Renewable Energy 201 (2022) 1076–1085

Contents lists available at ScienceDirect

Renewable Energy

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

Enhancing wind power forecast accuracy using the weather research and forecasting numerical model-based features and artificial neuronal networks

António Couto *, Ana Estanqueiro

LNEG - Laboratório Nacional de Energia e Geologia, Lisboa, Portugal

ARTICLE INFO

Keywords:
Wind power forecast
Wind power variability
Meteorological parameters
NWP model
Feature selection

ABSTRACT

Forecasting with accuracy the quantity of energy produced by wind power plants is crucial to enabling its optimal integration into power systems and electricity markets. Despite the remarkable improvements in the wind forecasting systems in recent years, large errors can still be observed, especially for longer time horizons. This work focuses on identifying new numerical weather prediction (NWP)-based features aiming to improve the overall quality of wind power forecasts.

The methodology also incorporates a sequential forward feature selection algorithm. This algorithm was designed to select iteratively the meteorological features which minimize the wind forecast errors.

The methodology was applied separately to seven wind parks in Portugal with different climate characteristics. The proposed approach allowed a reduction between 13% and 37% in the root mean square errors of wind power forecasts, compared with a baseline scenario. While the meteorological features identified for each wind park showed similarities within regions with analogous wind power generation profiles, each wind park required specific meteorological parameters as input data to obtain the best performance. Thus, the results show to be crucial to select the most relevant features of a specific site to maximize the accuracy of a wind power forecast.

1. Introduction

Accurate forecasting of wind power production is important for the efficient and safe operation of a power system, especially for large shares of embedded wind generation [1,2]. In addition, in most electricity markets, a better wind power forecast also enables to reduce the need of balancing energy from the reserve markets, which often incurs a high cost that reduces the profitability of the wind power producers [3].

Forecast approaches can be split into three main categories: i) statistical-based approaches using historical time series, ii) physical approaches based on the use of numerical weather prediction models (NWP), and iii) hybrid approaches that combine two or more models of similar or different nature [4]. For the forecast time horizon of interest in this work, numerous approaches have been developed in recent years to forecast wind energy based on NWP models [4]. Detailed reviews may be found in the literature [5,6]. Notwithstanding the advances observed on NWP outcomes, systematic errors still persist in their outputs due to the chaotic nature of the atmosphere, whereby small initial errors can grow within a deterministic system and eventually result in the failure of the forecast for a long time horizon [7]. To overcome some of the drawbacks of these models, statistical downscaling techniques of the NWP outputs have been proposed, providing location-specific wind power forecasts [8]. Several approaches have been developed that show a higher capability of improving forecast accuracy when compared with a purely NWP-based approach [9].

Also crucial for maximizing the accuracy of NWP are the meteorological parameters used as input data for a power forecast system [6,10]. In most of the existing literature, the focus is essentially on the development of statistical methods, often neglecting the potential of NWP accurate information to improve the accuracy of forecasts. While numerical models provide a variety of different meteorological parameters that can be used as inputs [11], most studies use only wind speed and direction as input variables in their forecast systems [10]. A few others include parameters such as air pressure, wind shear, temperature and humidity, which also influence the conversion efficiency [12,13]. However, recent works have shown that a careful selection of input variables for statistical methods can strongly improve the accuracy of such wind speed/power forecasts [11,14].

* Corresponding author.
E-mail addresses: antonio.couto@lneg.pt (A. Couto), ana.estanqueiro@lneg.pt (A. Estanqueiro).

https://doi.org/10.1016/j.renene.2022.11.022
Received 9 August 2022; Received in revised form 18 October 2022; Accepted 5 November 2022
Available online 12 November 2022

Methodologies based on principal component analysis (PCA) have proved to be useful for assessing wind power variability. In the region under study in this work, which is strongly influenced by climatic patterns such as the North Atlantic Oscillation (an important source of predictability for wind power [15,16]) some authors successfully used weather-type approaches based on PCA to understand and predict extreme events as wind power ramps [17] as well as the monthly wind power resource variability and production [18].

Another common approach in the literature is the use of PCA as a data dimensionality reduction technique [14]. The outputs from the PCA are used to feed the statistical methodologies enabling to obtain the wind power forecasts. A PCA approach coupled with a random forests filter algorithm to identify the spatial patterns of meteorological processes that are related to strong wind variability was used in Ref. [19]. The authors suggest that meteorological parameters, such as the geopotential height, are crucial to improving the characterization of wind speed variability. With a similar approach, based on PCA coupled with a mutual information approach, different degrees of meteorological connection with wind power ramps were detected [20]. They later identified the zonal and meridian wind components at different pressure levels, as well as the geopotential height, as the most important features that characterize wind power variability, especially during wind power ramp events.

An improvement of up to 20% in the hourly and daily wind speed forecast can be obtained when a feature selection approach to select the most meteorological input parameters was applied [11]. The model that ranked first in the European Energy Market 2020 wind power forecasting competition is based on a physics-oriented pre-processing system [21]. This model includes a NWP parameters selection approach that had a positive impact on the model performance. The authors identified 25 parameters as the most appropriate to improve the forecast accuracy up to 20% in hourly and daily wind speed accuracy from a set of 98 predictors extracted for a NWP single spatial point. These predictors include the meridional and zonal wind components at the surface and upper levels (500 hPa), as well as the vertical gradient of temperature, which is a measure of atmospheric instability.

As highlighted in some aforementioned works, another feature from NWP that has been explored is the use of spatial grid data information, in contrast with the use of data from just a single spatial point or an average of the points of the NWP domain that surrounds a wind park. Indeed, the variability of wind power observed in a specific location depends not only on local wind dynamics, but it is also strongly influenced by large-scale atmospheric patterns [22]. While these patterns may be correctly simulated for a larger area, the time series may show deviations in a specific location [23]. By using NWP spatial grid data information, it is possible to assimilate the dynamics of neighbouring regions to complement the current local information used for forecasting. Several spatial approaches to extract the most relevant information from a set of NWP grid data were implemented in Ref. [22]. In the case of wind power, the authors concluded the PCA approach helps to obtain the best performance.

1.1. Research gap

As described by several authors (e.g. Ref. [14]) one of the most relevant research directions for improving wind energy forecasting is the use of additional exogenous input parameters. Most of the existing works are limited to a small set of meteorological parameters such as wind speed and direction, air temperature and pressure to feed the wind power forecast systems. Even the works that explore up-to-date statistical approaches do not explore all the information available from the NWP models (e.g. Ref. [24]), which provide information regarding meteorological conditions with impact on the wind power output [14]. To the best knowledge of the authors, no previous research has investigated the benefits of combining optimally a large number of meteorological parameters with the use of PCA, neither highlight the importance of some meteorological parameters with impact in the wind power variability (e.g., vertical temperature gradient) to improve the accuracy of wind power forecasts. Nevertheless, if not properly selected, the use of a large amount of input data can lead to an underperformance of the forecasts [11]. In addition, the similarities in NWP meteorological features identified among regions/wind parks under distinct climate conditions were not addressed in the aforementioned works.

