Renewable Energy 205 (2023) 281–292

Contents lists available at ScienceDirect

Renewable Energy

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

A methodology for performance assessment at system level—Identification of operating regimes and anomaly detection in wind turbines

Jon Urmeneta [a,b], Juan Izquierdo [a,b], Urko Leturiondo [c,*]

[a] MIK Research Centre, Ibarra Zelaia 2, 20560, Oñati, Gipuzkoa, Spain
[b] Mondragon University, Faculty of Business Studies, Ibarra Zelaia 2, 20560, Oñati, Gipuzkoa, Spain
[c] Control and Monitoring Area, Ikerlan Technology Research Centre, Basque Research and Technology Alliance (BRTA), Pº J.M. Arizmendiarrieta, 2. 20500 Arrasate/Mondragón, Spain

A R T I C L E I N F O

Keywords:
Performance
Maintenance management
Wind energy
Anomaly detection
Machine learning

A B S T R A C T

In the growing wind energy sector, as in other high investment sectors, the need to make assets profitable has put the spotlight on maintenance. Efficient solutions which leverage from condition or performance based maintenance policies have been proposed during the last decades, but the proposed methods generally focus on individual components or stand for specific application areas. This paper aims to contribute to the development of performance based maintenance strategies within the wind energy sector by providing a condition monitoring based generic methodology for wind turbine performance assessment at system level. The proposed methodology is based on the detection of critical periods in which low performance is detected repeatedly. Multiple machine learning methods and models are applied to assess the wind turbine performance. This methodology has been applied in a case study with SCADA data of eight wind turbines. An analyst could benefit from the implementation of the methodology and the easy-to-interpret results shown in the proposed control chart, especially in cases in which there is less know-how about which variables have higher impact on systems performance.

1. Introduction

In the face of economic downturn, current global competition and increasing demands from stakeholders, there is a distinct need to improve manufacturing performance [1] and, in general, the performance of assets and processes. Due to the need to meet demanding requirements related to efficiency and effectiveness, the maintenance has stand out and it is no longer considered a necessary evil but a fundamental activity for the fulfillment of the strategic objectives of organizations [2]. The definition of operation and maintenance (O&M) strategies has even greater relevance in sectors in which large investments are required, such as wind industry.

The urgent need to switch from traditional energy sources to renewables has put the spotlight on the latter [3]. Wind power is recognized as one of the most attractive renewable energy sources, and it is expected to be the supplier of more than one-third of total electricity demand by 2050, leading the way in the transformation of the global electricity sector among with solar photovoltaic [4]. The decrease in the levelized cost of energy (LCOE), which reduced by 35% between 2010 and 2018 in the case of wind energy, promotes this transformation [4]. O&M costs are the main contributors to the variable costs of wind power plants, and they constitute a sizeable share of the total annual costs [5]. Off-shore wind farms are particularly difficult to maintain and faults lead to higher downtimes due to the challenging ambient conditions [6]. Therefore, the economic viability of wind farms depend on the success of the long term O&M [7], which should ensure system's reliability by performing sufficient maintenance actions, but also seek optimizing costs [8]. In a greater scale, the efficiency and effectiveness of the wind industry, as well as its leadership in the competence among other energy sources, depend on the reliability, availability, maintainability and safety of wind turbines (WTs) [9].

1.1. Performance-based maintenance

Maintenance can be classified as preventive maintenance (PM) and corrective maintenance (CM) [10]. Predetermined maintenance is the PM carried out in accordance with established intervals [10]. Although predetermined maintenance is based on statistical analysis of failure characteristics of equipment [11], this type of maintenance requires relatively predictable patterns of failure, which are harder to identify due to the complexity of newer equipment of the last decades [12].

* Corresponding author.
E-mail address: uleturiondo@ikerlan.es (U. Leturiondo).

https://doi.org/10.1016/j.renene.2023.01.035
Received 29 June 2022; Received in revised form 27 December 2022; Accepted 8 January 2023
Available online 10 January 2023
0960-1481/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Nomenclature

