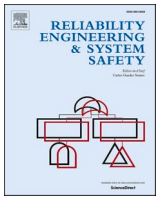
Reliability Engineering and System Safety 266 (2026) 111716

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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



Probabilistic evaluation for early wind turbine yaw misalignment detection

M.A. García-Vaca [a,*], J.E. Sierra-García [b], Matilde Santos [c]

[a] Computer Science Faculty, Complutense University of Madrid, Madrid, 28040, Spain
[b] Department of Digitalization, University of Burgos, Burgos, 09006, Spain
[c] Institute of Knowledge Technology, Complutense University of Madrid, Madrid, 28040, Spain

ARTICLE INFO

Keywords:
Reliability
Predictive maintenance
Anomaly detection
Probabilistic models
Power curve
Wind turbine
Yaw misalignment

ABSTRACT

Nowadays, one of the biggest challenges for wind turbines is to reduce operation and maintenance costs. Therefore, it is essential to develop predictive maintenance, anticipating failures early and thus avoiding unnecessary actions on the wind turbine. In this way, the uptime and performance of the turbine are maximized, and its useful life is extended. This work describes a general methodology for fault detection based on probabilistic models and its evaluation. This methodology combines a fault detection method based on the Fisher Test and the development of probabilistic models of wind turbine power curves. Several probabilistic models of power curves have been evaluated: Gaussian mixture model (GMM), Frank copula model, Gaussian mixture copula model (GMCM), Gaussian process regression (GPR) and epsilon-insensitive loss function support vector regression (ε-SVR). The results indicate that the Gaussian mixture copula model is the most efficient in terms of accuracy and computational cost. The detection of a wind turbine orientation misalignment error has been tested as a use case. It is shown how with this probabilistic approach it is possible to detect the fault in a short period of time from its appearance, 10–30 times faster than other techniques found in the literature with which it has been compared.

1. Introduction

In recent years, attention and interest in climate change-related issues have increased steadily. Society's growing energy demand together with greater environmental awareness have driven the transition and use of clean and sustainable energy sources [1,2]. A clear example of renewable alternatives is wind energy, capable of producing electricity in both, onshore and offshore wind farms, which has established itself as an energy option of great importance for the future, as demonstrated by its rapid growth in recent years [3].

To maximize the use and profitability of wind turbines, it is necessary to ensure their efficiency and reduce operation and maintenance (O&M) costs [4]. To date, the most common types of maintenance used in the industry are corrective and preventive. The disadvantages of corrective maintenance are clear: the action is carried out when the turbine has already stopped, which increases losses due to the inoperability of the turbines along with the costs of materials and personnel necessary to correct the unexpected failure. Preventive maintenance alleviates this disadvantage by performing interventions at fixed intervals to prevent failures before they occur. However, these interventions often involve performing unnecessary actions on components that would not need them, thus increasing the frequency of downtime and maintenance costs. In contrast, predictive maintenance overcomes the aforementioned disadvantages by anticipating failures prematurely and thus, avoiding unnecessary actions on the wind turbine, reducing O&M costs and maximizing uptime and performance [5,6].

To carry out this predictive maintenance, it is necessary to identify in advance when the turbine is operating in suboptimal conditions or shows some kind of anomaly. To evaluate these deviations from the ideal or expected behaviour, it is necessary to have an accurate model of the wind turbine power curve, since this curve is the best indicator of the performance of the wind device [7]. Modelling the shape and values of this power curve, which relates the electrical power generated as a function of the wind at that instant, is a non-trivial challenge. This is due, among other reasons, to the non-linear and complex nature of the system variables and the uncertainty associated with the measurements and values. Indeed, wind turbine's actual power curve is influenced by various factors (e.g., air density, wake effects) and that there is significant variability even between wind turbines of the same type and brand within the same wind farm.

* Corresponding author.
E-mail address: magvaca@ucm.es (M.A. García-Vaca).

https://doi.org/10.1016/j.ress.2025.111716
Received 23 December 2024; Received in revised form 14 July 2025; Accepted 11 September 2025
Available online 12 September 2025

There are various approaches to obtain the model of the operation of the wind turbine. One of the first options to consider is to obtain an analytical model of the system, based on the physical laws that govern its operation [8]. However, as will be discussed in the following sections, this option is not feasible in many real physical systems. Another option would be to use the power curve supplied by the manufacturer, which characterizes each particular turbine. But it is not a satisfactory solution either, among other things because the technical characteristics of the power converters in many cases differ from reality.

A possible alternative that may be more accurate and realistic in reflecting the behaviour of the wind turbine is the use of data-driven algorithms [9–14]. These methods, which rely on machine learning and AI techniques, have the ability to adapt to the particular conditions of the case under study and provide accurate and reliable results. In addition, some of these techniques provide confidence intervals that take into account the variability of the data, which is very beneficial as it can help estimate the probability of an imminent failure. This is why to deal with this inherent variability in the power curve we use data-driven algorithms rather than a theoretical or original equipment manufacturer (OEM) power curve.

Thus, in this work, we propose a novel approach for obtaining and analysing data-driven based power curve models that allow inferring the uncertainty associated with the fault prediction. That is, a probabilistic evaluation methodology is developed for fault detection. To do so, we chose five probabilistic models: Gaussian mixture model (GMM), Frank copula model, Gaussian mixture copula model (GMCM), Gaussian process regression (GPR), and support vector regression with epsilon-insensitive loss function (ε-SVR) for demonstrating and testing our fault detection proposal. These approaches not only model the behaviour of the power curve but, crucially, also consider the associated uncertainty [15]. While other, perhaps more modern models [16] may excel in prediction accuracy, they often lack the explicit uncertainty quantification that is fundamental to our probabilistic evaluation method. Furthermore, the selected probabilistic models offer a significant advantage in terms of computational cost compared to some more modern and complex alternatives. These different approximations to the power curve of a real turbine are evaluated and compared with each other to determine which curve model best fits the data provided by the wind device.

The second objective of this work is to develop a probabilistic assessment method that allows to statistically infer the early presence of faults, based on the detection of deviations from the probabilistic models obtained from the power curves of wind turbines. Among the possible faults that can occur in a turbine; in order to validate the effectiveness of the proposed method, the detection of a yaw misalignment is evaluated, since even small deviations in yaw angle can generate significant power losses. This probabilistic assessment method is used to identify which curve model (GMM, Frank, GMCM, GPR, ε-SVR) can best anticipate faults in a real turbine.

Therefore, to address the issues related to the application of theoretical power curve models, we propose the following contributions.

- Real-world SCADA data: We use historical SCADA data from the specific wind turbine being monitored. This data is filtered to select operating periods considered "healthy" or optimal. Consequently, the resulting probabilistic power curve model does not represent a theoretical ideal but rather captures the reference-specific normal behavior of that particular wind turbine, taking into account its operational and site-specific characteristics.
- Probabilistic modeling with uncertainty estimation for fault detection: assuming that power generation is not a deterministic function of wind speed, in our paper we evaluate various probabilistic models (GMM, Frank Copula, GPR, e-SVR, and GMCM). These models do not predict a single power value, but rather a probability distribution of power for a given wind speed. This allows for the quantification of natural uncertainty and variability in turbine operation.
- Anomaly detection as deviation from normal or expected values: An incipient fault, such as yaw misalignment, is detected as a statistically significant deviation from this learned normal behavior. That is, an alert is triggered when the turbine consistently operates outside the confidence limits for a given period of time.
- The methodology is validated for wind turbine yaw misalignment detection based on Fisher's Test, as a use case, allowing failure prediction.

The structure of the paper is as follows: after this introductory section, Section 2 reviews the actual state-of art and related works, Section 3 details the proposed methodology for data-driven model identification. Section 4 explains the use case of the wind turbine power curve. Section 5 shows and discusses the results obtained, comparing them with previous research results. Finally, Section 6 summarizes the conclusions drawn from the research and possible directions for future research.

2. Related works

There are different approaches for wind turbine power curve model. The paper by [17] reviews data-driven models, model-based models, and hybrid models, highlighting the combinations of machine learning methods and statistical approaches in data-driven models. Another review is presented in [18] where, once the vital importance of the power curve as condition indicator of the wind turbine has been stated, the paper presents a wide range of power curve-based applications, including anomaly and fault detection, data preprocessing and correction schemes. A more recent review can be found in [19], focused on research on the predictive and prescriptive maintenance of wind turbines based on the implementation of data-oriented models with the use of artificial intelligence tools. Also in 2024, [20] provide a framework for systematically understanding data-driven power curve models by utilizing explainable artificial intelligence techniques. They obtain more informed models that are applied in the context of wind turbine performance monitoring.

Most of the works address predictive maintenance based on the power curve. To do so, it is important to model and identify the specific relation of wind speed and power output of each turbine. For instance, [21] uses three performance curves (power, pitch angle, and rotor speed) from SCADA data to accurately describe normal wind turbine behaviour for performance monitoring and identification of anomalous signals. It introduces an unsupervised SVM-KNN model as an optimal outlier detection approach. The paper [22] presents a comprehensive approach to improving the accuracy of wind turbine power curve modeling. It combines outlier elimination using a KNN-estimated wind power curve and novel quantile regression models (e.g., decision tree quantile regression). In [23] a data-driven deep learning method is proposed to model wind turbine power curves. A novel multivariable power curve prediction modelling is proposed in [24]. It integrates stochastic gradient boosting regression tree with grey wolf optimization, along with data preprocessing and feature selection methods. In [25] compares nine machine learning methods for wind turbine power forecasting, identifying GNN-based models as outperforming others in power forecasting accuracy, while XGBoost provides optimal results concerning computational resources, supporting data-driven predictive maintenance strategies. Interestingly, the paper [26] makes a comparison of the performance of various metaheuristic optimization-based parametric methods in wind turbine power curve modeling.

