

Review

Evaluation Metrics for Wind Power Forecasts: A Comprehensive Review and Statistical Analysis of Errors

Paweł Piotrowski ^{1,*}, Inajara Rutyna ^{1,2}, Dariusz Baczyński ¹ and Marcin Kopyt ¹

¹ Electrical Power Engineering Institute, Warsaw University of Technology, Koszykowa 75 Street, 00-662 Warsaw, Poland

² Ideas NCBR Sp. z o.o., Nowogrodzka Street 47A, 00-695 Warsaw, Poland

* Correspondence: pawel.piotrowski@pw.edu.pl; Tel.: +48-22-234-7255

Abstract: Power generation forecasts for wind farms, especially with a short-term horizon, have been extensively researched due to the growing share of wind farms in total power generation. Detailed forecasts are necessary for the optimization of power systems of various sizes. This review and analytical paper is largely focused on a statistical analysis of forecasting errors based on more than one hundred papers on wind generation forecasts. Factors affecting the magnitude of forecasting errors are presented and discussed. Normalized root mean squared error (nRMSE) and normalized mean absolute error (nMAE) have been selected as the main error metrics considered here. A new and unique error dispersion factor (EDF) is proposed, being the ratio of nRMSE to nMAE. The variability of EDF depending on selected factors (size of wind farm, forecasting horizons, and class of forecasting method) has been examined. This is unique and original research, a novelty in studies on errors of power generation forecasts in wind farms. In addition, extensive quantitative and qualitative analyses have been conducted to assess the magnitude of forecasting error depending on selected factors (such as forecasting horizon, wind farm size, and a class of the forecasting method). Based on these analyses and a review of more than one hundred papers, a unique set of recommendations on the preferred content of papers addressing wind farm generation forecasts has been developed. These recommendations would make it possible to conduct very precise benchmarking meta-analyses of forecasting studies described in research papers and to develop valuable general conclusions concerning the analyzed phenomena.

Keywords: forecasting error; evaluation criteria metrics; wind power forecasting; wind turbine; wind farm; statistical analysis of errors; hybrid methods; ensemble methods; machine learning; deep neural network



Citation: Piotrowski, P.; Rutyna, I.; Baczyński, D.; Kopyt, M. Evaluation Metrics for Wind Power Forecasts: A Comprehensive Review and Statistical Analysis of Errors. *Energies* **2023**, *15*, 9657. <https://doi.org/10.3390/en15249657>

Academic Editor: Frede Blaabjerg

Received: 16 November 2022

Accepted: 16 December 2022

Published: 19 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The forecasting of power generation in wind farms has been an extensively explored research topic [1–8]. The growing significance of renewable energy sources (RES) and the remarkably dynamic growth of wind farms in most countries has highlighted the importance of accurate power generation forecasts due to, e.g., increased wind farm contribution to the overall power system. Cost-efficient and optimized management of a power system requires RES generation forecasts of the best possible accuracy. System operation optimization processes include scheduling the operation of fossil-based sources, scheduling maintenance works in the power grid, and preventive and remedial maintenance of RES themselves. In addition to obtaining forecasts with the best accuracy, the estimation of errors in these forecasts also proves to be important, as it translates into maintaining appropriate safety margins.

Forecasting purposes vary by time horizon [4,5,7]. Time horizon, also called planning horizon, is a fixed point in the future at which a certain process will be evaluated or assumed to have ended. In wind energy forecasting, time horizon affects the choice of forecasting

techniques, as they are classified into four types: very short-term (seconds), short-term (minutes to hours), medium-term (months), and long-term forecasts (months/years) [9]. Based on this classification, each approach is defined in Table 1.

Table 1. Forecasting horizons.

Horizon	Applications	Duration
Very short-term	Turbine regulation, control strategies, electricity market billing, real-time grid operation, network stability, voltage control	10 s
Short-term	Economic load dispatch planning, load amount, reversibility of power management, electric market reliability in day-ahead, load decisions for increments	5 min, 10 min, 15 min, 1 h, 24 h
Medium-term	Unit commitment, reserved requirement, generation operator, operation scheduling	1 month
Long-term	Maintenance scheduling, wind farm design, electricity market restructuring, optimization of operating costs	1 month, 1 year

Each time horizon tends to use different types of data. Very short-term forecasts are most often based on time series data. Short-term forecasts very often use online measurement data from a meteorological station, NWP, or combination of both as input data, with the expectation that weather conditions will remain the same in a short time.

1.1. Major Factors Affecting Wind Power Forecasts

Magnitudes of forecasting errors for wind power generation can vary widely. The quality of power generation forecasts is affected by a large number of independent factors based on quantities or classes. The final forecasting error can be seen as the output of a non-linear function that uses all the factors presented in Table 2 as inputs. Obviously, it is impossible to create a formal formula for such a function, although it is possible to verify how selected factors affect the magnitude of error. In some cases, it is possible to determine the functional relationship between the error and the factor, but, unfortunately, it requires a very large pool of research samples. In Table 2 modifiable and fixed factors are described. When the wind farm is already active and generating energy, the following factors are fixed: site, landscape, and size of the system.

Table 2. Description of major factors affecting wind power forecasts.

Factor	Description/Influence on Error Level
Time horizon	With increasing forecasting horizon, forecasting errors grow significantly, mainly due to the falling quality of NWP forecasts [5,6,8,10,11]
Forecasting method (complexity)	Complex (ensemble or hybrid) models have typically lower forecasting errors than single methods (more details in Section 2) [1–5,8]
Size of system	The inertia of power production usually grows with system size. This translates into more predictable production, especially in shorter horizons.
Site (onshore/offshore)	Forecasting errors for offshore wind farms should be less than for onshore wind farms due to distinctive characteristics of weather conditions (more stable and higher wind speeds).
Landscape	Forecasting errors for farms located in rough terrain can depend on the local landscape and meteorological features (e.g., a forest, hills, or a lake in a direct neighborhood). The best if the terrain has as little roughness as possible—this guarantees the optimal generation of wind energy [12]
Location of NWP forecasting points	NWP forecasts from points more distant from the wind farm can generate larger forecasting errors. The topic of optimum selection of NWP forecasting points and their location on the farm is subject to studies [12]

Table 2. Cont.

Factor	Description/Influence on Error Level
Types and quantities of input data	The more information related to power generation can be used in the model, the more accurate generation forecasts can be expected. In particular, using NWP as input data is especially important (if absent, generation forecast errors grow very sharply), and with horizons of more than several hours, NWPs are virtually indispensable [1,2,4,13]
NWP data sources	NWP data can vary by quality (forecasting accuracy)—the more accurate NWP forecasts, the lower power generation forecasting error [2]
Quantity of training sets	The model must use training data encompassing at least one year. With a growing number of years, the model uses more information and better represents the seasonality and daily variability of the process (a smaller forecasting error can be expected)
Data preprocessing	Properly conducted steps to clean up and process raw input data can reduce forecasting error [3]
Data postprocessing	Elimination of impossible situations, e.g., negative power forecasts or forecasts determined as unlikely for specific input data. Elimination of such cases reduces the error [13]
Measurement data availability lag	Using up-to-date, current data provides for identification and correction of errors “on the fly,” e.g., using switchable forecasting models

More attention should be paid to forecast model inputs due to the fact that, whereas some factors, such as location, size of the system, or forecasting horizon, cannot be changed, the biggest reduction in error can be achieved by appropriate selection of input data or such selection that encompasses as much information as possible related to the forecast power generation time series.

Regarding the selection of input data itself, both statistical analysis and a semi-machine approach can be applied. Statistical analysis using various tests can help to draw conclusions on data interdependencies, however, the time required for this makes it impractical for big datasets. If larger quantities of data are available, expert selection of a pool of input data combinations and their subsequent review can prove to be a more practical approach. The choice of solution depends significantly on the tool—the time required for statistical analysis can bring more benefits for tools that usually require greater optimization, e.g., ANN models.

Forecasting models require input data to predict wind power generation. The data format used by forecasting models need to be relevant to the model itself, i.e., it must consider which external phenomena have a direct impact on wind generation. This data can be divided into NWP and time series [14].

NWP is a multivariate dataset based on a set of physical models used to simulate conditions in the atmosphere; these models are available both on local and global scales. NWP dataset contains information generated by power metering and prediction of several meteorological variables, such as wind speed, wind direction, temperature, humidity, air pressure, time of the day, day of the year, etc. [15,16]. As NWP is a general dataset, the main factor that affects wind-generated power is wind speed [17,18].

The time series is a univariate dataset of a wind speed or wind power that is measured at timestamps over a certain period. To obtain a wind speed time series, a mast is usually installed at the wind farm, with an anemometer mounted at the hub height.

Decomposition methods often used during the forecasting process are based on the premise that the wind power time series contain different frequency signals with different characteristics, and that modeling each of the decomposed series separately can lead to an overall improvement in forecast quality [8]. Popular techniques are discrete wavelet transform (DWT), empirical mode decomposition (EMD) [3,8], ensemble empirical mode

decomposition (EEMD) [6], variational mode decomposition (VMD) [4,8], or wavelet packet transform (WPT).

Machine learning models can use NWP data as features, and/or time series as inputs, where NWPs are features with information related to the expected output wind speed. Input matrix X contains the following: historical wind inputs, weather information, and time data. Output vector y contains the time series and multiple values of predicted wind power with a changing prediction horizon [19].

When NWP data is not available to be used as input data, common practice is to use autoregression of the output variable. In some works, measured values of other variables from the recent past are also used—e.g., measured instantaneous rotor speed or measured weather parameters for a few hours in the past [20,21]. This allows us to take into account recent generation trends and a farm's short-term generation inertia. A practical drawback of this appeal would be that it cannot be used for forecasts with horizons longer than a few hours ahead. Other possible approaches would be using measures from one farm to forecast another one [22]. In this case, the final result would depend on the weather similarity between used and destination farms and tools used to translate one generation to another.

When only the wind speed time series is known, a technique called feature engineering can be used to fabricate new features. This technique's goal is to fabricate features by executing simple calculations based on the known feature, the wind speed time series. These calculations are described by standard deviation, average, minimum, or maximum wind speed for a period of time [20,23].

In practice, the choice of the source and types of NWP forecasts strongly depends on the *data cost-to-quality* ratio, nevertheless, it is worthwhile to maximize both the number of NWP models with various densities of forecasting points and the number of weather parameters derived from them [12]. In some cases, the application of various models, or even bundles of models, is recommended due to the diverse information content of different models. For instance, for long-term forecasts, NWP climatic models are different from each other, and identification of the best one can be not only difficult, but virtually impossible, and drawing any conclusion can require aggregation analysis of various scenarios. On the macro scale, equally important is to select a proper forecasting point, from which meteorological variables would be derived. Hence, the growing trend of extraction of spatial information using various tools, e.g., CNN, was applied by some research papers [24–26].

1.2. Objective and Contribution

The main objectives of this paper can be summarized as follows:

- classify wind power forecasting techniques;
- provide unique description of major factors affecting wind power forecasts;
- describe the performance of forecasting models;
- conduct comprehensive review (quantitative analysis) based on more than one hundred papers;
- conduct statistical analysis of errors (qualitative analysis).

Below are listed selected contributions of this paper:

- proposal for a novel, unique ratio, called EDF;
- analysis of variability of the new EDF ratio depending on selected characteristics (size of wind farm, forecasting horizon, and class of forecasting method) and the original, novel conclusions drawn from the analysis;
- development of a unique list of recommended content of papers addressing wind farm generation forecasts (the application of these recommendations would make it possible to conduct very accurate meta-analyses that would compare various forecasting studies).

The remainder of this paper is organized as follows: Section 2 presents the classification of wind power forecasting techniques. Section 3 describes the performance of the

forecasting model. Section 4 is the main part of the paper and includes comprehensive statistical analysis (quantitative and qualitative analysis). Discussion is provided in Sections 5 and 6 draws the main conclusions. References are listed at the close of this paper.

2. Classification of Wind Power Forecasting Techniques

The following alternative methodologies are applied to wind power forecasting: Naive, Physical, Statistical, and AI/ML methods (see Table 3).

Table 3. Classification of forecasting techniques.

Technique	Description	Advantage	Disadvantage
Naive/Persistence Methods	Simple approach that evaluates the last period t , benchmark for more complex methods	Fast results, applicable only for short-term forecasts	Does not consider correlations between input data and the results are not reliable for the subsequent steps in time
Physical Methods	Physical mathematical model of the power turbine	No need for training or historical data, enables understanding of physical behavior	Computationally complex, it relies on fresh data from several features, extensive setup effort
Statistical/Multivariate Methods (time series)	Mapping of relations between features, generally through recursive techniques, no NWP data as an input	Models can be built easily and with small computational effort	Quick loss of accuracy with time
AI/ML methods (LR, GBT, RF, BPNN, KNNR, LSTM, CNN, DBN, GRU, EDCNN, RBF, SVR)	Techniques that enable the computer to mimic human behavior without a predefined mathematical model ("black box")	Highest accuracy and can learn more complex nonlinear relations	Slower convergence speed, risk of overfitting, computational complexity

Statistical models have high precision in very short-term prediction [27]. The most used statistical model for wind forecasting is the times series model, due to the fact that future levels of wind power depend on weather features, but they also can depend on the prior value of wind power generated. The amount of wind power produced in the current hour affects the amount of wind power generation in the next hour. These models can determine conditions in time based on relationships between parameters. However, they depend on pre-set coefficient values.

AI and ML models are suitable for systems that are more complex to model, as they attempt to discover underlying relationships, and are widely used to accurately predict wind. Without an a priori structural hypothesis that relates wind power to several historical meteorological variables, they have a strong generalization and fast speed [18,28].

Each approach mentioned above can have a high forecasting error due to inherent weaknesses, especially when wind speeds have significant non-linear characteristics, as volatility causes complex fluctuations. In particular, the conventional single ANN model has the drawback of falling into local minimum and overfitting, and its performance can be influenced by the initial parameters. These weaknesses cannot be easily remedied with a single method. To reduce forecasting error and obtain advanced models that can achieve higher accuracy, a combination of methods described in Table 4 is incorporated.

Ensemble forecasting methods are generated through the application of various machine learning techniques and then by merging the outputs, which reduces the risk of overestimation and is aimed at preserving the diversity of models. The ensemble technique is known to be applied in both cooperative and competitive styles.

In a cooperative ensemble, the dataset is divided into data subsets, each subset being forecast individually and then aggregated with other sub-forecasts [29]. This technique is computationally lightweight due to less need for parameter tuning and is in general used for very short-term or short-term forecasting.

Table 4. Classification of complex forecasting techniques.

Technique	Description	Advantage	Disadvantage
Ensemble methods	Aggregate a combination of results from different methods	Usually perform better than a single method	Larger computational cost of running individual methods; the combination of results can obscure problems inherent in the methods
Hybrid methods	Combination of different methods connected in series to create hybrid prediction structures.	Usually perform better than a single method	Larger computational cost of running individual methods; require large quantities of data

Competitive ensembles build individual forecasting models with different parameters and initial values, and the results are obtained by aggregation of forecasts by different techniques, such as the Bayesian model average. This technique, used by [30], can cover a larger dataset and is used to achieve early detection of a large wind ramp before the changes in the wind speed propagate to other locations. However, it is considered computationally expensive and is mostly used in medium-term and long-term forecasting.

To obtain an advanced model with higher accuracy, hybrid forecasting models combine the advantages of different methods with individual superior features [31]. Overall forecasting effectiveness of hybrid methods can be improved, since hybrid methods can overcome the limitations and take advantage of the merits of individual models by integrating two or more types of models [28].

A neural network can be used in different steps of the algorithm, for example, a CNN-based model using transfer learning is used to address the problem of some newly constructed farms not having sufficient historical wind speed data to train a well-performing model by producing synthetic data [32]. In [26], the CNN is trained in layers to extract local features and relationships between the nodes, and the output layer of CNN is set in multiple dimensions to directly forecast future wind speed.

The most common approach is to adopt the machine learning algorithm as the main forecasting tool and to perform data treatment using general techniques as shown by [33], which consist of variational mode decomposition (VMD) of raw wind power series into a certain number of sub-layers with different frequencies; the K-means as a data mining approach being executed for splitting the data into an ensemble of components with a similar fluctuant level of each sub-layer; and LSTM is adopted as the principal forecasting engine for capturing unsteady characteristics of each component.

Some authors also combine both hybrid and ensemble approaches into one [34], using a hybrid technique of intelligent and heuristic algorithms that include neural networks, wavelet transform, diverse heuristic algorithms, and fuzzy logic. The hybrid technique uses wavelet transform to filter distortions and noise in wind power signals, the radial neural networks (RBF) technique being used as a preliminary predictor to find local solutions. With the local solution, an ensemble combining three neural networks of MLP using various learning methods along with heuristic WIPSO is used for the final prediction and modeling of the non-linear behavior of the wind power curve.

3. Performance of Forecasting Model

3.1. RMSE, MAE, and MAPE as Frequently-Used Metrics

The root mean square error (RMSE), given by Formula 1, is a quadratic scoring rule that estimates the average magnitude of error. It is the most standard function used to calculate the difference between predicted and observed values, since it reflects the level of differences between the actual and forecast values, in other words, the absolute magnitude

of prediction error [35]. However, RMSE is sensitive to outliers, so its outcome can be biased if the data is not clean [36].

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (1)$$

where \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. A smaller RMSE means that the proposed model performs better.

The mean absolute error (MAE), Equation (2), corresponds to the estimated level of absolute error. This level indicates the average magnitude of the actual value and the predicted value [37].