1.2. Novel contributions

The aim of this work is to identify new NWP-based meteorological features capable of improving the overall quality of a deterministic wind power forecast. Thus, this work addresses the identification and selection of the meteorological input data that proves to be most effective in improving the wind power forecast accuracy. Parameters traditionally not considered in the forecast systems were considered by exploring meteorological data derived directly from a NWP and others with a known influence on wind power variability [21], such as wind shear [13] or atmospheric instability, which can be computed based on the data from these models. The meteorological parameters include information from surface-level (e.g., mean sea level pressure) and vertical levels (e.g., air temperature). At the same time, this work contributes to identify if the use of spatial grid data instead of data from a single point enables to improve the forecast accuracy as discussed by several authors (e.g. Ref. [22]). Therefore, rather than explore new statistical approaches - an usual artificial neural network (ANN) technique is used - this methodology provides insights into how much the accuracy of a forecast can be improved by exploring all the capabilities of a NWP. The work considers seven wind parks located in distinct climatic regions in Portugal, thus enabling to provide some insights regarding the most adequate meteorological parameters for typical regions (e.g., coastal and mountain regions). The forecasting methodology uses a greedy sequential forward feature selection algorithm to choose iteratively the meteorological features that minimize an objective function. This algorithm embeds a statistical approach based on an artificial neuronal network, which is one of the most used approaches in wind power forecast systems.

The remainder of the paper is structured as follows: Section 2 describes the data and the methodology used in this work, Section 3 presents the results obtained for different wind parks, and, finally, in section 4 some final remarks are provided taking into consideration the main findings of this work.

2. Data and methodology

The period used to apply the methodology developed in this work is from January 1, 2015 to December 31, 2016. The data were split into two subsets: one for calibration, which includes training and testing; and another for validation. The validation dataset comprises two months, one in summer (August 2016) and the other in winter (December 2016). The remaining period (twenty-two months) is used to calibrate the methodology presented in this work.

2.1. NWP and wind power data

2.1.1. NWP data

NWP numerical models, such as the publicly available Weather Research and Forecasting (WRF) model [25], allow the grid output for a given limited region to be obtained without the need to use an extensive and expensive monitoring network. WRF models are able to accurately simulate the evolution of air masses using the Navier-Stokes equations, physical parameterizations, and boundary and lateral conditions data. For this reason, wind energy forecasting systems for time horizons above 6 h strongly rely on data from these models.

In this work, data from the WRF model provided by MeteoGalicia are used [26]. These data are freely available and comprise various historical meteorological forecast parameters. Three nested domains are used with the following spatial grid resolution: 36, 12 and 4 km. The 4 km domain does not cover all the locations analysed in this work, therefore, data from the 12 km domain was used. The longitude of the domain ranges from −21.33° to 6.16°, while the latitude is 33.78° to 49.46°. The time resolution is 1 h. Each WRF run is initialized at 00h UTC and the time horizon considered is 48 h ahead, comprising the time frames used by the Iberian Electricity Market. Here, only the last 24 h are analysed since bids in the electricity markets are based on this period and thus would directly benefit from improvements in the forecast accuracy.

In this work, most of the outputs provided by the WRF model are used and tested to identify new meteorological features that improve the accuracy of a wind power forecast (Table 1).

In addition to the outputs from the NWP model, new parameters that can influence the variability of wind speed and the wind-to-power conversion process were computed for the same grid-points. These variables are mean sea level pressure gradient ($MSLP_{Grad}$), vertical temperature variation ($T_{var.}$), wind shear (WindShear), and wind power density (WPD). $MSLP_{Grad}$ was calculated from the mean sea level pressure (MSLP) field using a first-order centred finite difference, Equation (1).

$$MSLP_{Grad} = \frac{\partial MSLP}{\partial x}\hat{i} + \frac{\partial MSLP}{\partial y}\hat{j} \approx \frac{MSLP_{i+1,j} - MSLP_{i-1,j}}{2\Delta x} + \frac{MSLP_{i,j+1} - MSLP_{i,j-1}}{2\Delta y} \tag{1}$$

In the previous equation *i-th* and *j-th* represent the spatial point of the domain, in the *x* and *y* coordinates of projection, respectively. Δx and Δy represent the spatial resolution in the *x* and *y* coordinate of projection, respectively.

The vertical temperature variation was determined using the temperature between two pressure levels from the NWP model at 2 m and 850 hPa, Equation (2):

$$T_{var.} = T_{850} - T_2 \tag{2}$$

The hourly wind shear parameter was determined based on the power law profile, which assumes that the wind speed varies with the height above ground, according to Equations (3) and (4):

$$WS(z) = \beta z^{WindShear} \tag{3}$$

$$\equiv \ln(WS(z)) = (WindShear \times \ln(z)) + \ln(\beta) \tag{4}$$

The previous expression is in the general form: $y = mx + b$. Therefore, the wind shear is assumed to be provided by the slope determined using a linear least squares algorithm that fits the NWP wind speed vertical profile. The wind speed heights used were $WS_{\sigma 0.994}$ and $WS_{\sigma 0.986}$, which correspond to nearly 50 m and 110 m above ground level, respectively. In the previous equations, β is a constant and z is the height above ground level. The wind power density is an indicator of the maximum recoverable power as previously used to improve the wind power forecast accuracy [21]. This parameter is determined as described in Equation (5):

$$WPD = 0.5 \times \rho_{air} \times WS^3_{\sigma 0.986} \tag{5}$$

Due to the data available, ρ_{air} was computed for 2 m above ground level. The barometric formula was used to extrapolate the pressure values using the MSLP and T_2 data.

2.1.2. Wind parks data

Instead of randomly selecting a set of wind parks to apply the methods presented in this work, it was decided to use the information presented in Ref. [27]. In the aforementioned work, the authors applied a clustering algorithm to identify regional wind power generation patterns across the locations of existing wind power plants in Portugal. The authors identified eight (spatial) regions and associated these regions with different weather climates and local characteristics (see Fig. 1). As described in Ref. [27], regions 1 to 4 are mostly located in central/northern Portugal and are mountain areas with highly complex orography. The remaining regions (5–8) are mainly located in coastal regions (centre and south of Portugal) with low-to-moderately complex orography.

In this work, seven of these regions are analysed, using one wind park located in each region. This approach enables to identify similarities in the NWP meteorological features among the different regions/wind parks. All wind parks have approximately 20 MW of nominal capacity. For each wind park (WP), a common period of two years of data was gathered. For Region 8, it was not possible to obtain data from a wind park and, therefore, this region was not analysed.

Table 1
Meteorological parameters obtained from the NWP model and their respective acronyms.

Type of information	Meteorological Parameter	Acronym
Single level	Cloud cover at high levels	CHF
	Convective available potential energy	CAPE
	Convective inhibition	CIN
	Humidity relative at 2 m	HR
	Mean sea level pressure	MSLP
	Planetary boundary layer height	PBL
	Surface downwelling longwave flux	LWF
	Surface downwelling shortwave flux	SWF
	Surface wind speed gust	Gust
	Total accumulated convective rainfall	Rainfall
	Atmospheric visibility	Vis
Multiple level	Air temperature at 850, 500, and 2 m	T_{500},T_{850}, T_2
	Geopotential height at 850, 500, sigma level 0.986	HGT_{500}, HGT_{850}, $HGT_{\sigma 0.986}$
	U-wind component at sigma levels 0.994 and 0.986, and 10 m	$U_{\sigma 0.986}$, $U_{\sigma 0.994}$, U_{10}
	V-wind component at sigma levels 0.994 and 0.986, and 10 m	$V_{\sigma 0.986}$, $V_{\sigma 0.994}$, V_{10}
	Wind speed at sigma levels 0.994 and 0.986	$WS_{\sigma 0.986}$, $WS_{\sigma 0.994}$

2.2. Methodology to obtain a deterministic wind power forecast

Fig. 2 depicts a flowchart of the proposed methodology to identify the most relevant NWP meteorological features. In this case, the most relevant are the ones that minimize the wind power forecast normalized root mean square error (NRMSE) for each wind park under analysis. Nevertheless, the methodology followed in this work could easily incorporate other objective functions. The NRMSE is a commonly used metric to evaluate forecasts against observations, enabling to assess the amplitude and phase errors.