Unlike conventional deterministic models, probabilistic models of the power curve allow for dealing with the uncertainly attained to the wind turbine performance. Focusing on the use of probabilistic models of the power curve of a wind turbine for anomaly or fault detection and predictive maintenance we found some remarkable works. In [27] proposes a novel probabilistic power curve model, expressed by the beta distribution, which is capable of estimating the uncertainty of power

output at a given wind speed. The research [28] contributes a new wind turbine performance degradation monitoring scheme based on the pairwise comparison of probabilistic power curves of different turbines within a wind farm. The recent paper [29], presents a specific kind of Generalized Linear Model (GLM) called beta regression that effectively handles the unique statistical features of wind power while retaining a manageable level of complexity. The paper [30] investigates the use of physically meaningful probabilistic power curve models derived from bounded Gaussian processes with a Beta likelihood for wind turbine structural health monitoring. In [12], for an onshore turbine, a Gaussian process regression is proposed to estimate the power curve and variables such as pitch angle and rotor speed as a function of wind. Also [31], contributes to wind turbine health monitoring by employing Gaussian Process Regression (GPR) to evaluate loads and predict turbine performance metrics, utilizing a hybrid database of simulation and SCADA data for robust assessment.

Finally, some papers like the one by [32], deals with both, wind uncertainty and wind turbine components condition to assess wind power system reliability.

3. Methodology for wind turbine yaw misalignment failure detection based on probabilistic models

The ultimate goal of this work is to generate alarms in the event of possible failures in the operation of wind turbines, as far in advance as possible. To this end, a methodology is proposed in which real data of the wind device are pre-processed, probabilistic models of the power curve are generated, and deviations from the model are detected, as shown in Fig. 1.

The methodology is divided into two stages: the probabilistic modelling of the actual power curve of the turbine, which is executed offline, and the probabilistic evaluation of possible deviations from it, which is executed online.

Probabilistic modelling. It includes:

1. *Preprocessing*. In this phase, the effect of the air density on wind speed is corrected, and anomalous data are removed from the data set (Section 3.1).
2. *Power curve model*: Using the power and wind data, the actual power curve of the wind turbine is identified according to the computational models considered in Section 3.2.

Probabilistic assessment, which consists of:

1. *Calculation of the deviation of the measured power from the estimated power*: For each time value of the case study, and using the probability distribution of the identified model, the deviation of the power generated at the current instant by the turbine from the predicted power is quantified. The specific details of this general procedure are presented in Section 3.3.
2. *Deviation monitoring*: To improve the efficiency of anomaly detection, the deviation calculated for a time instant is combined with the deviations calculated at previous times. In this way, the variation in the deviation is monitored to reduce false positives. The specific details are explained in Section 3.3.
3. *Fault alarm*: If the combined probability is below the significance level, an anomaly is considered to exist in the wind turbine, which triggers a fault alarm.
4. *Value update*: The probability value calculated in step 5 is stored as the current value, and steps 3 to 5 are repeated for each new time point.

3.1. Preprocessing

Power generation with turbines is cubically related to wind speed, as seen in (1). Therefore, even a small variation in wind speed can have a significant impact on the power generated. It is critical to have accurate wind speed measurements that do not depend on momentary atmospheric conditions.

$$P = 0,5\rho A C_p(\lambda,\beta)v^3, \tag{1}$$

where ρ is the air density (kg/m^3), A is the area swept by the blades (m^2), C_p is the power coefficient -which in turn depends on the tip speed ratio λ and pitch blade β (degrees), and v is the wind speed at the hub (m/s).

That is why first of all, and following the recommendations of the International Commission [33], a correction due to variations in air density is applied (2) [34]. This first preprocessing step is carried out at every instant of time applying the following expression,

$$v_c = v_m\left(\frac{\rho}{1.225}\right)^{\frac{1}{3}}, \tag{2}$$

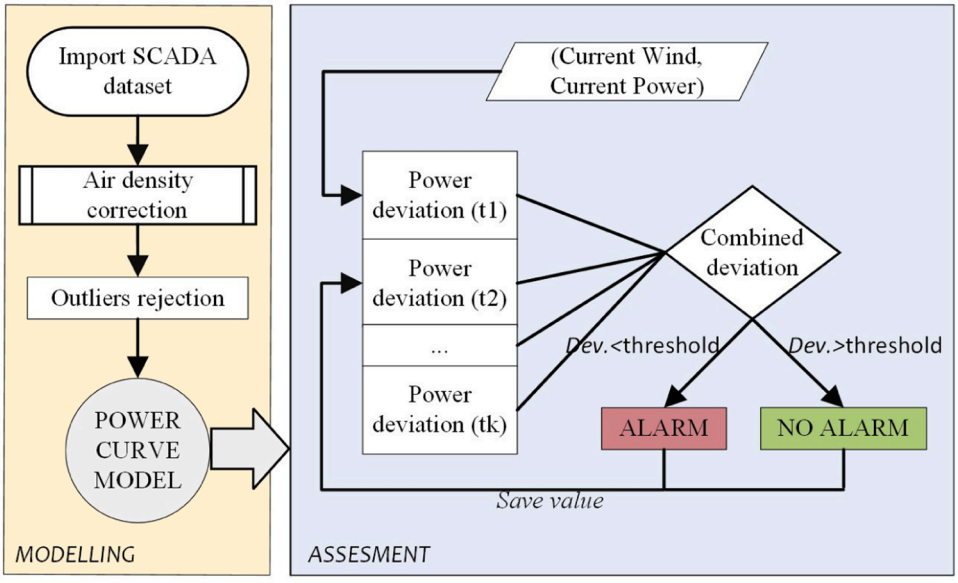


Fig. 1. Methodology for probabilistic failure detection, where t1 denotes the current instant, and tk denotes the previous k-th instant.

where v_c is the corrected air velocity (m/s), v_m is the measured air velocity (m/s), and the air density can be estimated by (3)

$$\rho = 1.225\left(\frac{288.15}{T}\right)\left(\frac{B}{1013.5}\right), \tag{3}$$

where T is the ambient temperature (Kelvin) and B is the atmospheric pressure (mbar)

Besides, typically, data collected by sensors in SCADA systems have outliers, so another preprocessing phase is often required to reject the outliers before using them to create a model. These outliers can be caused by a variety of reasons, but they do not indicate a fault in the turbine subsystems. For more details on outliers and their causes in wind turbines, a comprehensive and detailed study can be found in [22]. The most notable cases are summarized below.

- Turbine operational transitions: Since data is captured every 10 min, the turbine may have changed its operating state (on-off) during that time interval. This results in points with unreliable values. These outliers are often isolated and do not show a clear pattern (regions marked B in Fig. 2).
- Errors or noise in sensors or in the SCADA communication system (regions marked B in Fig. 2).
- Data loss: Occasionally, data may be lost due to failures in data acquisition or transmission.
- Negative powers below the start-up wind speed: When the wind speed is insufficient for the turbine to generate power, negative powers may sometimes be recorded. These outliers are located in region marked A in Fig. 2.

The processing applied to eliminate outliers is divided into two parts. First, data with 0 or negative generated power are eliminated, as well as those with indeterminate or missing values (Fig. 2, region A). In a second phase, the remaining outliers are eliminated (Fig. 2, region B). To eliminate these outliers, following the specifications of the international standard IEC 61,400–12–1:2017 [38], the power curve is divided into 0.5 m/s intervals. At each interval, the mean and standard deviation σ are calculated. A point that deviates more than ±2.5σ in its corresponding interval is considered an anomaly.

These are the two methods that are commonly applied for wind turbine data analysis and are widely recognized in the related literature [35–37].

3.2. Probabilistic power curve modelling

In order to develop a probabilistic fault assessment method using the power curve, it is necessary to obtain a reliable and accurate model of the wind turbine power curve. Among the different models that can be considered from data-driven algorithms, we will focus on those that not only provide a prediction of the generated power but also of the uncertainty associated with this prediction.

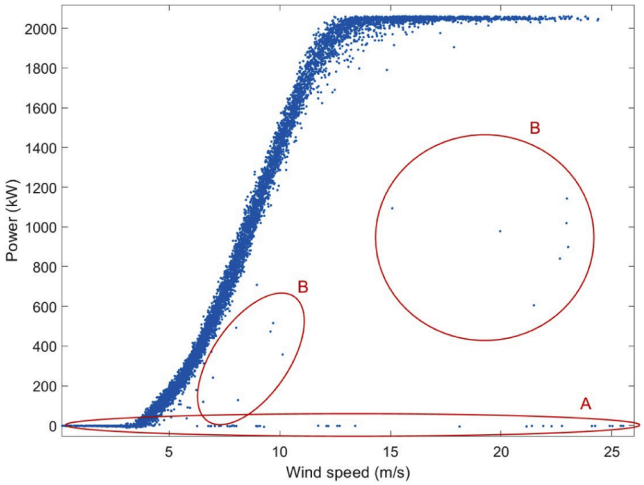


Fig. 2. Outliers in a 2050 KW Senvion MM82 wind turbine power curve.

The models used in this work can be divided into two categories:

- *Probabilistic models*, where we find the Gaussian mixture model (GMM), the Frank copula model (FC) and the Gaussian mixture copula model (GMCM). The models in this category provide the joint distribution density function of the power curve [29].
- *Regression models with uncertainty estimation*. These include Gaussian process regression (GPR) and epsilon-insensitive loss function support vector regression (e-SVR). These types of models give as output the power prediction and its confidence interval together with the distribution of this uncertainty [39].

3.2.1. Gaussian mixture model (GMM)

Finite mixture methods are used to describe the probability distribution of a population consisting of several groups or clusters with different underlying distributions. As the name implies, the total population is assumed to be a finite mixture of several independent components or modes. In the case of continuous variables, we can assume that all clusters are modeled as Gaussian distributions, each with different parameters (mean and covariance). This technique is known as Gaussian Mixture Model (GMM) [40].

For the bivariate case, let $\psi(x_1, x_2; \Theta)$ be the probability density function sought and Θ be the set of parameters of these Gaussian distributions; in this case Eq. (4) is satisfied:

$$\psi(x_1, x_2; \Theta) = \sum_{k=1}^{M} \alpha^{(k)} \phi\left(x_1, x_2; \theta^{(k)}\right), \tag{4}$$

where $\alpha^{(k)}$ is the mixing ratio of the M clusters (the sum of all $\alpha^{(k)}$ will be equal to 1) and $\theta^{(k)}$ the parameters (mean and covariance) of each cluster, with $\phi\left(x_1, x_2; \theta^{(k)}\right)$ the probability density of each mode.

