahead, respectively, while the medium-term forecast can be up to 1 day ahead. In comparison, the long-term and very long-term forecast can be up to 3 days and more ahead, respectively. However, this forecast classification is not universal (see Ref. [14]).
- Based on the spatial scale, WPF can be divided into 3 categories: point forecasting for a single wind turbine [15], wind farm forecasting for a cluster of wind turbines as a single entity [16,17], and regional forecasting for a specific region with several wind farms [18].

- Based on the method used [19], WPF can generally be classified as physical methods such as Numerical Weather predictions (NWP) and mesoscale models use physical conservation laws considering terrain obstacles, air pressure, humidity and temperature to forecast the weather, thus the wind power [20,21]. Traditional statistical methods such as autoregressive integrated moving average models (ARIMA) make the forecasts by establishing the relationship of the observed wind speed time series - in other words forecasting the wind speed based on observed historical data [22], and, more recently, the modern statistical or the so-called learning methods like artificial neural network (ANN) and support vector machines (SVM), which are getting wider attention among the forecasting community [10,23–25]. These approaches can also be combined to yield hybrid methods [11,26].
- Based on the input data, WPF can be classified as numerical weather predictions (NWP) based [17,27,28] or without utilizing the NWP [22,29].
- Based on the output type, WPF can be classified as either indirect forecast (which results from forecasting the wind speed first and then using this forecasted wind speed with the turbines' power curve for estimating the power) or direct forecast (which forecasts the wind power directly, without forecasting the wind speed [4,22,30,31].

WPF has been explored and researched intensively for around 20 years [13–19,22,23,32–34]. In the present study, we propose a new short-medium term point WPF approach, incorporating inputs from several NWPs, adopting an optimized machine learning pipeline. The proposed forecasting scheme is applied and compared in both direct and indirect approaches generating several WPFs. These forecasts are to be later weighted to establish a combined single value WPF merging the strengths of different NWPs based WPF.

The features of the proposed forecasting scheme, differentiating it from the previous approaches, are highlighted below:

- **The automated learning process:** In conventional wind power forecasting, the model type, its architecture and hyperparameters are selected based either on a trial-and-error basis or by applying the domain knowledge and experience as discussed in Refs. [35,36]. In contrast, the present study adopts a new approach to automate the process of selecting the most optimal model, its architecture and configuration in all the development phases. This approach uses a tree-based structure (see section 3.2.1) to represent the development process while performing genetic programming to achieve a stochastic global optimization of the selected pipeline process [37–39]. This could improve the efficiency of the proposed forecasting scheme.
- **Integrating inputs from several NWPs:** The performance of NWP models, which are used to forecast the input flow conditions in the WPF, significantly varies depending upon several factors [6,9,12]. An extensive literature review [14] concluded that for short- and medium-term WPF, NWP data is required to get accurate results. The author in Ref. [17] investigated the role of NWP data in WPF for the next 36 h using an ensemble model with 10 members known as MetCoOp Ensemble Prediction System (MEPS). Several models were built using the XGboost method based on different combinations of the latest observations and MEPS predictions while considering both MEPS generation time and members used. The author concluded that the quality of WPF based only on NWP does not vary much with lead time, and the latest observation would mainly impact the forecast up to 3 h ahead, while beyond 12 h ahead, it would have a negative impact. Regardless of the advances NWP models have reached, specific NWP models would perform well in some regions and worse in others for different reasons [20]. Therefore, to achieve reliable WPF, three NWP models are combinedly used in this study which is rarely reported in WPFs.
- **The point forecasting approach:** Most previous studies follow a wind farm level forecasting approach, while the point-wise level is rarely researched and has not been evaluated for different terrains and wind conditions, mainly due to the size and complexity of the modelling part. Here, we adopt a point-wise forecasting method in which the power output of each turbine in the farm, under a forecasted wind flow condition, is estimated. These are then aggregated to estimate the expected wind farm level power generation. This is expected to give better forecasting performance, as observed in Refs. [4,17]. In Ref. [4], the authors compared different techniques based on turbine level and farm level approaches applied on two wind farms in the UK. The results show the superiority of the turbine level based techniques (point approach) over the farm level approach. The authors concluded that simple aggregated individual turbine forecast and direction conditioned aggregated individual turbine forecast performed both better than the direct wind farm level forecast, where the root means squared error (RMSE) is reduced by 6–12% considering forecasts up to 48 h ahead.
- **The temporal resolution matches the Nordic energy market requirements:** The present study focuses on short-to medium-term forecasts with temporal resolutions corresponding to the Nord Pool's continuous intraday 'Elbas' and day-ahead 'Elspot' trade requirements. Hence, the proposed methods can support the efficient integration of wind energy in the Nordic power market
- **Comparison between the direct and indirect approaches for the same wind farm:** Some previous studies [22] reported that the direct WPF approach yields better accuracy. The authors compared the direct and indirect short-term WPF performance for a single 2 MW offshore turbine. The auto-regressive integrating moving average (ARIMA) models of both approaches were tested for 1-h ahead forecast (achieved at each 10 min), rendering the direct approach outperforming the indirect approach. In contrast, some other investigations favour the indirect approach [12,31,40]. The authors in Ref. [12] proposed a novel approach for an indirect wind farm level forecast by decomposing the wind speed into the mean trend and the stochastic component using the empirical mode decomposition method (EMD). The two components were later forecasted separately using two different techniques. In addition to that, the speed-power conversion was obtained from the historical data using polynomial fitting. This resulted in better performance for both direct power and direct speed forecasts for 1–4 h ahead. Similarly, the authors in Ref. [40] proposed an improved EMD to decompose wind speed measurements and then integrated it with a hybrid forecasting engine using bagging neural networks combined with the K-means clustering technique. The proposed approach was applied on three different wind farms and compared to several previously published approaches, where the proposed approach outperformed all of them for different forecast horizons. In the same track, research was carried out in Ref. [31] investigating wind farm power forecasts in the scope of energy markets (24–48 h) using both direct and indirect approaches while integrating NWP data. The authors studied several wind farms and applied several machine learning techniques, mainly tree-based ones. The ML techniques utilized data from the supervisory control and data acquisition system (SCADA) and weather data extracted from several grid points around the wind farms from the European Centre for Medium-Range Weather Forecasts (ECMWF) model. The results showed that the indirect approach performed better than the direct due to improving the wind forecast. The authors emphasized that having a good speed-power conversion in the indirect approach is essential for a better forecast. The previous studies do not clearly identify which WPF approach (direct/indirect) to best tackle the short-term forecast.

In this paper, we compare the accuracies of both the direct and indirect approaches at the turbine level and an aggregated farm level. The

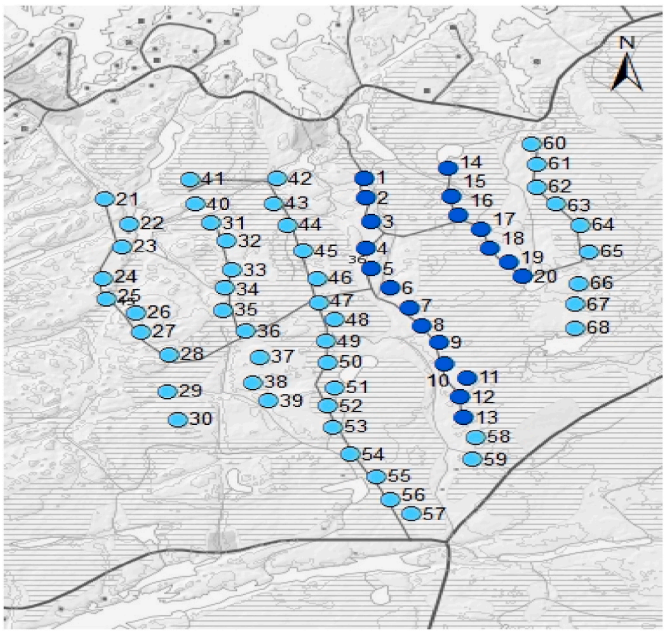

Fig. 2. Smøla wind farm layout.

WPF models are formulated in the scope of the main electricity markets the day-ahead and the intraday and optimized using an automated learning process while integrating data from several NWP models.

The remainder of the paper is organized as follows. The next section describes the data used to develop the forecast models and the pre-processing procedure to detect and process the outliers. This is followed by the descriptions of the methods used to develop and optimize individual turbine power forecast models and the weighing technique which is applied to combine the predictions. Thereafter, we discuss the performance of different models at both turbine and farm levels, which is finally summarized in the last section.

2. Data description and pre-processing

2.1. Wind farm and SCADA data

One of Norway's first commissioned wind farms, the 'Smøla' wind farm, was chosen for the proposed research. Smøla wind farm is composed of 68 wind turbines distributed in 6 rows. On average, the wind farm produces 356 GWh of electricity annually. It is situated in a flat and open terrain, which is 10–40 m above sea level. The wind farm was constructed in two phases. The first construction phase had 20 wind turbines (2 MW each) and became operational in September 2002, while the second phase had 48 wind turbines (2.3 MW each) and was commissioned in September 2005. As seen in Fig. 2, blue and green dots represent the two commissioning phases, respectively. The turbines have rotor diameters of 76–82.4 m, and the towers are 70 m high [41].