$\beta_{p,q}$	Weight of the pth eigenvector for the qth variable in V
Θ	Set of centroids for the operating regimes
$\theta_{k,m}$	Measure of the kth centroid of Θ for the mth variable in C
$\vec{\Theta}_k$	kth centroid in Θ
$\vec{B}_p$	pth eigenvector in B
$\vec{B}_p$	pth eigenvector in B
$\vec{C}_m$	mth variable in C
$\vec{f}_k$	Vector of E_k frequencies obtained for each of the unique values in $\vec{W}_k$
$\vec{H}_k$	Vector of cumulative relative contributions to the KPI of each of the unique values in $\vec{W}_k$
$\vec{h}_k$	Vector of relative contributions to the KPI done by each of the unique values in $\vec{W}_k$
$\vec{P}^j$	Vector of probabilities for transitions from G_i to each of the operating regimes
$\vec{V}_p$	pth variable in V
$\vec{W}_k$	Vector of E_k unique values of the KPI obtained in the kth operating regime
$\vec{X}_n$	nth variable in X
$\vec{Y}$	KPI
$\vec{Y}_{thr}$	The set of thresholds indicating lowest accepted values for the KPI in each operating regime
B	Set of P eigenvectors
C	Set of variables after dimensionality reduction
c_m	mth record in C
$c_{j,m}$	Value for the jth record of the mth principal component
d	Width of the time window in which critical records are assessed
E_k	Amount of unique values identified for the KPI in the kth operating regime
$f_{k,e}$	Relative frequency obtained
G	Set of operating regimes
G_i	In a transition between operating regimes, the old operating regime
g_j	Operating regime assigned to the jth record
G_l	In a transition between operating regimes, the new operating regime
$H_{k,e}$	Cumulative relative contribution to the KPI done by the eth unique value in the kth operating regime
$h_{k,e}$	Relative contributions to the KPI done by the eth unique value of the kth operating regime
$H_{k,q}$	The value of the cumulative contribution to the KPI done by the first q unique values. It is used to set the threshold $y_{thr,k}$
J	Number of records of the original data
K	Number of operating regimes
k	Operating regime indicator between 1 and K for operating regimes in G
M	Amount of principal components selected and amount of variables of C
m	Principal component indicator between 1 and M, for principal components in C
N	Number of variables of the original data
P	Number of variables of V
$p^j_{i,l}$	The probability at time j to make a transition from the ith operating regime in G, to the lth operating regime in G.
R_j	Amount of critical records within the time window of the jth record
R_{thr}	Threshold indicating the maximum amount of critical records accepted within the time window for a record
S	Covariance matrix of V
S_p	Covariance matrix of V
t_1	Threshold for the minimum probability required for the validation of the transition between operating regimes
T_k	Transition matrix composed by the probabilities for transitions between operating regimes in G
V	Set of variables which feed the model
$v_{j,p}$	The value of the pth variable in V for the jth record
$w_{k,e}$	eth unique value in the kth operating regime
X	Original set of variables
$x_{j,n}$	Value for the jth record of the nth variable
y_j	Value of the KPI for the jth record
$y_{thr,k}$	The lowest accepted value for the KPI for value in the kth operating regime

Besides, as it does not look after current equipment condition, some maintenance actions turn to be unnecessary and lead to higher maintenance costs.

Condition based maintenance (CBM) is the type of PM in which current equipment condition is monitored based on the premise posed by Bloch and Geitner [13] that certain signals may warn about immediate failures. Due to the benefits regarding failure detection at early stages and prevention of further deterioration, condition monitoring become increasingly important during the last decades [2]. Through the detection of degradation and deviations from a supposed normal behavior, CBM intends to anticipate maintenance actions to failures [14]. CBM involves a first process of diagnostics, in which faults are detected, isolated and identified, and a second process of prognostics, in which the impending failures and the remaining useful life are forecasted [15]. CBM can significantly reduce both the economic losses caused by system breakdown and the costs attributed to unnecessary repair and replacement of components [15], which is an advantage with respect to predetermined maintenance and CM, as maintenance actions are better planed and performed only when necessary [16].

The aforementioned types of maintenance are aimed at minimizing opportunity costs, as well as others such as replacement or repair costs. However, by focusing more in the first type of costs, it is interesting to study the impact of failures that do not involve a loss of functionality. Note that failures are detected when the components do not fulfill the required functions, which may be related to performance-based criteria. In fact, even if the system remains functional, under-performance can lead to considerable opportunity costs. Performance-based maintenance lies on the detection of instants when the system productivity differs from what is expected [17]. This type of criterion proposes a new approach that contrasts with the traditional view in which failure is described as a binary event. The present research is set in this context and aims to move forward in the performance based maintenance in the wind energy sector.

1.2. Related works

In the wind industry, the performance evaluation of WTs involves the estimation of the health state of a WT in comparison to a known or validated normal condition state [18]. However, there is still no consensus within the wind industry on the definition of key performance indicators (KPIs) during the O&M phase of WTs [19]. Some authors [20] claimed that monitoring WT performance is the main feature characterizing WT overall operation, also based on the fact that component malfunction may degrade the energy conversion efficiency, leading to performance deviations. The work of Kusiak et al. [21], which reviewed various approaches to condition and performance monitoring of WT components, concluded that research in WT condition monitoring focused more on individual components [22,23], and suggested that research on models for monitoring WT systems is needed. Gao and Liu [24] classified condition monitoring methods into model-based [25], signal-based [26], knowledge-based [27] and hybrid ones [28], and also provided a list of typical faults in WTs. Faults on blades and rotors, gearbox, generator, bearings, main shaft and hydraulic systems were among the most typical ones. Monitoring of the components housed in the nacelle is the most well-developed and market-penetrating type of WT monitoring [29]. Wymore et al. [29] provided a survey on existing monitoring systems and concluded that it is necessary to make a challenging transition from simple instrumentation system to a useful health monitoring system with easy-to-interpret results. The recent developments in sensors and signal processing systems, big data management, machine learning (ML) and improvements in computational capabilities have opened up opportunities for integrated and in-depth condition monitoring analytics, facilitating robust decision-making [30].