3.2.2. Model based on Frank's copula

Statistical copulas are families of mathematical functions capable of relating dependent variables with a complex correlation between them. The utility they provide is that, on the one hand, they differentiate the marginal distributions and, on the other hand, how they relate to each other. According to Sklar's theorem [41], which establishes the pillars of copula theory, the probability density function sought for the bivariate case can be decomposed according to the expression (5):

$$f(x_1, x_2) = c(u_1, u_2) f_1(x_1) f_2(x_2), \tag{5}$$

Where f_i denotes the marginal distribution function of the i th variable (x_i), u_i denotes the cumulative distribution function of these variables, and $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ is the copula function that relates these cumulative density functions.

In mathematical theory there are many families of parametric copulas such as Clayton, Frank, or Gumbel, among others [42]. The choice of one or another will depend on the data and the correlation at its extreme values; in other words, on the tail dependence they present. In the case of the power curve of a wind turbine, it shows a strong correlation in both extremes of the distribution (both, for high and low values). Therefore, a good choice could be the Frank family of copulas. These Frank copula functions are represented in Eq. (6):

$$c_{Frank}(u_1, u_2, \delta) = \frac{\delta\eta e^{-\delta(u_1+u_2)}}{\left[\eta - (1 - e^{-\delta u_1})(1 - e^{-\delta u_2})\right]^2}, \tag{6}$$

where $\eta = 1 - e^{\delta}$, and δ is the copula parameter that best fits the data.

The higher this value is, the stronger the dependence between the two variables is.

3.2.3. Copula model of Gaussian mixtures (GMCM)

In the case of the Gaussian Mixture Model (GMM), it is assumed that the distributions of each of the modes that make up the power curve are normally distributed and, therefore, have an elliptical shape. However, this assumption is not necessarily always met. In this case, when this condition is not met, the Gaussian Mixture Copula Model (GMCM) is used, which applies the copula theory to the GMM technique. Its main advantage lies in its ability to characterize non-elliptical (non-Gaussian) modes.

For the present case of a bivariate distribution, the copula function to be sought is expressed by Eq. (7) [43]:

$$c_{GMCM}(\mathrm{u}_1,\mathrm{u}_2;\theta)=\frac{\varphi\left(\phi_1^{-1}(u_1),\phi_2^{-1}(u_2);\theta\right)}{\varphi_1\left(\phi_1^{-1}(u_1)\right)\varphi_2\left(\phi_2^{-1}(u_2)\right)}, \tag{7}$$

where φ_i is the marginal density function, ϕ_i^{-1} is the inverse marginal distribution, and θ represents the parameters (mean and covariance) of all modes along with the mixing ratios of each mode.

3.2.4. Support vector regression with epsilon insensitive loss function (ε-SVR)

Among machine learning techniques, one of the most popular and widely used is Support Vector Machines (SVM), whose theoretical foundations were established in the 1990s. SVMs are based on the use of different types of kernels, selecting one or the other according to the characteristics of the problem under study. Its strength lies in producing accurate predictions without incurring significant computational costs. Although SVMs are widely used for classification tasks, we will focus on their regression version using an epsilon error-insensitive symmetric loss function based on [44].

Consider a training data set given by $\{(x_1,y_1),\ldots,(x_n,y_n)\}\subset X\times\mathbb{R}$, where X represents the input space (in this case, $X=\mathbb{R}$), and y_n represents the output data. The goal is to find a hyperplane that satisfies (8):

$$y=af(x)+b, \tag{8}$$

To find the coefficients a and b that define the line of separation in the hyperplane, Eq. (9) must be minimized:

$$\frac{1}{2}||w||^2+C\sum_{i=1}^{n}\left(\xi_i-\xi_i^*\right), \tag{9}$$

where $||w||$ represents the norm, C is a penalty constant, and ξ_i, ξ_i^* are "slack variables". This corresponds to dealing with a loss function called error-insensitive function, $||\xi_\epsilon||$, described by (10):

$$|\xi_\epsilon|:=\begin{cases}0 & \text{if } |\xi|\le\varepsilon\\ |\xi|-\varepsilon & \textit{otherwise}\end{cases}, \tag{10}$$

Specifically, the goal is to find the function that has a maximum deviation ε from the training data y_i, being as linear as possible. In other words, the idea is to use as a loss function an 'insensitive zone' with a width of ε, where errors are ignored as long as they are smaller than this parameter ε.

Several kernel functions can be used as a parameter to train the model. In this case we will use the radial basis function kernel (Gaussian), defined by Eq. (11):

$$K(x_i,x_j)=e^{-\gamma||x_i-x_j||^2}, \tag{11}$$

Where γ is the scale parameter.

Once the ε-SVR model is trained and the power curve function, $\bar{f}(x)$, has been obtained, a confidence interval, I, is determined around this function within which we can infer the probability that the model output belongs to I. That is, given a specific point x and the model (in our case, ε-SVR), we want to determine the probability $P(y|x.\bar{f}(x))$. To accomplish this, we will use a simple approximation that allows us to provide an estimate of these confidence intervals [45].

Let $\zeta_i=y_i-\bar{f}_j(x_i)$ be the representation of the residuals when predicting the point (x_i,y_i) for each k-fold cross-validation. It is verified that these residuals, ζ_i, fit a Laplacian distribution with mean 0. This distribution, which coincides with the desired probability $P(y|x,\bar{f}(x))$, follows (12):

$$P(\mathrm{z})=\frac{1}{2\sigma}e^{-\frac{|z|}{\sigma}}, \tag{12}$$

where σ is the scale parameter. Assuming that ζ_i are independent of each other, the scale parameter, σ, is calculated as follows (13):

$$\sigma^2=\frac{\sum_{i=1}^{n}|\zeta_i|}{n}, \tag{13}$$

3.2.5. Gaussian process regression model (GPR)

Gaussian Process Regression (GPR) models are nonparametric probabilistic models based on kernels [46]. A Gaussian Process (GP) is a collection of random variables such that any finite combination of them also follows a Gaussian distribution. In other words, if $f(x)$ is a GP, then if we take n observations of the process $x_1,\ x_2,\ldots x_n$, the joint distribution of the random variables $f(x_1),\ f(x_2),\ldots f(x_n)$ will be Gaussian.

A GP is defined by its mean function $m(x)$ and its covariance function $k(x,x')$. Therefore, if $f(x)$ is a GP, then $E(f(x))=m(x)$ and $Cov[f(x),\ f(x')]=k(x,x')$.

Based on the GPs, the output of the GPR models can be expressed by (14):

$$y(x)=h(x)^T\beta+f(x), \tag{14}$$

Where $f(x)$ is a GP with zero mean and covariance function $k(x,x')$, $h(x)$ is a set of basis functions that transform the input space, and β is a vector of coefficients. As a probabilistic model, the model output y_i at time i can be modeled by (15):

$$P(y_i|f(x_i),x_i)\sim N\left(y_i|h(x_i)^T\beta+f(x_i),\ \sigma^2\right), \tag{15}$$

Different basis functions and kernel types can be used. It is the responsibility of the model designer to select the kernel that best fits the process to be modeled. In our case, linear basis functions have been used, together with the squared exponential kernel given by (16):

$$k(x,x')=\sigma^2\exp\left(-\frac{(x-x')^2}{2l^2}\right), \tag{16}$$

Where l is the characteristic length scale and σ is the signal standard deviation.

3.3. Probabilistic failure detection

The different probabilistic techniques presented in the previous section are going to be applied to model the power curve of the turbine when it is operating optimally and without faults. Based on these models, a probabilistic method has been developed to evaluate in real time whether the turbine presents any anomaly or fault that causes a loss of performance.

For these, the deviation of the measured power with respect to the power predicted by the corresponding model is going to be quantified for each value of the time series, considering the uncertainty associated with the model. That is, it is desired to determine, in a probabilistic manner, whether the power deviation detected is statistically significant, and therefore the mismatch with respect to the probability distribution of the model can be attributed to the presence of an anomaly.

The null hypothesis is defined as the statement that any observed

deviation is due solely to chance or turbine failure. The p-value is defined as the probability that the observed power is less than that predicted power under the null hypothesis.

To calculate the p-value, probabilistic models directly provide the joint distribution function. In the case of regression models with uncertainty estimation, a power prediction curve is obtained along with its confidence interval and the corresponding probability distribution. In particular, the Gaussian regression (GPR) model generates a Gaussian distribution, while the ε-SVR model produces a Laplacian distribution.

The p-value corresponds to the probability that the wind turbine output power, P_{out}, given the wind speed v_{ctest}, is less than the measured power, P_{test}, what can be formalized by (17):

$$p_{value}(P_{test},\ v_{ctest}) = Prob(P_{out} \langle\ P_{test} | v_{ctest}), \tag{17}$$

This p-value can be obtained by the joint bivariate distribution function of the power curve obtained from the model, $Z(P_{out}, v_c)$ (18).

$$p_{value} = \frac{\int_0^{P_{test}} Z(P_{out}, v_{ctest}) dP_{out}}{\int_0^{\infty} Z(P_{out}, v_{ctest}) dP_{out}}, \tag{18}$$

As an example, Fig. 3 shows the joint bivariate distribution function of the power curve obtained with the GPR model for a wind speed of 4 m/s, where a power of 63 kW is obtained (black line). The p-value is represented in grey, which corresponds to the area under the curve between the observed power and the power equal to 0.

Considering a single p-value is not sufficient to establish strong evidence, without risk of having false positives or false negatives. However, if we calculate p-values for k previous consecutive instants and combine these p-values in a joint inference test, a more robust detection is obtained.

There are several ways to combine these probability values. In this work, the Fisher's combined probability test is proposed [47]. This test is based on the principle that several results that, individually, do not show statistical significance, can generate such significance when considered together. In other words, even if an observed power value of the wind turbine does not deviate significantly from the expected value, when combined with the values from the immediately preceding time instants, it is possible to identify an anomaly. The main advantage of the Fisher test, and the reason why it is used in this context, is its validity for small statistical samples.