Fig. 2 provides a flowchart of the main steps and a brief description of each step is provided in the following subsections.

The benchmark forecast used in this work consists in feeding the statistical approach with single point information of the most common meteorological input data in the wind power forecast systems: wind intensity ($WS_{\sigma 0.986}$) and information regarding the wind direction ($U_{\sigma 0.986}$ and $V_{\sigma 0.986}$).

2.2.1. Principal component analysis

The application of principal component analysis (PCA) aims to detect the most relevant spatial-temporal synoptic variability modes, while cancelling the impact of local effects that occur in principal components (PCs) with reduced percentage of variance [28]. PCA has been used by several authors as a dataset reduction dimension technique [14]. Mathematically, each of the aforementioned meteorological parameters that have *j-th* spatial points and a temporal dependency can be

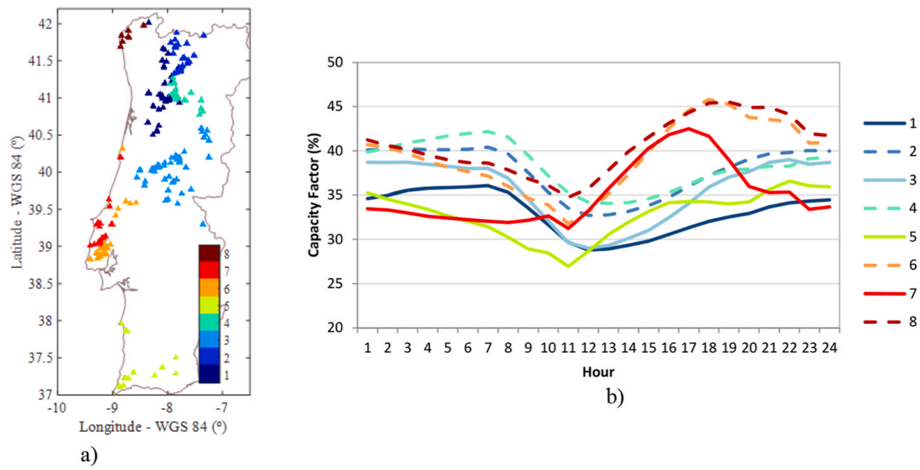

Fig. 1. a) Location of existing wind parks in Portugal (December 2018) and assignment of each wind park according to its geographical region, b) wind power daily average profile by region. Figure adapted from Ref. [27].

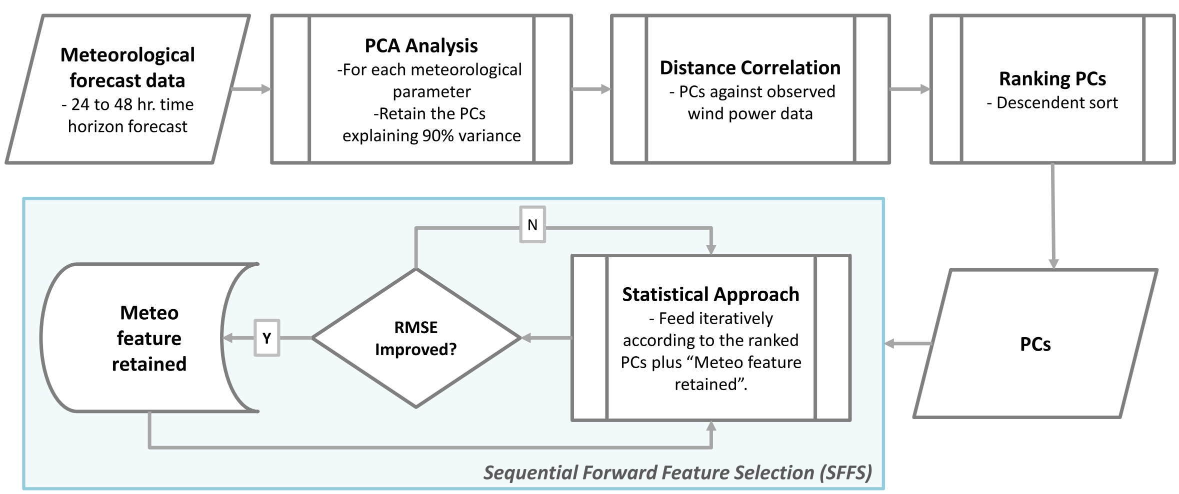

Fig. 2. Flowchart of the main steps to obtain the most relevant meteorological features. The steps in the blue box belong to the *K*-fold cross validation procedure.

decomposed through a linear combination of spatial and time weights, Equation (6).

$$Z = \sum_{j=1}^{N} a_{tj} e_j \tag{6}$$

where Z is the meteorological variable, a_{tj} are the PCs' scores, e_j is the eigenvector of the covariance matrix and N is the number of total spatial points [28]. The first PC explains the maximum amount of the variance for a specific parameter. From the residual variance, the second PC explains the maximum possible variance. This PC is uncorrelated with the first identified PC. The same happens until the *N-th* PC [28]. In a first step, in this work, the number of retained PCs for each meteorological parameter was determined according to the total explained variance. Typically, for forecast purposes, the total variance used ranges from 80% to 95% [22,29]. Therefore, since no predefined value was identified in the literature, the criterion to retain the number of PCs that explained 90% of the total variance for each meteorological parameter was used as a first approach.

Based on the NWP data, the PCA is applied for each meteorological variable after applying a data normalization process for each meteorological parameter, as shown in Equation (7).

$$Z_{t,j} = \frac{X_{t,j} - \overline{X}_j}{s_j} \tag{7}$$

In equation (7), $Z_{t,j}$ is the normalized value, $X_{t,j}$ is the original value, and $\overline{x}_j$ and s_j are the mean and standard deviations for *j-th* grid point, respectively.

2.2.2. Feature selection approach

Algorithms for the selection of the best input features aim to identify a subset of variables from a larger dataset, which can efficiently lead to more reliable predictions by reducing negative effects, from noise to irrelevant variables that might lead to an underperformance of the forecast model [30,31]. Thus, the benefits of this step are to: a) reduce the computational effort, b) simplify model selection procedures for accurate prediction, and c) improve the comprehension of complex dependencies between the predictor and the predicted variables [14].

Feature selection methods can be classified into three main types: filter, wrapper and embedded [32]. The filter methods remove the less significant variables *a priori*, and then a model is created based on this selection. In this case, for instance, the variables with low cross correlation values can be eliminated. The wrapping methods involve the entire training algorithm in the variable selection process, training several iterations (as many as there are variables) of the model by adding (or removing) variables to each one of them and evaluating the performance of the model obtained. The variables that help to improve the accuracy of the model are retained for the construction of the final model. On contrary, the variables that do not allow improving the accuracy of the model are discarded. Lastly, embedded methods introduce the variable selection process directly into the training process, in order to avoid the complete search that happens in integrated methods, thus reducing computational complexity.