The turbine's production, operating conditions, and nacelle wind speed were retrieved from the wind farm's (SCADA) system with a frequency of 10-min from January 2017 to December 2020. The retrieved data, particularly production and wind speed measurements, are converted to an hourly average to be compatible with the scope of the studied electricity markets.

2.2. NWP' weather forecasts

Historical weather forecasts from 3 NWPs' archives, with different spatial resolutions, were used in this study. The NWP source selection was based upon the temporal resolution (1-h scale) and the availability of the model's coverage at the wind farm's terrain.

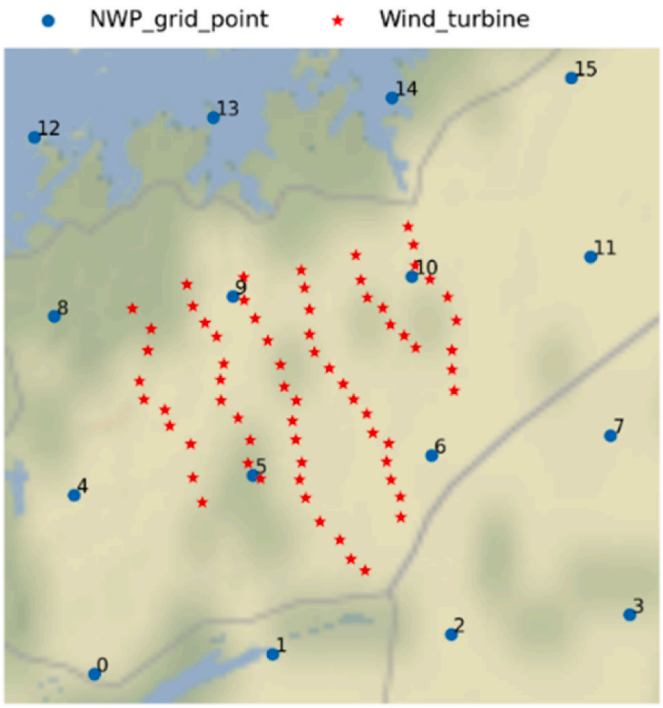

Fig. 3. NWP grid points from the high-resolution model.

Details of the selected NWP sources are summarized below:

- ECMWF-ifs (IFS): Global high-resolution integrated forecast system. This model is operated by (ECMWF). It has a temporal resolution of 1 h and a spatial resolution of 7.6 km.
- Ukmo-euro 4 (EURO): European model run by the UK MetOffice. It has a temporal resolution of 1 h and a spatial resolution of 3.05 km.
- MEPS: This model is Norway's operational weather forecast model. It has a temporal resolution of 1 h and a spatial resolution of 2.5 km. This model is an ensemble model with 10 members [42]. However, in this study, we used data from only one member (the control member) as recommended by MET Norway.

It should be noted that the NWP sources have an update rate of 4 times per day at 00, 06, 12 and 18 UTC, except for IFS, which is updated twice a day since 2018.

As seen in Fig. 3, the NWP grid coordinates, within and surrounding the wind farm, were chosen according to the highest NWP spatial resolution (MEPS). Therefore, the data from IFS and EURO sources were interpolated to the chosen coordinates. Both IFS and EURO data were retrieved from a third-party database, while MEPS data were retrieved directly from the MET Norway database.

Moreover, the air temperature (Temp) was retrieved at 2 m above the ground from the prementioned NWPs' archives. Wind speed, gust, and direction (WS, WG, WD) were extracted at 10 m above the ground from the MEPS source due to availability, while they were extracted at the hub height (70 m) from the IFS and the EURO sources.

2.3. Data pre-processing

Initially, the data were cleaned by removing the outliers. There were four types of outliers which are indicated in Fig. 4 and listed below.

- Type 1 represents a horizontal cluster at the bottom of the power curve due to downtime for several reasons, such as maintenance and error in the log.

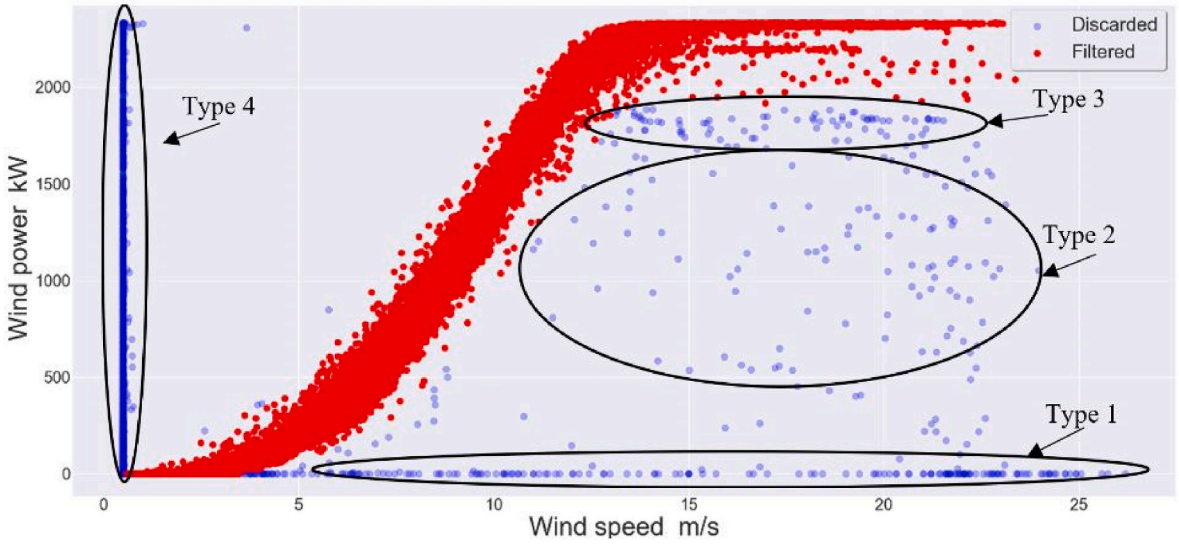

Fig. 4. Filtered data versus discarded data.

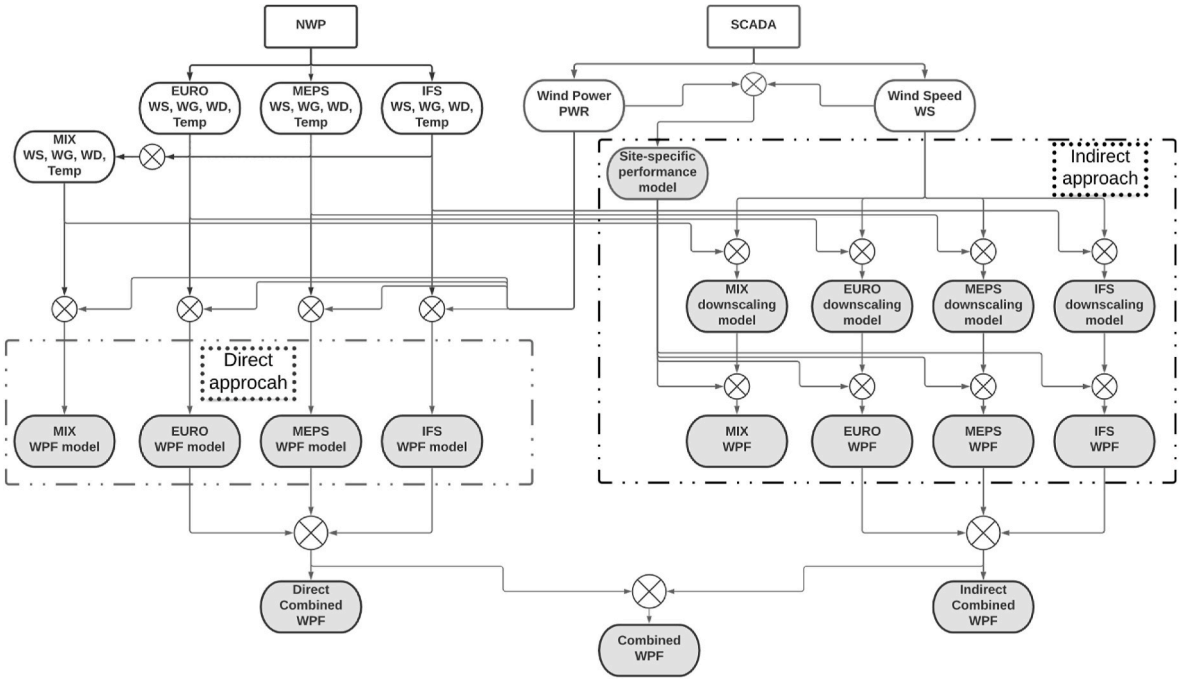

Fig. 5. Direct and indirect wind power forecast models.