The Fisher's test defines the T statistic calculated by (19)

$$T = -2\sum_{i=1}^{k} ln(P_{valuei}) \sim \mathscr{X}^2(2k), \tag{19}$$

This T-statistics approximates a chi-square distribution with 2k degrees of freedom [48], where k is the number of consecutive p-values considered, and P_{valuei} is the p-value of time i calculated according to (18).

The first part of the test is the calculation of the T-statistic value with (19). T is used as input of the cumulative chi-square distribution. Subsequently the inverse probability (1- χ^2 in Fig. 4) is compared with the significance level α.

Then, in order to trigger the alarm if an anomaly is detected, a threshold is set. This threshold is equivalent to the significance level α used in the context of hypothesis testing. In our case, an α value or threshold of 0.05 was selected for two reasons. First, because this value is commonly adopted in most research fields and is often considered a standard for statistical significance. Second, because it also represents a good balance, minimizing false positives without excessively increasing false negatives.

For triggering alarms, the calculated p-value is compared with the threshold to determine whether an alarm is statistically significant. If the p-value is less than α (i.e., inverse probability below the dotted line in Fig. 4), it can be inferred that it is a statistically significant result, which would indicate a failure in the turbine: the alarm is considered true; otherwise, it is attributed to chance.

The probabilistic failure detection process here described can be formalized by Algorithm 1.

As for the number of time values considered, k, if we select a high value, the detection becomes more robust, minimizing the presence of false positives due to noisy signals. However, a high value of k also increases the detection time window. It is necessary to find a balance between a shorter detection time (with a low k) and the reduction of false positives (with a high k). In our case, a value of $k = 2$, corresponding to 20 min, has proven to be a satisfactory trade-off.

This value of k is also motivated for other reasons. Our goal is not to detect instantaneous operational deviations in the yaw angle, which are a normal part of operation due to the yaw system's slow response to sudden wind changes. The typical adjustment frequency of the yaw control system is on the order of a few minutes [49,50], and a response speed ranging from 0.5°/s to 2.0°/s [51]. The method here proposed focuses on detecting true systematic yaw misalignment: a persistent failure (causing continuous and sustained production loss).

That is why to isolate this systematic failure from noisy normal operation, our methodology employs two key filters:

- We deliberately use SCADA data averaged over 10-minute intervals. This average acts as a low-pass filter, smoothing out rapid fluctuations and momentary deviations that do not indicate an underlying failure.
- Our detection algorithm incorporates a time window parameter (in our case, $k = 2$). A failure alert is only triggered if the turbine

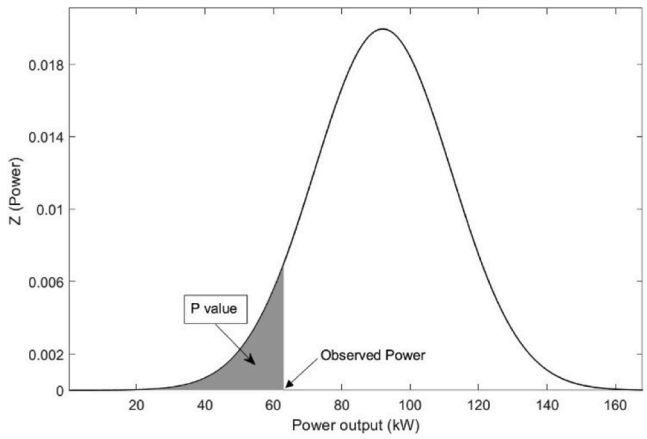


Fig. 3. Example of p-value obtained from the power curve probabilistic model.

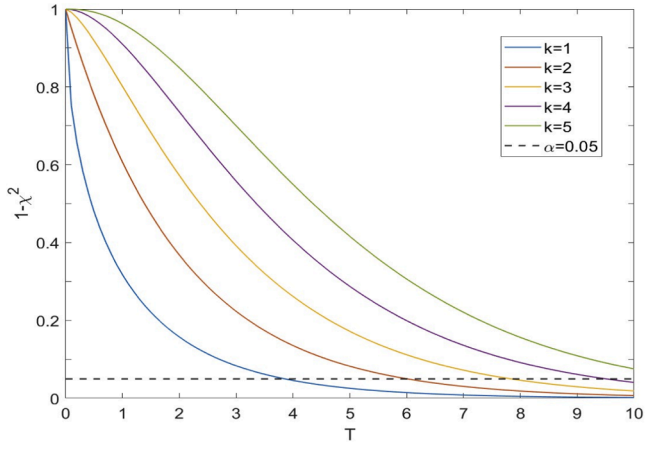


Fig. 4. Example of use of T-statistics with a chi-square distribution with 2k degrees of freedom, and significance level α=0.05.

Algorithm 1
Probabilistic failure detection.

$k \leftarrow 2,\ \alpha \leftarrow 0.05$ //Configuration parameters
For $i = 1$ to $k - 1$ **do**
FIFO.push(0)
End for
While {true } {
$[P_{test}, v_{c\,test}] \leftarrow getData()$
$p_{value} \leftarrow \frac{\int_0^{P_{test}} Z(P_{out}, v_{c\,test}) dP_{out}}{\int_0^{\infty} Z(P_{out}, v_{c\,test}) dP_{out}}$
FIFO.push(p_{value})
$T \leftarrow 0$,
For p **in** FIFO
$T \leftarrow T - 2\ln(p)$
End for
if $1 - f_{\chi^2}(T) < \alpha$
Activate_Alarm()
Else
Reset_Alarm()
End if
FIFO.pop
FIFO.push($1 - f_{\chi^2}(T)$)
} **endWhile**

operates below its expected performance consecutively for k 10-minute intervals.

These two key factors create a minimum detection window of 20 min. As mentioned above, this window is significantly longer than the typical time for yaw control system actions. Therefore, this methodological design ensures that only persistent deviations from efficient turbine behavior, which are indicative of a systematic failure, are triggered.

4. Use case: Yaw misalignment in a wind turbine

To validate the proposed methodology, the detection of yaw error in a wind turbine has been chosen as a use case. For optimum power output, a wind turbine's blades should remain perpendicular to the wind as often as possible. Every turbine is outfitted with technology to measure the wind direction, usually a sonic anemometer which measures wind direction and speed and adjusts accordingly. Any measured difference between the wind direction and the nacelle position of the turbine is known as the yaw error. That is, yaw misalignment means that the wind turbine is not fully facing the wind [52]. Analysis and experimentation have shown that the loss of power typically varies with the square of the cosine of the yaw misalignment. As such, it is essential to monitor yaw misalignment continuously and to prevent it. The objective is to identify genuine systematic failures (resulting in ongoing and persistent output loss) rather than instantaneous deviations in the yaw angle.

To obtain the models with the techniques presented, it is necessary to have a sufficient data set, representative of all sections of the power curve, which reflects its normal behavior. On the other hand, to test the methodology presented, it is necessary to have data that reflects some type of error in order to detect it.

To the best of our knowledge, there is no public dataset with these desired characteristics. Therefore, we have created a new dataset by merging two existing open datasets. The dataset of the onshore wind farm located in the city of Penmanshiel, United Kingdom [53] has been used, which contains sufficient data to obtain the power curve model of a turbine. And the dataset included in [12], which contains a real case of yaw misalignment error, which has also been used to validate the failure prediction methodology.

To estimate the power curve model, first open access data from an onshore wind farm located in the town of Penmanshiel, United Kingdom have been used [53]. Of all the turbines in this park, turbine WT01 has been selected for the study. This choice is motivated because it is the most isolated turbine with respect to the others in the farm, in order to try to reduce the turbulence effects generated by the wakes of adjacent turbines [54]. Fig. 5 shows the location of the wind farm with its latitude and longitude, geographical north and distances between wind turbines.

The WT01 turbine is a Senvion MM82 pitch regulated wind turbine; its technical specifications are presented in Table 1.

The WT01 turbine is equipped with an integrated SCADA system that records data for more than 100 different variables, such as output power, wind speed and direction, temperature, nacelle position, rotor speed, blade angle, …, among others. These variables are sampled every 10 min, and statistical values of mean, maximum value, minimum value and standard deviation are collected for each variable.

Only some of the available variables are used in this work, specifically the following:

- Time stamp (Year-Month-Day HH:mm:ss).
- Wind speed (m/s): Average wind speed every 10 min, measured by an anemometer at the height of the nacelle, corrected for air density.
- Power generated (kW): Average electrical power generated by the turbine every 10-minutes.

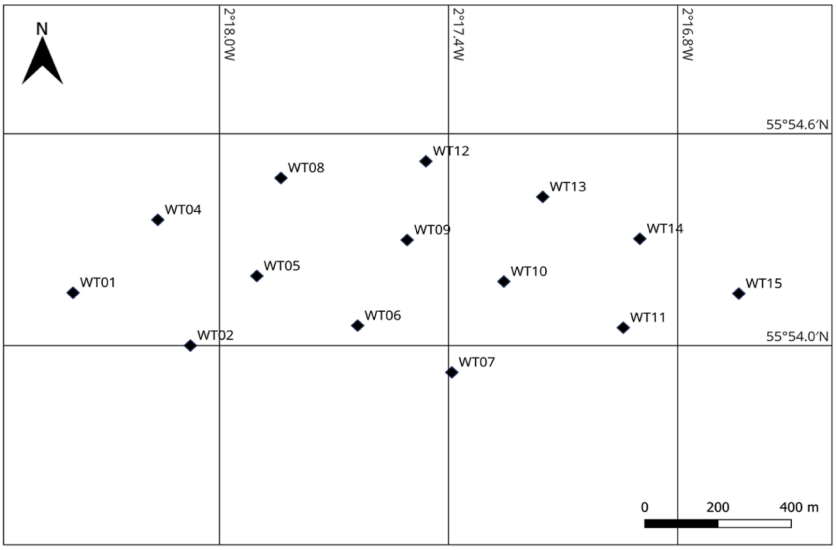


Fig. 5. Geographical layout of the wind farm.

Table 1
Technical characteristics of the wind turbine.