In this study, to identify the most relevant meteorological features, a wrapper technique was implemented. Specifically, the Sequential Forward Feature Selection (SFFS) algorithm was selected [30]. The SFFS implemented is a greedy algorithm that chooses the "most attractive" solution in each iteration [11]. In this case, the algorithm attempts to find the optimal feature subset by selecting, iteratively, the meteorological PC that reduces the NRMSE value for the calibration period. Note that this type of approach may not be as effective compared with a meta-heuristic approach, in which different sets of features are evaluated together, as discussed in Ref. [11]. However, a greedy algorithm guarantees a good enough solution with reduced computational cost.

Considering the robustness of the SFFS results, i.e., the consistent production of a reliable feature subset, a ten-fold cross validation was employed by splitting the calibration period into training and testing data [30]. This step consists in removing randomly two months of data from the calibration period, which constitute the testing data, Fig. 3. Then, the SFFS algorithm is applied using the remaining training data corresponding to a total of twenty months. Afterwards, the NRMSE performance is evaluated using the data from the testing period. Ultimately, only the PCs that repeat eight or more times are used in the final results (validation dataset) to assess the benefits of the feature selection algorithm to improve the wind power forecast.

2.2.3. Ranking the PCs

The distance correlation [33] between each retained PC and the observed wind power is computed to rank in descending order the PCs that will interactively feed the SFFS algorithm, Fig. 4. The main advantage of the distance correlation, compared with the common Pearson's correlation, is its capability to detect both the linear and nonlinear associations between two random variables [33]. Distance correlation ranges from 0 to 1. A value of zero indicates independence between the PC and the observed wind power production from a given wind park, while a value of one implies linearity between the two time-series. The PC with the highest correlation value is used to start the SFFS algorithm.

2.2.4. Artificial neural network – statistical approach to obtain wind power forecast

An artificial neural network (ANN) methodology capable of assessing nonlinear relationships is used to forecast the wind power for a site [10]. Due to its capability to "learn by examples and experience", the ANN is a highly interesting technique for solving nonlinear [34]. Thus, the main advantage of the ANN is its learning capability, which enables it to approximate nonlinear functions to solve problems where the input-output connexion is not well defined or not easily computable [4].

The use of ANN involves two main steps: training and learning. One of the most effective learning algorithms in ANN is the backpropagation algorithm [35,36] This type of algorithm uses supervised learning to adjust the weights and biases in each unit to produce the desired response to the given inputs. Thus, the algorithm aims to reduce this error until the ANN learns the training data [35]. The training begins with random weights that are adjusted iteratively so that the error (the difference between estimated and expected results) based on input and output data will always be minimized to an acceptable value. Within backpropagation algorithms, the Levenberg-Marquardt (LM) algorithm is one of the most efficient training algorithms for methodologies based on artificial neural networks [34,36]. LM is an appropriate choice when it comes to the minimization of nonlinear functions, where a fast but efficient and stable convergence is required.

2.3. Metrics to assess the wind power forecast performance

Currently, a single unique metric that can describe or measure the performance of a forecasting methodology does not exist. For this reason, it is usual to analyse the performance of a variety of metrics in order to assess the methodology's strengths and weaknesses. In this work, the following metrics are used: NRMSE, Equation (8) and the Pearson correlation coefficient (r), Equation (9).

$$NRMSE = \frac{\sqrt{\frac{\sum_{t=1}^{T}\left(P_{For.}(t) - P_{Obs.}(t)\right)^2}{T}}}{NominalPower} \tag{8}$$

$$r = \frac{\sum\left(P_{For.}(t) - \overline{P_{For.}(t)}\right)\left(P_{Obs.}(t) - \overline{P_{Obs.}(t)}\right)}{\sqrt{\sum\left(P_{For.}(t) - \overline{P_{For.}(t)}\right)^2}\sqrt{\sum\left(P_{Obs.}(t) - \overline{P_{Obs.}(t)}\right)^2}} \tag{9}$$

where $P_{For.}$ and $P_{Obs.}$ are the wind power forecast and observed for each t-th time, respectively. NominalPower corresponds to the nominal power of each wind park. The optimal value for NRMSE is zero. The Pearson correlation coefficient measures the similarities between observed and the forecast data. This coefficient varies between −1 and 1. A value close to zero means a poor forecast, while the unit value represents perfect forecast without phase errors. A high negative value means that the forecast is in phase opposition.

To quantify the improvement of including additional meteorological parameters, the approach followed in Ref. [22] is used for each metric and wind park as follows, Equation (10).

$$Forecast\ improvement(\%) = \left(1 - \frac{Forecast_{Methodology}}{Forecast_{Benchmark}}\right) \times 100 \tag{10}$$

where $Forecast_{Methodology}$ represents the results of a specific metric for one wind park using the different forecast methodologies analysed in this work, and $Forecast_{Benchmark}$ represents the forecast results for the benchmark methodology. A positive *Forecast improvement* value indicates an improvement of the proposed forecast methodology. A negative value corresponds to an underperformance of the forecast methodology, with the benchmark showing the highest performance.

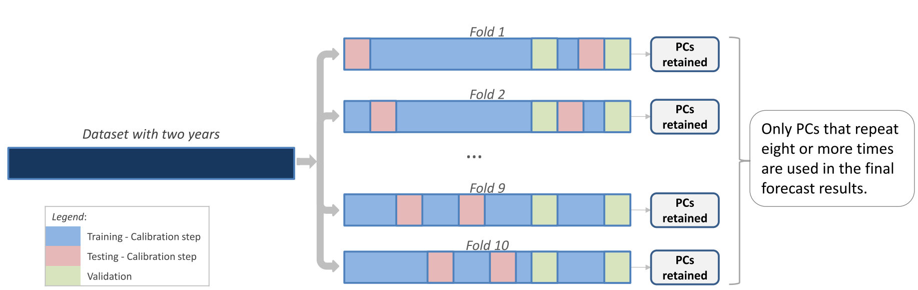

Fig. 3. Visual representation of the different subsets defined: calibration and validation.

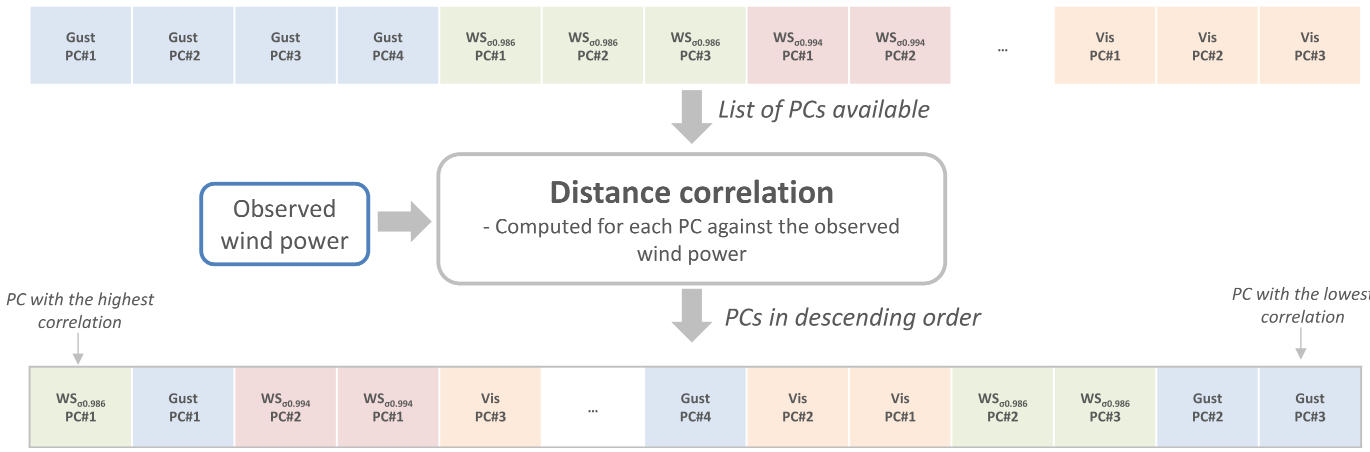

Fig. 4. Visual representation of the PCs ranking based on the distance correlation.