A literature review has been done and some aspects regarding the aforementioned topics are highlighted in the following lines. Most ML-driven condition monitoring models use SCADA (Supervisory Control And Data Acquisition) systems as data source [30]. Mérigaud and Ringwood [6] stated that SCADA system's outputs may not allow accurate damage detection or severity assessment, but can evidence the presence of faults. This system provides large volumes of data through operational data, availability data and alarm system data, which is integrated with maintenance logs and work orders to create complete databases. Major efforts have been made to drive the automation of data collection, but as stated by Leahy et al. [31], still most of cases require a significant amount of transformation or manual data entry before being useful, and some analysis cannot be done due to issues related to data quality.

Due to the high amount of variables and size of the databases, analyzing relationships among them can be hard [32]. Besides, not all data provided by SCADA systems are of interest for condition monitoring purposes [33] and data captured through sensing systems are typically not direct indicators of failure occurrence, so the processing of measurements is necessary in most of the cases [34].

Marti-Puig et al. [32] reviewed feature selection algorithms for WT failure prediction, and compared their performance against the method proposed for automatic feature selection [32]. In addition, Stetco et al. [30] referred to the need to combine and reduce the amount of the selected variables, which is performed by means of dimensionality reduction techniques as auto-encoders or principal component analysis (PCA), although other authors also included polynomial feature generation [35].

Condition monitoring solutions seek the definition of validated state, which can be obtained through classification-based or regression-based approaches. The first ones determine, for each instant, if the WT is or not operating normally by means of ML classifiers. For such purpose, classification-based models find relationships between each of the predefined categories and the explanatory features. Categories can be set in different levels of granularity, from generic "faulty" and "healthy" to deeper states associated to specific failure modes. As stated by Stetco et al. [30], this process can be time consuming, error prone and likely to result in unbalanced number of classes. These three issues were addressed by Leahy et al. [35], and a framework was presented for the identification of faulty operation based on alarm combinations, classification of data as healthy or pre-fault, and real time fault prediction. In this case, support vector machines and decision trees were the best performing classification algorithms, but others such as artificial neuronal networks [23,36,37] or Bayesian classification models [38] are also found in literature.

Alternatively, regression-based approaches focus on modeling the normal behavior to predict numeric outputs. The difficulty in identifying nominal operating periods was mentioned in [31]. Authors noted that filtering abnormal periods through alarms and fault logs or statistical methods requires significant processing, and claimed that such efforts could be avoided through improved turbine data. Korkos et al. [39] developed a method for fault detection in the hydraulic pitch system using adaptative neuro fuzzy inference system. Great efforts were required for the pre-processing and labeling the data, including the definition of boundaries based on the power curve. The use of power curves stands out in literature for performance monitoring purposes. In [30] parametric and non-parametric modeling techniques were distinguished for power curves, but a deeper survey was provided by Lydia et al. [40], focusing on each modeling technique. Huang et al. [41] proposed a WT health assessment framework based on power analysis through ideal power curve estimation. In this case, K-means and density-based clustering methods were applied for outlier elimination, obtaining better results by the latter. Power curves are also used for segmenting normal behavior in [42]. Xu et al. [42] proposed an adaptive fault detection scheme based on normal behavior modeling through power curves and random forest, and adaptive residual monitoring by means of cumulative sum control chart.

Several of the reviewed works considered failures as a binary state [35], or were developed for specific failure modes [22,23,26,27,39]. Other proposals [24,25] considered the performance of the system, but finally transformed it to a binary output. In this paper, a new methodology which considers the performance of the WT at system level is proposed, as an alternative to existing ones [18,20,33,42]. In the proposals for WT condition monitoring and fault detection of Schlenchtingen et al. [33] and Xu et al. [42], as well as in the proposal based on high frequency SCADA data of Gonzalez et al. [20], normal and abnormal behaviors are distinguished, but no operating regimes are considered within normal behavior. In contrast, Lapira et al. [18] proposes a framework that utilizes multi-regime modeling approach to consider the dynamic working conditions of WTs. However, the operating regimes are implicitly identified by the employed models, whereas in the proposed methodology they are identified explicitly. Besides, in the proposed methodology the probability to switch between operating regimes is analyzed, and, in order to handle uncertainties, transitions between regimes with low probabilities are identified.