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2)$$

where \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. MAE is not susceptible to outliers and can better reflect the actual status of predicted errors [38]. The model is deemed to be accurate when MAE is close to zero.

The mean absolute percentage error (MAPE), Equation (3), calculates the percentage error relative to the actual value, which is stated as the average ratio, and is also commonly used to compare different models [36,39].

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\% \quad (3)$$

where \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples.

Although RMSE is usually used to express the dispersion of the results, MAE and MAPE can indicate the deviation of the prediction [17]. The smaller the values of RMSE, MAE, and MAPE, the more accurate the forecasting model.

3.2. MSE, nMAE, nRMSE, and R2 As Occasionally Used Metrics

The mean squared error (MSE), Equation (4), simply averages the mean squared difference between the estimated and original parameters [40], which can avoid the problem that the errors cancel each other out, and accurately reflects the actual prediction error [35].

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4)$$

where \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples.

Sometimes authors need to normalize the MAE and RMSE to quantitatively examine the prediction performances of some models, their norms being given by the normalized mean absolute error (nMAE), Equation (5), and normalized root mean squared error (nRMSE), Equation (6).

$$nRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{y_i - \hat{y}_i}{C_i} \right)^2} \quad (5)$$

$$nMAE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{C_i} \right|, \quad (6)$$

where C_i is the operating capacity of time point i , \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. In general, smaller

values of these metrics indicate that the corresponding solution offers less deviation of prediction performance [41].

The R-square or coefficient of determination (R2), Equation (7), is the proportion of the variance in the dependent variable that is predictable from independent variable(s) [42,43]. It indicates the level of correlation between predicted value and the actual value, and it helps to select the best model with highest forecasting accuracy [44]. It is mostly used in datasets of large amplitudes [17].

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \text{ for } \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (7)$$

where \bar{y} is the average of actual values, \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. An R^2 closer to one indicates more accurate forecasting. It can also be displayed as negative to denote an arbitrarily worse predicting model [20].

3.3. R, PICP, PINAW, sMAPE, MRE, and TIC As Seldom Used Metrics

The Pearson linear correlation coefficient (R or CC), Equation (8), is a metric that determines the relationship between inputs and outputs by determining the linear dependence between results and observations [9,37].

$$R = \frac{\sum_{i=1}^N (x_i - \hat{x}_i)(y_i - \hat{y}_i)}{\sqrt{\sum_{i=1}^N (x_i - \hat{x}_i)^2 \sum_{i=1}^N (y_i - \hat{y}_i)^2}} \quad (8)$$

where \hat{y}_i and \hat{x}_i are predicted values, y_i and x_i are the actual values, and N is the number of prediction points or number of samples. The possible R score range can vary between 1 and -1 , with 1 representing the biggest correlation, and -1 the lowest one [45].

Prediction interval coverage probability (PICP), Formula (9), measures the ability of the constructed confidence interval to cover the target values for prediction intervals.

$$\text{PICP} = \frac{1}{N} \sum_{i=1}^N \rho_i, \text{ for } \rho_i = \begin{cases} 1, & \text{if } y_i \in [L_i, U_i] \\ 0, & \text{else} \end{cases} \quad (9)$$

where L_i and U_i are the lower bound and the upper bound, respectively, of the prediction values, y_i is the actual value, and N is the number of prediction points or number of samples. The greater the PICP, the more reliable the prediction values [46,47].

Prediction interval normalized average width (PINAW), Equation (10), is used to measure the width of the PIs for a given length of the prediction interval.

$$\text{PINAW} = \frac{1}{N} \sum_{i=1}^N \frac{(U_i - L_i)}{t_{max} - t_{min}} \quad (10)$$

where t_{min} and t_{max} are the maximum and minimum values of the predicted values, and L_i and U_i are the lower bound and the upper bound, respectively, of the prediction values [48].

The Symmetric Mean Absolute Percentage Error (sMAPE) metric, Formula (11), a variation of MAPE, is used to describe the relative error of a set of forecasts and their labels as a percentage [36,37].

$$\text{sMAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i + \hat{y}_i} \right| \cdot 200\% \quad (11)$$

where \bar{y} is the average of actual values, \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples.

Theil's inequality coefficient (TIC), Equation (12), is used to measure the predictive performance of the model [49].

$$\text{TIC} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N \bar{y}_i^2} + \sqrt{\frac{1}{N} \sum_{i=1}^N \hat{y}_i^2}} \quad (12)$$

where \bar{y} is the average of actual values, \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. The smaller the TIC value, the stronger the prediction ability [50].

The mean relative error (MRE), Equation (13), calculates the magnitude of the difference between predicted and actual values [51].

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (13)$$

where \bar{y} is the average of actual values, \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples.

The mean bias error (MBE), Equation (14), gives the average bias error of prediction. It is used to determine if the predicted value is underestimated <0 or overestimated >0 [1,2,52].

$$\text{MBE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i) \quad (14)$$

where \hat{y}_i is the predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. This metric is useful to identify the need to add extra steps to calibrate the model.

3.4. Interesting Usage of Other Metrics

Some studies use prediction accuracy metrics such as MSE, combinations of MAE, and RMSE to create the fitness function, as the fitness function directly affects the convergence of the algorithms and the optimal solution [17,35,49].

The MAPE is a commonly used evaluation metric that generates infinite values when the actual value y_i is zero or close to zero. To avoid this problem, mean arctangent absolute percentage error (MAAPE) is used, Equation (15).

$$\text{MAAPE} = \frac{1}{N} \sum_{i=1}^N \arctan \left| \frac{y_i - \hat{y}_i}{y_i + \hat{y}_i} \right| \quad (15)$$

where MAAPE ranges from 0 to $\frac{\pi}{2}$, \bar{y} is the average of actual values, \hat{y}_i is predicted value, y_i is the actual value, and N is the number of prediction points or number of samples. A smaller MAAPE indicates smaller forecasting error [9].

To compare the predictive performance of the models, promoting percentages (P) are applied in different metrics, Equation (16).

$$\text{PMETRIC} = \left| \frac{\text{METRIC}_1 - \text{METRIC}_2}{\text{METRIC}_1} \right| \quad (16)$$

where METRIC₁ and METRIC₂ are the error metrics calculated for two different prediction models. The promoting percentages are called PMAE, PMAPE, PRMSE, PNMAE, PNRMSSE, and PSMAPE [26,33,48,53–55].

The metrics MAE, MSE, and RMSE are usually used in deterministic forecasting methods. As for probabilistic forecasting, the process can be more complicated, due to the influence of external factors leading to a better analysis based on the verification of the quantile forecasts given by PICP and PINAW [56].

For a comparative assessment of the performance test of the analyzed methods, the skill score (SS) metric is useful. The skill score metric uses one nRMSE (Equation (17)) or nMAE metric (Equation (18)) or two error metrics—nRMSE and nMAE—and in this case, it is calculated by Equation (19) [2]. Higher SS values are an indication of superior prediction quality. An advantage of using a skill score is the ability to compare the forecasting qualities of various systems, using the level of reduction in forecasting error relative to the reference method as the quality indicator (persistence method—naive model).

$$SS_{RMSE} = \left(1 - \frac{nRMSE_{forecast}}{nRMSE_{reference}} \right) \quad (17)$$

where $nRMSE_{forecast}$ is the error of the analyzed method, and $nRMSE_{reference}$ is the error of the reference method (persistence method—naive model).

$$SS_{MAE} = \left(1 - \frac{nMAE_{forecast}}{nMAE_{reference}} \right) \quad (18)$$

where $nMAE_{forecast}$ is the error of the analyzed method and $nMAE_{reference}$ is the error of the reference method (persistence method—naive model).

$$SS_{RMSE, MAE} = \frac{1}{2} \left[\left(1 - \frac{nMAE_{forecast}}{nMAE_{reference}} \right) + \left(1 - \frac{nRMSE_{forecast}}{nRMSE_{reference}} \right) \right] \quad (19)$$

4. Comprehensive Statistical Analysis

Out of 106 papers, statistical analysis was conducted on those which applied nRMSE and nMAE errors and which could calculate these two error metrics based on the rated power of the system and the levels of RMSE and MAE errors. In addition, based on the content of those papers, crucial details (factors) of studies were selected to enable statistical quantitative analysis and error analysis and their relationship with other factors. Table 5 (onshore systems, data from 60 papers) and Table 6 (offshore systems, data from six papers) contain sets of selected information from the studies presented in the papers.

Table 5. Summary of errors obtained for onshore wind power forecasting models.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
1.	Wind power forecasting based on daily wind speed data using machine learning algorithms [23] Proposed Method—	1 year	1 MW	RF (ensemble)	0.0302 *	0.0070 *	Daily wind speed, mean wind speed, standard deviation, total generated wind power values—Turkish State Meteorological Service	
				XGBoost (ensemble)	0.0344 *	0.0065 *		
2.	Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm [17] Proposed Method—IDA-SVM	48 h	2050 KW	IDA-SVM (hybrid)	0.0524	0.0404	Wind power, wind speed, wind direction, and temperature—La Haute Borne wind farm	
				DA-SVM (hybrid)	0.0617	0.0515		
3.	A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition [24] Proposed Method—EMD-VMD-CNN-LSTM	1, 2, 3 steps	? MW	EMD-VMD-CNN-LSTM (Dataset #1, avg, hybrid)	0.0329 *	0.0227 *	Three sets of hourly averaged wind power, wind speed, and wind direction time series, Sotavento Galicia wind farm	
				EMD-VMD-CNN (Dataset #1, avg, hybrid)	0.0396 *	0.0321 *		
4.	A Model Combining Stacked Auto Encoder and Back Propagation Algorithm for Short-Term Wind Power Forecasting [57] Proposed Method—SAE_BP	1, 2, . . . , 9 steps	1500 MW	SAE_BP (avg, hybrid/ensemble)	0.0941 *	0.0738 *	Wind generation from Ireland all island—EirGrid Group	
				SVM (avg, single)	0.1396 *	0.1185 *		
5.	Deep belief network based k-means cluster approach for short-term wind power forecasting [16] Proposed Method—DBN	10 min	17.56 MW	DBN (single)	0.0322	0.0236	Wind speed, wind direction, temperature, humidity, pressure, history wind speed, history wind power—NWP meteogalicia and Sotavento wind farm	
				BP (single)	0.0580	0.0446		
6.	Multi-distribution ensemble probabilistic wind power forecasting [29] Proposed Method—MDE	1, 6, 24 h	16 MW	Q-learning (6 h, single)	0.2355	0.1841	Meteorological information (many features), synthetic actual wind power, and wind power forecasts generated by the Weather Research and Forecasting (WRF) model—MDE probabilistic forecasting framework, from the Wind Integration National Dataset (WIND) Toolkit	
				NWP (24 h, single)	0.1817	0.1337		
7.	Forecasting energy consumption and wind power generation using deep echo state network [36] Proposed Method—DeepESN—Stacked Hierarchy of Reservoirs	10 steps	6910 million KWh	DeepESN (hybrid/ensemble)	0.0326 *	0.0247 *	?—historical WPG data in Inner Mongolia	
				BP (single)	0.0957 *	0.0764 *		
8.	Feature Extraction of NWP Data for Wind Power Forecasting Using 3D-Convolutional Neural Networks [25] Proposed Method—3D-CNN—Three Dimensional Convolutional Neural Network	30 min—72 h	Normalized data	NAÏVE (single)	0.2033 *	0.1593 *	Wind power from every 10 s, from April 2015 to July 2017 + NWP—wind farm in Tohoku region	
				3D-CNN (20 h, single)	0.210 **	0.150 **		
				2D-CNN (20 h, single)	0.212 **	0.156 **		
9.	A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting [58] Proposed Method—Gated Recurrent Unit Neural Networks	24 h	?	Proposed (single)	0.1345	0.0687	Wind speed, wind power, and NWP wind speed data, sampled at a period of 15 min—wind farm in Sichuan Province	
				ANN (single)	0.1350	0.0680		

Table 5. Cont.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
10.	Medium-term wind power forecasting based on multi-resolution multilearner ensemble and adaptive model selection [9] Proposed Method—MRMLE-AMS—Multilearner Ensemble and Adaptive Model Selection	6, 12, 18 h Step = 6 h	200 MW	MRMLE (dataset #4, step 3, hybrid)		0.1254 *	0.0943 *	Actual wind power series in Canada as experimental datasets
				Hao's model (dataset #4, step 3, hybrid/ensemble)		0.1554 *	0.1245 *	
				NAÏVE (dataset #4, step 3, single)		0.4218 *	0.3226 *	
11.	Short-term wind power prediction based on improved small-world neural network [59] Proposed Method—SWBP—Small-World BP neural network	15 min	Normalized data	SWBP (single)	0.0820	0.0601	Wind speed and power, air temperature, barometric pressure, relative humidity, wind velocity, and direction at the height of 10, 30, 50, and 70 m—Jiangsu wind farm	
				BP (single)	0.1072	0.0841		
12.	Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network [31] Proposed Method—OTSTA-SSA-HLNN—opposition transition state transition algorithm-SSA-HLNN	1, 3, 6, 12 steps	Approx. 1500 KW (graph)	OTSTA-SSA-HLNN (experiment1, hybrid)	0.0612	0.0923	?—Wind farm in Xinjiang	
				STA-SSA-HLNN (experiment1, hybrid/ensemble)	0.0727	0.1101		
13.	Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and Correntropy Long Short-term memory neural network [50] Proposed Method—IVMD-SE-MCC-LSTM—Improved VMD-SE-MCC-LSTM	?	1402.7 KW	IVMD-SE-MCC-LSTM (hybrid)	0.0347 *	0.0241 *	Wind power series from 1 to 4 parameters—Gansu and Dingbian wind farm	
				EMD-SE-MCC-LSTM (hybrid)	0.0370 *	0.0270 *		
				VAPWCA-VMD-RMKPINN (3h, 150 samples, hybrid)	0.0310 *	0.0261 *		
14.	Short term wind power forecasting using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network [44] Proposed Method—VAPWCA-VMD-RMKPINN—VMD based kernel regularized pseudo inverse neural network	10, 30 min and 1, 3 h	Approx. 30 MW (graph)	VAWC-A-EMD-RMKPINN (3h, 150 samples, hybrid)	0.0337 *	0.0274 *	Wind power—Wyoming wind farm	
				LSTM-RNN (24 h, single)	0.1043	0.0765		
15.	Short-term wind power forecasting using long-short term memory based recurrent neural network model and variable selection [27] Proposed Method—LSTM-RNN	1 to 24 h	17.56 MW				Surface pressure, temperature, wind speed at 10 M, 35 M, 100 M, 170 M, wind direction at 35 M, 170 M—Sotavento Galicia wind farm NWP from Weather & Energy PROGnoses (WEPROG)	

Table 5. Cont.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
16.	A novel hybrid model for short-term wind power forecasting [35] Proposed Method—The Proposed Hybrid Model	1, 2, 3 steps	2.5 KW	The Proposed Hybrid Model (set 1, 3 steps, single)		0.0773 *	0.0570 *	Wind speed—Sotavento Galicia wind farm
				EMD-ENN (set 1, 3 steps, single)		0.0793 *	0.0625 *	
				NAÏVE (set 1, 3 steps, single)		0.1280 *	0.1020 *	
17.	A novel hybrid technique for prediction of electric power generation in wind farms based on WIPSO, neural network and wavelet transform [34] Proposed Method—RBF + HNN + WT + WIPSO	24, 72 h	254 MW	RBF + HNN + WT + WIPSO (72 h, winter, hybrid)		0.0201	0.0173	Wind power, temperature, wind direction, wind speed, humidity, and air pressure—wind farms in Alberta, Canada
				RBF + HNN + WT + MPSO (72 h, winter, hybrid)		0.0250	0.0200	
				LD-PCA-IPSO-BPNN (hybrid)		0.0095 *	0.0082 *	
18.	Wind Speed Prediction of IPSO-BP Neural Network Based on Lorenz Disturbance [43] Proposed Method—LD-PCA-IPSO-BPNN	10 min ?	?	PCA-IPSO-BPNN (hybrid)		0.0269 *	0.0235 *	Wind speed, wind direction, temperature, air pressure, specific volume, specific humidity, and surface roughness—Sotavento wind farm
				LSTM (wind speed correction step *, hybrid)		0.0181 *	0.0090 *	
19.	Multi-step Ahead Wind Power Forecasting Based on Recurrent Neural Networks [53] Proposed Method—LSTM/GRU with wind speed correction	1, 2, 3, 4 steps step: 10	3 MW	GRU (wind speed correction step *, hybrid)		0.0185 *	0.0096 *	Wind turbine state, rotation rate, generating capacity, sine of wind direction, cosine of wind direction, pitch angle, temperature, and wind power—wind turbine in North China and NWP
				Proposed (7 days, hybrid)		0.1942 *	0.2762 *	
				MLP (7 days, single)		0.4704 *	0.3519 *	
20.	A New Prediction Model Based on Cascade NN for Wind Power Prediction [60] Proposed Method—Proposed	10, 20, ..., 50, 60 min 1, 2, ..., 7 days	17.56 MW	NAÏVE (7 days, single)		0.7204 *	0.5108 *	Hourly wind speed and power forecast—Sotavento wind farm