3. Results

The results for the different metrics are shown for the validation period, which comprises two months: a summer month (August 2016) and a winter month (December 2016). Using the criteria described in section 2.2.1 (minimum number of PCs that explain a total variance above 90%) the number of PCs obtained can be significant for some meteorological parameters such as rainfall that presents a high spatial and temporal variability. After preliminary tests, it was verified the retained PCs of order equal to or greater than five were unable to improve the prediction and were therefore discarded by the SFFS algorithm. This result is explained by the reduced variability explained by these higher order PCs that tend to represent only local effects that have no impact on the final results. As a consequence these PCs were removed from the analysis to reduce the computational effort. A total of 123 PCs were attained.

Regarding the ANN, there is no predefined rule to define its configuration. Therefore, general rules of thumb and sensitivity tests were followed for determining a proper configuration. The following parameters were imposed in the ANN: 1) *The number of hidden layers*: It was considered only one hidden layer; 2) *The different numbers of units in each layer*: In this case, the number of hidden units starts with two. Then, two-thirds the size of the input layer plus the size of the output layer is used [37]. When necessary, rounding up is performed. Thus, this number increases as the number of PC that helps to reduce the NRMSE are identified; 3) *The transfer function*: In this case, a sigmoid function was considered for each unit from the input layer to the hidden layer. From the hidden layer to the output layer, a linear function was used; and 4) *The learning algorithm: is* the Levenberg-Marquardt [34]. The methodology was carried out using the Neural Network Toolbox [38], which is available in the MATLAB software package.

3.1. Impact of the meteorological features in the NRMSE

As an example, Fig. 5 presents the NRMSE for the first-fold cross validation, for four of the wind parks under analysis. This figure illustrated the impact of testing all PCs identified. To easily compare the behaviour of the algorithm for different WPs, the data were normalized by the forecast value achieved using the PC that ranked in the first place.

From Fig. 5, it is possible to verify that the most accentuated decrease in the NRMSE is verified in the first ten components introduced in the SFFS algorithm. For the best case (WP 5), the reduction in the NRMSE is more than 30% of the initial value. The smallest decrease, around 20% of the initial value, was observed in WP 3. The remaining PCs allowed for further reduction of the errors, but in a less expressive way. WP 5 is the one that presents the most significant improvements as PCs with low distance correlation values are introduced in the algorithm.

After the PC ranked in 33rd place was introduced, it is possible to verify that WP 7 does not benefit from more information to improve its forecast results. On the other hand, in the case of WP 3, it is possible to verify that the PC ranked in 108th place with a low distance correlation value allows a slight improvement in the forecast.

This result highlights the importance of applying this type of feature selection algorithm, which does not filter *a priori* any PC that exhibits low distance correlation values with the observed wind power production. Thus, one can conclude that the first features are much more important to reduce the NRMSE of the forecasting model than the other remaining features. However, if only the first features are used as input, the overall performance of the forecasting will be worse than the one with all features combined, as shown in Fig. 5.

Despite the sharp decrease in the NRMSE values in the first ten principal components, it is possible to verify that there are some meteorological features that degrade the quality of the forecast and,

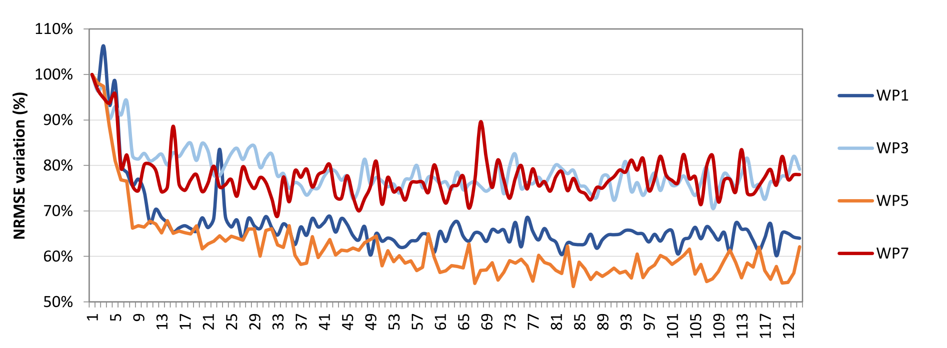

Fig. 5. NRMSE variation according to the number of PCs used in the SFFS algorithm for wind parks 1, 3, 5 and 7.

therefore, should not be considered in the final solution. Table 2 presents the top ten PCs for each wind park highlighting those that were selected through the SFFS algorithm. In the supplementary material, the list of the meteorological parameters used for each WP is presented.

According to this table, it is possible to verify that, in all regions, the PCs that present the highest distance correlation values with the observed production are those parameters related to the primary source of wind power – wind speed or wind gust. Specifically, the parameter wind gust, *Gust*, usually not applied in forecasting systems, is the one with the highest distance correlation values in wind parks located in the interior of central and northern Portugal in mountainous areas. In the other locations, the average wind speed at different heights is the parameter with the highest distance correlation values. Other parameters such as wind shear, planetary boundary layer and visibility also appear in the top 10 for several wind parks. These parameters are associated with atmospheric stability, which has an influence on wind power production. In some cases, they can have an active role in improving the forecast accuracy as depicted in Table 2.

3.2. Feature selection results

Table 3 shows quantity of PC retained for each parameter after applying the SFFS approach presented in this work.

From Table 3, it is possible to verify that only three variables are used in all regions: wind gust, and wind speed at two different heights. WP 5 is the one that uses more input features, 20 in the total. On the other hand, WP 7 only needs 11 to achieve its best performance.

Although it can vary at the vertical level selected, the V-wind component of the wind is another variable retained to reduce the NRMSE in all regions under analysis. By contrast, the U-wind component of the wind is most useful in wind parks located in coastal regions and with reduced topographical complexity (WPs 5–7). For these wind parks, the identification of vertical temperature variation (T_{var}) as a parameter is particularly important to reduce errors. The identification of these later two variables can be partially explained by the sea/land breezes that develop in these regions. This phenomenon is characterized by atmospheric instability caused by imbalances in the warming and cooling of the land surface and the sea.