Parameter	Senvion MM82
Rated power (kw)	2050
Rotor diameter (m)	82
Hub height (m)	59
Rated wind speed (m/s)	14.5
Cut-out wind speed (m/s)	25
Cut-in wind speed (m/s)	2.5

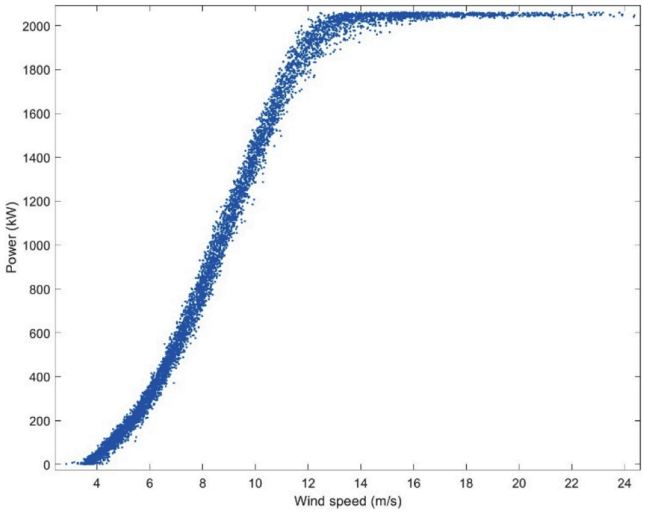


Fig. 6. Wind turbine output power data without outliers.

The dataset already includes corrections for air density of wind speed (2–3). For simplicity, we will refer to this variable simply as 'wind speed' from now on.

The data from the WT01 turbine from January to February 2017 have been used. A total of 8352 points are available. Fig. 2 shows in blue the power curve resulting from representing these raw data.

Fig. 6 shows the clean data that constitutes the set on which the models will be identified with the considered techniques. A total of 7479 points are available, after removing all outliers. Fig. 7 shows the power (left) and wind speed (right) histograms. It can be seen how most of the wind points belong to the range (4–10 m/s), and how the probability of winds above 8 m/s decreases with wind speed. In the case of power, most of the values are slightly above 2 MW.

As stated above, this paper proposes the detection of yaw misalignment as a case study. To do so, it is necessary to have the following variables, both during the failure and in the moments before and after, in order to also consider possible false positives.

- Time stamp (Year-Month-Day HH:mm:ss)
- Wind speed (m/s)
- Power generated (kW).
- Yaw error (°). Difference between nacelle position measured with respect to geographical north (°) and wind direction (°).

Since the dataset [53] used to obtain the power curves did not contain data on failures of the nacelle yaw control system, information from the dataset [12] with a real case of yaw error was used to validate the failure prediction methodology, as said before.

The dataset in [12] contains the yaw error and the corresponding wind speed when the nacelle control system failed and caused the misalignment with respect to the wind direction. In Fig. 8 the yaw error is shown in red, while the wind speed is represented in blue. The turbine failure occurs between minutes 360 and 1230.

The wind turbine with real data of the yaw error and the data of the turbine used to obtain the power curve models do not have the same nominal power. To solve this problem, the power generated by the Senvion MM82 wind turbine has been simulated as a function of the wind turbine's own power curve and the yaw error (20).

$$P(\varphi_e, v_c) = P_0(v_c)(\cos(\varphi_e))^{\gamma}, \tag{20}$$

where $P(\varphi_e)$ is the wind turbine power as a function of the misalignment angle φ_e; $P_0(v_c)$ is the power when the wind is v_c and there is no misalignment, i.e. $\varphi_e = 0$, and γ is a parameter varying between 1.8 and 5.14, extracted from [54]. In our case, considering the location of the turbine within the farm and the distance between it and adjacent turbines (Fig. 5), we assume $\gamma = 2$, following [54]. In [54] an approximate value of 2 is suggested for our scenario: an onshore wind turbine that considers the effects of turbulence with an inter-turbine spacing (distance normalized by the wind turbine diameter) of 3 or more.

From the expression $P(\varphi_e, v_c)$ in (20) it is possible to extract all the variables needed in the dataset: power generated, wind speed, and yaw error. The power generated is P in $P(\varphi_e, v_c)$, wind speed is v_c, and φ_e denotes the yaw error.

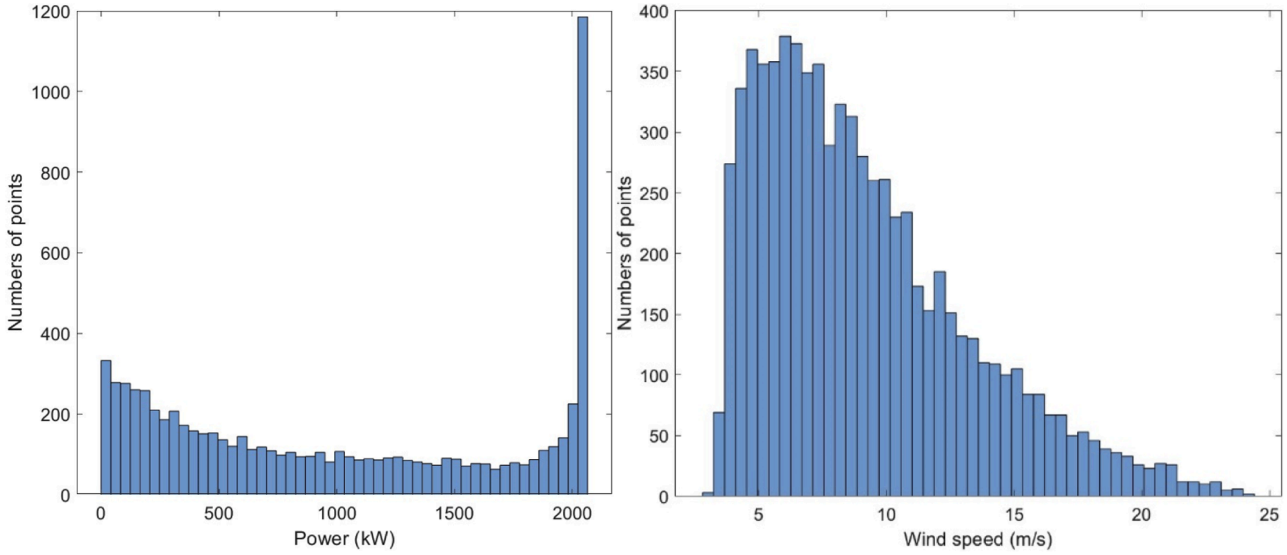


Fig. 7. Histogram of power (left). Histogram of wind speed (right).

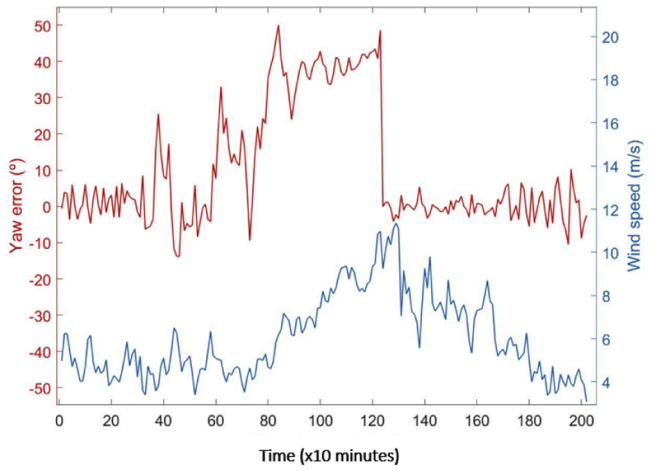


Fig. 8. Yaw misalignment and wind speed.

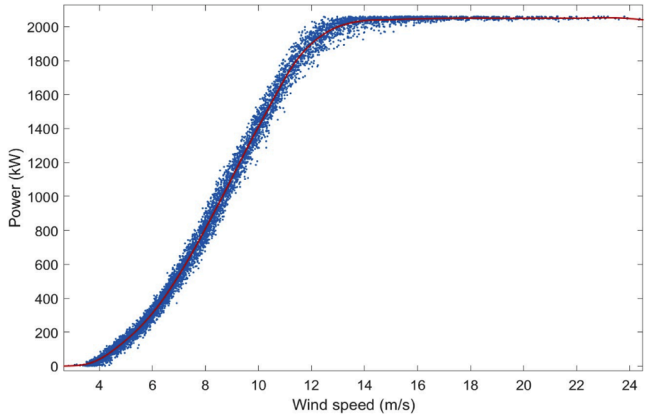


Fig. 9. Simulated power without yaw error.

To obtain the power curve under non-misalignment conditions, $P_0 = P(\varphi_e = 0)$ in (20), a different method is used than those used in the modelling section. Following the international standard IEC 61,400–12–1 [33], the data are grouped into 0.5 m/s intervals and the average is calculated in each of these intervals. From these averages, a spline interpolation function is generated, which is represented by the red line in Fig. 9, while the blue points correspond to the values of the dataset [53].

5. Discussions of the results

For the wind turbine whose characteristics are described in Table 1, different probabilistic models are obtained by applying the techniques presented above. These models are first validated with the turbine data and evaluated and compared with each other. Five models have been obtained: GMM, Frank copula, GMCM, GPR, and ε-SVR.

For the GPR and SVR models, the hyperparameter was tuned using Bayesian optimization. This optimization process uses the acquisition function, the Lower Confidence Bound, to identify the optimal values obtained.

5.1. Comparison of the models of the wind turbine power curve

In the GMM model, the number of clusters or modes that make up the power curve is $k = 3$, equivalent to the number of operating regimes of the turbine according to the wind speed. Fig. 10 shows the joint probability density obtained with the GMM model and, with blue dots, the real data set.

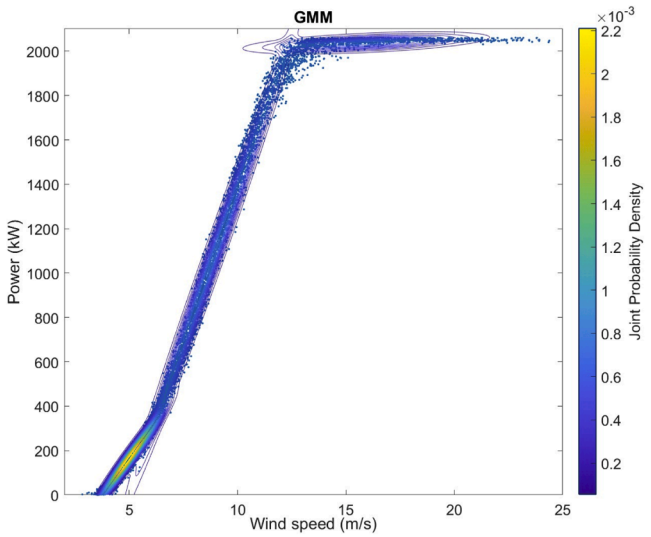


Fig. 10. Power curve with the Gaussian mixture model (GMM).