Table 5. Cont.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
21.	A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets [61] Proposed Method—Optimized WT-FS-LSTM by CSA and by PSO	1 step Step: approx. 240 (graph)	Normalized data	Optimized WT-FS-LSTM by CSA (hybrid)	0.1536	0.1217		Hourly wind speed—Sotavento wind farm
				Optimized WT-FS-LSTM by PSO (hybrid)	0.1621	0.1306		
				WT-FS-LSTM (hybrid)	0.1655	0.1311		
22.	Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network [54] Proposed Method—EWT-LSTM-Elman	1, 2, 3 steps	?	EWT-LSTM-Elman (3 steps, hybrid)	0.065 *	0.049 *		Wind speed series
				EWT-BP (3 steps, hybrid)	0.077 *	0.058 *		
23.	A Comparative Study of ARIMA and RNN for Short Term Wind Speed Forecasting [62] Proposed Method—	?	?	RNN (dataset 1, single)	0.0942 *	0.0750 *		Hourly mean wind speed—Coimbatore meteoblue weather website
				ARIMA (dataset 1, single)	0.1058 *	0.0917 *		
24.	Can we improve short-term wind power forecasts using turbine-level data? A case study in Ireland [63] Proposed Method—VMD-ELM with WT-data	1,2, . . . , 8 h	24,923 KW	VMD-ELM with WT-data	0.0773	0.0564		Wind power, WT-data, or WF-data—wind farm in West and Southwest of Ireland
				VMD-ELM with WF-data	0.0811	0.0605		
25.	Short term forecasting based on hourly wind speed data using deep learning algorithms [64] Proposed Method—	?	Normalized data	LSTM (October–December, single)	1.236	0.8742		Hourly wind speed, air temperature, pressure, dew point, and humidity—metrological wind dataset from the location
				ELM (October–December, single)	0.9215	0.3978		
26.	Hybrid Approach for Short Term Wind Power Forecasting [65] Proposed Method—Proposed Method	24 h	2100 KW	Proposed method (December, hybrid)	0.0162 *	0.0077 *		Wind speed—Jodhpur wind farm
27.	Very Short-Term Spatial and Temporal Wind Power Forecasting: A Deep Learning Approach [66] Proposed Method—CST-WPP—convolution-based spatial-temporal wind power predictor	5–30 min	Normalized data	CST-WPP (30 min)	0.0749	0.0494		5 min mean wind power data of 28 wind farms—Australian Energy Market Operator
				VAR (30 min, single)	0.0861	0.0651		
				NAÏVE (30 min, single)	0.1174	0.0974		

Table 5. Cont.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	-	nRMSE	
28.	Short-Term Wind Speed Forecasting Using Ensemble Learning [67] Proposed Method—	48 h	?	Bagged trees (December, ensemble)	0.138 *	0.107 *	10 min wind speed data—modern-era retrospective analysis for research and applications, version 2 (MEERA-2)	
				Boosted trees (December, ensemble)	0.147 *	0.113 *		
29.	Seasonal Self-evolving Neural Networks Based Short-term Wind Farm Generation Forecast [68] Proposed Method—SSEN—Seasonal Self-Evolving Neural Networks	10 min, 1 h	300.5 MW	SSEN (10 min, ensemble)	0.0302	0.0170	Power outputs, wind speed and direction—National Renewable Energy Laboratory (NREL) and Xcel Energy	
				LSTM (10 min, single)	0.0307	0.0180		
30.	Markov Models based Short Term Forecasting of Wind Speed for Estimating Day-Ahead Wind Power [69] Proposed Method—	24 h	2100 KW	Markov chain model (Group 1, single)	0.0714 *	0.0620 *	Wind speed—Dark Sky API	
31.	Day-Ahead Wind Power Forecasting Based on Wind Load Data Using Hybrid Optimization Algorithm [49] Proposed Method—VMD-mRMR-FA-LSTM	24 h	Approx. 85 MW (graph)	VMD-mRMR-FA-LSTM (hybrid)	0.0358 *	0.0348 *	Wind power, load, temperature, air pressure, humidity, wheel height, wind speed and wind direction at 10m, 30m, 50m and 70m—Beijing Lumingshan Wind Power Plant	
				mRMR-FA-LSTM (hybrid)	0.0388 *	0.0379 *		
32.	Short-Term Wind Power Prediction Using GA-BP Neural Network Based on DBSCAN Algorithm Outlier Identification [70] Proposed Method—DBSCAN-GA-BP	10 min, 1 h	?	DBSCAN-GA-BP (single)	0.0646	0.0524	Wind speed, temperature, humidity, cosine value of wind direction, air pressure and power	
				GA-BP (single)	0.0978	0.0798		
33.	Ultra-Short-Term Wind-Power Forecasting Based on the Weighted Random Forest Optimized by the Niche Immune Lion Algorithm [71] Proposed Method—WD-NILA-WRF	5 min	49.5 MW	WD-NILA-WRF (WF A, hybrid)	28.5797	24.2484	Wind power, wind speed, wind direction, temperature, humidity, air density, air pressure, ground roughness, and other factors—wind farm in Inner Mongolia A	
				NILA-RF (WF A, hybrid)	70.9289	54.4785		
34.	Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform [72] Proposed Method—DWT_LSTM	1, 2, 3, 4, 5 steps	Approx. 1500 MW (graph)	DWT_LSTM (WF 3–5 steps)	0.024 *	0.016 *	Wind power—wind farm in Yunnan, China	
				DWT_BP (WF 3–5 steps)	0.027 *	0.018 *		
35.	Short-Term Wind Power Forecasting: A New Hybrid Model Combined Extreme-Point Symmetric Mode Decomposition, Extreme Learning Machine and Particle Swarm Optimization [73] Proposed Method—ESMD-PSO-ELM	15 min?	49.5 MW	ESMD-PSO-ELM (hybrid)	2.7	2.23	Wind power—wind farm in Yunnan, China	
				EMD-PSO-ELM (hybrid)	3.25	2.88		

Table 5. Cont.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
36.	Hourly Day-Ahead Wind Power Prediction Using the Hybrid Model of Variational Model Decomposition and Long Short-Term Memory [74] Proposed Method—Direct-VMD-LSTM	24 h	Approx. 30 MW (graph)	D-VMD-LSTM (hybrid)	0.1287 *	0.1070 *		Hourly wind power—wind farm located in Henan
				VMD-BP (hybrid)	0.1487 *	0.1927 *		
37.	Wind Power Prediction Based on Extreme Learning Machine with Kernel Mean p-Power Error Loss [51] Proposed Method—ELM-KMPE	Short-term?	6 MW	ELM-KMPE (hybrid)	1.3643 *	0.6369 *	Wind speed, wind direction, temperature, atmospheric humidity, and atmospheric pressure + weather data—NWP and SCADA	
				ELM (single)	1.3648 *	0.6703 *		
38.	TRSWA-BP Neural Network for Dynamic Wind Power Forecasting Based on Entropy Evaluation [75] Proposed Method—TRSWA-BP	1, 4, 6, 24 h	?	TRSWA-BP (24h, hybrid)	0.1703 *	0.1018 *	Wind power	
				BP (24h, single)	0.1747 *	0.1132 *		
39.	Ultra Short-Term Wind Power Forecasting Based on Sparrow Search Algorithm Optimization Deep Extreme Learning Machine [76] Proposed Method—SSA-DELML	Ultra-short-term?	3600 KW	SSA-DELML (hybrid)	0.0321 *	0.0194 *	Wind Power, wind Speed, wind Direction—SCADA + wind farm data	
				WA-DELML (hybrid)	0.0323 *	0.0201 *		
40.	Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast [77] Proposed Method—? Hybrid	> 6 h	?	Hybrid (hybrid)	0.163	0.105	?—NWP	
				Baseline (single)	0.214	0.137		
41.	A Hybrid Framework for Short Term Multi-Step Wind Speed Forecasting Based on Variational Model Decomposition and Convolutional Neural Network [26] Proposed Method—VMD-CNN	16 h	?	VMD-CNN (hybrid)	0.0655 *	0.0515 *	Wind speed—Sotavento Galicia wind farm	
				VMD-NN (hybrid)	0.0959 *	0.0764 *		
42.	Forecasting Models for Wind Power Using Extreme-Point Symmetric Mode Decomposition and Artificial Neural Networks [55] Proposed Method—Proposed model	24 h	49.5 MW	Proposed model (hybrid)	0.0251	0.0191	Wind power, humidity (30 m), pressure (50 m), temperature (100 m), wind direction (50 m), wind speed (50 m)—wind farm, China	
				EEMD (ensemble)	0.0281	0.0218		
43.	Multi-Step Ahead Wind Power Generation Prediction Based on Hybrid Machine Learning Techniques [41] Proposed Method—mRMR-IVS	1, 2, 3, 4 h	94 MW	mRMR-IVS (Xuqiao, hybrid)	0.0971	0.0762	Power generation + wind speed, trigonometric wind direction, temperature, humidity, and atmospheric pressure—Anhui, Xuqiao wind farm + NWP	
				PSR-IVS (Xuqiao, hybrid)	0.1041	0.087		
44.	An improved residual-based convolutional neural network for very short-term wind power forecasting [37] Proposed Method—Ours	1, 2, 3 h	2.5 MW	Ours (jan, single)	0.0326	0.0234	Wind power, wind speed, and wind direction—wind farm located in southeastern Turkey	
				VGG-16 (jan, single)	0.0942	0.07		
45.	A novel hybrid approach based on variational heteroscedastic Gaussian process regression for multi-step ahead wind speed forecasting [45] Proposed Method—CEEMDAN-VHGPR (CVHGPR)	1, 2, 3 steps	?	CVHGPR (W1, 3-step)	0.0244 *	0.0189 *	Wind speed—Sotavento Galicia wind farm	
				CGPR (W1, 3-step)	0.0250 *	0.0193 *		

Table 5. Cont.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
46.	A novel reinforced online model selection using Q-learning technique for wind speed prediction [78] Proposed Method—OMS-QL	1 h	?	OMS-QL (hybrid)	0.0635 *	0.0348 *		Wind speed—Colorado wind farm
				CLSTM (M9, hybrid)	0.0937 *	0.0670 *		
47.	Hybrid and Ensemble Methods of Two Days Ahead Forecasts of Electric Energy Production in a Small Wind Turbine [1] Proposed Method—The ensemble method based on hybrid methods	48 h	5 KW	The ensemble method based on hybrid methods (ensemble/hybrid)	0.0936 *	0.0563 *		Generated power, wind speed at different levels, blades pitch angle, nacelle orientation, etc.—SCADA and Levenmouth Demonstration Turbine (LDT)
				Physical model+LSTM (hybrid)	0.0940 *	0.0565 *		
				NAIVE (single)	0.2005 *	0.1267 *		
48.	Advanced Ensemble Methods Using Machine Learning and Deep Learning for One-Day-Ahead Forecasts of Electric Energy Production in Wind Farms [2] Proposed Method—Ensemble Averaging Without Extremes (many single and ensemble methods)	24 h	50 MW	Ensemble Averaging Without Extremes (many single and ensemble methods) (ensemble/hybrid)	0.1618	0.1130		Many input data, lagged values of forecasted time series, meteo forecast—wind speed, Indictor of variability of the daily electric energy production—NWP UM, south of Poland
				GBT (XGBOOST) (ensemble)	0.1636	0.1185		
				NAIVE (single)	0.3833	0.2877		
49.	A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting [79] Proposed Method—Proposed	Short-term ?	1112.40 KW	Proposed (hybrid)	0.0386 *	0.0318 *		Wind speed—Shaanxi Dingbian wind farm
				PSO-DBN (hybrid)	0.1565 *	0.1004 *		
50.	Ultra-short term wind power prediction applying a novel model named SATCN-LSTM [80] Proposed Method—SATCN-LSTM	1 step step: 16	Normalized data	SATCN-LSTM (Q1st, hybrid)	0.9070	0.4970		Wind speed, air density, wind direction, temperature, and surface pressure—wind farm California
				TCN-LSTM (Q1st, hybrid)	0.9160	0.5060		

* recalculated from RMSE or MAE and nominal system power, ** approximated values from graphs, *** nominal system power only applies to the error for a specific case rather than all cases.

Table 6. Summary of errors obtained for offshore wind power forecasting models.

No.	Title/Reference	Horizon	System Nominal Power ***	Method Details		Error Metric		Input Data Details
				-	-	nRMSE	nMAE	
1.	Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network [20] Proposed Method—	?	7 MW	Tensor Flow with feature engineering (single)	0.07789 *	0.0572 *	Active power, wind speeds, blade pitch angles, nacelle orientation, ambient temperature, and yaw error, from 1 July 2018 to 30 June 2019—SCADA database from Offshore Renewable Energy (ORE) Catapult	
				Tensor Flow with ReLU activation (single)	0.07390 *	0.0534 *		
2.	A hybrid deep learning-based neural network for 24-h ahead wind power forecasting [42] Proposed Method—CNN-RBFNN-DGF	24 h	2,000,800 KW	CNN-RBFNN-DGF (hybrid)	0.0385 *	-	Wind power generation—Changgong	
				CNN-RBFNN (hybrid)	0.0697 *	-		
3.	Exploiting Deep Learning for Wind Power Forecasting Based on Big Data Analytics [15] Proposed Method—EDCNN	24 h	900 MW	EDCNN (ensemble)	0.085	-	Wind speed, dew point temperature, dry bulb temperature, past lagged values of wind power, and wavelet packet decomposed wind power—NWP and 100 weather station in West Texas Mesonet (WTM)"	
				SELU CNN (ensemble)	0.12	-		
4.	Wind power forecasting using attention-based gated recurrent unit network [19] Proposed Method—AGRU + Attention	5, 15, 13 min, 1h, 2 h	16 MW	AGRU + Attention (2 h, ensemble)	0.1007	Historical wind power sequences and NWP data—National Renewable Energy Laboratory (NREL)		
				AGRU + RFE (2 h, ensemble)	0.1147			
5.	Offshore wind power forecasting-A new hyperparameter optimization algorithm for deep learning models [81] Proposed Method—LSTM optimized by Optuna	?	6.5 MW	LSTM optimised by Optuna (Data1, single)	0.0950 *	-	Generated power, wind speed at different levels, blades pitch angle, nacelle orientation, etc.—SCADA and wind farm, Levenmouth Demonstration Turbine (LDT)	
				LSTM (Data1, single)	0.0963 *	-		
				NAÏVE (Data1, single)	0.0979 *	-		
6.	Short-term multi-hour ahead country-wide wind power prediction for Germany using gated recurrent unit deep learning [82] Proposed Method—RNN-GRU, RNN-LSTM	1, 3, 5, 12 h	33,626 MW	RNN-GRU (1 h)	0.0017 *	0.0009 *	Wind speeds, air density, air pressure, power price index, wind speed V1, V10, and V50 (2 m above displacement height), h2 (10 m above displacement height), and h3 (50 m above ground), surface roughness length, temperature 2 m—OPSD Time Series, 2019	
				RNN-LSTM (1 h)	0.0026 *	0.0012		
				SVR-RBF (1 h)	0.2968 *	0.2333 *		

* recalculated from RMSE or MAE and nominal system power, *** nominal system power only applies to the error for a specific case rather than all cases.

In addition to the papers mentioned in Tables 5 and 6, in Section 4.1 (*Comprehensive Quantitative Review*) below we also use in certain analyses information from those papers which did not provide nRMSE and nMAE error values or these values could not be calculated due to absence of information on the rated power of the system. These papers are the following: [26,83–122].

4.1. Comprehensive Quantitative Review

Based on data from 116 papers, statistical analysis was conducted to determine, among other things: the frequency of use of various error metrics, classes of forecasting methods, distinct types of input variables for forecasting models, scopes of rated powers of the systems subject to forecasting, location of the systems subject to forecasting, and typical forecasting horizons. The analysis in this subparagraph excludes papers that provided less reliable or no data.

Figure 1 presents the outcome of statistical analysis of the number of forecasting studies concerning wind power generation in particular regions of the world based on the research papers analyzed here. What is remarkable is the very uneven distribution of studies across regions of the world. Special attention must be drawn to China—by far the largest number of papers addressing wind farm generation forecasting. The second best is the United States.

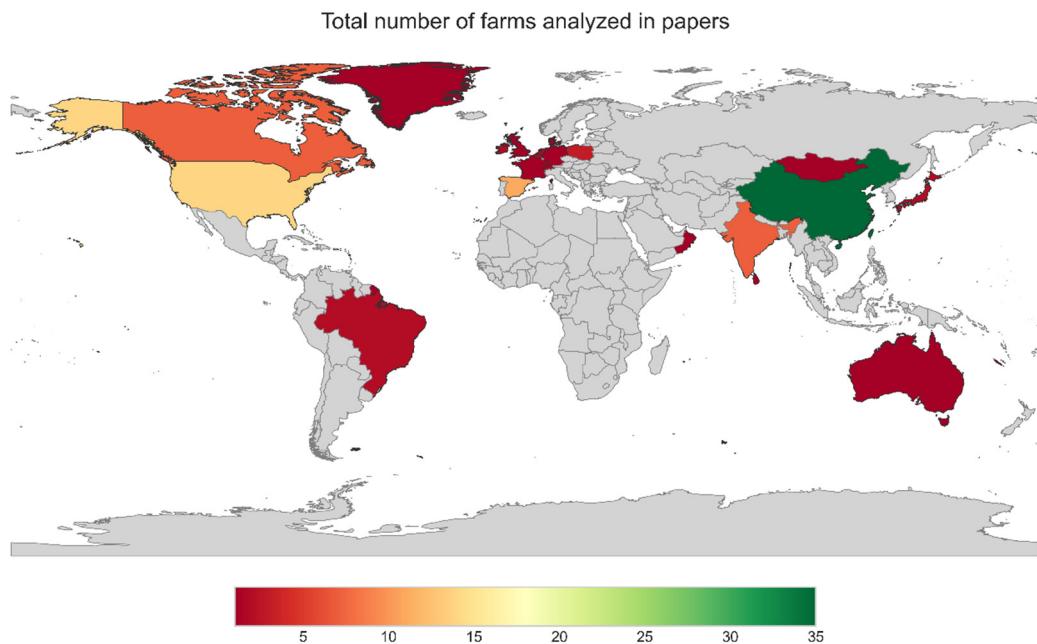


Figure 1. Total number of farms analyzed in 116 papers.