From the methodological point of view, the main contribution of this paper is the way in which performance is assessed, explicitly identifying operating regimes, analyzing the probability to switch between operating regimes and identifying transitions between operating regimes with low probabilities to handle uncertainty. ML methods that are well known in literature have been employed to evaluate the suitability of the methodology by implementing it in the case study, but identifying the best ML methods for the implementation of the methodology in the case study is considered for further research.

1.3. Motivating industry problem

Existing maintenance policies consider the failure as a binary state, which has led to the development of failure detection methods for individual components of the system, complicating the assessment and the maintenance decision making process. In contrast, the need to define maintenance policies which consider the non-binary failures and the degradation of performance has been stated in literature. This promoted the development of the proposed methodology, which considers the WT as a unique system for performance assessment and relies on ML models. The novelty of the work resides in the manner in which performance assessment is proposed in the methodology, taking into account the degradation of the WTs systems as a unit. The methodology can support the definition of performance based maintenance strategies, which can be combined with traditional maintenance strategies based on individual component failures.

The implementation of this methodology in offshore case studies is of great interest. This lies in the fact that preventive maintenance interventions are performed less frequently and require more planning in offshore wind farms. Therefore, opportunity costs related to WTs that do not produce at top performance have more impact in this case. However, the methodology is also applicable to onshore WTs, in which the higher frequency of maintenance actions leads to higher reliability of the components of the systems and higher performance. From this point of view, onshore scenarios are more challenging for the implementation of the proposed methodology. That is why, even if the methodology was thought to be implemented in offshore wind farms, real data about 8 onshore WTs was used in the case study presented in this paper. It is expected that, if the methodology is capable of providing good results in this case study, it can be effectively extended to offshore scenarios.

Analysts facing the practice application of the proposed methodology can benefit from the high interpretability of the results. Besides, it can be adapted to different risk levels to be assumed while detecting low performance intervals in new practice applications, by modifying the criteria to identify them.

The proposed methodology is described in Section 2, and its implementation in the case study, as well as the results, are shown in Section 3. Finally, Section 4 summarizes the conclusions about the methodology and gives some future lines worth of further research.

2. Methodology for anomaly detection in WTs

The aforementioned motivation promoted the development of the performance based maintenance methodology presented in this paper. This approach is based on the detection of critical periods in which, through condition monitoring, low performance is detected repeatedly. It is composed by two modules preceded by historical data pre-processing. The first module is oriented towards system performance characterization based on historical data and the second one to the system performance assessment based on real time measurements. Proved and widely used algorithms and techniques were selected for the implementation of the methodology in the case study.

Based on historical data of the set of variables obtained by the condition monitoring system, and once it is adequately cleaned, it is intended to define a series of operating status or operating regimes of the WTs. Each of these operating regimes is related to a range of power output and a minimum acceptable value based on the aforementioned range, which is compared to the real output once the current operating regime is identified in real time. Besides, the identification of the transitions between operating regimes within the historical data allows to test if it is likely to switch from the previous operating regime to the currently identified one.

The proposed methodology is graphically shown in Fig. 1. As it can bee seen in the figure, the methodology consists on various steps split in the following modules:

- **Previous step of pre-processing:** A previous process involving data cleaning issues is required for the implementation of the methodology. It is common to have problems regarding missing data or quality of the data.
- **First module:** The information provided by some of the measured variables is used in order to define operating regimes with different production outputs. Then, the transitions between the operating regimes are analyzed.
- **Second module:** Based on the information provided by the variables, the operating regime corresponding to the current time is identified, and it is validated depending on the likelihood to switch from the previous operating regime to the new one. Finally, the real power output is compared to the minimum acceptable one, and instants with low performance are identified. The methodology proposes a time window frame for the identification of critical periods in which low performance is identified repeatedly.

2.1. Proposed methodology

The methodology is based on the historical data study. These are set by the variables $X = (\vec{X}_1, \vec{X}_2, \dots \vec{X}_N)$, having for each variable the x_j measurements recorded in J temporally equidistant instants $j = (1, 2, \dots, J)$. Thus, $x_{j,n}$ indicates the measurement of the jth record for the nth variable $\vec{X}_n$ being $n = (1, 2, \dots, N)$ and N total amount of variables. By the methodology, it is intended to identify, among the totality of variables X, the subset of variables V which feeds the model, so that $V \subset X$, where $V = (\vec{V}_1, \vec{V}_2, \dots \vec{V}_P)$ and $P < N$. The jth record is defined as v_j and the pth variable of the subset as $\vec{V}_p$ from which $v_{j,p}$ measurements are obtained. For the dimensionality reduction, M principal components are selected in order to form the new data $C = (\vec{C}_1, \vec{C}_2, \dots \vec{C}_M)$, where c_j stands for the jth record and $\vec{C}_m$ for the mth variable, being $m = (1, 2, \dots, M)$. In this way, $c_{j,m}$ represents the jth measurement of the mth new variable. Based on these new variables, the K operating regimes $G = \{G_1, G_2, \dots, G_K\}$ are defined. Besides, the operating regime of each jth record is set as $g_j \in G$, and for the transitions between operating regimes, leaving regimes and arrival regimes are defined as $G_i \in G$ and $G_l \in G$ respectively, as well as the threshold t_1 establishing the minimum likelihood to validate the transitions. Finally, the KPI $\vec{Y}$ and its value for each record y_j are defined. The lowest accepted value for the KPI in each operating regime is also set as a threshold vector $\vec{Y}_{thr} = (y_{thr,1}, y_{thr,2}, \dots, y_{thr,K})$.