- Type 2 represents random scatters due to sensor malfunction, noise in signal processing or power curtailment.
- Type 3 represents a dense cluster below the rated power where the wind turbine production is limited for reasons like power curtailment.
- Type 4 represents a vertical cluster at zero wind speed due to icing or malfunction of the anemometer.

Missing values and outliers represent a fraction between 5% and 30% of the data for each turbine. This data was discarded before developing the forecasting models. However, it should be mentioned that not all turbines suffered from all types of outliers.

3. Methods

3.1. Forecasting models

Models for direct and indirect WPF are developed under the study. Forecasts on wind speed (WS), wind gust (WG), wind direction (WD) and temperature (Temp) from the 4 NWP grid cells nearest to the turbines are directly correlated with the turbine's power output in the direct forecasting approach. On the other hand, in the case of indirect WPF, these forecasts from NWPs were first correlated with the nacelle's wind speed measurements for developing the downscaling model, which estimates the velocity 'felt' by the turbines at its hub coordinates. Then, the power corresponding to these velocities are estimated using the site-specific performance model of the turbine. In both cases, power estimates from individual turbines are accumulated together to arrive at the

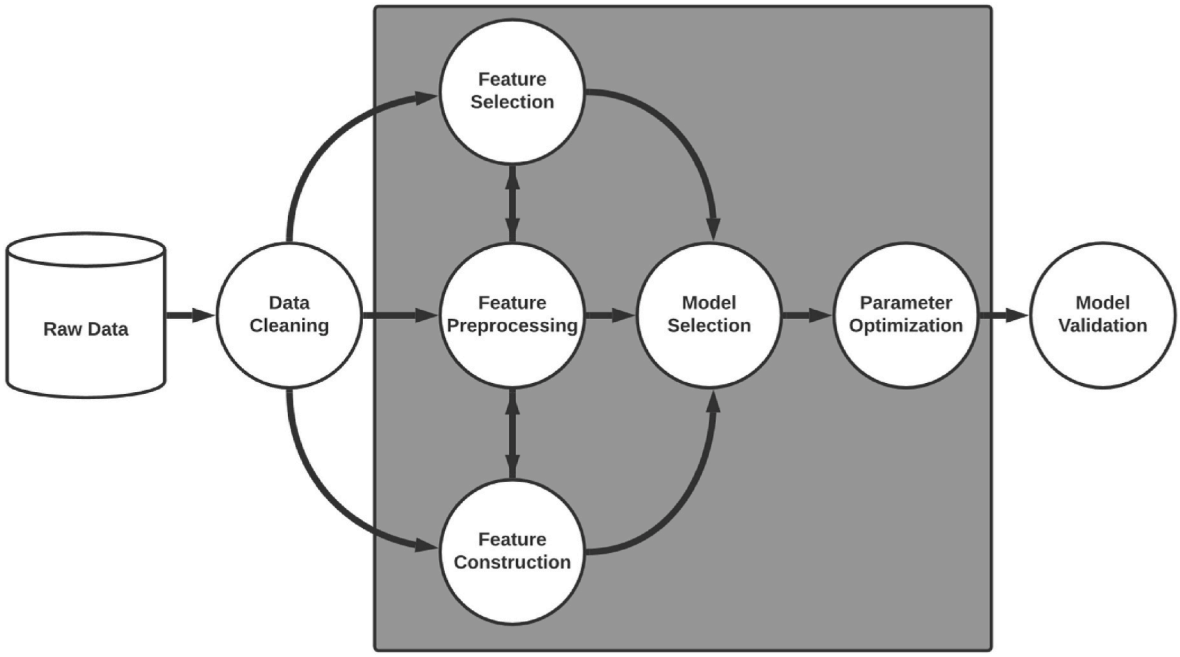

Fig. 6. Overview of the adapted pipeline search [38].

farm level power estimates.

Wind forecasts from the three NWP sources considered under the study (IFS, EURO and MEPS) were used individually and in combination (MIX) to develop each wind turbine's direct and indirect WPF. The MIX models were built upon using the three NWP's data together as input. Thus, as shown in Fig. 5, each wind turbine has 4 direct WPF models, 4 wind speed downscaling models and 1 turbine's site-specific performance model. In addition to this, the above-discussed WPFs based on direct and indirect approaches are individually combined to form single WPFs based upon direct and indirect approaches and both together, as shown in Fig. 5. More details are described in section (3.3).

It should be noted that, for forecasting the wind power at the time (t+1), we used weather data from the NWP sources at the start and the end of the hour, i.e., at time (t) and (t+1).

The historical data records for each turbine were divided into two sets, one for model development and calibration (train and validation) and the other for testing (unseen). The data division was done by preserving the first 10 days of each month from the years 2017–2019 for testing while using the rest for model calibration. The test set was almost 1 year out of the 4 years used in this study.

3.2. Model development process

The model development process had mainly three phases summarized as follows:

- The pre-processing phase which consisted of feature selection, feature pre-processing (division, scaling), and feature construction. The main objective of this phase was to select and prepare the inputs for the model in proper form and shape.
- The model selection phase involved model architecture development (selection and calibration) and optimization of hyperparameters. This phase aimed at achieving high accuracy associated with high generalization ability.
- The model evaluation phase was focused on the performance evaluation of the developed models by applying several error metrics and statistical methods.

Table 1
List of models and data techniques used.

	Model/Ensemble method	Processing techniques
1	Elastic Net model with Cross-Validation	Independent Component Analysis (ICA)[a]
2	Cross-validated Lasso, using LARS algorithm	Principal Component Analysis (PCA)[a]
3	Ridge Regression with built-in cross-validation	Feature Agglomeration Clustering[a]
4	Extremely Randomized Trees	RBF Kernel approximation[a]
5	Gradient Tree Boosted	Feature Normalizer[b]
6	AdaBoost	Maximum Minimum Scaler[b]
7	Decision Tree	Standard Scaler[b]
8	Random Forrest Regressor	Robust Scaler[b]
9	K-Nearest Neighbors	Polynomial Feature Generation[b]
10	Support Vector Machines	Feature Selection based on Variance Threshold[c]
11	Extreme Gradient Boosting XGBoost	Feature Selection based on Extreme Tree Regressor[c]
12	Stochastic Gradient Descent Regressor	Feature Selection based Family wise error[c]

[a] Constructor.
[b] Preprocessor.
[c] Selector.

3.2.1. Model's architecture (pre-processing, model selection and optimization)

In the present study, we adopted a new approach to fully automate the process of selecting the most optimal techniques and configurations for all the development phases. This approach was developed by researchers from the University of Pennsylvania which was initially applied in the biomedical field [37–39]. It uses a tree-based structure to represent the development process for a predictive modelling problem, including data preparation, model selection and hyperparameters tuning while performing genetic programming (GP) to optimize the selected pipeline process. In other words, this approach tries a pipeline, evaluates its performance, and randomly changes parts of it using a genetic search (natural selection) to find a better-performing pipeline. Fig. 6 shows the end-to-end pipeline development process, where the pipeline search elements include feature selection, feature processing, feature construction, model selection, and hyperparameter optimization.

The GP used in this study searches over several well-known regression models, data processing techniques (pre-processing, construction and selection) to generate the search space, and optimizing each pipeline. So, the population consists of a list of randomly generated pipelines. For each generation, copies of the best-performing pipelines are created and imposed with random changes (e.g., the addition or removal of an operation or the parameter tuning of an operation), enabling the development of new pipelines that were never explored before. The worst-performing pipelines are removed from the population at the end of each generation before starting the next generation. Table 1 presents the models and the data processing techniques utilized in the genetic search. However, each of these models/techniques has several hyper-parameters which were tuned during the search. A full description of the models/techniques used can be seen in Refs. [43,44].

Although GP is faster and more efficient than traditional optimization methods, the utilized GP had to be configured to maintain a good diversity to achieve an optimal or near-optimal solution in a reasonable time. The adopted GP had to iterate for 10 generations having a population size of 70, off-spring size of 10, mutation and crossover rates were set to 90% and 10%, respectively. Moreover, the cross-validation method was implemented, dividing the calibration set into 3 folds, 2 for training and 1 for validating. In other words, each pipeline in the genetic search had to be trained and evaluated 3 times using different combinations of folds. Nevertheless, an automatic termination mechanism was implemented to terminate the search if no improvements were detected during the last 3 generations. Furthermore, a scoring function was implemented to rank the established pipelines depending on the mean absolute error (MAE) as it is robust to outliers.

3.2.2. Model's performance evaluation (post-processing)

The accuracy of the developed forecasts was quantified through different metrics, including (RMSE), normalized RMSE (NRMSE), MAE, normalized MAE (NMAE), coefficient of determination (R^2) and the index of agreement (IoA), which are expressed in the equations given below.

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}[P_k - \widehat{P}_k]^2}{n}} \tag{1}$$

$$NRMSE = \frac{RMSE}{Rated\ power} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|P_k - \widehat{P}_k| \tag{3}$$

$$NMAE = \frac{MAE}{Rated\ power} \tag{4}$$

Where, P_k and $\widehat{P}_k$ represent the observed and the predicted values of the studied parameter at the time instant k, whereas n represents the number of data used in the test stage. Note that the rated power is either 2 or 2.3 MW depending on the turbine (see section 2.1).