The performance of a wind power forecasting model is measured with different statistical metrics. These metrics quantify the prediction error of a model, providing the accuracy between the predicted values and the measured data [65]. It is difficult to make a comprehensive evaluation using the single error index, and, as Figure 2 shows, studies can consider up to six different metrics to evaluate, compare performance and quantify forecasting errors [20,40]. However, in general, only 2 or 3 statistical metrics are used in model validation. In some cases, authors also do not specify the metric used to evaluate the performance of the model. The combination of statistical metrics, presented by Figure 2, varies by study.

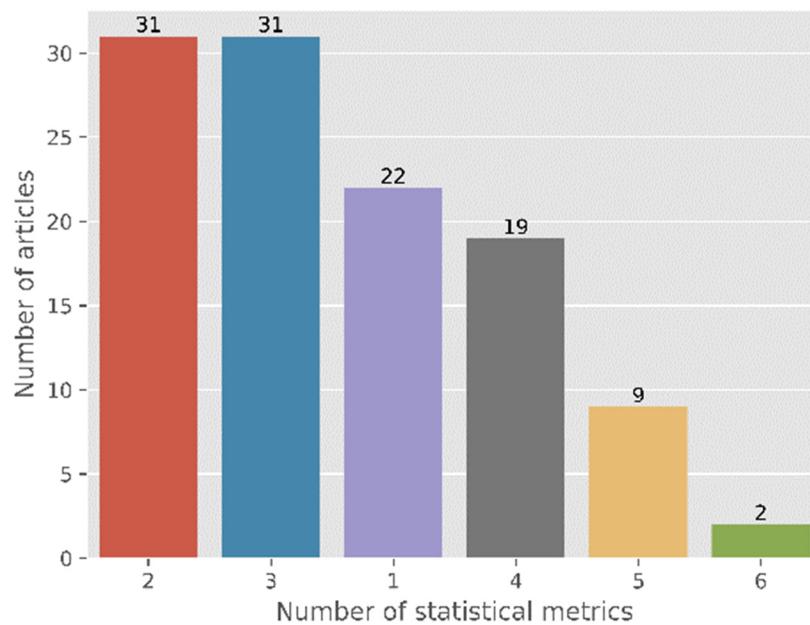


Figure 2. Number of statistical metrics used per article to evaluate the performance of forecasting models (evaluation of 114 articles).

It means that the metrics used to evaluate the performance of each model are different for each study. Figure 3 shows a summary of quantifiers used in the studies analyzed, and these metrics are split into four groups (see Table 7).

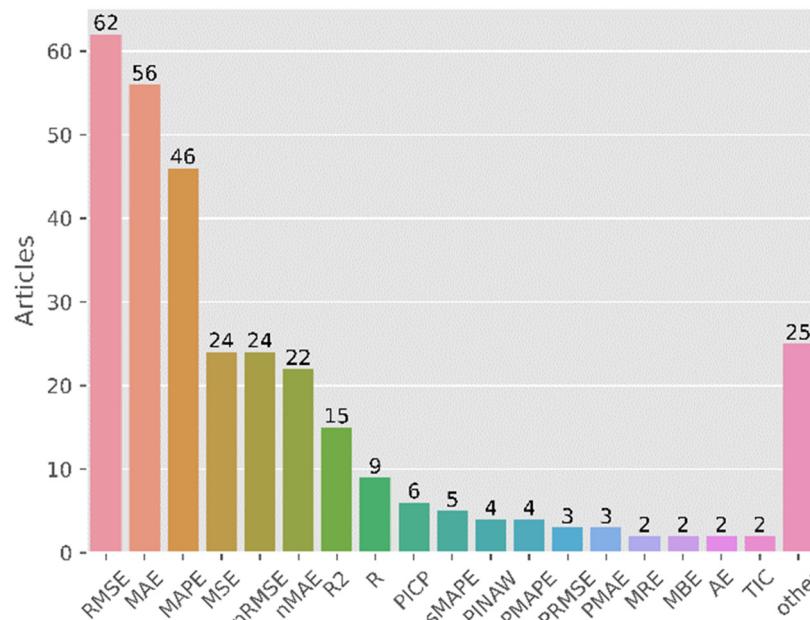


Figure 3. Common ways to measure and evaluate the error of models predicting quantitative data.

Table 7. Frequency of use of model performance evaluation metrics.

Use	Metric	Number of Times Used
Frequent	RMSE, MAE, and MAPE	46–62
Occasional	MSE, nRMSE, nMAE, and R2	15–24
Seldom	R, PICP, sMAPE, MBE, MRE, . . . , TIC	2–9
Rare	Other Metrics	1

RMSE, MAE, and MAPE are popular accuracy metrics due to their ease of interpretability by decision-makers and participants of energy markets. Those metrics, unlike mean bias errors, do not falsify the average quality of forecasts by compensating over-forecasting with under-forecasting. Moreover, they give a decent estimation of the average error one can expect from forecasts for each prediction step [123]. Because of squaring the error, RMSE is more sensitive to the detection of high values of errors in error time series, which make it a good metric for detecting extreme error values. In turn, MAE does not additionally magnify extreme values of errors and is the closest to the most naturally expected type of error—mean error. Unlike the two previous metrics, MAPE is not dependent on the scale of values in the data, which makes it useful for comparing data of different scales—e.g., errors for prosumer wind turbines and very big wind farms. It also allows us to find how an accurate model is in the scale of changing the momentary real value. This metric is, however, susceptible to zero/small values of generation appearing in its denominator [124]. The result can be either an indefinite expression or a substantial error at a given step, and as a consequence, its value is reflected in the final average value. Because of the aforementioned, we recommend not using MAPE. In Figure 3 one can see the rare use of metrics other than the three most frequent ones. Usually, they are root or derivative metrics of those three and are used to solve previous ones' drawbacks, e.g., nRMSE and nMAE add an aspect of comparability between objects of different scales, which cannot be easily conducted without normalization of time series, used in MAPE for example. A coefficient of determination is also relatively frequently used. It serves as a means of describing not how well a model predicts but how much of the modeled process is actually modeled.

Forecasting methods were classified into single methods, ensemble methods, and hybrid methods, and calculations were conducted on how frequently each of those methods was the best method (lowest nRMSE error) in each of the studies described in the papers reviewed here. Figure 4 presents the outcome of our analysis. A hybrid method was usually the best class of forecasting methods (almost 44%). Quite a substantial percentage (almost 25%) of studies in which a single method was the best is surprising. This can be explained by the fact that some papers proposed single methods only, without comparing them to other classes (ensemble, hybrid). Additionally, note that, in some cases, ensemble and hybrid methods have the characteristics of both classes. For instance, the general hybrid model also contains model(s) from the ensemble class. Comparison of forecast quality of the single, ensemble, and hybrid methods are presented in Section 4.2.2. Analysis of errors and EDF depending on the class of forecasting methods.

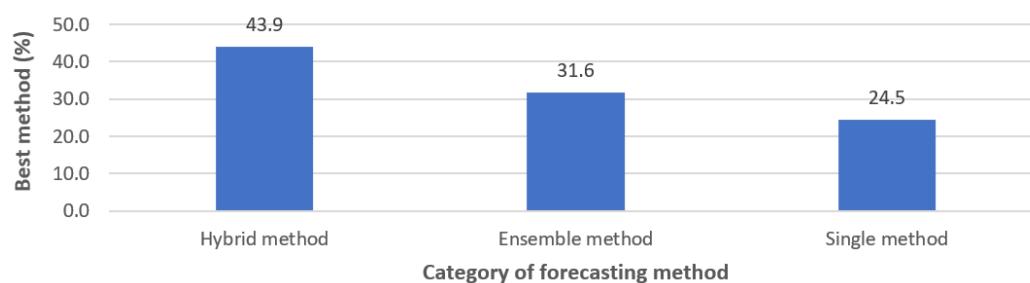


Figure 4. Frequency of using forecasting methods (quantity statistics).

Based on information from papers (for which rated powers were provided), the percentage of farms for which generation forecasts were conducted was calculated for specific ranges of rated powers. By far most frequently studied were systems sized from more than 10 MW to 100 MW, with the second largest group of systems being those sized up to 10 MW (Figure 5). Domination of the former range is probably due to the fact that it is the most frequent range of powers in wind farms, and, on the other hand, for very small (prosumer) systems, power generated from wind turbines is forecast much less frequently.

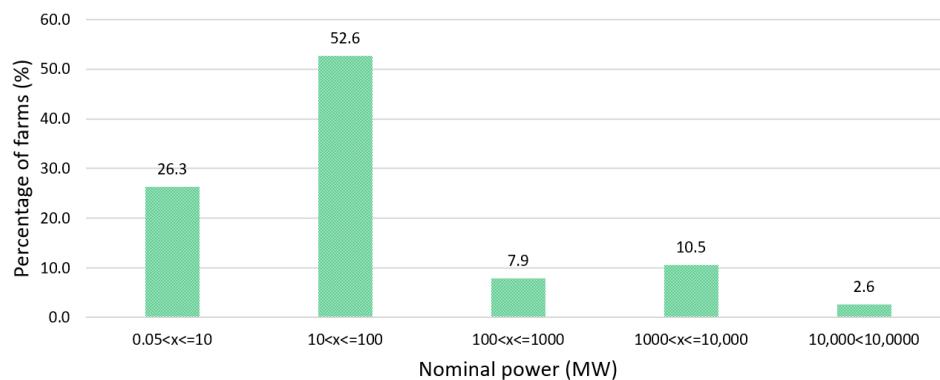


Figure 5. Ranges of rated powers of wind farms.

The percentage of farms for which generation forecasts have been conducted was calculated by forecasting horizon (Figure 6). By far the most frequent forecasting horizon is “24 h” and “few steps”, with “one step” (5 min, 10 min, 15 min, or 1 h) being slightly less frequent. Forecasts with horizons of more than 24 h are clearly rare. On the one hand, the reason may be more difficult access to NWP with such horizons and the awareness of the loss of quality of such forecasts, especially as compared to horizons with few steps ahead.

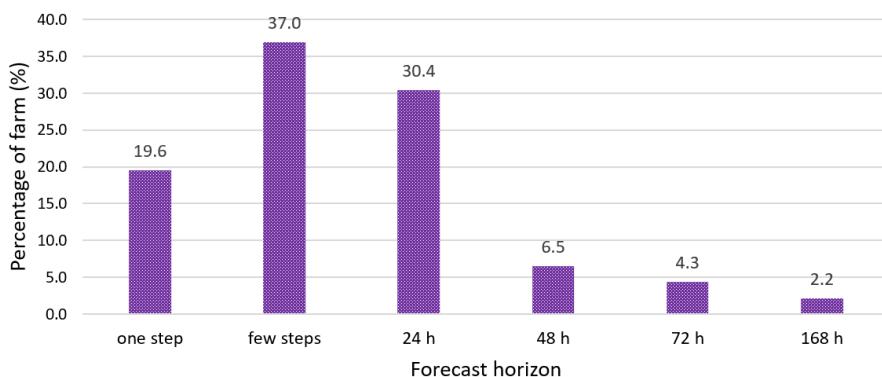


Figure 6. Frequency of forecasts from different forecasting horizons.

The frequency of use of various sets of input data in the forecasting models described in the papers was calculated (Figure 7). Lagged generation values of the forecast time series are used clearly most frequently. NWP and weather measurements are used only slightly less frequently. Other types of input data are used at least ten times less frequently than the three input data mentioned above (or incidentally). Such infrequent use of input data such as lagged NWP, time variables and generation stats (statistics on the forecast times series) in forecasting models is surprising.

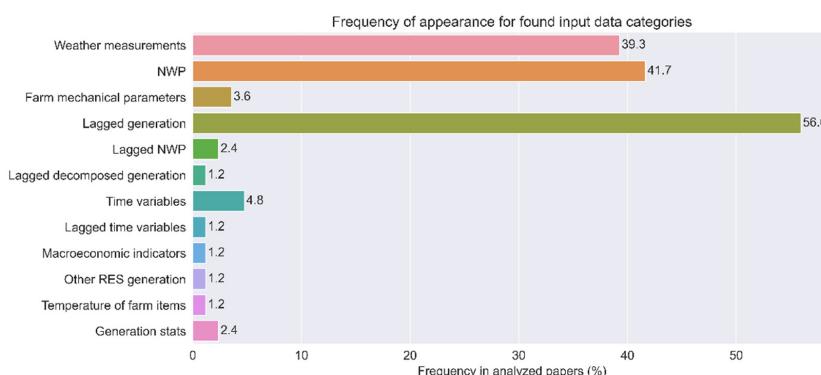


Figure 7. Input data categories by frequency of use.

4.2. Comprehensive Error Analysis

Out of 116 papers, error analysis was conducted on those which applied both nRMSE and nMAE errors and which could calculate these two error metrics based on the available rated power of the system and the values of RMSE and MAE errors (errors pre-normalization). In addition, quotients of nRMSE to nMAE errors were calculated. A new, unique EDF (error dispersion factor) metric has thus been introduced to analyses, described by Formula (20). Therefore, EDF is a combination of two frequently used error metrics. Statistical analyses in Section 4.2, Section 4.2.1 and Section 4.2.2 apply, among others, to the potential usefulness of EDF in analyses of wind power forecasts.

$$EDF = \frac{RMSE}{MAE} = \frac{nRMSE}{nMAE} = \sqrt{\frac{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}{\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|}} \quad (20)$$

The analysis in this subparagraph excludes papers that provided less reliable data (abnormal errors, abnormal error quotients)—abnormal phenomena are addressed in the Section 5. Table 8 presents basic statistics, and Figure 8 visualizes selected statistics.

Table 8. Descriptive statistics of errors and error quotients.

Descriptive Statistics	nRMSE	nMAE	EDF
Mean	0.0787	0.0566	1.4160
Standard deviation	0.0578	0.0407	0.2307
Minimum	0.0162	0.0077	1.0283
Maximum	0.2762	0.1942	2.1129
Quotient of maximum/minimum	17.05	25.22	1.75
The 10th percentile	0.0244	0.0174	1.1841
The 25th percentile (lower quartile)	0.0324	0.0236	1.2740
The 50th percentile (median)	0.0654	0.0504	1.3603
The 75th (upper quartile)	0.1007	0.0756	1.5019
The 90 percentile	0.1618	0.1070	1.7741
Variance	0.0033	0.0017	0.0562
Skewness	1.3448	1.2811	1.2939
Kurtosis	1.9858	1.8489	1.5423

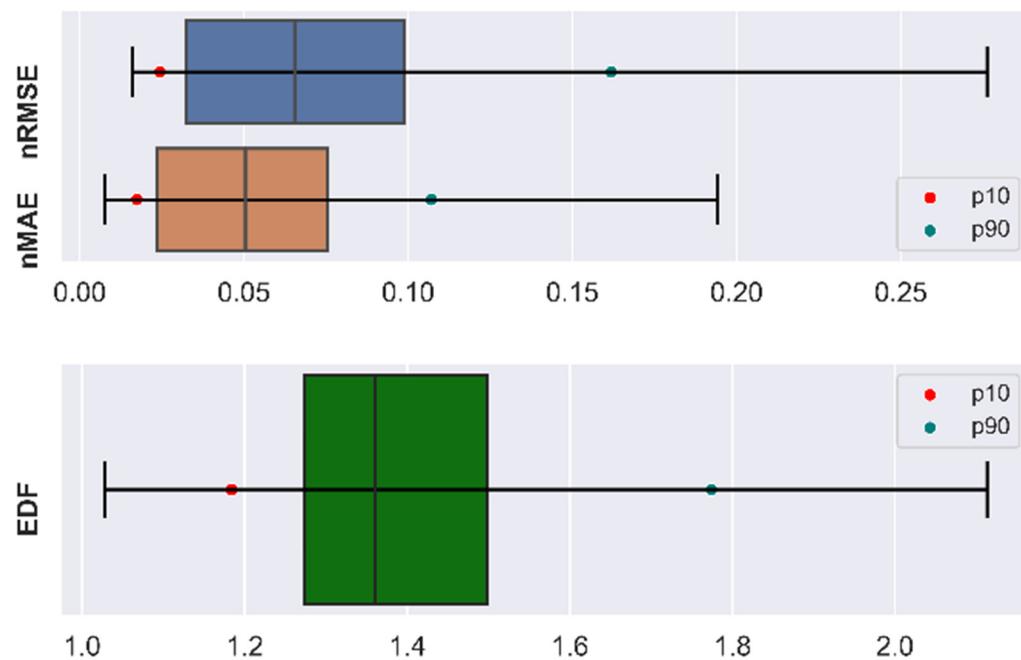


Figure 8. Selected statistics of errors and error quotients.

The averages in Table 8 are slightly larger than the medians both for nRMSE, nMAE, and EDF. The dispersion of errors is remarkably high—the maximum/minimum quotient for nRMSE error metric is more than 17, and for nMAE errors, the quotient is more than 25. Such large dispersion of values can be partly justified by different forecasting horizons (from 10 min to 72 h).

4.2.1. Analysis of Errors by Forecasting Horizon

Figures 9 and 10 present nRMSE and nMAE errors, respectively, in ascending order, based on the papers in which these error metrics were provided (also considering those which did not provide the rated power of the system). In addition, information on the forecasting horizon is provided. Forecasts with longer horizons display significantly much larger nRMSE errors, which is unsurprising (the accuracy of wind speed forecasts decreases with increasing forecasting horizon).