Similar to the proposals of some of the reviewed works, a time-window based framework for performance assessment is proposed. The r_j value indicates for the jth record if it is a critical record or not. R_j stands for the sum of critical records within the time lapse corresponding to the last d records. Low performance periods are defined when R_j exceeds the threshold R_{thr}.

2.1.1. Previous process

The efficacy of the model depends on both the quality of the collected data and the goodness with which the selected KPI represents the real performance of the system. Therefore, it is necessary to carry out a prior process which has been also represented in Fig. 1, oriented to the identification of the appropriate KPIs for the specific application environment, as well as to the preprocessing of the historical data.

2.1.2. First module: Performance characterization

The first module is aimed at describing a range of minimum acceptable values for the selected KPI by means of descriptive statistics, taking into account the conditions in which the system is operating. For such purpose, the subset of explanatory variables $V \subset X$ is defined, and then, the dimensionality of V is reduced to obtain C. The operating regimes G are defined based on the information of C and one of the operating regimes is assigned to each record. After these three steps, two parallel processes are carried out as shown in Fig. 1. On the one hand, and based on the values of the KPI for the records of the different operating regimes, a set of minimum acceptable values $\vec{Y}_{thr}$ is defined for the KPI.The criteria for the definition of the minimum acceptable values depends on the specific case and the type of KPI. On the other, the transition sequence is evaluated and the probability of operating in a current regime is defined taking into account the operating regime of the previous record.

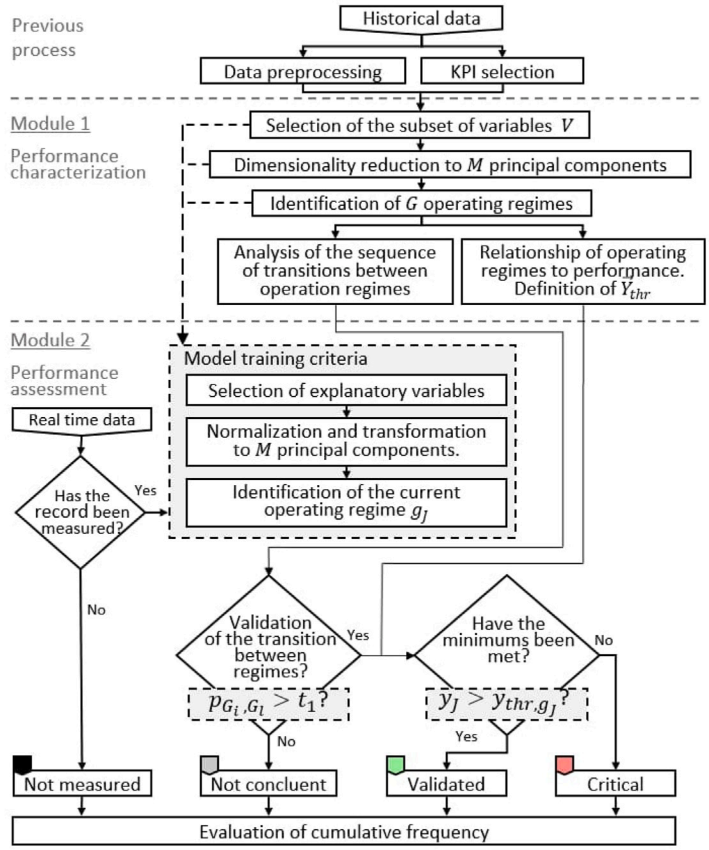


Fig. 1. Process of the proposed methodology.

2.1.3. Second module: Performance assessment

The second module is oriented towards performance assessment. For each new register x_j, the subset of measurements v_j corresponding to explanatory variables are selected, normalized and transformed to obtain c_j. Based on these values, the operating regime g_j is assigned. As it might be seen in Fig. 1, the model training criteria is used to perform these three steps. The register is validated or not based on the likelihood to change from the previous operating regime G_i and $G_l = g_j$. If the new register is accepted, then its measure for the KPI is compared against the minimum acceptable value for the current operating regime y_{thr,g_j}. Both the likelihood to switch from previous to current operation regime $p^j_{i,l}$ and the minimum acceptable value for the assigned operating regime are previously calculated as shown in Fig. 1. As a result of the process, the new register is classified as either not measured, not conclusive, validated or critical.