For WPs 1–5, part of the atmospheric instability is introduced in the forecast model through variables such as the planetary boundary layer (PBL) height. The PBL varies throughout the day due to factors such as large-scale dynamics, cloudiness, convective mixing, and the diurnal cycle of solar radiation. The importance of carefully selecting the PBL scheme for wind power application was analysed in several works, e.g. Refs. [39,40], since this parameterization can have an important role in the characterization of the low-level wind structure [41]. In the specific case of the model used in this work, the PBL is computed using Yonsei State University (YSU) non-local closure parameterization [42]. According to the authors, the PBL height parameter is given by the height where the minimum turbulent flux (heat, momentum, moisture) occurs. Visibility also has an impact on the forecast performance for four of the seven wind parks analysed. This parameter also reflects atmospheric stability, since it takes into account aspects such as water vapour, cloud and rain water mixing ratios, as well as air pressure.

Regarding the variables that were calculated based on the model outcomes, as expected, power density is the parameter most used for different parks. The results also suggest that, with the methodology applied in this work, parameters at surface level such as mean sea level pressure, relative humidity, cloud cover at high levels, surface downwelling longwave flux, as well as information on the highest levels of the atmosphere (e.g. temperature at 500 hPa) are not useful in reducing forecast error.

3.3. Wind power forecast results

Fig. 6 presents the results of improvements in the NRMSE of the proposed forecast solution in comparison with the benchmark methodology used in this work. The Pearson's correlation results are provided in appendix B. The forecast improvement is computed for each WP using equation (10).

The results demonstrate an improvement in the performance of the forecast for all wind parks (WP). This improvement varies from 13% (WP1) to 37% (WP4) in the case of the NRMSE. As shown in Table 3, for WP4, it is possible to verify that the most common parameters in wind power forecast systems (associated with wind direction) are, in practice, discarded and only $V_{\sigma 0.994}$ is used. Another peculiarity of this park is the selection of information at high atmospheric levels, geopotential height at 500 hPa (HGT 500). The second best improvement in the forecast was observed for WP7 with a value of 33%. In this case, the most common variables are used as well as parameters associated with atmospheric stability, such as vertical temperature variation, $T_{var.}$. Although not very expressive, the results suggest that wind parks located in mountainous regions tend to benefit more from the forecast methodology proposed to reduce the NRMSE.

Although the results are not shown, it should be noted that the final weather features obtained for each park were tested in the remaining parks. In all cases, the results showed a lower performance than the one obtained for the model calibrated specifically for each wind park.

3.4. Understanding the benefit from using spatial grid data and the SFFS algorithm

In order to clearly understand the benefits of the proposed methodology - a feature selection with NWP grid data – identified by the acronym *PCA with SFFS*, two other forecast methods, that cover the most common approaches in the wind power forecast systems, are also analysed: *PCA without SFSS*, and P*oint with SFFS* (Table 4).

The first method (*PCA without SFFS*) uses the same meteorological parameters as the benchmark method (see section 2.2). In this case, the PCA technique is applied and all PCs that explain 90% of the total variance are retained, and used to feed the ANN approach. For the second model *Point with SFFS*", the feature selection algorithm is applied using only the meteorological information from one NWP grid-point (as in the benchmark methodology). The nearest point between the NWP

Table 2
List of the first ten meteorological parameters and the number of the PC for each wind park. The parameters in bold text are the ones used for the final forecast.

PC ranking	WP 1	WP 2	WP 3	WP 4	WP 5	WP 6	WP 7
1	Gust PC#1	Gust PC#1	Gust PC#1	Gust PC#1	$WS_{\sigma 0.994}$ PC#1	$WS_{\sigma 0.986}$PC#1	$WS_{\sigma 0.994}$ PC#1
2	$WS_{\sigma 0.986}$PC#1	$WS_{\sigma 0.986}$PC#1	$WS_{\sigma 0.986}$PC#1	$WS_{\sigma 0.986}$PC#1	$WS_{\sigma 0.986}$PC#1	$WS_{\sigma 0.994}$ PC#1	Gust PC#1
3	**$MSLP_{Grad}$** PC#1	$WS_{\sigma 0.994}$ PC#1	$WS_{\sigma 0.994}$ PC#1	$WS_{\sigma 0.994}$ PC#1	Gust PC#1	Gust PC#1	$WS_{\sigma 0.986}$PC#1
4	$WS_{\sigma 0.994}$ PC#1	**WPD** PC#1	**WPD** PC#1	**$MSLP_{Grad}$** PC#1	**WPD** PC#1	**WPD** PC#1	**WPD** PC#1
5	**WPD** PC#1	**$MSLP_{Grad}$** PC#1	**WindShear** PC#1	**WPD** PC#1	$WS_{\sigma 0.994}$ PC#2	**WindShear** PC#1	PBL PC#1
6	**WindShear** PC#2	**WPD** PC#2	$V_{\sigma 0.986}$PC#3	PBL PC#3	Gust PC#2	HGT $_{\sigma 0.986}$ PC#2	$V_{\sigma 0.986}$PC#3
7	PBL PC#3	$WS_{\sigma 0.994}$ PC#2	Vis PC#1	Vis PC#1	$WS_{\sigma 0.986}$PC#2	PBL PC#3	SWF PC#2
8	WS $_{\sigma 0.986}$PC#2	**WindShear** PC#2	PBL PC#3	ConvPrec PC#1	**WPD** PC#2	Gust PC#2	$U_{\sigma 0.994}$ PC#1
9	ConvPrec PC#1	$WS_{\sigma 0.986}$PC#2	**SWF PC#3**	**WindShear** PC#2	PBL PC#3	$V_{\sigma 0.986}$PC#3	U_{10} PC#1
10	Vis PC#1	Vis PC#1	PBL PC#1	**WPD** PC#2	HGT $_{\sigma 0.986}$ PC#2	$WS_{\sigma 0.986}$PC#2	$U_{\sigma 0.986}$PC#1

Table 3
Features selected with the SFFS algorithm for the different wind parks.Colours represent the total number of PCs used for each meteorological parameter..

Parameter	WP 1	WP 2	WP 3	WP 4	WP 5	WP 6	WP 7
CAPE	1	1					
CHF							
CIN	1						
Gust	1	1	1	1	2	2	1
HGT $_{\sigma 0.986}$			1			1	
HGT_{850}					1	1	
HGT $_{500}$				2			
HR							
LWF							
MSLP							
$MSLP_{Grad}$		1		1		1	
PBL	1	1	2	1	1		
Rainfall							
SWF	1	1	1		2		1
T_2							
T_{850}							
T_{500}							
T_{var}					1	1	1
U $_{10}$					1		1
U $_{\sigma 0.994}$					2		1
U $_{\sigma 0.986}$	2					1	1
V $_{10}$	1	1	4			1	1
V $_{\sigma 0.994}$		1	1	2			
V $_{\sigma 0.986}$	1	1	1		1	2	1
Vis	1			1	1		
WindShear	1		2		2	1	
WPD		2		2	2	2	1
WS $_{\sigma 0.994}$	1	1	1	1	2	1	1
WS $_{\sigma 0.986}$	2	1	1	1	2	1	1

Nr. of PCs used	0	1	2	3	4

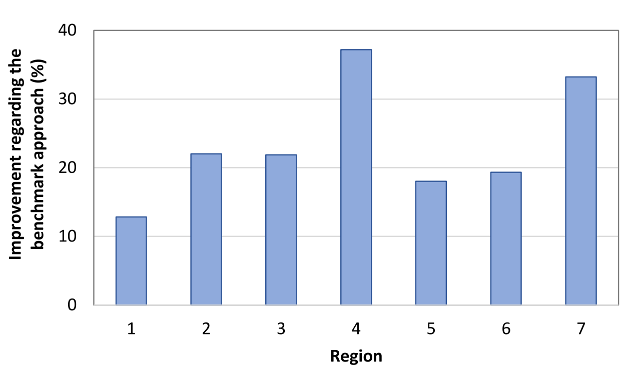

Fig. 6. Forecast improvements of RMSE over the benchmark by using the forecast approach proposed in this work for the different wind parks (WP).