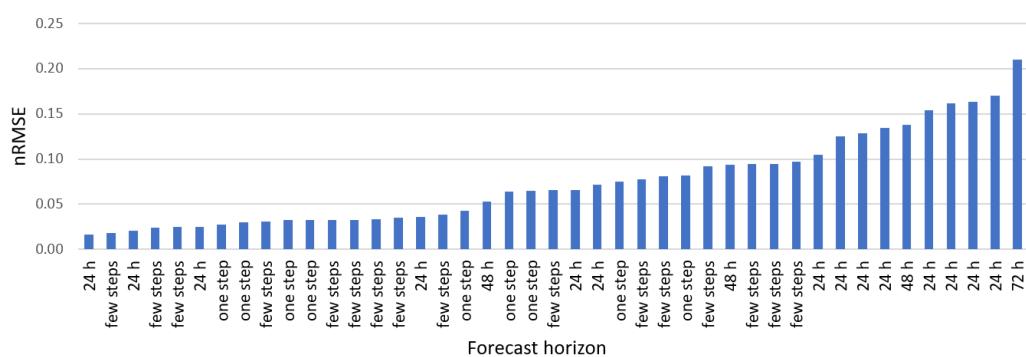


Figure 9. nRMSE errors with a note on forecasting horizon, in ascending order.

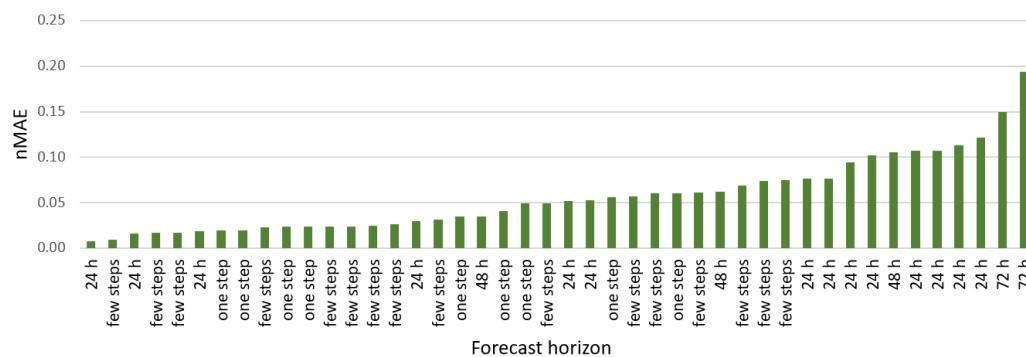


Figure 10. nMAE errors with a note on forecasting horizon, in ascending order.

Figure 11 presents how the amount of error depends on the forecasting horizon. This figure summarizes information from Figures 9 and 10—average values for both metrics were calculated for selected forecasting horizons. In general, average errors grow with increasing forecasting horizon, although, for the 24 h horizon, average errors are slightly more than for the 48 h horizon. This is probably due to the fact that there were significantly fewer papers describing forecasts with a 48 h horizon than with a 24 h horizon (random element of lower errors from a small number of samples). By far the largest were the average errors for the 72 h horizon—more than two-and-a-half larger than for 24 h and 48 h horizons. For the “one step” horizon, average errors are two times smaller than average errors for the 24 h horizon. This information has large practical significance—it shows what magnitude of normalized errors should be expected from the respective forecasting horizon. Please note that the averages calculated for the 48 h and 72 h horizons may not be fully representative due to a small number of samples.

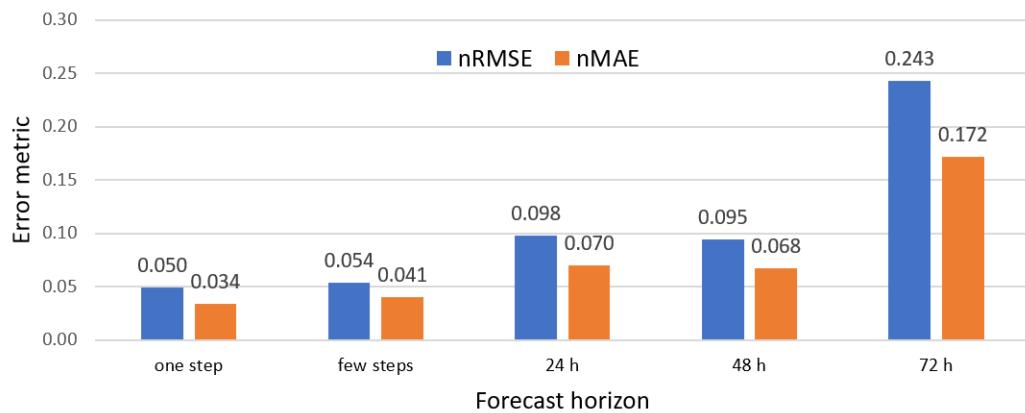


Figure 11. Magnitudes of error by forecasting horizon.

To determine precisely whether there is a statistically significant relationship between the forecasting horizon and error magnitudes, numerical forecasting horizons were selected (1/6 h, 1/4 h, 1/2 h, 1 h, 6 h, 12 h, 24 h, 48 h, and 72 h), which enabled us to calculate Pearson linear correlation. The statistical analysis concluded a statistically significant (5% level of significance) positive linear correlation between the forecasting horizon (multiples of 1) and the magnitude of nRMSE error ($R = -0.347$). nRMSE error grows with increasing forecasting horizon. Figure 12 presents how the magnitude of nRMSE error depends on the forecasting horizon.

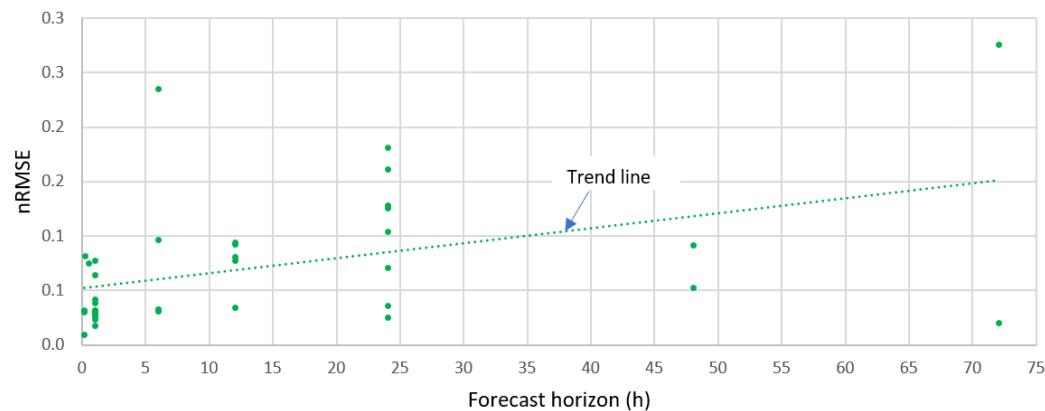


Figure 12. Dependence of nRMSE on forecasting horizon.

The statistical analysis concluded a statistically significant (5% level of significance) positive linear correlation between the forecasting horizon (multiple of 1) and the magnitude of nRMSE error ($R = 0.410$). nMAE error grows with increasing forecasting horizons. Figure 13 presents how the magnitude of nMAE error depends on the forecasting horizon. It is worthwhile to emphasize that the linear correlation between the forecasting horizon and the magnitude of error is slightly larger for the nRMSE error metric than for nMAE.

The statistical analysis concluded a statistically insignificant (5% level of significance) negative linear correlation between the forecasting horizon and EDF ($R = -0.196$). The EDF slightly decreases with increasing forecasting horizon. Figure 14 presents how EDF depends on the forecasting horizon.

For forecasts with very short horizons (from 10 min to 1 h), the average EDF is 1.422, and 1.3163 for 6 h, and it falls to 1.2724 for the 24 h horizon. For the 48 h and 72 h horizons, the samples are too few to calculate reliable averages.

In addition, statistical analysis omitting the 48 h and 72 h horizons concluded a negative correlation. Not a very large one, but statistically significant (15% level of significance), between the forecasting horizon and EDF ($R = -0.283$).

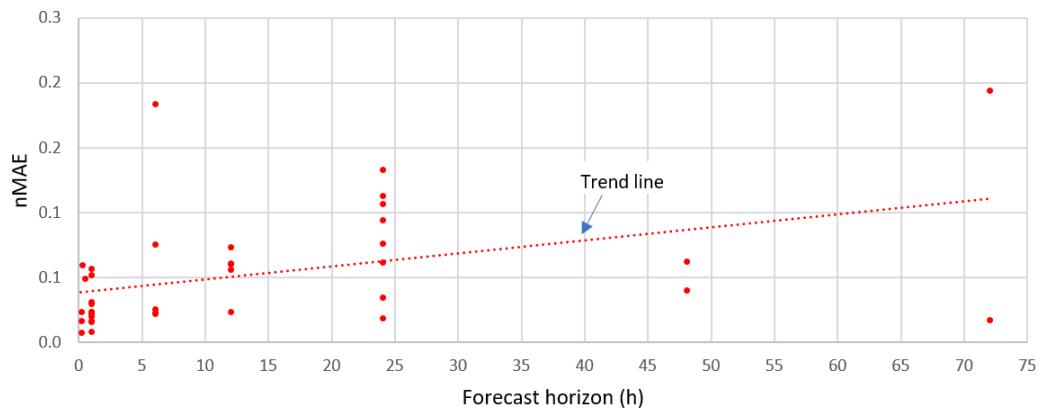


Figure 13. Dependence of nMAE errors on forecasting horizon.

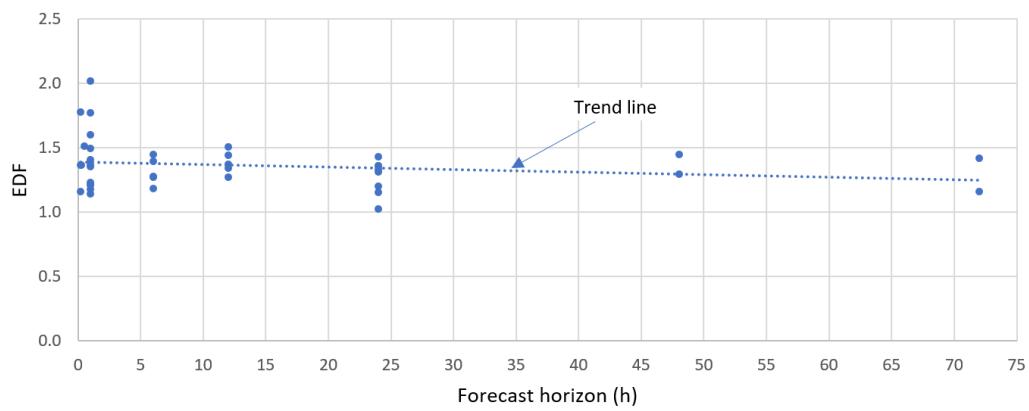


Figure 14. Dependence of EDF on forecasting horizon.

The EDF (Figure 14) and (Formula (20)) show the average variability of the moduli of error regardless of the magnitude of the error. If absolute errors on all samples are the same, this ratio reaches its minimum value of 1. The larger the error deviation on particular samples from the average, the larger the ratio. This resembles the behavior of standard deviation determined for the moduli of error for samples, however, with the difference that standard deviation reaches a minimum value equal to zero, and the dynamics of that ratio is much larger—significantly dependent on particular samples. For the EDF, the dynamics of values are smaller, which better illustrates the variability of errors across the sample pool. It should also be mentioned that the EDF in fact shows the ratio of the second moment of error to the first moment of error.

The decreasing levels of EDF with a rising forecasting horizon means that the variability of error decreases with an increasing forecasting horizon. On the one hand, it is probably due to the growing error, and, on the other, the averaging nature of the forecasting models for longer horizons, which stabilizes errors around certain values.

It is worthwhile to note that statistical analysis of hourly values of wind speed presented in [1] concluded that the variance of wind speed forecasts for horizons ranging from 1 to 24 h was 3.121, and for 25- to 48-h horizons, it was 3.063, which is less.

4.2.2. Analysis of Errors and EDF Depending on the Class of Forecasting Methods

Some of the 116 papers analyzed here provide the forecasting error of a method from the “single method” class. The primary objective of the analysis was to investigate percentage error reduction achieved by the best (proposed) method from the ensemble or hybrid class relative to the single method with the largest forecasting error (excluding the outcome of the naive method). Figure 15 presents, in descending order, percentage reductions of nRMSE and percentage reductions of nMAE of the best methods relative to single methods. What is remarkable is a very wide dispersion of percentage reductions of

error. For nRMSE, the largest percentage reduction of error is 80.02%, and the smallest is 2.76%. Similar observations apply to the dispersion of nMAE.

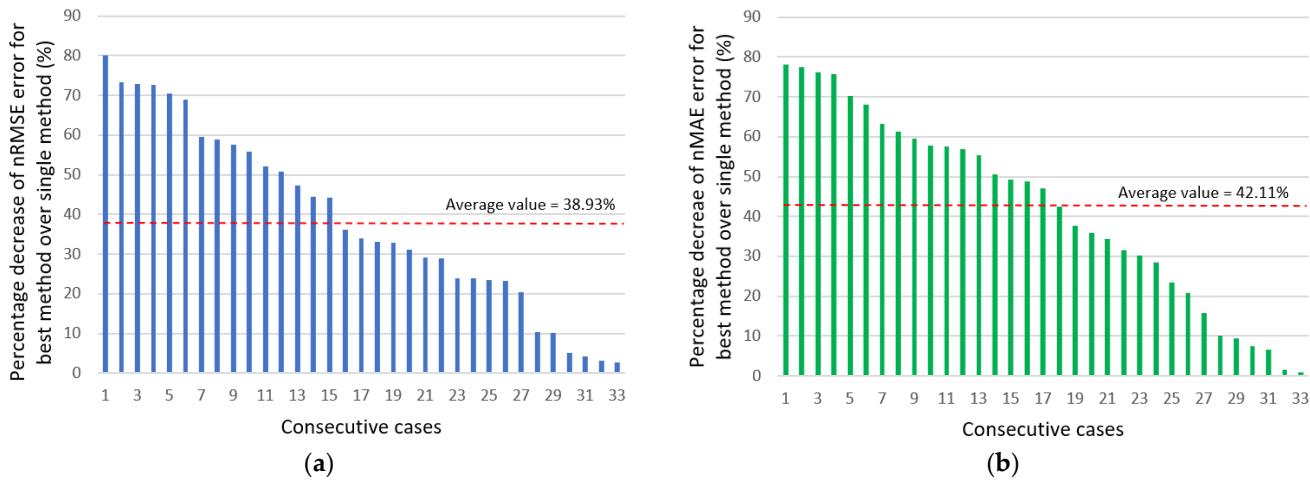


Figure 15. (a) Percentage reduction in nRMSE for the best method relative to the single method; (b) percentage reduction in nMAE for the best method relative to the single method.

Figure 16 presents the average percentage improvement of the hybrid methods and ensembles method relative to the single method for nRMSE and nMAE error metrics. The percentage improvement of error metrics is much bigger for hybrid methods in comparison to ensemble methods however the number of cases (19 for hybrid methods and 14 for ensemble methods) is too small to generalize this fact.

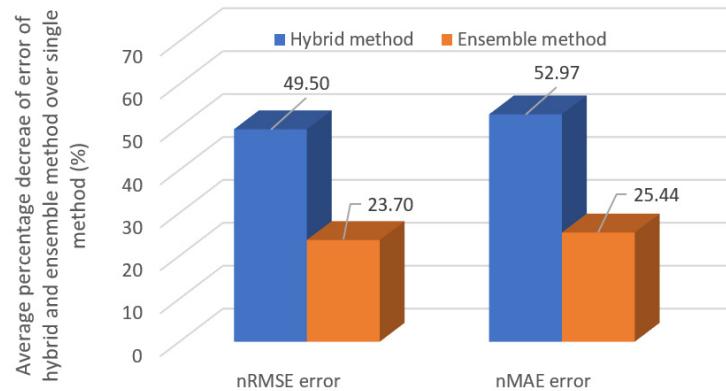


Figure 16. Average percentage improvement of the hybrids method and ensembles method relative to the single method for nRMSE and nMAE error metrics.

Unfortunately, a small proportion of the papers reviewed here provide forecasting error using a naive (persistence) method—such error would be the best benchmark for the level of improvement achieved by other proposed methods, including single methods. The forecasting methodology assumes that a forecasting method is valuable if its error is less than the error of the naive method. Six papers provide errors for the naive method. Figure 17 presents, in descending order, percentage reductions of nRMSE relative to the naive method for six cases (pairs of nRMSE and nMAE)

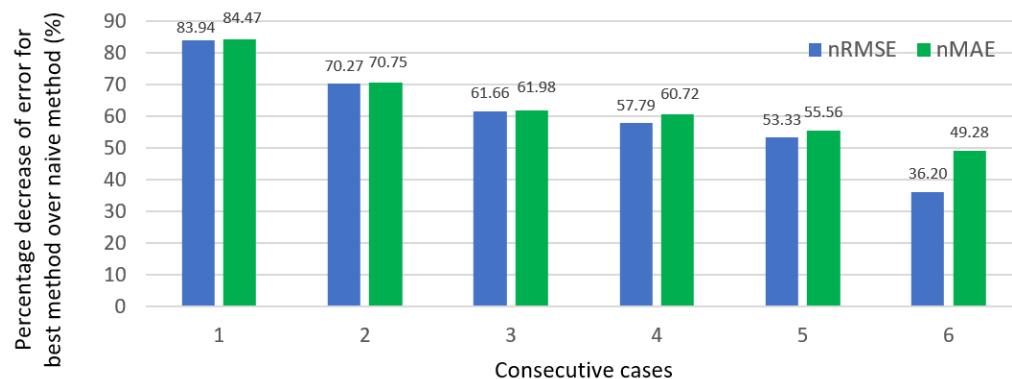


Figure 17. Percentage improvement of the best method relative to the naive method for nRMSE and nMAE error metrics.