The proposed methodology is not only oriented to the classification of instants and detection of critical values, but to the evaluation of their cumulative frequency. For each record, as it may be seen in Eq. (1), R_j stands for the cumulative frequency of critical records within the time lapse defined by the d previous records of the corresponding time interval. If R_j exceeds R_{thr}, then a critical period of low performance is detected.

$$R_j = \sum_{z=j-d}^{j} r_z \tag{1}$$

where

$$r_z = \begin{cases} 1, & \text{if } y_j < y_{thr,g_j} \\ 0, & \text{if } y_j \geq y_{thr,g_j} \end{cases} \tag{2}$$

It is important to note that a good balance on the definition of d and R_{thr} for the time window framework allows avoiding the effect of punctual performance losses on the identification of false positives, as well as problems related to wrong or missing values. On the one hand, time windows with small d are more likely to obtain more false positives, whereas too large d can lead to less faithful representations of the current performance. On the other hand, lower R_{thr} can drive to false positives, and higher ones to false negatives. The balance between these two values needs to be assessed based on the specific considerations of the case study, and the performance of the rest of the methods used in previous steps of the methodology.

2.2. Methods for the implementation of the methodology

Next, the methods that have been used in this research work to implement the methodology will be described. PCA and K-means are well known unsupervised methods that have been used in literature for feature selection, dimensionality reduction and clustering purposes. They were considered appropriate for the implementation of the methodology in the case study as they provide simple and straightforward results.

2.2.1. Principal component analysis

Dealing with high volumes of multivariate data makes it difficult to obtain and interpret information. Dimensionality reduction techniques, such as PCA, facilitate this task. By means of this technique, an original set of variables is reduced to a set of principal components defined by linear combinations. Let the original variables be $V = (\vec{V}_1, \vec{V}_2, \dots \vec{V}_P)$, where $\vec{V}_p$ stands for the pth variable, being $p = (1, 2, \dots, P)$, and v_j the jth record, where $j = (1, 2, \dots, J)$. This way, the measure of the jth record for the pth variable is defined as $v_{j,p}$. Let the first principal component be defined by the linear combination in Eq. (3), where $\vec{B}_1 = (\beta_{1,1}, \beta_{1,2}, \dots, \beta_{1,P})$ stands for the set of weights of the first principal component, and $\beta_{1,p}$ is the weight for the pth variable in the first principal component.

$$\sum_{p=1}^{P} \vec{V}_p \beta_{1,p} = V\vec{B}_1 \quad \forall\, (\vec{B}_1 = \beta_{1,1}, \beta_{1,2}, \dots, \beta_{1,P}) \tag{3}$$

It is crucial that, when dimensionality reduction is done, most of the information of the original variables is captured. Variance is the characteristic of the data which provides the information, so in order to minimize the loss of information, the variance of the principal components calculated must be maximized. To this aim, the optimization of the function $var(V\vec{B}_1) = \vec{B}_1' S \vec{B}_1$ is posed, where S represents the covariance matrix of V. The constraint given by $\vec{B}_1\vec{B}_1' = 1$ is also set to avoid the increase in the variance as a consequence of the increase in the coefficients of the linear combination. The problem is solved using the Lagrange multiplier method. Eq. (4) is optimized deriving for B_1 and equaling 0, from which, applying the Roché-Frobenius theorem, Eq. (5) is obtained.

$$L(B_1) = V\vec{B}_1 - \lambda_1(\vec{B}_1\vec{B}_1' - 1) \tag{4}$$

$$\begin{aligned} (S - \lambda_1 I)\vec{B}_1 &= 0 \\ |S - \lambda_1 I| &= 0 \end{aligned} \tag{5}$$

Therefore, λ_1 is the eigenvalue for S, and its eigenvector is obtained applying Eq. (5). Note that Eq. (5) stands the same when values multiply by -1, so the sings of the coefficients of the eigenvectors, or weights of the variables in the linear combinations, are arbitrary.

From Eq. (5), $S\vec{B}_1 = \lambda_1 I \vec{B}_1$ can be deduced, and therefore, as shown in Eq. (6), the variance of each principal component is given by the eigenvalue of the corresponding linear combination.

$$var(V\vec{B}_1) = \vec{B}_1' S \vec{B}_1 = \vec{B}_1' \lambda_1 \vec{B}_1 = \lambda_1 \tag{6}$$

The next eigenvalues and eigenvectors are set by adding constraints related to the linear independence among previous eigenvectors. For a set of data with P variables, with a S_p matrix of covariance, P eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_P)$ are obtained, with their corresponding eigenvectors $B = (\vec{B}_1, \vec{B}_2, \dots, \vec{B}_P)$.