Table 4
Characteristics of the scenarios analysed.

Forecast method	NWP data	All meteorological parameters	PCA	SFFS algorithm
Benchmark	Grid-point	No	No	No
PCA with SFFS	Grid-data	Yes	Yes	Yes
Spatial – PCA	Grid-data	No	Yes	No
Point with SFFS	Grid-point	Yes	No	Yes

grid points and the wind park is used. In Fig. 7, the quantification of the potential benefit of the different solutions against the benchmark approach considered in this work is presented.

From Fig. 7, it is possible to verify that the results are consistent across all wind parks studied. *PCA with SFFS* has the lowest NRMSE value, followed by the *Point with SFFS* method. On average, the difference between the *PCA with SFFS* and *Point with SFFS* is approximately 8%, with the highest values being shown in WPs 2 and 6. The *PCA without SFFS* method that uses only the most common meteorological variables and spatial information also enables a reduction in the error values. This result is consistent with those identified by other authors, e.g., Ref. [22]. However, this method without SFFS in WPs 2, 4 and 5 shows an improvement of less than 5% compared with the benchmark results. The results in Fig. 7 highlight that the identification of the most relevant meteorological parameters with the application of the SFFS

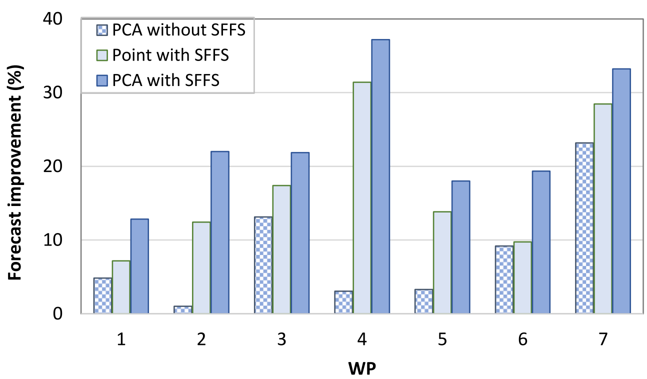

Fig. 7. Forecast improvements of RMSE over the benchmark by using different forecast methods: *Spatial – PCA*, *Point with SFFS* and *PCA with SFFS* (proposed in this work) for the different wind parks.

algorithm, even without PCA, enables to clearly improve the performance of the forecast and reduces the NRMSE always above 7% (results for WP1 using the *Point with SFFS* method).

4. Conclusions

New meteorological features to improve the accuracy of wind power forecasts were investigated in this work. A wrapper feature selection approach was adopted by using numerical weather prediction (NWP) data that is publicly available. The meteorological feature selection was executed using a common statistical algorithm in the wind energy sector, the artificial neural network (ANN). The proposed methodology was applied to seven wind parks located in Portugal. Those wind parks present distinctive climatic conditions, enabling to identify similarities in the NWP meteorological features that contribute to improve the forecast at the same time they enable to compare the performance of the forecast among the wind parks available.

It was shown that the use of key meteorological parameters such as wind gusts, wind power density, wind shear, and planetary boundary layer height can be used to improve a forecast of wind power. Using the methodology proposed, it was possible to observe an improvement in the forecast performance ranging between 13% (wind park 1) and 37% (wind park 4) in the case of the normalized root mean square error.

This paper also analyses the benefits of using spatial information from the NWP grid compared with the use of time series for a specific grid point. It was possible to verify that the impact of the feature selection algorithm is superior, when compared with the use of spatial information. However, combining the use of spatial information from the NWP with a rigorous choice of input meteorological data revealed to be the approach that produces the most accurate results. Future works should focus on understanding the impact of using meteorological features that explain a reduced variance as well as the use of other objective functions in the feature selection algorithm. (e.g., maximization of the remuneration in the wholesale electricity market).The results suggest that an appropriate selection of meteorological features can improve the characterization of wind power variability, reducing the uncertainties in a wind power forecast. Thus, this selection constitutes a crucial step towards improving the accuracy of wind energy forecasting systems, which will play a crucial role in supporting stable, robust, and near 100% renewable power systems.

Funding

This work has received funding from the EU Horizon 2020 research and innovation program under project TradeRES (grant agreement No 864276)

CRediT authorship contribution statement

António Couto: performed the literature, Writing – review & editing, collected, processed all the, Data curation, implemented all the necessary computational routines, Writing – original draft, the manuscript, All authors have read and agreed to the published version of the manuscript. **Ana Estanqueiro:** discussed the work to be developed, supported the identification of the most relevant results to be shown, also, Writing – review & editing, the manuscript, All authors have read and agreed to the published version of the manuscript.

Data availability

Part of data is publicly available as mentionded in the data availability section.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2022.11.022.

References

[1] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, C. Draxl, The state-of-the-art in short-term prediction of wind power A literature overview, in: Technical Report ANEMOS.Plus, 2011, p. 109.
[2] K.F. Forbes, E.M. Zampelli, Accuracy of wind energy forecasts in Great Britain and prospects for improvement, Util. Pol. 67 (Dec. 2020), 101111.
[3] H. Algarvio, F. Lopes, A. Couto, A. Estanqueiro, Participation of wind power producers in day-ahead and balancing markets: an overview and a simulation-based study, Wiley Interdiscip. Rev. Energy Environ. 8 (5) (2019).
[4] H. Liu, C. Chen, X. Lv, X. Wu, M. Liu, Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods, Energy Convers. Manag. 195 (May) (2019) 328–345.
[5] B. Yang, et al., State-of-the-art one-stop handbook on wind forecasting technologies: an overview of classifications, methodologies, and analysis, J. Clean. Prod. 283 (2021), 124628.
[6] M.S.H. Lipu, et al., Artificial intelligence based hybrid forecasting approaches for wind power generation: progress, challenges and prospects, IEEE Access 9 (2021) 102460–102489.
[7] E.N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20 (2) (1963) 130–141.
[8] C. Sweeney, R.J. Bessa, J. Browell, P. Pinson, The future of forecasting for renewable energy, WIREs Energy Environ 9 (2) (Mar. 2020).
[9] H. Wang, S. Han, Y. Liu, J. Yan, L. Li, Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system, Appl. Energy 237 (2019) 1–10.
[10] S. Hanifi, X. Liu, Z. Lin, and S. Lotfian, "A critical review of wind power forecasting methods—past, present and future," Energies, vol. 13, no. 15, p. 3764, Jul. 2020.
[11] S. Salcedo-Sanz, L. Cornejo-Bueno, L. Prieto, D. Paredes, R. García-Herrera, Feature selection in machine learning prediction systems for renewable energy applications, Renew. Sustain. Energy Rev. 90 (Jul. 2018) 728–741.
[12] P. Lu, L. Ye, Y. Zhao, B. Dai, M. Pei, Z. Li, Feature extraction of meteorological factors for wind power prediction based on variable weight combined method, Renew. Energy 179 (2021) 1925–1939.
[13] Q. Han, T. Wang, F. Chu, Nonparametric copula modeling of wind speed-wind shear for the assessment of height-dependent wind energy in China, Renew. Sustain. Energy Rev. 161 (Jun. 2022), 112319. February.
[14] H. Liu, C. Chen, Data processing strategies in wind energy forecasting models and applications: a comprehensive review, Appl. Energy 249 (May) (2019) 392–408.
[15] S. Jerez, et al., The impact of the North Atlantic oscillation on renewable energy resources in southwestern europe, J. Appl. Meteorol. Climatol. 52 (10) (Oct. 2013) 2204–2225.
[16] S. Bayo-Besteiro, M. García-Rodríguez, X. Labandeira, J.A. Añel, Seasonal and subseasonal wind power characterization and forecasting for the Iberian Peninsula and the Canary Islands: a systematic review, Int. J. Climatol. 42 (5) (2022) 2601–2613.
[17] A. Couto, P. Costa, L. Rodrigues, V.V. Lopes, A. Estanqueiro, Impact of weather regimes on the wind power ramp forecast in Portugal, IEEE Trans. Sustain. Energy 6 (3) (2015) 934–942.
[18] J.M. Correia, A. Bastos, M.C. Brito, R.M. Trigo, The influence of the main large-scale circulation patterns on wind power production in Portugal, Renew. Energy 102 (2017) 214–223.
[19] R.J. Davy, M.J. Woods, C.J. Russell, P.a. Coppin, Statistical downscaling of wind variability from meteorological fields, Boundary-Layer Meteorol. 135 (1) (Jan. 2010) 161–175.
[20] C. Gallego-Castillo, E. Garcia-Bustamante, A. Cuerva, J. Navarro, Identifying wind power ramp causes from multivariate datasets: a methodological proposal and its application to reanalysis data, IET Renew. Power Gener. 9 (8) (2015) 867–875.