In this Fig. 10, three clusters can be clearly observed: one centered around 5 m/s, for low wind speeds; another centered about 10 m/s in the MPPT (maximum power point tracking) region, and finally, a cluster centered around 16 m/s in the wind turbine nominal power region [55]. The center of this last cluster is close to the turbine rated speed, which is 14.5 m/s.

For the case of modelling with the Frank copula method, Fig. 11 shows the results for the joint probability density, and the real data set represented by blue dots.

It can be observed in Fig. 11 that the model based on Frank's copula fits the data much better than the GMM. Unlike the previous one, there are no separate clusters but a single cluster. In the cluster there are two areas with higher probability: one for low speeds, around 5 m/s (coinciding with the maximums of the probability distribution in Fig. 7, right), and another around the power of 2 MW (coinciding with the maximums of the probability distribution in Fig. 7, left).

The copula parameter, denoted as δ in Eq. (6), indicates the intensity of the dependence between variables, the higher this value, the greater the dependence. In this case, a value of δ=73.56 was obtained, which is equivalent to the Spearman correlation coefficient, ρ=0.997. Therefore, by choosing this family of copulas we observe a high correlation between the variables.

For the GMCM model, as in the GMM case, the number of clusters is $k = 3$. Fig. 12 shows the GMCM results for the joint probability density and, with blue dots, the real data set.

As a summary of the results obtained with these three probabilistic models, it is evident that the GMM fails to accurately fit the shape of the power curve; furthermore, it does not present a smooth transition between the different modes that compose it. This suggests that the power curve is not formed by elliptical (Gaussian) modes. In contrast, the Frank copula model shows a marked improvement in the ability to capture the shape of the power curve with greater fidelity. However, this model still presents certain deficiencies in accuracy, especially noticeable in the final section of the power curve (nominal power). The Gaussian copula mixture model (GMCM) overcomes these inaccuracies, fitting the curve to the real data with greater fidelity and precision.

Now, for the two regression techniques, Fig. 13 shows the power curve model obtained with GPR, representing the resulting regression curve with a red line; the black lines correspond to the 95 % confidence interval and, superimposed on it, the real data set is shown with blue dots.

This model has been obtained with the squared exponential kernel and the values of the parameter defined in (16) are $l = 2.95$, σ=307.68.

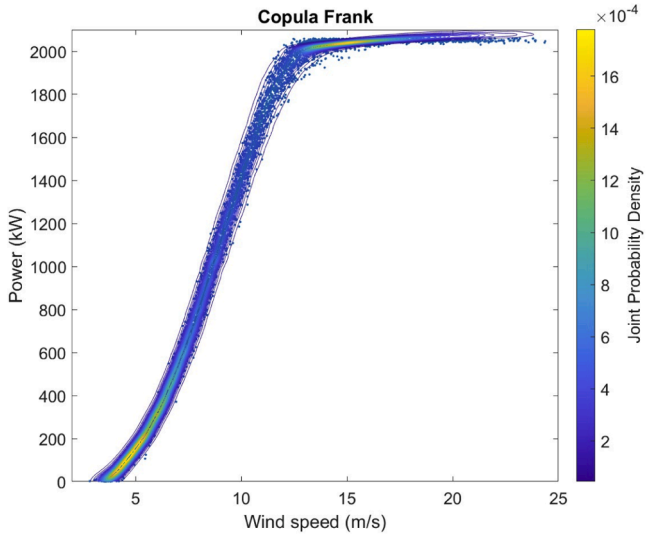


Fig. 11. Power curve with Frank's copula model.

It can be observed how the confidence interval is narrower in the central area of the curve and wider at the ends.

Fig. 14 shows the power curve model of the turbine obtained with the ε-SVR technique, represented with a red line; the black lines correspond to the 95 % confidence interval and, overlapped, the real data set is shown with blue dots. The results obtained for the scale parameter of (13) are σ=29.62.

As can be seen in the two figures above, with both the GPR and ε-SVR models the regression curve fits the dataset accurately over the entire range of the turbine power curve. Fig. 15 presents a histogram with the residuals of the regression fit in both models, where the red line corresponds to the GPR model and the yellow line to the ε-SVR model. As said before, the residuals of the GPR model follow a Gaussian distribution, while the residuals of the ε-SVR model fit better to a Laplacian distribution.

Furthermore, it can be observed that the ε-SVR model produces a better fit to the residuals compared to the GPR model. This indicates that the ε-SVR provides a more accurate fit to the uncertainty associated with the regression curve, resulting in more reliable confidence intervals. However, as shown in Figs. 13 and 14, both models tend to overestimate the confidence intervals as we approach and enter the nominal power

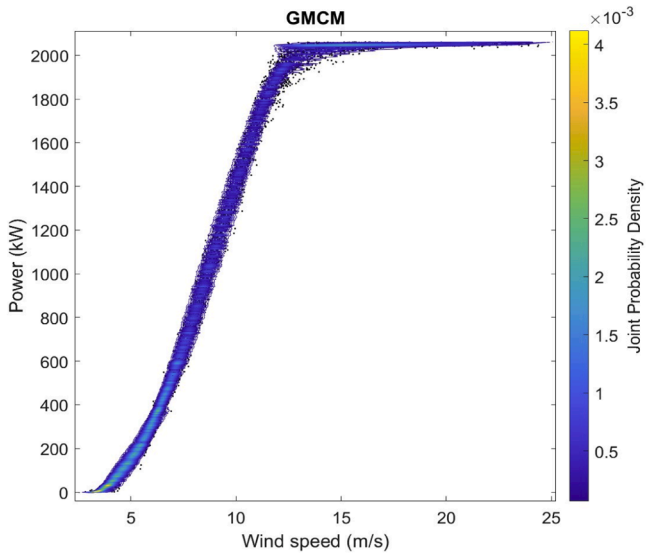


Fig. 12. Power curve with GMCM model.

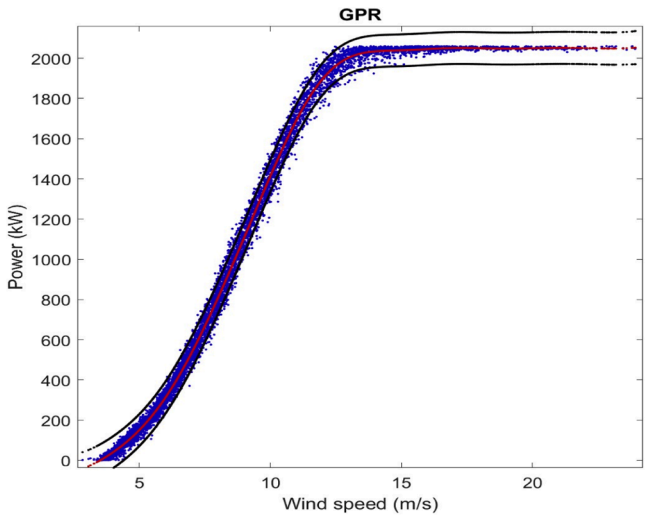


Fig. 13. Wind power curve obtained with the GPR model.

region of the curve. This tendency of GPR and ε-SVR models to overestimate confidence limits is one the main and known limitation of these methods. This overestimation is due to the fact that both models assume a constant dispersion of the data throughout the range of the variable, an assumption that is not met in the real data set used.

To quantify the goodness of fit and assess which is the best model for our purpose, different metrics were calculated. Table 2 summarizes the values obtained for the root mean square error (RMSE), the symmetric mean absolute percentage error (SMAPE), the mean square error (MSE) and the mean absolute error (MAE).

The computational cost required to obtain the models has also been obtained, both in terms of memory and processing time, which are summarized in Table 3. The best results in both tables are marked in bold.

A priori, one might think that the best model, taking into account all the aspects and metrics evaluated, and assuming a compromise between accuracy, associated confidence interval and computational cost, is the GMCM. Quantitatively, according to Table 2 showing the accuracy, the best models are GPR and ε-SVR. However, the difference in value with the GMCM model is practically negligible. But it is important to note that, as noted above, the regression models, GPR and ε-SVR, present deficiencies in the estimation of the uncertainty associated with the prediction. The ability to adequately capture this uncertainty is crucial for a good maintenance strategy, and the GMCM model has proven to be able to meet this requirement more effectively. In addition, the GMCM

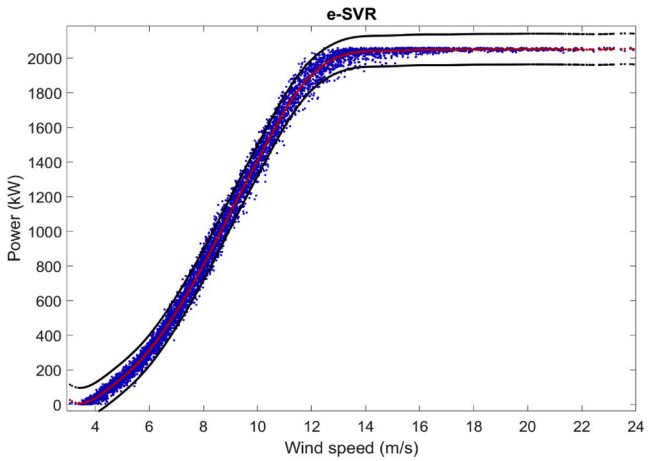


Fig. 14. Wind power curve obtained with the ε-SVR model.