The average percentage reduction calculated for six cases is 60.53% for nRMSE and 63.79% for nMAE. Therefore, both average percentages are much larger than similar values calculated for nRMSE and nMAE reductions when errors of best methods are compared to single methods. Nevertheless, in a small number of cases, percentage reductions of nRMSE and nMAE for the best method relative to single methods are large and similar to the best method compared to the naive method. It means that some single methods referred to in literature are only marginally better than naive methods.

The second objective of our analysis is to compare EDF for the best (proposed: ensemble or hybrid) method and a single method. Based on 33 cases (pairs of ratios), we have determined that in 77% of cases, the EDF for the best method is larger than the EDF for the single method used in the respective study—this is more frequently observed for larger values of those ratios. The Pearson coefficient of linear correlation (R) between the ratios for the best method and the ratios for the single method is 0.737.

Figure 18 presents pairs of EDF sorted in descending order by the level of ratios for the best method.

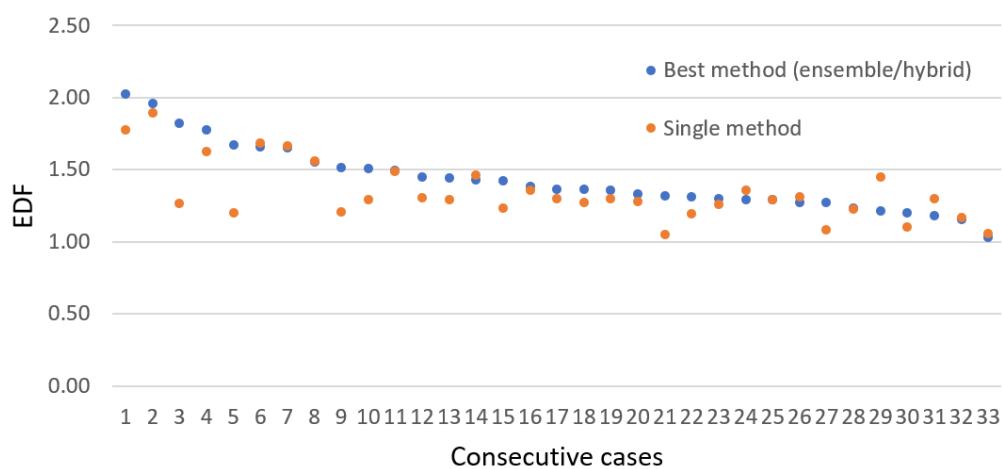


Figure 18. Pairs of EDF sorted in descending order by the level of quotients for the best method.

The average EDF for the best method is 1.432 and the median EDF is 1.364. The average EDF for the single method is 1.352, and the median EDF is 1.294. Therefore, both the average and median levels are clearly larger for the best method. Both series do not have normal distribution—Shapiro–Wilk test was conducted. Wilcoxon signed-rank test has been therefore applied to the analysis of pairs, which concluded that there are statistically significant differences between pairs in both series (they have different expected values). Therefore, differences between medians are statistically significant, and not without reason.

An interesting conclusion can be drawn based on our analysis—the variability of the moduli of errors in the best methods (smallest forecasting errors) is typically larger than for the “single method” class (much larger forecasting errors). The moduli of errors in the “single method” class are much larger and much closer to each other than in the best (hybrid or ensemble) method. In some studies, a single method could also use slightly less information (a different set of input data), which can also affect the characteristics of errors (magnitude and variability level).

4.2.3. Analysis of Errors Based on System Size

Our analysis covered the studies which provided nRMSE and nMAE and the size of the system. Statistical analysis did not reveal a statistically significant (5% level of significance) linear correlation between the size of the system (rated power) and nRMSE and nMAE errors ($R = -0.110$, $R = -0.111$, respectively). In theory, errors should grow with increasing size of the system due to much less uniform weather conditions (wind speed) in wind farms occupying extensive areas, the fact of using usually point-based meteorological forecasts, the wake effect, and other factors affecting the farm which are more difficult to represent if they overlap in the same space. The Pearson coefficient of linear correlation (R) between nRMSE and nMAE is 0.994 (5% level of significance). It means that these error metrics are very similar to each other. The details are presented in Figure 19. In addition, there is a large dispersion of the magnitudes of error for systems of similar sizes. This can be due to different sets of input data (different quality of information).

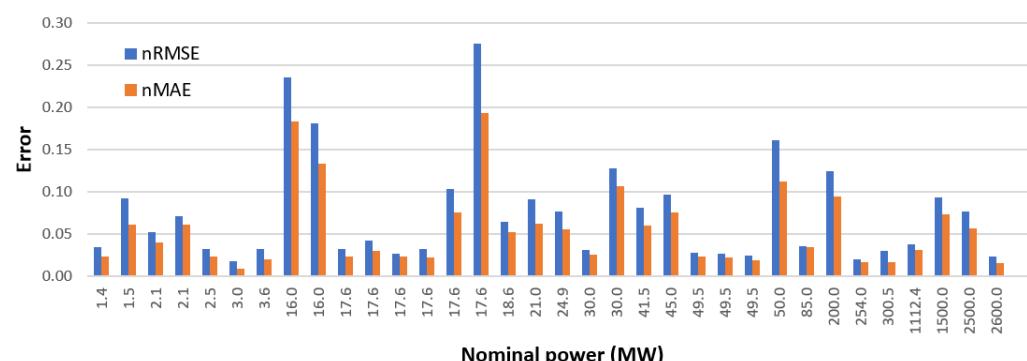


Figure 19. Magnitude of error depending on rated power of the system.

The statistical analysis concluded an insignificant (5% level of significance), marginally positive linear correlation between the size of the system (rated power) and EDF ($R = 0.039$). The EDF usually varies between 1 and 2. The average value of the EDF is 1.35. The details are presented in Figure 20.

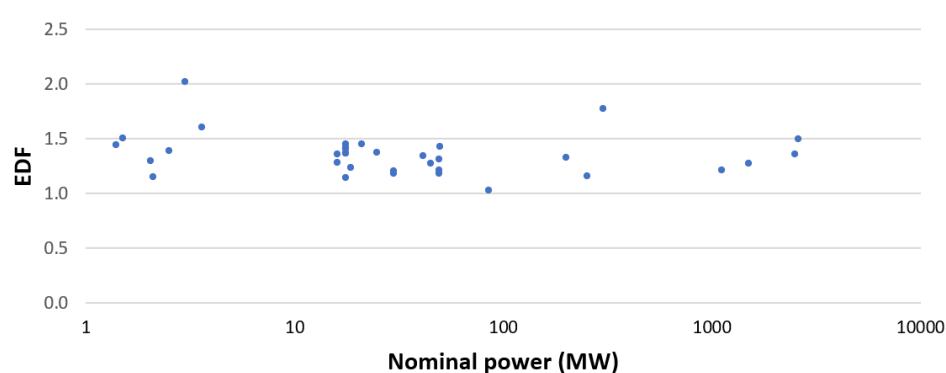


Figure 20. Magnitude of EDF depending on rated power of the system.

4.2.4. Analysis of Error Based on System Location (Onshore v. Offshore)

The number of papers addressing forecasting for offshore farms is small, as they constitute less than 6% of the 116 papers subject to this analysis. Only six papers (Table 6) provide nRMSE or nMAE, which is too little to conduct an accurate statistical analysis. Figure 21 compares nRMSEs for two forecasting horizons (average of the errors provided in the papers) for offshore and onshore farms.

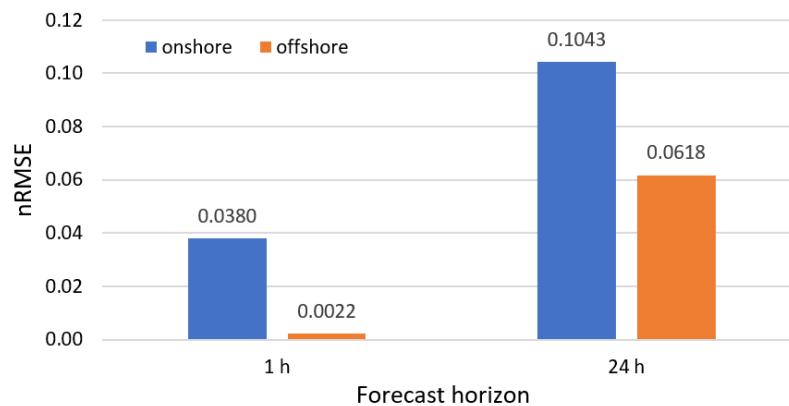


Figure 21. Magnitude of nRMSEs depending on farm location and forecasting horizon.

Offshore farms have smaller forecasting errors than onshore farms. This is expected, as it results from more stable and stronger winds at offshore farms. In addition, these are typically very large systems. For a 1 h horizon, such a sizable difference may result from the fact that the average for offshore farms was based on only two values of error and the fact that some onshore forecasts did not use meteorological forecasts of wind speed (larger forecasting errors occur in such cases). In real terms, at 24 h horizon, nRMSE at onshore farms can be about twice as large (assuming that meteorological forecasts are used in both locations).

5. Discussion

A comprehensive review and statistical analysis of errors based on an extensive selection of 116 papers allowed us to conclude, using actual figures, a correlation between the magnitude of error and selected factors. The quantitative analysis is provided for the aggregate assessment of how frequently various categories of (quite diverse) forecasting methods are applied, what typical input data are (meteorological forecasts are typically used for horizons above 6 h), how often various forecasting horizons have been used (typical horizons being in the range of 1 h to 24 h).

The analyses concluded that some papers used incomplete data that prevented them from being used in an aggregate meta-analysis of studies, which applies, in particular, to error metrics (nRMSE and nMAE).

In addition, several untypical (extreme) nRMSE and nMAE error levels have been identified, which, due to extreme dissimilarity to the remaining data of the same class (forecasting error being too large or too small) by expert judgment have been excluded from the analyses presented in Section 4.2. *Comprehensive Error analysis*. Figure 22 presents the variability of nRMSE error in the papers reviewed here.

A novel, unique ratio called EDF has been explored. The EDF shows the average variability of the moduli of error regardless of the magnitude of error. The analysis of variability of the new EDF ratio depending on selected characteristics (size of wind farm, forecasting horizon, and class of forecasting method) has been performed. There is a small negative correlation but statistically significant between the forecasting horizon and EDF. Additionally, the EDF for the best forecasting method is larger than the EDF for the single forecasting method. The analysis concluded an insignificant, marginally positive linear correlation between the size of the system (rated power) and EDF.

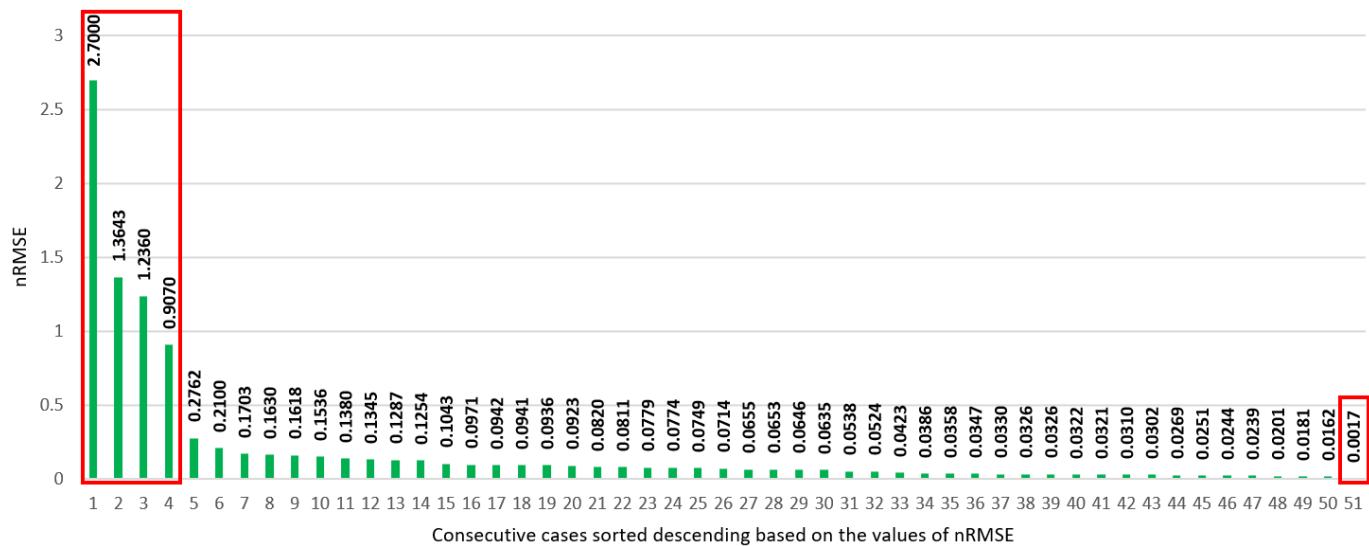


Figure 22. Variability of nRMSEs in the reviewed papers, with red-marked less reliable values.

Statistical analysis in one paper concluded, in addition, an untypical value of EDF. Statistical data from the reviewed papers for which EDF could be calculated shows that EDF levels range from 1.028 to 7.478, although a vast majority of EDF levels range between 1 and 2 (this range seems to be most credible—minimum value of the ratio is 1). The outcome of our analysis is presented in Figure 23.

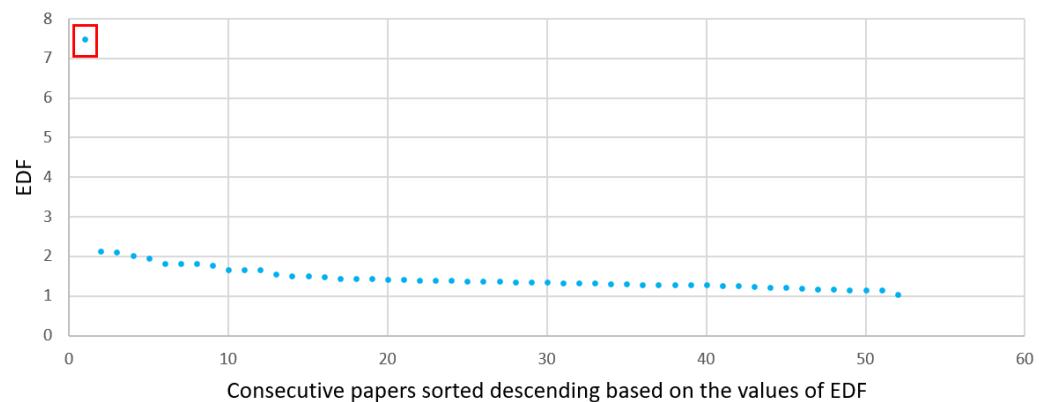


Figure 23. Variability of EDF in reviewed papers, with red-marked less credible value.

Based on our analysis of papers, in our subjective assessment, to maximize the quality of aggregate meta-analysis of studies addressing power generation forecasting in wind farms, a research paper should contain the following items (our recommendations):

- mandatory use of normalized error metrics for the assessment of forecasting quality: nRMSE Formula (5) and nMAE Formula (6), accompanied by a description of the normalization method (recommended normalization using the rated power of the system), which would enable comparative assessments of the quality of studies of systems of various sizes (or regardless of how big the wind farm is);
- we do not recommend the use of MAPE metric, which is susceptible to substantial error for small or zero generations;
- mandatory use of a forecast conducted by naive (persistence) method to enable the assessment of the quality of the best model proposed by the authors of the paper relative to the reference model; in such case calculation of the skill score metric by Formula (17), (18), or (19) is recommended in addition;

- provide strict, precise information on forecasting horizon(s) and information if the forecast is, e.g., “One Step,” e.g., from 1 h to 24 h, or one step ahead with follow-up predictions (in which case forecasting errors are typically smaller);
- provide strict, precise information on the set(s) of input data for the proposed model(s);
- provide the source of meteo forecasts (GFS, ECMWF, other sources) if used in the forecasting model;
- explicit statement on whether the forecasting model uses meteo forecasts and/or on-site measurements of weather conditions at the wind farm;
- provide the range of training, validation, and testing data, and how the range of data is divided into the identified subsets;
- provide details of the location, unless confidential (location and size of the wind farm, and landscape features prevalent at the location).

6. Conclusions

This paper is the outcome of a comprehensive review and statistical analysis of errors using more than one hundred research papers. The quantitative analyses allowed us to assess the distribution of frequency of application of selected parameters in research studies (including the number and type of error metrics, forecasting horizon, rated power of the system, classes of forecasting methods, and location of the forecast systems).

Our qualitative analyses allowed us to provide an aggregate assessment of power generation forecasting in wind farms, including how selected factors affect the magnitude of forecasting errors. In addition, the rationale for using complex (ensemble, hybrid) forecasting methods instead of single methods was verified, by examining how this improves the quality of forecasts.

Notably, only 6 of 116 papers addressed power generation forecasts in offshore farms—it means that such research should intensify going forward, although it is in part due to a significantly smaller number of such systems than of onshore farms. The offshore location of a farm involves a number of distinct characteristics (such as a surface with exceptionally low roughness, significantly higher wind speeds, and more stable power generation). The magnitude of forecasting errors is significantly smaller. Due to a small number of offshore-related papers, our analysis was much more constrained.