RMSE and MAE are most commonly used for measuring the variations between predicted and measured values [28,45–47]. However, normalized versions of both metrics are used generally for comparisons with other study cases. Both RMSE and MAE express average model prediction error in units of the variable of interest. However, the errors are squared before being averaged in RMSE. Therefore, it gives high weights to significant errors, which is useful in case large errors are undesirable. In comparison, MAE is more robust to large errors due to the way it is calculated.

R^2 indicates how well the recorded data can be forecasted by the proposed model (goodness of fit) [48].

$$R^2 = 1 - \frac{\sum_{k=1}^{n}[P_k - \widehat{P}_k]^2}{\sum_{k=1}^{n}\left[P_k - \frac{1}{n}\sum_{k=1}^{n}(P_k)\right]^2} \tag{5}$$

The highest possible score can be 1, but it can also display negative scores, indicating an arbitrarily worse predicting model than a linear one.

The metric IoA is a ratio between the mean squared error and the potential error [49]. It is a standardized measure of the degree of model prediction error which varies between 0 and 1. This metric evaluates the agreement between the prediction and the observed values by detecting additive and proportional differences in the observed and predicted means and variances. However, it is overly sensitive to extreme values due to the squared differences.

$$IoA = 1 - \frac{\sum_{k=1}^{n}[P_k - \widehat{P}_k]^2}{\sum_{k=1}^{n}[|\widehat{P}_k - \overline{P}| + |P_k - \overline{P}|]^2} \tag{6}$$

Where, $\overline{P}$ is the mean of the observed values over the test set for the studied parameter.

3.3. Multiple forecast weighing (combined forecast)

As mentioned previously, each wind turbine has 6 wind power predictions depending on both the forecast approaches (direct and indirect) and the NWP model used (IFS, EURO and MEPS). In addition to those, 2 WPFs are created based on merging all the NWPs parameters together, denoted as the MIX models as seen in Fig. 5. In general, The NWPs' performance varies depending on the period of the year, location and other model's configurations. Therefore, it is essential to have a single combined wind power prediction to represent the wind power predictions from the developed models. This is achieved by using a weighting multiplier for the developed models' predictions according to each forecast's short-term historical performance. The predicted value of multiple forecasting models for the single turbine is determined by:

$$\widehat{\boldsymbol{P}}_{t+\Delta t} = \sum\left(\boldsymbol{W}_{App_j}^{NWP_i} * \widehat{\boldsymbol{P}}_{App_j}^{NWP_i}\right)_{t+\Delta t} \tag{7}$$

where $NWP_i \in [IFS, EURO, MEPS]$, $App_j \in [Direct, Indirect]$ and $\boldsymbol{W}^{NWP}$ are the weighting multipliers of different forecasting models, which were determined using RMSE as the performance index. The performance index for each forecasting model was calculated using the last 6 h' observations of the NWPs updates. Thus,

$$\boldsymbol{W}_{App_j}^{NWP_i}(t+\Delta t) = \frac{\left(\boldsymbol{RMSE}_{App_j}^{NWP_i}(t-6)\right)^{-2}}{\sum\left(\boldsymbol{RMSE}_{App_j}^{NWP_i}(t-6)\right)^{-2}} \tag{8}$$

So, the less the magnitude of RMSE for a particular model, the more would be the weightage multiplier W in Eq. (7), and vice versa. It should be noted that the Δt is chosen to be 2 h as this study focuses on the intraday market (in Norway) to reduce the balancing costs.

To illustrate more, the 3 WPFs (IFS, EURO and MEPS) developed using the direct approach are combined, producing a forecast denoted as 'Direct Combined', whose performance is compared to the direct MIX forecast. The same is applied to the WPFs using the indirect approach scheme producing a forecast denoted as 'Indirect Combined'. Moreover, both approaches' forecasts (in total 6) are combined in a single forecast denoted as the 'Combined', as illustrated in Fig. 5, whose performance is compared with both the MIX and the combined ones resulting from both approaches.

Table 2
Individual turbines' wind speed downscaling and NWP's raw wind speed forecasts evaluation results (averaged across the turbines).

Metric	Wind Speed Downscaling				Raw Wind Speed Forecast			
	MIX	IFS	EURO	MEPS	MIX	IFS	EURO	MEPS
RMSE [m/s]	1.30	1.41	1.51	1.66	–	1.67	1.71	2.13
NRMSE [%]	8.2	8.8	9.4	10.4	–	10.4	11.0	13.3
MAE [m/s]	0.98	1.07	1.13	1.27	–	1.28	1.31	1.66
NMAE [%]	6.1	6.7	7.1	8.0	–	8.0	8.2	10.4
IoA [%]	96.9	96.3	95.9	94.7	–	94.6	94.2	91.2
R^2 [%]	88.9	87.0	85.3	82.1	–	80.5	79.6	68.1

Table 3
Site-specific turbine's performance models evaluation (averaged across the turbines).

Metric	RMSE [kW]	NRMSE [%]	MAE [kW]	NMAE [%]	IoA [%]	R^2 [%]
Avg	54.76	2.5	33.27	1.5	99.8	99.4
Std	11.45	0.4	7.77	0.3	0.05	0.2

4. Results and discussion

The turbine as well as farm-level performances of the proposed wind power forecasting models, under both the direct and indirect forecasting approaches, are discussed in this section. As 8 forecasting models were developed for each studied turbine, for concisely presenting the results, the average performances of the models across all the turbines and the corresponding standard deviation are presented.

4.1. Turbine level wind downscaling models for indirect WPF

As mentioned previously, the indirect forecast scheme comprises two models, the wind downscaling and the turbine's performance model. In this subsection, the wind downscaling results will be presented and discussed.

Table 2 shows the performance of the wind downscaling models, which enhances the spatial resolution of the wind predictions from the NWPs to the hub coordinates of the wind turbines. Compared with the models based on a single NWP, the downscaling accuracy (reduced error) is improved using the three NWPs together (MIX model). These improvements varied between 8%, 14% and 22%, on average across all the turbines, compared to individual IFS, EURO and MEPS downscaling models, respectively. Moreover, in comparison with the raw wind forecast from the NWP models used, it is observed that the downscaling models improve the NWP raw wind forecasts from IFS, EURO and MEPS by around 15%, 11% and 22%, respectively. The increased improvement performance found for the MEPS model indicates that the downscaling enhances the wind forecast in vertical coordinates as well, as the MEPS data were extracted at 10 m above the ground. It should be noted that the rated wind speed of 16 m/s is used to normalize both MAE and RMSE.

4.2. Turbine level site-specific performance models for the indirect WPF

Performance of the site-specific wind speed-power characteristics, computed for all the turbines and then averaged, are shown in Table 3. The relatively low values of the error indices and the high values of IoA and R2 imply the proposed algorithms' capability to model the turbine's performances. This is further illustrated in Fig. 7, where the power production of a representative turbine predicted using the developed model is compared with the actual observations.

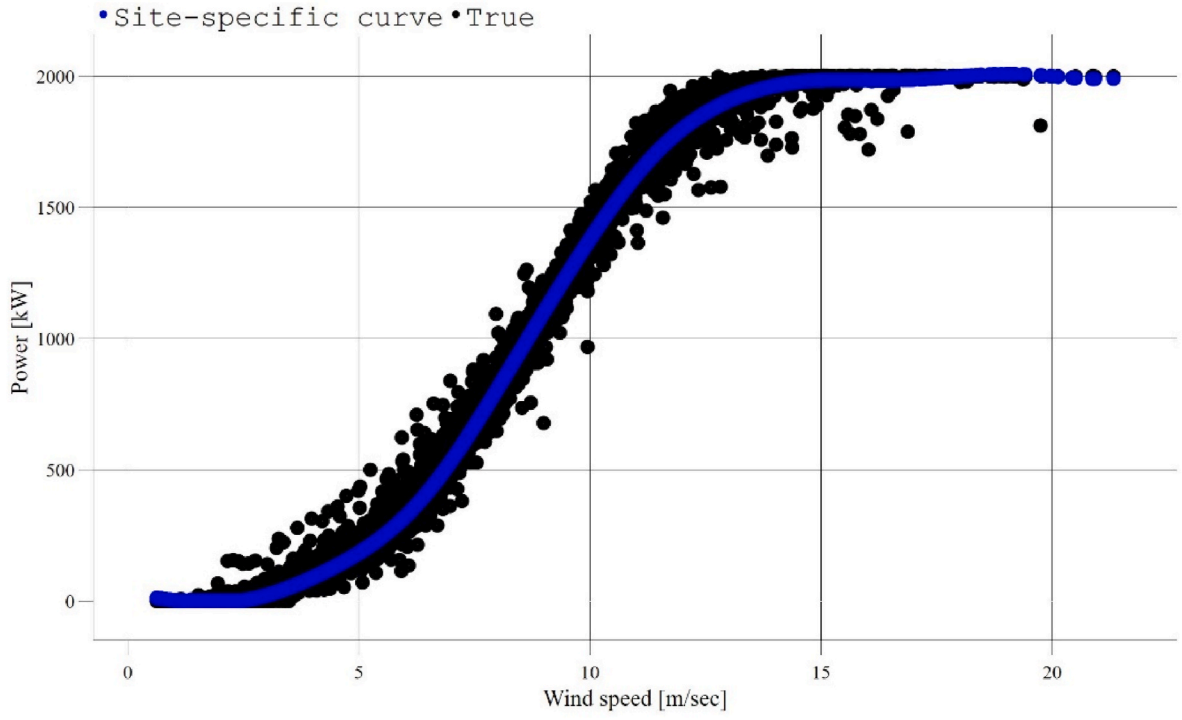

Fig. 7. Site-specific performance model versus observations for turbine number 1.