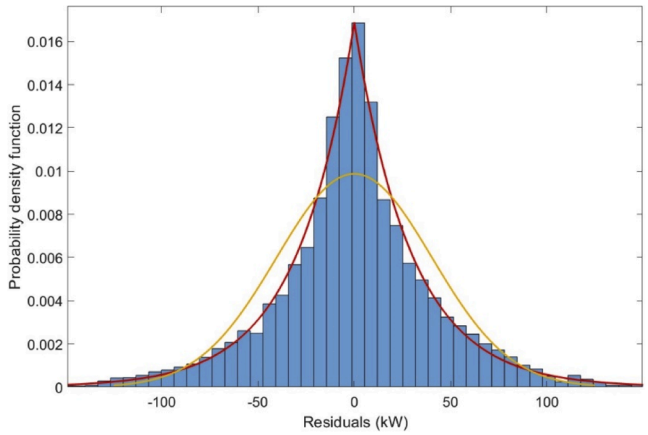


Fig. 15. Residuals of the GPR (red line) and ε-SVR (yellow line) models.

Table 2
Results of the evaluation metrics with the five models.

	RMSE (kW)	S-MAPE (%)	MSE (kW^2)	MAE (kW)
GMM	51.16	8.57	2618	33.08
Frank	41.86	8.55	1752	31.06
GMCM	41.44	8.10	1717	30.08
ε-SVR	41.44	**7.96**	**1716**	**29.62**
GPR	**41.43**	8.28	**1716**	29.79

model is computationally more efficient than the other models evaluated, making it the most viable option for applications that require efficient use of resources.

5.2. Probabilistic fault detection

The results of the previously proposed probabilistic evaluation method for detecting a yaw control failure between the minutes 360 and 1230 are presented below (Fig. 8). Fig. 16 shows the results obtained: the blue line is the absolute value of the yaw misalignment, the red dashed line indicates the minutes at which the turbine failure occurs, and the yellow line represents the alarm activation using the GMCM model. This model has been used since it has given the best overall results.

It is observed that the first alarm is triggered 20 min after the failure, which implies an early detection of the anomaly. After both this first and a second alarm occur, the alarm is reset because the wind speed has changed and thus the misalignment of the orientation is reduced. This is expected since the proposed method only detects the suboptimal behaviour of the turbine.

Subsequently, the third alarm remains activated while the failure persists, without registering any false negatives in this time range. It is also important to note that no false positives occur at any time, neither during the failure nor before or after it.

Fig. 17 presents the results obtained using the same probabilistic evaluation method, but applying the other power curve models considered. As can be seen, these models present a lower performance compared to the GMCM, detecting the turbine failure later. In particular, the GMM model even presents false positives.

Table 4 summarizes the results obtained from the minutes elapsed since the failure occurred until its first detection (first column), the presence or absence of false positives (second column) and the number of alarms activated (third column) for the five models considered.

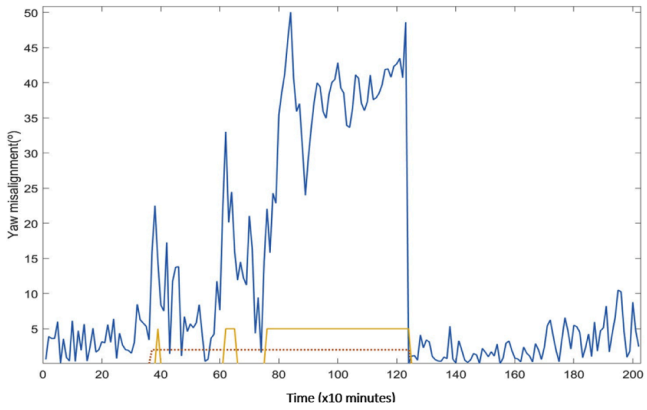


Fig. 16. Detection of yaw misalignment with the GMCM model.

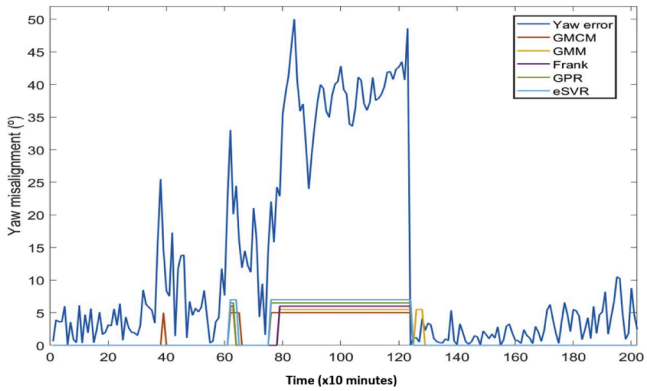


Fig. 17. Detection of yaw misalignment with the five model.

Table 4, which shows that for a fault resulting in a loss of turbine performance, such as when yaw misalignment occurs over a period of time, probabilistic assessment developed using power curve models from the wind turbine's own real data provides an effective and reliable approach for early fault detection. Particularly, it confirms the effectiveness of the probabilistic detection method with the GMCM model of the turbine power curve in terms of fault response time, detection accuracy, and alarm reliability.

Indeed, Table 4 offers a directly comparable analysis among those probabilistic models as the dataset, the wind turbine, and the anomalous conditions, are the same. In this context, we can conclude with some confidence that detection is, in fact, 10 times faster when using our probabilistic detection method with the GMCM than with the same data and evaluation criteria using the other models studied.

The results obtained here have been compared in Table 5 with other results found in the literature, which use real cases of yaw misalignment. These data should be interpreted with due caution as the datasets used in those papers are not available to apply our methodology to them and make a fairer comparison. Table 5 shows a summary of the studies considered and the number of minutes elapsed from the occurrence of the fault until its first detection.

From these studies, in the one by [56], a probabilistic evaluation of a GPR model with the inclusion of rotor speed is used, and the first alarm

Table 3
Computational costs of the different models.

	GMM	Frank	GMCM	GPR	ε-SVR
Time (s)	**1**	**1**	5	1465	425
Memory(bytes)	300	59,832	**168**	690,176	479,232

Table 4
Results of fault detection and alarm generation.

	First alarm (minutes)	False positives	N° alarms
GMCM	20	No	54
GMM	250	Yes	51
Frank	250	No	48
GPR	250	No	51
ε-SVR	250	No	52

Table 5
Comparative with other studies on yaw misalignment detection.

Work	Model used	First alarm detected (minutes)
[56]	GP model with rotor speed	280
[12]	IEC binning model	470
[57]	Copulas model	600
This work	GMCM	20

is triggered at 280 min. In reference [12], the IEC clustering method is used, and the first alarm is triggered at 470 min. In the work of [57], an empirical copulas method is used, and the first alarm is triggered at 600 min. From our results, using the proposed model and method, it can be concluded that the efficiency is improved compared to these methods described in the literature. The methodology presented here allows to detect yaw misalignment at earlier stages of failure, even when this failure is not so clear.

6. Conclusions and future works

The power curve is the best indicator of the state and performance of a wind turbine, and therefore identifying accurate power curve models is necessary for designing control strategies and for operation and maintenance tasks. Data-driven models can be obtained using data measured by sensors in SCADA acquisition systems, which provide a real-time monitoring for wind turbine conditions. In addition, having an estimate of the uncertainty associated with the model is a significant advantage for reliable and accurate power curve modelling and prediction.

In this work, five probabilistic models have been applied to identify the power curve of a real turbine, namely GMM, Frank copula, GMCM, GPR, and ε-SVR. Of these, the GMCM has proven to be the best model in this case. Its main advantage lies in its ability to faithfully fit the shape of the power curve over its entire range, while providing an accurate estimate of the associated uncertainty. All this is achieved while keeping the computational cost low.

These models have been used to apply a methodology that allows early detection of turbine failures. Specifically, it has been applied to a real case of failure in the yaw control, and the proposed probabilistic evaluation method together with the GMCM model of the power curve have achieved the detection of the 20-minute yaw misalignment. This result represents a significant improvement with respect to other methods in the literature.

It is also important to highlight that, like the other models considered in this work, the GMCM does not produce false positives on the data set used. All this makes the proposed probabilistic evaluation method fast and reliable for early fault detection, providing a credible tool for predictive maintenance and improving the operating costs and performance of wind turbines.

We cannot ignore that one of the main limitations of regression models such as GPR and ε-SVR is their tendency to overestimate confidence intervals in estimation and prediction-related uncertainties. Possible options to address this limitation would be to explore calibrating GPR hyperparameters using more extensive cross-validation or optimizing the kernel function to better account for data heteroskedasticity. For ε-SVR, alternative loss functions could be investigated that penalize interval overestimation differently.

As future lines of research, the suitability of the proposed method could be tested on other types of failures that also imply a reduction in turbine performance, which can be identified from the power curve, such as blade failures, failures in the pitch system, rigidities in the mechanical parts, etc. It could also be studied whether it is possible to identify which of all the systems that make up a turbine is the one that causes the failure. Other variables of the SCADA system could be also included as input parameters to the model to further refine it, such as rotor speed, wind direction or pitch blade. Finally, it would be desirable a large-scale validation with multiple turbine models and wind farms to confirm the method universality, considering this study as a proof-of-concept validation.

Data availability

Data are available through the links provided by the public databases.

CRediT authorship contribution statement

M.A. García-Vaca: Writing – original draft, Investigation, Data curation, Conceptualization. **J.E. Sierra-García:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Matilde Santos:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This research has been partially supported by the Spanish Ministry of Science and Innovation under the MCI/AEI/FEDER project number PID2021–123543OB-C21.