In our view, research on topics related to aggregate statistical analyses (meta-analyses) should continue. We are planning to increase the number of reviewed papers at least two- or three-fold in the future. Such a number will enable us to conduct a more precise statistical assessment of a large number of factors affecting the magnitude of forecasting error, and expand the analyses related to the EDF factor proposed by us. In our view, it is crucial that published papers on generation forecasts in wind farms contain information from our recommended list, to enable conducting the necessary analyses.

Author Contributions: Conceptualization, P.P. and D.B.; methodology, P.P., D.B., I.R. and M.K.; investigation, P.P., D.B., I.R. and M.K.; supervision, P.P.; validation, P.P. and M.K.; writing, P.P., D.B., I.R. and M.K.; visualization P.P., I.R., M.K. and D.B.; project administration, P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the 2021 edition of the competition for grant of the Scientific Council for the Discipline Automatic Control, Electronics and Electrical Engineering of the Warsaw University of Technology (to P.P., D.B., M.K.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used throughout this manuscript:

ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
BP	Back Propagation
BPNN	BP Neural Network
CNN	Convolutional Neural Network
CSA	Crow Search Algorithm
DA	Dragonfly Algorithm
DBN	Deep Belief Network
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DGF	Double Gaussian Function
DWT	Discrete Wavelet Transform
EDCNN	Efficient Deep Convolution Neural Network
EDF	Error Dispersion Factor
EEMD	Ensemble Empirical Mode Decomposition
EMD	Empirical Mode Decomposition
EWT	Empirical Wavelet Transformation
FS	Feature Selection
GBT	Gradient Boosted Trees
GRU	Gated Recurrent Unit
HLNN	Laguerre Neural Network
HNN	Hybrid Neural Network
IPSO	Improved Particle Swarm Optimization
KNNR	K-Nearest Neighbors Regression
LD	Lorenz Disturbance System
LR	Linear Regression
LSTM	Long-Short-Term Memory
MAAPE	Mean Arctangent Absolute Percentage Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MBE	Mean Bias Error
MCC	Maximum Correntropy Criterion
MKRPINN	Multi-Kernel Regularized Pseudo Inverse Neural Network
MPSO	Modified Particle Swarm Optimization
MRE	Mean Relative Error
mRMR	Maximum Relevance and Minimum Redundancy Algorithm
MSE	Mean Squared Error
nMAE	Normalized Mean Absolute Error
NN	Neural Network
nRMSE	Normalized Root Mean Squared Error
NWP	Numerical Weather Prediction
OMS-QL	Online Model Selection using Q-learning
P	Promoting Percentages
PCA	Principal Component Analysis
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval Normalized Average Width
PSO	Particle Swarm Optimization
R or CC	Pearson Linear Correlation Coefficient
R2	R-square or Coefficient of Determination
RBF	Radial Basis Function
RES	Renewable Energy Sources
RF	Random Forest
RMSE	Root Mean Square Error
SAE	Stacked Auto-Encoders
SE	Sample Entropy
sMAPE	Symmetric Mean Absolute Percentage Error
SS	Skill Score
SSA	Singular Spectrum Analysis
STA	State Transition Algorithm
SVM	Support Vector Machine

SVR	Support Vector Regression
TIC	Theil's Inequality Coefficient
VMD	Variational Mode Decomposition
WF	Wind Farm
WIPSO	Weight Improved Particle Swarm Optimization
WPT	Wavelet Packet Transform
WT	Wavelet Transform
xGBoost	eXtreme Gradient Boost

References

- Piotrowski, P.; Kopyt, M.; Baczyński, D.; Robak, S.; Gulczynski, T. Hybrid and Ensemble Methods of Two Days Ahead Forecasts of Electric Energy Production in a Small Wind Turbine. *Energies* **2021**, *14*, 1225. [[CrossRef](#)]
- Piotrowski, P.; Baczyński, D.; Kopyt, M.; Gulczynski, T. Advanced Ensemble Methods Using Machine Learning and Deep Learning for One-Day-Ahead Forecasts of Electric Energy Production in Wind Farms. *Energies* **2022**, *15*, 1252. [[CrossRef](#)]
- Wang, Y.; Zou, R.; Liu, F.; Zhang, L.; Liu, Q. A review of wind speed and wind power forecasting with deep neural networks. *Appl. Energy* **2021**, *304*, 117766. [[CrossRef](#)]
- Hanifi, S.; Liu, X.; Lin, Z.; Loffian, S. A Critical Review of Wind Power Forecasting Methods-Past, Present and Future. *Energies* **2020**, *13*, 3764. [[CrossRef](#)]
- Santhosh, M.; Venkaiah, C.; Vinod Kumar, D. Current advances and approaches in wind speed and wind power forecasting for improved renewable energy integration: A review. *Eng. Rep.* **2020**, *2*, e12178. [[CrossRef](#)]
- González-Sopeña, J.; Pakrashi, V.; Ghosh, B. An overview of performance evaluation metrics for short-term statistical wind power forecasting. *Renew. Sustain. Energy Rev.* **2021**, *138*, 110515. [[CrossRef](#)]
- Nazir, M.S.; Alturise, F.; Alshmrany, S.; Nazir, H.M.J.; Bilal, M.; Abdalla, A.N.; Sanjeevikumar, P.; Ali, Z.M. Wind Generation Forecasting Methods and Proliferation of Artificial Neural Network: A Review of Five Years Research Trend. *Sustainability* **2020**, *12*, 3778. [[CrossRef](#)]
- Tawn, R.; Browell, J. A review of very short-term wind and solar power forecasting. *Renew. Sustain. Energy Rev.* **2022**, *153*, 111758. [[CrossRef](#)]
- Chen, C.; Liu, H. Medium-term wind power forecasting based on multi-resolution multi-learner ensemble and adaptive model selection. *Energy Convers. Manag.* **2020**, *206*, 112492. [[CrossRef](#)]
- Mishra, S.P.; Dash, P.K. Short-term prediction of wind power using a hybrid pseudo-inverse Legendre neural network and adaptive firefly algorithm. *Neural Comput. Appl.* **2019**, *31*, 2243–2268. [[CrossRef](#)]
- Wang, H.; Liu, Y.; Zhou, B.; Li, C.; Cao, G.; Voropai, N.; Barakhtenko, E. Taxonomy research of artificial intelligence for deterministic solar power forecasting. *Energy Convers. Manag.* **2020**, *214*, 112909. [[CrossRef](#)]
- Piotrowski, P.; Baczyński, D.; Kopyt, M.; Szafranek, K.; Helt, P.; Gulczynski, T. Analysis of forecasted meteorological data (NWP) for efficient spatial forecasting of wind power generation. *Electr. Power Syst. Res.* **2019**, *175*, 105891. [[CrossRef](#)]
- Piotrowski, P.; Parol, M.; Kapler, P.; Fetliński, B. Advanced Forecasting Methods of 5-Minute Power Generation in a PV System for Microgrid Operation Control. *Energies* **2022**, *15*, 2645. [[CrossRef](#)]
- Yousuf, M.U.; Al-Bahadly, I.; Avci, E. Current Perspective on the Accuracy of Deterministic Wind Speed and Power Forecasting. *IEEE Access* **2019**, *7*, 159547–159564. [[CrossRef](#)]
- Mujeeb, S.; Alghamdi, T.A.; Ullah, S.; Fatima, A.; Javaid, N.; Saba, T. Exploiting Deep Learning for Wind Power Forecasting Based on Big Data Analytics. *Appl. Sci.* **2019**, *9*, 4417. [[CrossRef](#)]
- Wang, K.; Qi, X.; Liu, H.; Song, J. Deep belief network based k-means cluster approach for short-term wind power forecasting. *Energy* **2018**, *165*, 840–852. [[CrossRef](#)]
- Li, L.L.; Zhao, X.; Tseng, M.L.; Tan, R.R. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. *J. Clean. Prod.* **2020**, *242*, 118447. [[CrossRef](#)]
- Lopez, E.; Valle, C.; Allende, H.; Gil, E.; Madsen, H. Wind Power Forecasting Based on Echo State Networks and Long ShortTerm Memory. *Energies* **2018**, *11*, 526. [[CrossRef](#)]
- Niu, Z.; Yu, Z.; Tang, W.; Wu, Q.; Reformat, M. Wind power forecasting using attention-based gated recurrent unit network. *Energy* **2020**, *196*, 117081. [[CrossRef](#)]
- 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. [[CrossRef](#)]
- Kisvari, A.; Lin, Z.; Liu, X. Wind power forecasting A data-driven method along with gated recurrent neural network. *Renew. Energy* **2021**, *163*, 1895–1909. [[CrossRef](#)]
- Ju, Y.; Sun, G.; Chen, Q.; Zhang, M.; Zhu, H.; Rehman, M.U. A Model Combining Convolutional Neural Network and LightGBM Algorithm for Ultra-Short-Term Wind Power Forecasting. *IEEE Access* **2019**, *7*, 28309–28318. [[CrossRef](#)]
- Demolli, H.; Dokuz, A.S.; Ecemis, A.; Gokcek, M. Wind power forecasting based on daily wind speed data using machine learning algorithms. *Energy Convers. Manag.* **2019**, *198*, 111823. [[CrossRef](#)]

24. Yin, H.; Ou, Z.; Huang, S.; Meng, A. A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition. *Energy* **2019**, *189*, 116316. [\[CrossRef\]](#)
25. Higashiyama, K.; Fujimoto, Y.; Hayashi, Y. Feature Extraction of NWP Data for Wind Power Forecasting Using 3D-Convolutional Neural Networks. *Energy Procedia* **2018**, *155*, 350–358. [\[CrossRef\]](#)
26. Pan, J.S.; Hu, P.; Chu, S.C. Novel Parallel Heterogeneous Meta-Heuristic and Its Communication Strategies for the Prediction of Wind Power. *Processes* **2019**, *7*, 845. [\[CrossRef\]](#)
27. Cali, U.; Sharma, V. Short-term wind power forecasting using long-short term memory based recurrent neural network model and variable selection. *Int. J. Smart Grid Clean Energy* **2019**, *8*, 103–110. [\[CrossRef\]](#)
28. Jiang, P.; Li, R.; Li, H. Multi-objective algorithm for the design of prediction intervals for wind power forecasting model. *Appl. Math. Model.* **2019**, *67*, 101–122. [\[CrossRef\]](#)
29. Sun, M.; Feng, C.; Zhang, J. Multi-distribution ensemble probabilistic wind power forecasting. *Renew. Energy* **2020**, *148*, 135–149. [\[CrossRef\]](#)
30. Du, P. Ensemble Machine Learning-Based Wind Forecasting to Combine NWP Output with Data from Weather Station. *IEEE Trans. Sustain. Energy* **2019**, *10*, 2133–2141. [\[CrossRef\]](#)
31. Wang, C.; Zhang, H.; Ma, P. Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network. *Appl. Energy* **2020**, *259*, 114139. [\[CrossRef\]](#)
32. Wang, Z.; Zhang, J.; Zhang, Y.; Huang, C.; Wang, L. Short-Term Wind Speed Forecasting Based on Information of Neighboring Wind Farms. *IEEE Access* **2020**, *8*, 16760–16770. [\[CrossRef\]](#)
33. Sun, Z.; Zhao, S.; Zhang, J. Short-Term Wind Power Forecasting on Multiple Scales Using VMD Decomposition, K-Means Clustering and LSTM Principal Computing. *IEEE Access* **2019**, *7*, 166917–166929. [\[CrossRef\]](#)
34. Nourani Esfetang, N.; Kazemzadeh, R. A novel hybrid technique for prediction of electric power generation in wind farms based on WIPSO, neural network and wavelet transform. *Energy* **2018**, *149*, 662–674. [\[CrossRef\]](#)
35. Du, P.; Wang, J.; Yang, W.; Niu, T. A novel hybrid model for short-term wind power forecasting. *Appl. Soft Comput.* **2019**, *80*, 93–106. [\[CrossRef\]](#)
36. Hu, H.; Wang, L.; Lv, S.X. Forecasting energy consumption and wind power generation using deep echo state network. *Renew. Energy* **2020**, *154*, 598–613. [\[CrossRef\]](#)
37. Yildiz, C.; Acikgoz, H.; Korkmaz, D.; Budak, U. An improved residual-based convolutional neural network for very short-term wind power forecasting. *Energy Convers. Manag.* **2021**, *228*, 113731. [\[CrossRef\]](#)
38. Zheng, H.; Wu, Y. A XGBoost Model with Weather Similarity Analysis and Feature Engineering for Short-Term Wind Power Forecasting. *Appl. Sci.* **2019**, *9*, 3019. [\[CrossRef\]](#)
39. Bhatt, G.A.; Gandhi, P.R. Statistical and ANN based prediction of wind power with uncertainty. In Proceedings of the 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 23–25 April 2019; pp. 622–627. [\[CrossRef\]](#)
40. Shahid, F.; Zameer, A.; Mehmood, A.; Raja, M.A.Z. A novel wavenets long short term memory paradigm for wind power prediction. *Appl. Energy* **2020**, *269*, 115098. [\[CrossRef\]](#)
41. Dong, W.; Yang, Q.; Fang, X. Multi-Step Ahead Wind Power Generation Prediction Based on Hybrid Machine Learning Techniques. *Energies* **2018**, *11*, 1975. [\[CrossRef\]](#)
42. Hong, Y.Y.; Rioflorido, C.L.P.P. A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. *Appl. Energy* **2019**, *250*, 530–539. [\[CrossRef\]](#)
43. Zhang, Y.; Chen, B.; Zhao, Y.; Pan, G. Wind Speed Prediction of IPSO-BP Neural Network Based on Lorenz Disturbance. *IEEE Access* **2018**, *6*, 53168–53179. [\[CrossRef\]](#)
44. Naik, J.; Dash, S.; Dash, P.; Bisoi, R. Short term wind power forecasting using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network. *Renew. Energy* **2018**, *118*, 180–212. [\[CrossRef\]](#)
45. Zhang, C.; Peng, T.; Nazir, M.S. A novel hybrid approach based on variational heteroscedastic Gaussian process regression for multi-step ahead wind speed forecasting. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107717. [\[CrossRef\]](#)
46. Shen, Y.; Wang, X.; Chen, J. Wind Power Forecasting Using Multi-Objective Evolutionary Algorithms for Wavelet Neural Network-Optimized Prediction Intervals. *Appl. Sci.* **2018**, *8*, 185. [\[CrossRef\]](#)
47. Yuan, X.; Chen, C.; Jiang, M.; Yuan, Y. Prediction interval of wind power using parameter optimized Beta distribution based LSTM model. *Appl. Soft Comput.* **2019**, *82*, 105550. [\[CrossRef\]](#)
48. Li, R.; Jin, Y. A wind speed interval prediction system based on multi-objective optimization for machine learning method. *Appl. Energy* **2018**, *228*, 2207–2220. [\[CrossRef\]](#)
49. Qin, G.; Yan, Q.; Zhu, J.; Xu, C.; Kammen, D.M. Day-Ahead Wind Power Forecasting Based on Wind Load Data Using Hybrid Optimization Algorithm. *Sustainability* **2021**, *13*, 1164. [\[CrossRef\]](#)
50. Duan, J.; Wang, P.; Ma, W.; Tian, X.; Fang, S.; Cheng, Y.; Chang, Y.; Liu, H. Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and Correntropy Long Short-term memory neural network. *Energy* **2021**, *214*, 118980. [\[CrossRef\]](#)
51. Li, N.; He, F.; Ma, W. Wind Power Prediction Based on Extreme Learning Machine with Kernel Mean p-Power Error Loss. *Energies* **2019**, *12*, 673. [\[CrossRef\]](#)