Table 4
Individual turbines WPF evaluation results using the direct approach (averaged across the turbines).

Direct Wind Power Forecast										
Metric	MIX		IFS		EURO		MEPS		Direct Combined	
	Avg	Std	Avg	Std	Avg	Std	Avg	Std	Avg	Std
RMSE [kW]	248.03	20.36	270.76	24.10	283.52	23.80	320.49	36.94	255.82	20.44
NRMSE [%]	11.21	0.50	12.23	0.55	12.80	0.54	14.45	1.31	11.56	0.53
MAE [kW]	158.39	13.27	174.37	15.24	181.13	16.45	214.68	34.31	165.06	14.88
NMAE [%]	7.16	0.40	7.88	0.42	8.18	0.45	9.69	1.34	7.46	0.45
IoA [%]	96.34	0.33	95.59	0.34	95.18	0.37	93.35	1.48	95.98	0.43
R^2 [%]	87.04	1.06	84.59	1.10	83.08	1.22	78.16	4.79	86.19	1.31

Table 5
Individual turbines WPF evaluation results using the indirect approach (averaged across the turbines).

Indirect Wind Power Forecasts										
Metric	MIX		IFS		EURO		MEPS		Indirect Combined	
	Avg	Std	Avg	Std	Avg	Std	Avg	Std	Avg	Std
RMSE [kW]	250.11	21.85	272.68	24.71	287.76	25.17	318.59	30.87	255.22	21.09
NRMSE [%]	11.28	0.55	12.29	0.64	12.97	0.64	14.37	0.92	11.54	0.55
MAE [kW]	153.14	13.14	167.70	15.01	175.21	15.83	196.26	20.38	156.26	13.30
NMAE [%]	6.91	0.39	7.56	0.45	7.902	0.46	8.85	0.67	7.06	0.40
IoA [%]	96.41	0.33	95.71	0.33	95.23	0.36	93.96	0.84	96.18	0.38
R^2 [%]	86.83	1.11	84.38	1.11	82.59	1.20	78.60	2.52	86.27	1.23

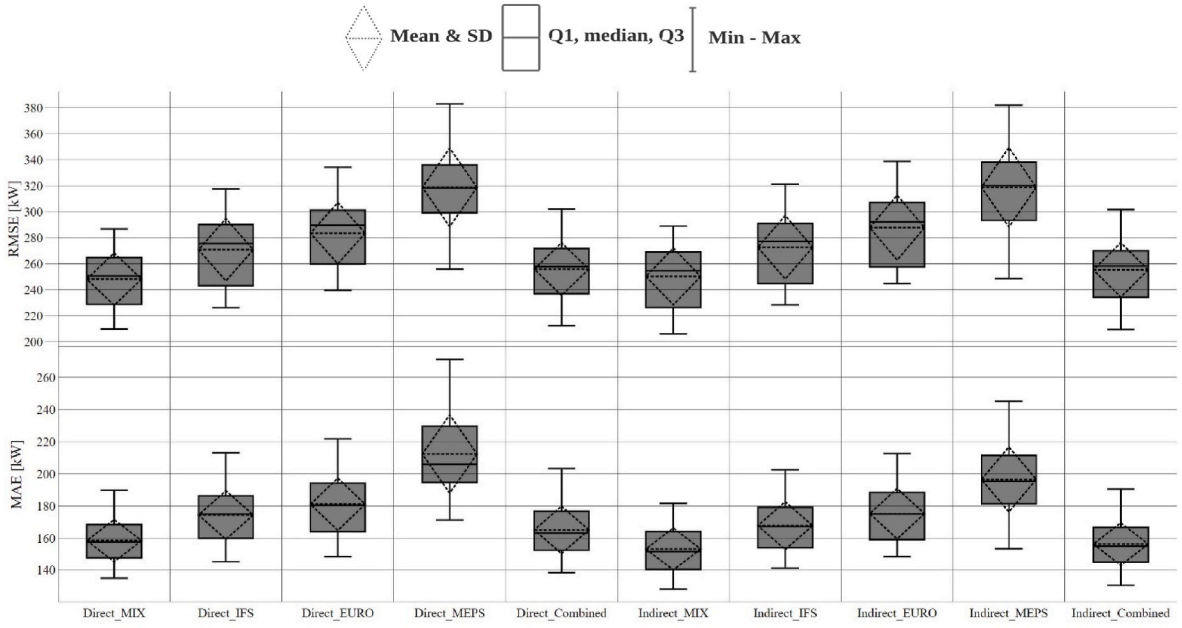

Fig. 8. Average RMSE and MAE across all the turbines associated with the best and poorest turbines.

4.3. Turbine level wind power forecasts (indirect versus direct)

Performances of both the direct and indirect forecasting models, implemented at the turbine level, are presented in Tables 4 and 5, respectively. It is evident that the models' performances are significantly improved by using several NWP sources together (MIX). For example, RMSE and MAE of direct forecasting models could be enhanced, depending on the NWP used, by (8%–22%) and (9%–25%), respectively, by using the MIX. Corresponding improvements in the indirect forecasts were (8%–21%) and (8% 10%) respectively. Similar improvements (but in lesser magnitude) could be seen in the case of IoA and R^2 as well. Additionally, the combined forecast in each approach resulted in improved performance compared to the 3 individual WPFs (IFS, EURO and MEPS) to the same degree as for the MIX forecast. However, the latter performs slightly better with an improved RMSE and MAE below 4%. This confirms that the use of several NWPs yields better overall forecasts.

Despite its lowest spatial resolution among NWP sources used, WPF using IFS data achieved better performances than EURO and MEPS sources. The improvements achieved are in the range of 4%–18% in both approaches. In contrast, the highest spatial resolution of MEPS is not reflected in the performances of the forecasting models in which it is integrated. One possible reason for this could be that the data from MEPS were only available at 10 m, whereas the turbines' hub height was at 70 m. It should be noted that the performances of WPF using different NWP sources could be site-dependent, and further research with varying site conditions is required for arriving to a more generalized conclusion.

The standard deviation of the error metrics is also found to be

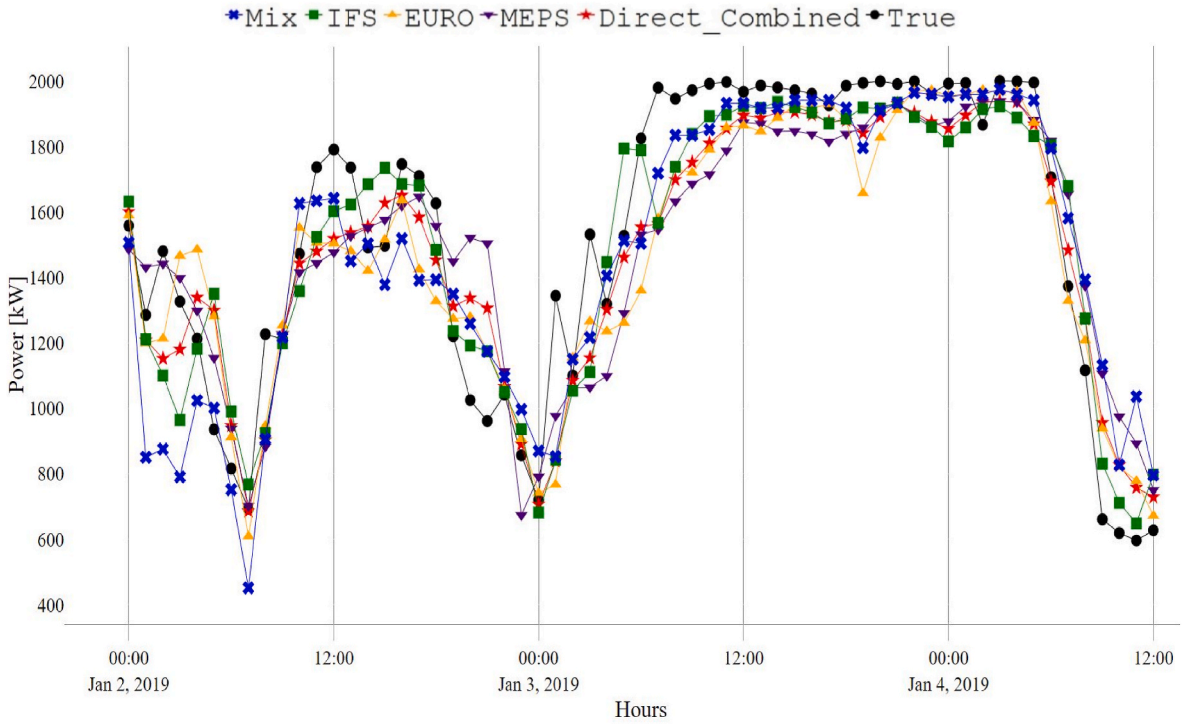

Fig. 9. Direct WPFs versus true power for turbine number 1 between the 2nd and January 4, 2019.