[21] K. Bellinguer, V. Mahler, S. Camal, G. Kariniotakis, Probabilistic forecasting of regional wind power generation for the EEM20 competition: a physics-oriented machine learning approach, in: 17th International Conference on the European Energy Market (EEM), 2020, 2020, p. 6.

[22] J.R. Andrade, R.J. Bessa, Improving renewable energy forecasting with a grid of numerical weather predictions, IEEE Trans. Sustain. Energy 8 (4) (Oct. 2017) 1571–1580.

[23] T.T. Warner, Quality assurance in atmospheric modeling, Am. Meteorol. Soc. 92 (12) (2011) 1601–1610.

[24] H. Chen, Y. Birkelund, Q. Zhang, Data-augmented sequential deep learning for wind power forecasting, Energy Convers. Manag. 248 (2021), 114790.

[25] W.C. Skamarock, et al., A Description of the Advanced Research WRF Version 3," *NCAR Tech*, Note NCAR/TN-556+STR, 2019, p. 145.

[26] MeteoGalicia, "MeteoGalicia THREDDS [Online]. Available: https://www.meteogalicia.gal, 2021. (Accessed 2 July 2021). Accessed.

[27] A. Couto, A. Estanqueiro, Assessment of wind and solar PV local complementarity for the hybridization of the wind power plants installed in Portugal, J. Clean. Prod. 319 (August) (Oct. 2021), 128728.

[28] A. Couto, P. Costa, L. Rodrigues, V.V. Lopes, A. Estanqueiro, Impact of weather regimes on the wind power ramp forecast in Portugal, IEEE Trans. Sustain. Energy 6 (3) (2015) 934–942.

[29] F. Davò, S. Alessandrini, S. Sperati, L. Delle Monache, D. Airoldi, M.T. Vespucci, Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting, Sol. Energy 134 (2016) 327–338.

[30] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40 (1) (2014) 16–28.

[31] N. Huang, E. Xing, G. Cai, Z. Yu, B. Qi, L. Lin, Short-Term wind speed forecasting based on low redundancy feature selection, Energies 11 (7) (Jun. 2018) 1638.

[32] S. Solorio-Fernández, J.A. Carrasco-Ochoa, J.F. Martínez-Trinidad, A review of unsupervised feature selection methods, Artif. Intell. Rev. 53 (2) (Feb. 2020) 907–948.

[33] G.J. Székely, M.L. Rizzo, Brownian distance covariance, Ann. Appl. Stat. 3 (4) (2009) 1236–1265.

[34] G. Nazaré, R. Castro, L.R.A. Gabriel Filho, Wind power forecast using neural networks: tuning with optimization techniques and error analysis, Wind Energy 23 (3) (2020) 810–824.

[35] J.P.S. Catalão, H.M.I. Pousinho, V.M.F. Mendes, Short-term wind power forecasting in Portugal by neural networks and wavelet transform, Renew. Energy 36 (4) (2011) 1245–1251, Apr.

[36] B. Bochenek et al., "Day-ahead wind power forecasting in Poland based on numerical weather prediction," Energies, vol. 14, no. 8, p. 2164, Apr. 2021.

[37] M.T. Hagan, H.B. Demuth, M.H. Beale, O. De Jesús, Neural Network Design, second ed., Oklahoma State University, ebook), 2014.

[38] Matlab, "MATLAB and Neural Network Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States.".

[39] D. Carvalho, A. Rocha, M. Gómez-Gesteira, C. Silva Santos, Sensitivity of the WRF model wind simulation and wind energy production estimates to planetary boundary layer parameterizations for onshore and offshore areas in the Iberian Peninsula, Appl. Energy 135 (2014) 234–246.

[40] D. Siuta, G. West, R. Stull, WRF hub-height wind forecast sensitivity to PBL scheme, grid length, and initial condition choice in complex terrain, Weather Forecast. 32 (2) (2017) 493–509.

[41] C. Draxl, A.N. Hahmann, A. Peña, G. Giebel, Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes, Wind Energy 17 (1) (2014) 39–55.

[42] S.-Y. Hong, Y. Noh, J. Dudhia, A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev. 134 (9) (2006) 2318–2341.

List of acronyms

ANN: Artificial neural network
CAPE: Convective available potential energy
CHF: Cloud cover at high levels
CIN: Convective inhibition
$Forecast_{Benchmark}$: Forecast for benchmark methodology
$Forecast_{Methodology}$: Forecast methodology
Gust: Surface wind speed gust
HGT: Geopotential height
HR: Humidity relative
LM: Levenberg-Marquardt
LWF: Surface downwelling longwave flux
MSLP: Mean sea level pressure
$MSLP_{Grad}$: Mean sea level pressure gradient
NominalPower: Wind park nominal power
NRMSE: Normalized root mean square error
NWP: Numerical Weather Prediction
P_{For}: Wind power forecast
P_{Obs}: Wind power observed
PBL: Planetary boundary layer
PC: Principal Component
PCA: Principal component analysis
Rainfall: Total accumulated convective rainfall
SFFS: Sequential forward feature selection
SWF: Surface downwelling shortwave flux
T: Air temperature
T_{var}: Vertical temperature variation
U: U-wind component
UTC: Universal time coordinated
V: V-wind component
Vis: Atmospheric visibility
WindShear: Wind shear
WP: Wind park
WPD: Wind power density
WRF: Weather Research and Forecasting
WS: Wind speed
YSU: Yonsei State University