At this point, it is important to consider the trace of the covariance matrix $tr(S)$ from two different points of view. On the one hand, as the diagonal values of the covariance matrix are the variances of the original variables, $tr(S) = \sum_{p=1}^{P} var(\vec{V}_p)$. On the other hand, by applying the spectral decomposition to the covariance matrix, Eq. (7) is posed, where Λ is a diagonal matrix composed by the eigenvalues of the covariance matrix and B is the orthonormal matrix composed by the eigenvectors of S_p. From these two statements, it can be deduced that the sum of the variances of the variables in the original data V equals the sum of the eigenvalues of S_p and the sum of the variances of the principal components, as shown in Eq. (8). Further explanation about this method and the mathematical properties can be found in the work of Jolliffe [43].

$$tr(S) = tr(B\Lambda B') = tr(B'B\Lambda) = \sum_{p=1}^{p} \lambda_p \tag{7}$$

$$\sum_{p=1}^{p} \lambda_p = \sum_{p=1}^{P} var(\vec{V}_p) = \sum_{p=1}^{P} var(V\vec{B}_1) \tag{8}$$

Therefore, the percentage of information of original data captured by each eigenvector can be measured by the relative value of the eigenvalue. Eq. (9) gives the percentage variance captured by the mth principal component.

$$\frac{\lambda_m}{\sum_{p=1}^{P} \lambda_p} \cdot 100 \tag{9}$$

For dimensionality reduction the principal components defined by the eigenvectors corresponding to the highest eigenvalues are selected. Ordering the eigenvalues decreasingly allows to calculate the amount of information captured by the first M principal components by Eq. (10). The M number of principal components to be kept in further analysis is defined so that an acceptable percentage of variance of the original data is preserved. The new set of data $C = (\vec{C}_1, \vec{C}_2, \dots \vec{C}_M)$ is obtained by Eq. (11).

$$\frac{\sum_{p=1}^{M} \lambda_p}{\sum_{p=1}^{P} \lambda_p} \cdot 100 \tag{10}$$

$$\vec{C}_m = V\vec{B}_m \quad \forall m \in (1, M) \tag{11}$$

Prior to the application of PCA, normalizing data is highly recommended in order to avoid the effect of the difference among magnitudes. Once PCA is applied, the coefficients of each variable in the eigenvectors indicate their weight in the principal components. By means of biplots it is possible to represent this information in a easy-to-interpret way.

2.2.2. K-means clustering

K-means is a clustering method belonging to unsupervised ML. It is generally accepted and easy to implement and understand, which widespreads its use in the literature. Morissette and Chartier [44] explain the method and present the most common algorithms. K-means method consists on assigning one of a known number of clusters to each observation by iterative relocation of centroids. The objective is to group similar observations within each cluster and set clusters of different characteristics regarding the data variables.

The method takes as input the data set and the number of clusters. Let the set of data be $C = (\vec{C}_1, \vec{C}_2, \dots \vec{C}_M)$ where $\vec{C}_m$ stands for the mth variable and $(c_1, c_2, \dots, c_J)$ for the records measured in $j = (1, 2, \dots, J)$ instants, being $c_{j,m}$ the measure of the mth variable for the jth record. Based on the similarities with regard to the set of variables C, the J records are clustered in K number of groups $G = \{G_1, G_2, \dots, G_K\}$. For this purpose, a set of centroids $\Theta = (\vec{\Theta}_1, \vec{\Theta}_2, \dots \vec{\Theta}_K)$ is located randomly, and each one is linked to a group of G. This way, $\vec{\Theta}_k$ represents the kth centroid and it is defined by a set of measures $\vec{\Theta}_k = (\theta_{k,1}, \theta_{k,2}, \dots \theta_{k,M})$ related to the variables. Then, one of the groups is assigned to each record, and the centers are calculated again. This process is repeated until shutdown criteria related to number of iterations, number of reassigned cases or improvement in the last iterations are met.

2.2.3. Markov chain

Discrete time Markov chains are used for representing stochastic processes in which the likelihood that certain event g_j happens at time j depends on the previous one g_{j-1}. Let a set of events be defined by a discrete variable with K states $G = \{G_1, G_2, \dots, G_K\}$, so the likelihood in the time j to move from $g_{j-1} = G_i$ to $g_j = G_l$ is given by Eq. (12).

$$p_{i,l}^{(j)} = P(g_j = G_l | g_{j-1} = G_i) \quad \forall\, i, l = (1, 2, \dots, K) \tag{12}$$

Time homogeneous Markov chains pose the specific case in which this likelihood remains constant in time $p_{i,l}^{(j-1)} = p_{i,l}^{(j)} \quad \forall\, j \in (1, J)$. By calculating this probability for all the combinations of i and l the transition matrix T_K shown in Eq. (13) is obtained, in which each $p_{i,l}$ stands for the probability to move from the ith state to the lth. Note that, as each time a state is leaved another one arises, the sum of the probabilities in each row is equal to 1, which means $\sum_{l=1}^{K} p_{i,l} = 1 \quad \forall i \in \{1, K\}$.