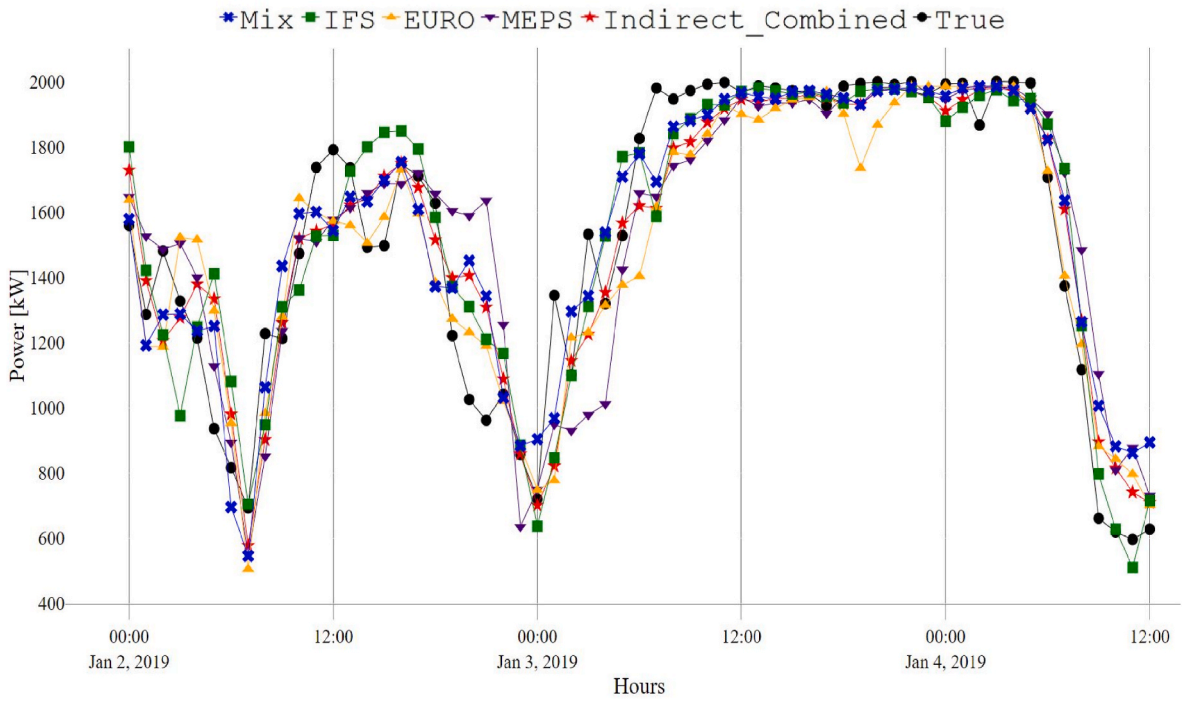

Fig. 10. Indirect WPFs versus true power for turbine number 1 between the 2nd and January 4, 2019.

reduced while using several NWPs together, indicating the consistency of the models while predicting the performances of different turbines. These are further evident in Fig. 8, where the medians, ranges and the percentiles of these matrices are illustrated.

Tables 4 and 5 and Fig. 8 show that the WPF models based on direct and indirect methods performed somewhat equally well for the wind turbines under the study. Interestingly, this is in contrast to the findings of some of the previous studies. For example, the conclusions of [22] indicate that the forecasting based on the direct approach is more reliable than the indirect one, whereas the findings of [31] are contrary. The performance differences for all the models across both approaches are almost identical, where the differences are around 2% for different error metrics. This is evident in the comparison between Figs. 9 and 10.

The 'Combined' forecast resulted from combining the six developed WPFs for each turbine through Eqs. (7) and (8), which are evaluated in Table 6. Combined forecast performances are comparable with the

Table 6
Combined forecast performance (averaged across the turbines).

Metric	RMSE [kW]	NRMSE [%]	MAE [kW]	NMAE [%]	IoA [%]	R^2 [%]
Avg	252.42	11.41	157.58	7.12	96.19	86.56
Std	20.29	0.53	13.45	0.40	0.39	1.23

mixed NWP forecasting scheme (MIX) and the combined forecasts based on each approach. However, in some hours, the combined forecast yielded better predictions, especially in comparison with the direct predictions approach, as seen in Fig. 11.

It should be mentioned that the results had been tested and validated for several hours ahead. In particular, they are valid for up to 12 h ahead for the IFS WPFs and 6 h ahead for the remaining WPFs. This is mainly because WPFs models are trained with historical NWPs, which were updated according to the earlier mentioned update rate.

4.4. Wind farm-level power forecasts

The performance of the models in estimating the power generated from the whole wind farm by aggregating the individual power estimates from all the turbines is shown in Table 7. This is further shown in Fig. 12, in which the percentile variations from the predictions are also illustrated. It should be mentioned that the prediction interval is a quantitative representation of several developed forecasts. It does not reflect the confidence interval or the probability of a specific value associated with certain conditions. As in the case of turbine level forecasts, the mixed and combined modelling approaches outperformed all other options with individual NWP inputted models. The aggregated forecasts have shown lower overall errors than those for the single turbines. This is because the over and under predictions of the turbine's power output could have been cancelled out to some extent while aggregated. Similar improvements can also be observed in the case of accuracy measures like R^2 and IoA.

More in-depth error quantification for the wind farm aggregated forecasts is presented in Table 8 and visualized in Fig. 13. The forecasts' errors have, generally, slightly negative medians and positive means, which means having slightly more than 50% of the predictions being overestimated (negative median). In contrast, the rest of the predictions were underestimated with a higher magnitude (positive average). It should be noted that the error is calculated based on the observed minus the prediction. In addition, all the developed forecasts are right-skewed, i.e., medians are closer to the first quartile (Q1) than to the third quartile (Q3). However, by comparing the interquartile range (IQR), which contains 50% of the predictions, the MEPS forecasts show the highest IQR ranges, 9.1 and 10.4 MW, indicating that the predictions are more dispersed than other forecasts, which is also valid by comparing the standard deviations (SD). This strongly reflects having a wide fence range. Nevertheless, approximately 85% of the error falls into the fence ranges across all the developed forecasts, as shown in Fig. 13.

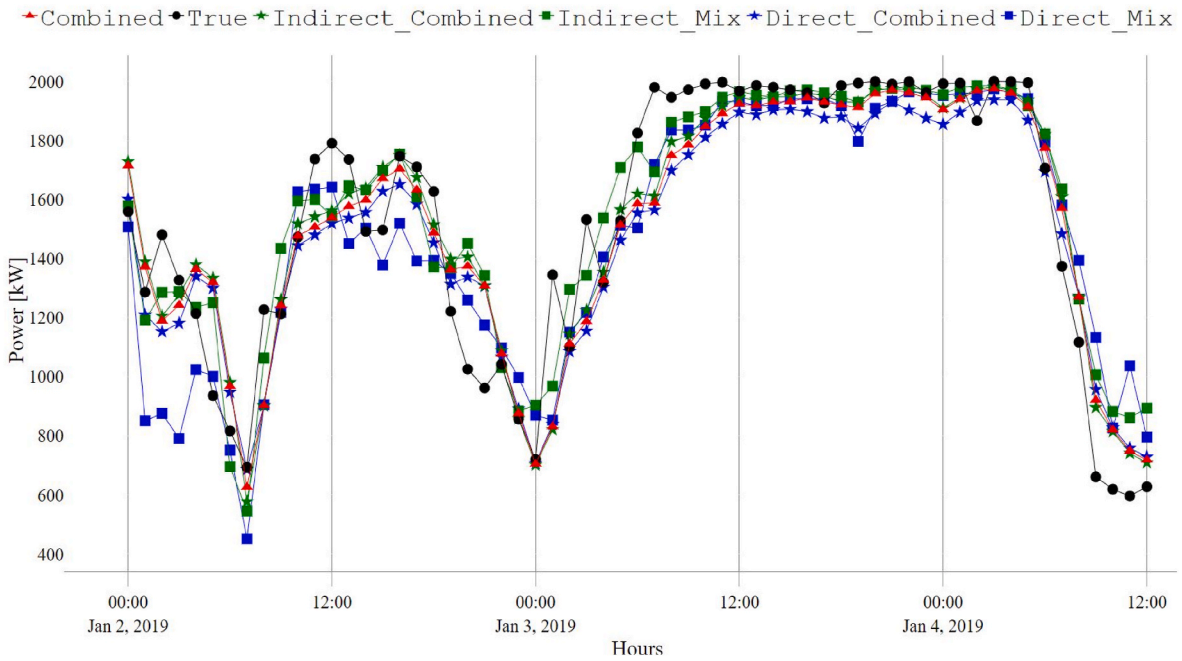

Fig. 11. Combined WPFs and MIX forecasts versus true power for turbine number 1 between the 2nd and January 4, 2019.