52. Amila, T.; Peiris, J.J.; Rathnayake, U. Forecasting Wind Power Generation Using Artificial Neural Network: “Pawan Danawi” A Case Study from Sri Lanka. *J. Electr. Comput. Eng.* **2021**, *2021*, 5577547. [[CrossRef](#)]
53. Fu, Y.; Hu, W.; Tang, M.; Yu, R.; Liu, B. Multi-step Ahead Wind Power Forecasting Based on Recurrent Neural Networks. In Proceedings of the IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Kota Kinabalu, Malaysia, 7–8 October 2018; pp. 217–222. [[CrossRef](#)]
54. Liu, H.; Mi, X.W.; Li, Y.F. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. *Energy Convers. Manag.* **2018**, *156*, 498–514. [[CrossRef](#)]
55. Zhou, J.; Xu, X.; Huo, X.; Li, Y. Forecasting Models for Wind Power Using Extreme-Point Symmetric Mode Decomposition and Artificial Neural Networks. *Sustainability* **2019**, *11*, 650. [[CrossRef](#)]
56. Bazionis, I.K.; Georgilakis, P.S. Review of Deterministic and Probabilistic Wind Power Forecasting: Models, Methods, and Future Research. *Electricity* **2021**, *2*, 13–47. [[CrossRef](#)]
57. Jiao, R.; Huang, X.; Ma, X.; Han, L.; Tian, W. A Model Combining Stacked Auto Encoder and Back Propagation Algorithm for Short-Term Wind Power Forecasting. *IEEE Access* **2018**, *6*, 17851–17858. [[CrossRef](#)]
58. Ding, M.; Zhou, H.; Xie, H.; Wu, M.; Nakanishi, Y.; Yokoyama, R. A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting. *Neurocomputing* **2019**, *365*, 54–61. [[CrossRef](#)]
59. Wang, S.X.; Li, M.; Zhao, L.; Jin, C. Short-term wind power prediction based on improved small-world neural network. *Neural Comput. Appl.* **2019**, *31*, 3173–3185. [[CrossRef](#)]
60. Torabi, A.; Mousavy, S.A.K.; Dashti, V.; Saeedi, M.; Yousefi, N. A New Prediction Model Based on Cascade NN for Wind Power Prediction. *Comput. Econ.* **2019**, *53*, 1219–1243. [[CrossRef](#)]
61. Memarzadeh, G.; Keynia, F. A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets. *Energy Convers. Manag.* **2020**, *213*, 112824. [[CrossRef](#)]
62. Shivani; Sandhu, K.; Ramachandran Nair, A. A Comparative Study of ARIMA and RNN for Short Term Wind Speed Forecasting. In Proceedings of the 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6–8 July 2019; pp. 1–7. [[CrossRef](#)]
63. Sopena, J.M.G.; Pakrashi, V.; Ghosh, B. Can we improve short-term wind power forecasts using turbine-level data? A case study in Ireland. In Proceedings of the IEEE Madrid PowerTech, Madrid, Spain, 28 June–2 July 2021; pp. 1–6. [[CrossRef](#)]
64. Saini, V.K.; Kumar, R.; Mathur, A.; Saxena, A. Short term forecasting based on hourly wind speed data using deep learning algorithms. In Proceedings of the 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), Jaipur, India, 7–8 February 2020; pp. 1–6. [[CrossRef](#)]
65. Reddy, V.; Verma, S.M.; Verma, K.; Kumar, R. Hybrid Approach for Short Term Wind Power Forecasting. In Proceedings of the 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; pp. 1–5. [[CrossRef](#)]
66. Hu, T.; Wu, W.; Guo, Q.; Sun, H.; Shi, L.; Shen, X. Very short-term spatial and temporal wind power forecasting A deep learning approach. *CSEE J. Power Energy Syst.* **2020**, *6*, 434–443. [[CrossRef](#)]
67. Karthikeyan, M.; Rengaraj, R. Short-Term Wind Speed Forecasting Using Ensemble Learning. In Proceedings of the 7th International Conference on Electrical Energy Systems (ICEES), Chennai, India, 11–13 February 2021; pp. 502–506. [[CrossRef](#)]
68. Liu, Y.; Ghasemkhani, A.; Yang, L.; Zhao, J.; Zhang, J.; Vittal, V. Seasonal Self-evolving Neural Networks Based Short-term Wind Farm Generation Forecast. In Proceedings of the IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Tempe, AZ, USA, 11–13 November 2020; pp. 1–6. [[CrossRef](#)]
69. Verma, S.M.; Reddy, V.; Verma, K.; Kumar, R. Markov Models Based Short Term Forecasting of Wind Speed for Estimating Day-Ahead Wind Power. In Proceedings of the International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 22–23 February 2018; pp. 31–35. [[CrossRef](#)]
70. Zhang, P.; Wang, Y.; Liang, L.; Li, X.; Duan, Q. Short-Term Wind Power Prediction Using GA-BP Neural Network Based on DBSCAN Algorithm Outlier Identification. *Processes* **2020**, *8*, 157. [[CrossRef](#)]
71. Niu, D.; Pu, D.; Dai, S. Ultra-Short-Term Wind-Power Forecasting Based on the Weighted Random Forest Optimized by the Niche Immune Lion Algorithm. *Energies* **2018**, *11*, 1098. [[CrossRef](#)]
72. Liu, Y.; Guan, L.; Hou, C.; Han, H.; Liu, Z.; Sun, Y.; Zheng, M. Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform. *Appl. Sci.* **2019**, *9*, 1108. [[CrossRef](#)]
73. Zhou, J.; Yu, X.; Jin, B. Short-Term Wind Power Forecasting: A New Hybrid Model Combined Extreme-Point Symmetric Mode Decomposition, Extreme Learning Machine and Particle Swarm Optimization. *Sustainability* **2018**, *10*, 3202. [[CrossRef](#)]
74. Shi, X.; Lei, X.; Huang, Q.; Huang, S.; Ren, K.; Hu, Y. Hourly Day-Ahead Wind Power Prediction Using the Hybrid Model of Variational Model Decomposition and Long Short-Term Memory. *Energies* **2018**, *11*, 3227. [[CrossRef](#)]
75. Wang, S.; Zhao, X.; Li, M.; Wang, H. TRSWA-BP Neural Network for Dynamic Wind Power Forecasting Based on Entropy Evaluation. *Entropy* **2018**, *20*, 284. [[CrossRef](#)]
76. An, G.; Jiang, Z.; Chen, L.; Cao, X.; Li, Z.; Zhao, Y.; Sun, H. Ultra Short-Term Wind Power Forecasting Based on Sparrow Search Algorithm Optimization Deep Extreme Learning Machine. *Sustainability* **2021**, *13*, 10453. [[CrossRef](#)]
77. Donadio, L.; Fang, J.; Porté-Agel, F. Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast. *Energies* **2021**, *14*, 338. [[CrossRef](#)]

78. Kosana, V.; Teeparthi, K.; Madasthu, S.; Kumar, S. A novel reinforced online model selection using Q-learning technique for wind speed prediction. *Sustain. Energy Technol. Assess.* **2022**, *49*, 101780. [[CrossRef](#)]
79. Duan, J.; Wang, P.; Ma, W.; Fang, S.; Hou, Z. A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting. *Int. J. Electr. Power Energy Syst.* **2022**, *134*, 107452. [[CrossRef](#)]
80. Xiang, L.; Liu, J.; Yang, X.; Hu, A.; Su, H. Ultra-short term wind power prediction applying a novel model named SATCN-LSTM. *Energy Convers. Manag.* **2022**, *252*, 115036. [[CrossRef](#)]
81. Hanifi, S.; Lotfian, S.; Seyedzadeh, S.; Zare-Behtash, H.; Cammarano, A. Offshore wind power forecasting—A new hyperparameter optimisation algorithm for deep learning models. *Energies* **2022**, *15*, 6919. [[CrossRef](#)]
82. Farah, S.; David, A.W.; Humaira, N.; Aneela, Z.; Steffen, E. Short-term multi-hour ahead country-wide wind power prediction for Germany using gated recurrent unit deep learning. *Renew. Sustain. Energy Rev.* **2022**, *167*, 112700. [[CrossRef](#)]
83. Kosovic, B.; Haupt, S.E.; Adriaansen, D.; Alessandrini, S.; Wiener, G.; Delle Monache, L.; Liu, Y.; Linden, S.; Jensen, T.; Cheng, W.; et al. A Comprehensive Wind Power Forecasting System Integrating Artificial Intelligence and Numerical Weather Prediction. *Energies* **2020**, *13*, 1372. [[CrossRef](#)]
84. Yu, R.; Gao, J.; Yu, M.; Lu, W.; Xu, T.; Zhao, M.; Zhang, J.; Zhang, R.; Zhang, Z. LSTM-EFG for wind power forecasting based on sequential correlation features. *Future Gener. Comput. Syst.* **2019**, *93*, 33–42. [[CrossRef](#)]
85. Yu, R.; Liu, Z.; Li, X.; Lu, W.; Ma, D.; Yu, M.; Wang, J.; Li, B. Scene learning: Deep convolutional networks for wind power prediction by embedding turbines into grid space. *Appl. Energy* **2019**, *238*, 249–257. [[CrossRef](#)]
86. Oh, E.; Wang, H. Reinforcement-Learning-Based Energy Storage System Operation Strategies to Manage Wind Power Forecast Uncertainty. *IEEE Access* **2020**, *8*, 20965–20976. [[CrossRef](#)]
87. Wang, X.; Li, P.; Yang, J. Short-term Wind Power Forecasting Based on Two-stage Attention Mechanism. *IET Renew. Power Gener.* **2019**, *14*, 297–304. [[CrossRef](#)]
88. Dorado-Moreno, M.; Navarin, N.; Gutierrez, P.; Prieto, L.; Sperduti, A.; Salcedo-Sanz, S.; Hervas-Martinez, C. Multi-task learning for the prediction of wind power ramp events with deep neural networks. *Neural Netw.* **2020**, *123*, 401–411. [[CrossRef](#)]
89. Ma, Y.J.; Zhai, M.Y. A Dual-Step Integrated Machine Learning Model for 24h-Ahead Wind Energy Generation Prediction Based on Actual Measurement Data and Environmental Factors. *Appl. Sci.* **2019**, *9*, 2125. [[CrossRef](#)]
90. Khosravi, A.; Machado, L.; Nunes, R. Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil. *Appl. Energy* **2018**, *224*, 550–566. [[CrossRef](#)]
91. Wang, R.; Li, C.; Fu, W.; Tang, G. Deep Learning Method Based on Gated Recurrent Unit and Variational Mode Decomposition for Short-Term Wind Power Interval Prediction. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 3814–3827. [[CrossRef](#)] [[PubMed](#)]
92. Han, L.; Jing, H.; Zhang, R.; Gao, Z. Wind power forecast based on improved Long Short Term Memory network. *Energy* **2019**, *189*, 116300. [[CrossRef](#)]
93. Sharifian, A.; Ghadi, M.J.; Ghavidel, S.; Li, L.; Zhang, J. A new method based on Type-2 fuzzy neural network for accurate wind power forecasting under uncertain data. *Renew. Energy* **2018**, *120*, 220–230. [[CrossRef](#)]
94. He, Y.; Li, H. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. *Energy Convers. Manag.* **2018**, *164*, 374–384. [[CrossRef](#)]
95. Shao, H.; Deng, X.; Jiang, Y. A novel deep learning approach for short-term wind power forecasting based on infinite feature selection and recurrent neural network. *J. Renew. Sustain. Energy* **2018**, *10*, 043303. [[CrossRef](#)]
96. Dong, D.; Sheng, Z.; Yang, T. Wind Power Prediction Based on Recurrent Neural Network with Long Short-Term Memory Units. In Proceedings of the 2018 International Conference on Renewable Energy and Power Engineering (REPE), Toronto, ON, Canada, 24–26 November 2018; pp. 34–38. [[CrossRef](#)]
97. Yang, X.; Zhang, Y.; Yang, Y.; Lv, W. Deterministic and Probabilistic Wind Power Forecasting Based on Bi-Level Convolutional Neural Network and Particle Swarm Optimization. *Appl. Sci.* **2019**, *9*, 1794. [[CrossRef](#)]
98. Ak, R.; Li, Y.F.; Vitelli, V.; Zio, E. Adequacy assessment of a wind-integrated system using neural network-based interval predictions of wind power generation and load. *Int. J. Electr. Power Energy Syst.* **2018**, *95*, 213–226. [[CrossRef](#)]
99. Sun, G.; Jiang, C.; Cheng, P.; Liu, Y.; Wang, X.; Fu, Y.; He, Y. Short-term wind power forecasts by a synthetical similar time series data mining method. *Renew. Energy* **2018**, *115*, 575–584. [[CrossRef](#)]
100. Sitharthan, R.; Parthasarathy, T.; Sheeba Rani, S.; Ramya, K.C. An improved radial basis function neural network control strategy-based maximum power point tracking controller for wind power generation system. *Trans. Inst. Meas. Control* **2019**, *41*, 3158–3170. [[CrossRef](#)]
101. Abedinia, O.; Lotfi, M.; Bagheri, M.; Sobhani, B.; Shafeikhah, M.; Catalao, J.P.S. Improved EMD-Based Complex Prediction Model for Wind Power Forecasting. *IEEE Trans. Sustain. Energy* **2020**, *11*, 2790–2802. [[CrossRef](#)]
102. Hammerschmitt, B.; Lucchese, F.C.; Silva, L.N.; Abaide, A.R.; Rohr, J.A.D.A. Short-Term Generation Forecasting Against the High Penetration of the Wind Energy. In Proceedings of the 8th International Conference on Modern Power Systems (MPS), Cluj-Napoca, Romania, 21–23 May 2019; pp. 1–6. [[CrossRef](#)]
103. Abesamis, K.; Ang, P.; Bisquera, F.I.; Catabay, G.; Tindogan, P.; Ostia, C.; Pacis, M. Short-Term Wind Power Forecasting Using Structured Neural Network. In Proceedings of the IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines, 29 November–1 December 2019; pp. 1–4. [[CrossRef](#)]

104. Bhushan Sahay, K.; Srivastava, S. Short-Term Wind Speed Forecasting of Lelystad Wind Farm by Using ANN Algorithms. In Proceedings of the International Electrical Engineering Congress (iEECON), Krabi, Thailand, 7–9 March 2018; pp. 1–4. [CrossRef]
105. Vaitheswaran, S.S.; Ventrapragada, V.R. Wind Power Pattern Prediction in time series measurement data for wind energy prediction modelling using LSTM-GA networks. In Proceedings of the 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6–8 July 2019; pp. 1–5. [CrossRef]
106. Wang, S.; Sun, Y.; Zhai, S.; Hou, D.; Wang, P.; Wu, X. Ultra-Short-Term Wind Power Forecasting Based on Deep Belief Network. In Proceedings of the Chinese Control Conference (CCC), Guangzhou, China, 27–30 July 2019; pp. 7479–7483. [CrossRef]
107. Medina, S.V.; Ajenjo, U.P. Performance Improvement of Artificial Neural Network Model in Short-term Forecasting of Wind Farm Power Output. *J. Mod. Power Syst. Clean Energy* **2020**, *8*, 484–490. [CrossRef]
108. Pang, D.; Li, G. Research on Wind Power Output Reconstruction Technology Based on Multiple Time Scales. In Proceedings of the 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20–22 October 2018; pp. 1–5. [CrossRef]
109. Prabha, P.P.; Vanitha, V.; Resmi, R. Wind Speed Forecasting using Long Short Term Memory Networks. In Proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur, India, 5–6 July 2019; Volume 1, pp. 1310–1314. [CrossRef]
110. Musbah, H.; Aly, H.H.; Little, T.A. A Novel Approach for Seasonality and Trend Detection using Fast Fourier Transform in Box-Jenkins Algorithm. In Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), London, ON, Canada, 30 August–2 September 2020; pp. 1–5. [CrossRef]
111. Viet, D.T.; Phuong, V.V.; Duong, M.Q.; Tran, Q.T. Models for Short-Term Wind Power Forecasting Based on Improved Artificial Neural Network Using Particle Swarm Optimization and Genetic Algorithms. *Energies* **2020**, *13*, 2873. [CrossRef]
112. Liu, T.; Huang, Z.; Tian, L.; Zhu, Y.; Wang, H.; Feng, S. Enhancing Wind Turbine Power Forecast via Convolutional Neural Network. *Electronics* **2021**, *10*, 261. [CrossRef]
113. Tang, B.; Chen, Y.; Chen, Q.; Su, M. Research on Short-Term Wind Power Forecasting by Data Mining on Historical Wind Resource. *Appl. Sci.* **2020**, *10*, 1295. [CrossRef]
114. Liu, B.; Zhao, S.; Yu, X.; Zhang, L.; Wang, Q. A Novel Deep Learning Approach for Wind Power Forecasting Based on WD-LSTM Model. *Energies* **2020**, *13*, 4964. [CrossRef]
115. Bochenek, B.; Jurasz, J.; Jaczewski, A.; Stachura, G.; Sekuła, P.; Strzyżewski, T.; Wdowikowski, M.; Figurski, M. Day-Ahead Wind Power Forecasting in Poland Based on Numerical Weather Prediction. *Energies* **2021**, *14*, 2164. [CrossRef]
116. Li, Y.; Yang, F.; Zha, W.; Yan, L. Combined Optimization Prediction Model of Regional Wind Power Based on Convolution Neural Network and Similar Days. *Machines* **2020**, *8*, 80. [CrossRef]
117. Gouveia, H.T.V.; De Aquino, R.R.B.; Ferreira, A.A. Enhancing Short-Term Wind Power Forecasting through Multiresolution Analysis and Echo State Networks. *Energies* **2018**, *11*, 824. [CrossRef]
118. Ahmadi, A.; Nabipour, M.; Mohammadi-Ivatloo, B.; Amani, A.M.; Rho, S.; Piran, M.J. Long-Term Wind Power Forecasting Using Tree-Based Learning Algorithms. *IEEE Access* **2020**, *8*, 151511–151522. [CrossRef]
119. Liu, H.; Chen, C.; Lv, X.; Wu, X.; Liu, M. Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods. *Energy Convers. Manag.* **2019**, *195*, 328–345. [CrossRef]
120. Al-Zadjali, S.; Al Maashri, A.; Al-Hinai, A.; Al Abri, R.; Gajare, S.; Al Yahyai, S.; Bakhtvar, M. A Fast and Accurate Wind Speed and Direction Nowcasting Model for Renewable Energy Management Systems. *Energies* **2021**, *14*, 7878. [CrossRef]
121. Ding, C.; Zhou, Y.; Ding, Q.; Li, K. Integrated Carbon-Capture-Based Low-Carbon Economic Dispatch of Power Systems Based on EEMD-LSTM-SVR Wind Power Forecasting. *Energies* **2022**, *15*, 1613. [CrossRef]
122. Yang, H.; Jiang, P.; Wang, Y.; Li, H. A fuzzy intelligent forecasting system based on combined fuzzification strategy and improved optimization algorithm for renewable energy power generation. *Appl. Energy* **2022**, *325*, 119849. [CrossRef]
123. Madsen, H.; Pinson, P.; Kariniotakis, G.; Nielsen, H.A.; Nielsen, T.S. A Protocol for Standardizing the Performance Evaluation of Short Term Wind Power Prediction Models. *Wind Eng.* **2005**, *29*, 475–489. [CrossRef]
124. Hyndman, R.J.; Koehler, A.B. Another Look at Measures of Forecast Accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]