References

[1] Zhou B, Zhang Z, Li G, Yang D, Santos M. Review of key technologies for offshore floating wind power generation. Energies (Basel) 2023;16(2):710.
[2] Qadir SA, Al-Motairi H, Tahir F, Al-Fagih L. Incentives and strategies for financing the renewable energy transition: a review. Energy Reports 2021;7:3590–606.
[3] IEA (2024) https://www.iea.org/reports/offshore-wind-outlook-2019, International Energy Agency. [En línea] Last accessed: 10/2024.
[4] Yan R, Dunnett S, Jackson L. Impact of condition monitoring on the maintenance and economic viability of offshore wind turbines. Reliab Eng Syst Saf 2023;238:109475.
[5] Florian E, Sgarbossa F, Zennaro I. Machine learning-based predictive maintenance: a cost-oriented model for implementation. Int J Prod Econ 2021;236:108114.
[6] Pandit R, Astolfi D, Hong J, Infield D, Santos M. SCADA data for wind turbine data-driven condition/performance monitoring: a review on state-of-art, challenges and future trends. Wind Eng 2023;47(2):422–41.
[7] Marti-Puig P, Hernández JÁ, Solé-Casals J, Serra-Serra M. Enhancing reliability in wind turbine power curve estimation. Appl Sci 2024;14(6):2479.
[8] Elkodama A, Ismaiel A, Abdellatif A, Shaaban S, Yoshida S, Rushdi MA. Control methods for horizontal axis wind turbines (HAWT): state-of-the-art review. Energies (Basel) 2023;16(17):6394.
[9] Pandit R, Santos M, Sierra-García JE. Comparative analysis of novel data-driven techniques for remaining useful life estimation of wind turbine high-speed shaft bearings. Energy Sci Eng 2024;12:4613–23.
[10] Pandit R, Infield D, Santos M. Accounting for environmental conditions in data-driven wind turbine power models. IEEE Trans Sustain Energy 2022;14(1):168–77.
[11] Wang Y, Duan X, Zou R, Zhang F, Li Y, Hu Q. A novel data-driven deep learning approach for wind turbine power curve modeling. Energy 2023;270:126908.
[12] Pandit RK, Infield D. SCADA-based wind turbine anomaly detection using gaussian process models for wind turbine condition monitoring purposes. IET Renew Power Gener 2018;12(11):1249–55.
[13] Zhang C, Hu D, Yang T. Research of artificial intelligence operations for wind turbines considering anomaly detection, root cause analysis, and incremental training. Reliab Eng Syst Saf 2024;241:109634.

[14] Zheng M, Man J, Wang D, Chen Y, Li Q, Liu Y. Semi-supervised multivariate time series anomaly detection for wind turbines using generator SCADA data. Reliab Eng Syst Saf 2023;235:109235.
[15] Abdelsattar M, Ismeil MA, Menoufi K, AbdelMoety A, Emad-Eldeen A. Evaluating Machine Learning and Deep Learning models for predicting wind turbine power output from environmental factors. PLoS One 2025;20(1):0317619.
[16] Alagha N, Khairuddin ASM, Haitaamar ZN, Al-Khatib O, Kanesan J. Artificial intelligence in wind turbine fault detection and diagnosis: advances and perspectives. Energies (Basel) 2025;18(7):1680.
[17] Zhang W, Vatn J, Rasheed A. A review of failure prognostics for predictive maintenance of offshore wind turbines. In J Phys: Conf Series 2022;2362(1): 012043.
[18] Bilendo F, Meyer A, Badihi H, Lu N, Cambron P, Jiang B. Applications and modeling techniques of wind turbine power curve for wind farms—A review. Energies (Basel) 2022;16(1):180.
[19] Santiago RADF, Barbosa NB, Mergulhão HG, Carvalho TFD, Santos AAB, Medrado RC, Nascimento EGS. Data-driven models applied to predictive and prescriptive maintenance of wind turbine: a systematic review of approaches based on failure detection. Diagnosis, and Prognosis Energies 2024;17(5):1010.
[20] Letzgus S, Müller KR. An explainable AI framework for robust and transparent data-driven wind turbine power curve models. Energy and AI 2024;15:100328.
[21] Zhang S, Robinson E, Basu M. Wind turbine condition monitoring based on three fitted performance curves. Wind Energy 2024;27(5):429–46.
[22] Mushtaq K, Waris A, Zou R, Shafique U, Khan NB, Khan MI, Khan MI. A comprehensive approach to wind turbine power curve modeling: addressing outliers and enhancing accuracy. Energy; 2024, 131981.
[23] Wang Y, Duan X, Zou R, Zhang F, Li Y, Hu Q. A novel data-driven deep learning approach for wind turbine power curve modeling. Energy 2023;270:126908.
[24] Yin W, Jia M, Liu L, Li M, Guo Y, Lei G, Zhu JG. Advanced power curve modeling for wind turbines: a multivariable approach with SGBRT and grey wolf optimization. Energy Convers Manag 2025;332:119680.
[25] Dhungana H. A machine learning approach for wind turbine power forecasting for maintenance planning. Energy Inform 2025;8(1):2.
[26] Yesilbudak M, Ozcan A. Performance comparison of metaheuristic optimization-based parametric methods in wind turbine power curve modeling. IEEE Access 2024;12:99372–81.
[27] Qian GW, Ishihara T. A novel probabilistic power curve model to predict the power production and its uncertainty for a wind farm over complex terrain. Energy 2022; 261:125171.
[28] Li Y, Wang P, Wu Z, Su Y. Collaborative monitoring of wind turbine performance based on probabilistic power curve comparison. Renew Energy 2024;231:120919.
[29] Capelletti M, Raimondo DM, De Nicolao G. Wind power curve modeling: a probabilistic beta regression approach. Renew Energy 2024;223:119970.
[30] Mclean JH, Jones MR, O'Connell BJ, Maguire E, Rogers TJ. Physically meaningful uncertainty quantification in probabilistic wind turbine power curve models as a damage-sensitive feature. Struct Health Monit 2023;22(6):3623–36.
[31] Haghi R, Stagg C, Crawford C. Wind turbine damage equivalent load assessment using gaussian process regression combining measurement and synthetic data. Energies (Basel) 2024;17(2):346.
[32] Tian Z, Wang H. Wind power system reliability and maintenance optimization considering turbine and wind uncertainty. J Qual Maint Eng 2022;28(1):252–73.
[33] IEC (International Electrotechnical Commission). Wind energy generation systems—Part 12-1: power performance measurements of electricity producing wind turbines. Int Electr Commission (IEC), IEC Central Office 2017;3. 2017-03.
[34] Jargalsaikhan N, Ueda S, Masahiro F, Matayoshi H, Mikhaylov A, Byambaa S, Senjyu T. Exploring influence of air density deviation on power production of wind energy conversion system: study on correction method. Renew Energy 2024;220: 119636.
[35] Zou M, Djokic SZ. A review of approaches for the detection and treatment of outliers in processing wind turbine and wind farm measurements. Energies (Basel) 2020;13:4228.
[36] Shokrzadeh S, Jafari Jozani M, Bibeau E. Wind turbine power curve modeling using advanced parametric and nonparametric methods. *Sustainable energy*. IEEE Transactions on 2014;5:1262–9.
[37] Zuo C, Dai J, Li G, Li M, Zhang F. Investigation of data pre-processing algorithms for power curve modeling of wind turbines based on ECC. Energies (Basel) 2023;16 (6):2679.
[38] Villena-Ruiz R, Honrubia-Escribano A, Jiménez-Buendía F, Sosa-Avendaño JL, Frahm S, Gartmann P, Gómez-Lázaro E. Extensive model validation for generic IEC 61400-27-1 wind turbine models. Int J Electr Power Energy Syst 2022;134:107331.
[39] Simankov V, Buchatskiy P, Teploukhov S, Onishchenko S, Kazak A, Chetyrbok P. Review of estimating and predicting models of the wind energy amount. Energies (Basel) 2023;16(16):5926.
[40] Bouguila N, Fan W, editors. Mixture models and applications. Berlin: Springer; 2020.
[41] Geenens G. (*Re-*) Reading Sklar (1959)—A personal view on Sklar's theorem. Mathematics 2024;12(3):380.
[42] Yan J. Multivariate modeling with copulas and engineering applications. Springer handbook of engineering statistics; 2023. p. 931–45.
[43] Tewari A. On the estimation of gaussian mixture copula models. Int Conf Machine Learn 2023;202:34090–104.
[44] Zhang F, O'Donnell LJ. Support vector regression. Machine learning; 2020. p. 123–40.
[45] García-Vaca MÁ, Sierra-García JE, Santos M. Prediction and uncertainty estimation in power curves of wind turbines using ε-SVR. In: International Conference on Intelligent Data Engineering and Automated Learning; 2023. p. 509–17.
[46] Wang J. An intuitive tutorial to Gaussian process regression. Comput Sci Eng 2023; 25(4):4–11.
[47] Yu X, Li D, Xue L. Fisher's combined probability test for high-dimensional covariance matrices. J Am Stat Assoc 2024;119(545):511–24.
[48] Yoon S, Baik B, Park T, Nam D. Powerful p-value combination methods to detect incomplete association. Sci Rep 2021;11(1):6980.
[49] Rott Andreas, Höning Leo, Hulsman Paul, Lukassen Laura, Moldenhauer Christof, Kühn Martin. Wind vane correction during yaw misalignment for horizontal-axis wind turbines. Wind Energy Sci 2023;8:1755–70.
[50] Starke GM, Meneveau C, King JR, Gayme DF. A dynamic model of wind turbine yaw for active farm control. Wind Energy 2024;27(11):1302–18.
[51] Zhang Jianwei, Wang Jianwen, Yan Sijia. The effect of yaw speed and delay time on power generation and stress of a wind turbine. Int J Green Energy 2022;20: 1–13.
[52] Sacie M, Santos M, López R, Pandit R. Use of state-of-art machine learning technologies for forecasting offshore wind speed, wave and misalignment to improve wind turbine performance. J Mar Sci Eng 2022;10(7):938.
[53] Plumley, C. (2022) https://doi.org/10.5281/zenodo.5946808. Last accessed: 10/ 2024.
[54] Urbán AM, Liew J, Dellwik E, Larsen GC. The effect of wake position and yaw misalignment on power loss in wind turbines. In Journal of Physics: Conference Series 2019;1222(1):012002.
[55] Muñoz-Palomeque E, Sierra-García JE, Santos M. Técnicas de control inteligente para el seguimiento del punto de máxima potencia en turbinas eólicas. Revista Iberoamericana de Automática e Informática industrial 2024;21(3):193–204.
[56] Pandit R, Infield D, Dodwell T. Operational variables for improving industrial wind turbine yaw misalignment early fault detection capabilities using data-driven techniques. IEEE Trans Instrum Meas 2021;70:1–8.
[57] Wang Y, Infield DG, Stephen B, Galloway SJ. Power curve based online condition monitoring for wind turbines. In: COMDEM 2013 conference. 10; 2013. p. 4492–9928.