Table 7
Performance evaluation of the aggregated turbines' forecasts.

Wind Farm Aggregated Power Forecasts											
Metric	MIX		IFS		EURO		MEPS		Combined		
	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Both
RMSE [MW]	11.76	11.90	13.12	13.24	13.89	14.12	15.04	15.11	12.08	12.05	11.98
NRMSE [%]	7.82	7.91	8.73	8.80	9.23	9.39	10.0	10.05	8.03	8.01	7.97
MAE [MW]	6.99	6.81	7.84	7.58	8.23	8.02	9.48	8.82	7.34	6.98	7.06
NMAE [%]	4.65	4.53	5.21	5.04	5.48	5.33	6.31	5.86	4.88	4.64	4.69
IoA [%]	97.5	97.5	96.9	96.9	96.5	96.5	95.6	95.8	97.3	97.4	97.4
R^2 [%]	91.0	90.8	88.8	88.6	87.4	87.0	85.3	85.1	90.5	90.6	90.1

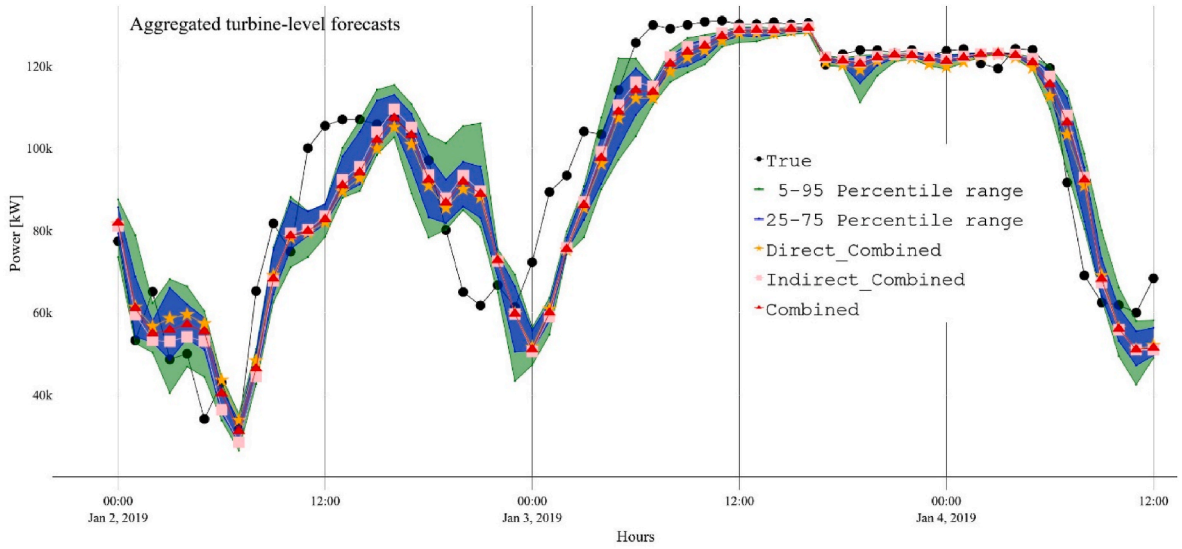

Fig. 12. Wind farm power forecast versus true power with a prediction interval.

Table 8
Error quantification of the aggregated WPFs.

Wind Farm-level error quantification											
Error [MW]	MIX		IFS		EURO		MEPS		Combined		
	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Both
Mean	0.9	1.9	0.7	1.8	0.8	2.0	0.8	2.4	1.0	2.0	1.6
Median	−0.4	0.02	−0.5	0	−0.6	0	−1.4	−0.1	−0.7	0	−0.2
SD	11.7	11.8	13.1	13.1	13.9	14.0	15.0	14.9	12.0	11.9	11.9
Q1	−3.2	−2.0	−3.7	−2.4	−3.8	−2.4	−6.0	−3.4	−3.6	−2.2	−2.7
Q3	3.6	4.5	4.0	4.8	4.2	5.1	4.4	5.7	4.1	4.8	4.4
IQR	6.8	6.5	7.7	7.2	8.0	7.5	10.4	9.1	7.7	7.0	7.1
Upper fence	13.8	14.3	15.5	15.5	16.1	16.2	20.0	19.3	15.6	15.3	15.0
Lower fence	−13.4	−11.8	−15.2	−13.1	−15.7	−13.5	−21.5	−16.9	−15.0	−12.7	−13.4

5. Conclusions

In this study, we have explored the relevance of using several NWP sources in power forecasting from the perspective of individual wind turbine and aggregated wind farm level forecasts. The direct and indirect forecast approaches were also compared. The forecast models were built by implementing a genetic optimization search to find the most optimum machine learning pipeline utilizing various possible algorithms and configurations. The conclusions can be summarized as follows:

- Using several NWPs could improve the accuracies in WPF irrespective of the direct or indirect approaches used. By this, normalized RMSE and MAE of various forecasts were reduced by (8%–22%), depending on the NWP and approach used.
- In the present study, there is no clear evidence proving that the direct WPF outperforms the indirect WPF, nor on the contrary.
- The wind downscaling model, which extrapolates the wind flow conditions to the hub coordinates of the turbine, could improve the WPF accuracies.
- The combined WPF, which resulted from either direct or indirect approaches, performed as accurately as the MIX WPF. This further confirms that using several NWP models yields a better forecast, irrespective of whether they are integrated into one “MIX” model or the forecasts based on individual NWPs are combined together as in the case of the discussed “combined” models.
- The point forecasts aggregated to the wind farm level yielded a lower-normalized error, which could be mainly due to the cancellation of positive and negative errors in the forecasts of individual turbines.

One specific advantage of the proposed point forecasting approach, in contrast with the farm level forecasts, is that it could be successfully implemented even if some of the turbines in the farm are down or are curtailed due to various reasons and operational settings.

The point (turbine level) forecasting approach proposed under this study is currently being compared with the farm level approach and the results of this comparison would be shared with the wind power forecasting community through a later publication. Further, the present study, which was on a land based wind farm over a relatively flat terrain, has to be extended to other operating environments like complex terrains and offshore conditions.

Credit authors statement

Ghali Yakoub: Conceptualization, Methodology, Software, Visualization, Investigation, Validation, writing the original draft.**Sathyajith Mathew:** Supervision, Conceptualization, Validation, Writing review and editing.**Joao Leal:** Co-Supervision, Conceptualization, Validation, Writing review and editing.

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.

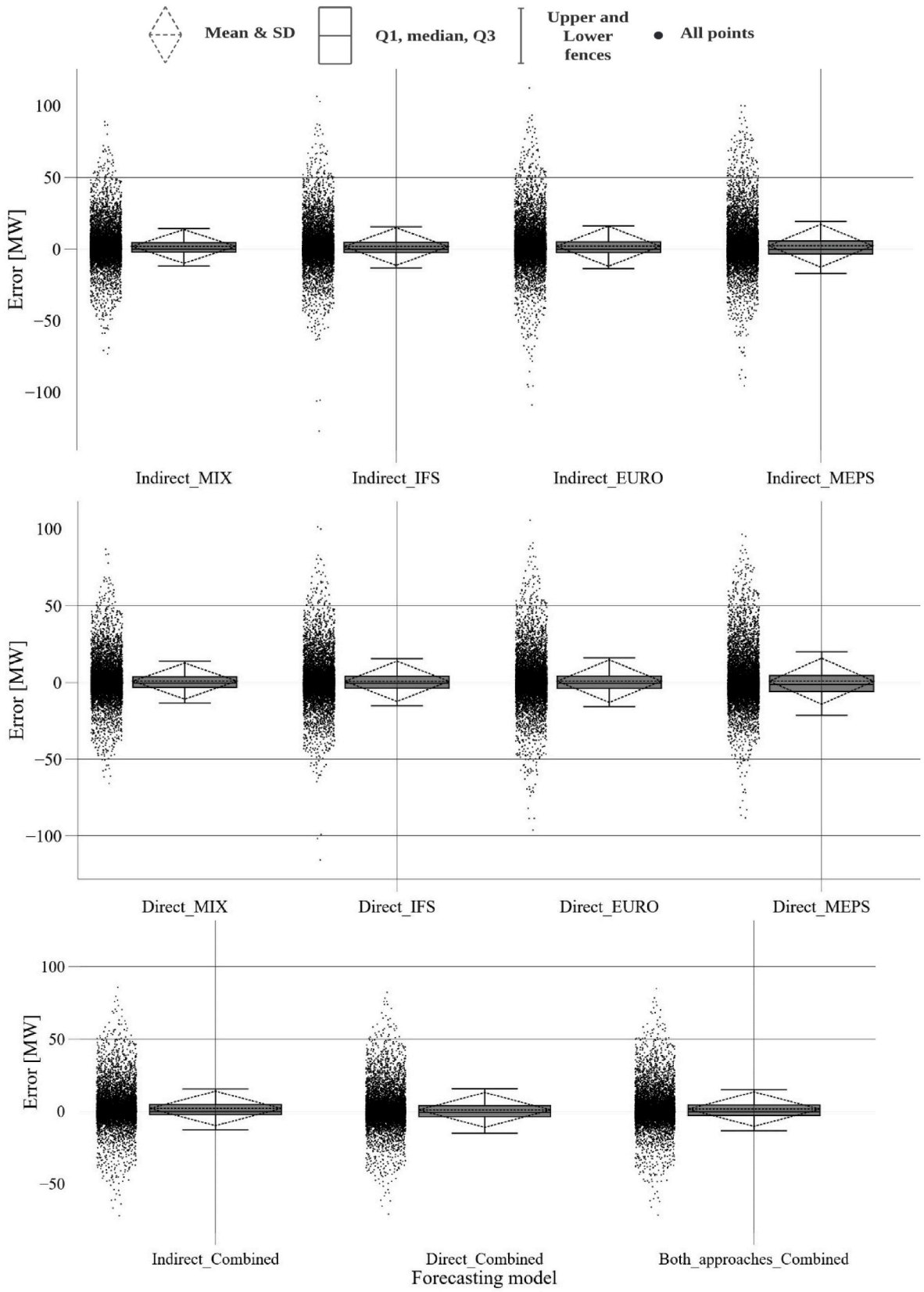

Fig. 13. Aggregated wind farm WPF's error quantification.

Data availability

The data that has been used is confidential.

Acknowledgement

The work presented in this paper is funded by the University of Agder, Norway (project number: 19/04543–1). The authors acknowledge Jørgen Olsen for his contributions to this research through fruitful interactions and discussions. Thanks are also due to Statkraft for providing the data required for this study.

References

[1] Monteiro C, et al. Wind power forecasting : state-of-the-art 2009. 2009. United States.
[2] Skajaa A, Edlund K, Morales JM. Intraday trading of wind energy. IEEE Trans Power Syst 2015;30(6):3181–9.
[3] The Power Market, Energy facts Norway 2021,[cited 2021]. The Power Market 2021 [Online], https://energifaktanorge.no/en/norsk-energiforsyning/kraftmarkedet/.
[4] Browell J, Gilbert C, McMillan D. Use of turbine-level data for improved wind power forecasting. In: 2017 IEEE Manchester PowerTech; 2017.
[5] Zhao X, Wang S, Li T. Review of evaluation criteria and main methods of wind power forecasting. Energy Proc 2011;12:761–9.
[6] Pinson P. Wind energy: forecasting challenges for its operational management. Stat Sci 2013;28(4):564–85. 22.
[7] Zhang Y, Wang J, Wang X. Review on probabilistic forecasting of wind power generation. Renew Sustain Energy Rev 2014;32:255–70.
[8] Yan J, et al. Reviews on uncertainty analysis of wind power forecasting. Renew Sustain Energy Rev 2015;52:1322–30.
[9] Hanifi S, et al. A critical review of wind power forecasting methods-past, present and future. Energies 2020;13.
[10] Liu H, et al. Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods. Energy Convers Manag 2019;195:328–45.
[11] Foley AM, et al. Current methods and advances in forecasting of wind power generation. Renew Energy 2012;37(1):1–8.
[12] Hong DY, et al. An indirect short-term wind power forecast approach with multi-variable inputs. In: 2016 IEEE innovative smart grid technologies - Asia (ISGT-Asia); 2016.
[13] Okumuş İ, Dinler A. Current status of wind energy forecasting and a hybrid method for hourly predictions. Energy Convers Manag 2016;123:362–71.
[14] Soman SS, et al. A review of wind power and wind speed forecasting methods with different time horizons. In: North American power symposium 2010; 2010.
[15] Lin Z, Liu X. Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network. Energy 2020;201:117693.
[16] Duarte Jacondino W, et al. Hourly day-ahead wind power forecasting at two wind farms in northeast Brazil using WRF model. Energy 2021;230:120841.
[17] Bremnes JB. On the use of NWP forecasts in wind power forecasts for the next few hours. In: METeports. The Norwegian Meteorological Institute; 2018. p. 11.
[18] Revheim PP. Improving and enhancing NWP based wind power forecasts under Norwegian conditions. In: Faculty of engineering and science. University of Agder Norway; 2015. p. 162.
[19] Lei M, et al. A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 2009;13(4):915–20.
[20] Bauer P, Thorpe A, Brunet G. The quiet revolution of numerical weather prediction. Nature 2015;525(7567):47–55.
[21] James EP, Benjamin SG, Marquis M. Offshore wind speed estimates from a high-resolution rapidly updating numerical weather prediction model forecast dataset. Wind Energy 2018;21(4):264–84.
[22] Shi J, Qu X, Zeng S. Short-term wind power generation forecasting: direct versus indirect Arima-based approaches. Int J Green Energy 2011;8:100–12.
[23] Sfetsos A. A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renew Energy 2000;21(1):23–35.
[24] Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Appl Energy 2010;87(7):2313–20.
[25] Marugán AP, et al. A survey of artificial neural network in wind energy systems. Appl Energy 2018;228:1822–36.
[26] Mana M, Burlando M, Meissner C. Evaluation of two ANN approaches for the wind power forecast in a mountainous site. Int J Renew Energy Resour 2017;7:1629–38.
[27] Piotrowski P, et al. Analysis of forecasted meteorological data (NWP) for efficient spatial forecasting of wind power generation. Elec Power Syst Res 2019;175:105891.
[28] Chen N, et al. Wind power forecasts using Gaussian processes and numerical weather prediction. IEEE Trans Power Syst 2014;29(2):656–65.
[29] Korprasertsak N, Leephakpreeda T. Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models. Energy 2019;180:387–97.
[30] Bokde N, et al. A novel and alternative approach for direct and indirect wind-power prediction methods. Energies 2018;11:1–19.
[31] Dione M, Matzner-Løber E. Short-term forecast of wind turbine production with machine learning methods: direct and indirect approach. In: Theory and applications of time series analysis. Cham: Springer International Publishing; 2019.
[32] Kariniotakis GN, Stavrakakis GS, Nogaret EF. Wind power forecasting using advanced neural networks models. IEEE Trans Energy Convers 1996;11(4):762–7.
[33] Landberg L. Short-term prediction of the power production from wind farms. J Wind Eng Ind Aerod 1999;80(1):207–20.
[34] Landberg L, Watson SJ. Short-term prediction of local wind conditions. Boundary-Layer Meteorol 1994;70(1):171–95.
[35] Yakoub G, Mathew S, Leal J. Downscaling and improving the wind forecasts from NWP for wind energy applications using support vector regression. J Phys Conf 2020;1618:062034.
[36] Maier HR, et al. Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. Environ Model Software 2010;25(8):891–909.
[37] Olson RS, et al. Automating biomedical data science through tree-based pipeline optimization. In: Applications of evolutionary computation. Cham: Springer International Publishing; 2016.
[38] Olson R, et al. Evaluation of a tree-based pipeline optimization tool for automating data science. 2016. p. 485–92.
[39] Le TT, Fu W, Moore JH. Scaling tree-based automated machine learning to biomedical big data with a feature set selector. Bioinformatics 2020;36(1):250–6.
[40] Abedinia O, et al. Improved EMD-based complex prediction model for wind power forecasting. IEEE Trans Sustain Energy 2020;11(4):2790–802.
[41] Statkraft. Smøla Wind Farm. 2021 [cited 2021 4/11]; Available from, https://www.statkraft.com/about-statkraft/where-we-operate/norway/smola-wind-farm/.
[42] The MetCoOp team CaUA. The MetCoOp ensemble MEPS. ALADIN-HIRLAM Newsletter No 2017;8(8):98–103.
[43] Pedregosa F, et al. Scikit-Learn: machine learning in Python. J Mach Learn Res 2011;12:2825–30.
[44] Buitinck L, et al. API design for machine learning software: experiences from the scikit-learn project. In: European conference on machine learning and principles and practices of knowledge discovery in databases; 2013. Prague, Czech Republic.
[45] Luo X, et al. Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy. IEEE Trans Ind Inf 2018;14(11):4963–71.
[46] Sideratos G, Hatziargyriou ND. An advanced statistical method for wind power forecasting. IEEE Trans Power Syst 2007;22(1):258–65.
[47] Zhao Y, et al. Correlation-constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting. IEEE Trans Power Syst 2018;33(5):5029–40.
[48] Mashaly AF, et al. Predictive model for assessing and optimizing solar still performance using artificial neural network under hyper arid environment. Sol Energy 2015;118:41–58.
[49] Willmott C, Robeson S, Matsuura K. A refined index of model performance. Int J Climatol 2012;32.