$$T_K = \begin{bmatrix} p_{1,1} & \cdots & p_{1,K} \\ \vdots & \ddots & \vdots \\ p_{K,1} & \cdots & p_{K,K} \end{bmatrix} \tag{13}$$

Taking into account the initial probability row vector $\vec{P}^{(0)} = (p_1^{(0)} \dots p_K^{(0)})$ and the transition matrix T_K, the likelihood to achieve each of the G_l states at time $j = 1$ can be calculated by $\vec{P}^{(1)} = \vec{P}^{(0)} \cdot T_K$. The generalization for any jth time is given by Eq. (14).

$$\vec{P}^{(j)} = \vec{P}^{(j-1)} \cdot T_K = \vec{P}^{(j-2)} \cdot T_K^2 = \dots = \vec{P}^{(0)} \cdot T_K^j \tag{14}$$

Table 1
Positioning of the case study according to criteria in [3].

Criteria	Description	Present research positioning
System configuration	The type of wind power asset and the level of system modeling	3 MW WT at WT level.
Decision-making attribute	Planning horizon, the decision-maker and the availability of field data	Finite time horizon of 3 years considering time as discrete-time states of 10 min.
		The decision-maker are considered to be the wind farm owners and operators or independent service providers. Supplementary and production data through operational SCADA data.
System failure modeling	Include the type of damage/failure and the failure modeling approach	Minor failures are considered with degradation model from statistical inference.
Optimization model	Optimality criterion and the solution technique	Out of scope.
Maintenance strategy	The maintenance policy and the effectiveness of the repair actions	Out of scope.

3. Case study

This section presents and describes the case study in which the methodology is implemented. In order to introduce it, it was classified according to the framework proposed by Shafiee and Sørensen [3] which presented several criteria to classify case studies related to the wind energy sector. The features of the case study corresponding to each of the criteria defined by Shafiee and Sørensen [3] can be found in the Table 1.

The case study is related to the wind energy sector and aims for performance loss detection in WTs. As the modeling relies on supplementary and production data containing operational variables measured by SCADA systems, having access to this data is a necessary feature of the decision makers. Consequently, wind farm owners and operators are considered to be the decision makers, even if independent service providers might also be considered in case SCADA data were available. A planning horizon length of three years was considered for the case study, with discrete-time states of 10 min. Regarding the system failure modeling, minor failures causing performance loss are detected in the case study through degradation models based on statistical inference.

This proposal would allow the implementation of performance based maintenance policies with minimum production loss optimization criterion. However, the case study is focused on the previous phase of modeling the system behavior. Therefore, as the optimization framework is out of the scope, so is the definition of maintenance policies and effectiveness (see Table 1).

The details about the development of the case study are grouped, and those regarding to data description, followed strategies and anomaly detection are given in the following subsections, as well as the results of the case study. Finally, a display is proposed for the visualization of the results of the performance evaluation.

3.1. Data description

In the case study, real data about eight 3 MW WTs was available. It consisted on ten-minute operational data recorded by SCADA systems during more than two years. No information about availability or alarm system data was available, as well as other interesting information regarding installation aging or maintenance background. A set of 23 variables $X = (\vec{X}_1, \vec{X}_2, \dots \vec{X}_2 3)$ about operational data corresponding to external or internal variables were measured, but their meaning was not taken into account in order to maintain the holistic point of view of the methodology and avoid hasty conclusions. Table 2 provides a descriptive summary of the quantitative variables composing the data. Another variable corresponding to the time stamp of each record was available, with format $yyyy - mm - dd\ hh : mm : ss$. The variable indicating the power generation within the last 10 min was selected as target KPI $\vec{Y}$. A subset of 8 explanatory variables V was defined attending to correlations among variables and power output, as well as their weights in the principal components. From each high-correlated group of variables of X, one of the variables was selected as representative of the group, and the representative variables with higher weights in the two first principal components were selected to be included in V.

Data preprocessing was principally oriented towards solving format problems, ordering data temporally and adding empty instances for missing values. Negative power outputs were reset to 0 and some anomalous measures were identified and removed, but in depth data cleaning was difficult because labeling the data was not possible. Data was split, and the first year of data composed by more than 52 000 records was set as historical data. This approach ensures that the possible seasonality-related differences in the behavior of the systems will not affect the model training and posterior anomaly detection [45]. The remaining records, which were around 60 000, were used for performance evaluation.

3.2. Training strategies

Once the subset of explanatory variables was selected, PCA was applied in order to reduce the dimensionality of V. The weights of the explanatory variables for the first principal components were similar in some of the WTs, and identifying these similarities led to the definition of three strategies for the implementation of the methodology: training by WT, training by group and unique training.

1. **Training by WT** consist in characterizing the performance of each WT and using this information as a reference for its performance assessment.
2. **Training by group** consist in using data of all WTs within a group for performance characterization, which is taken into account for performance assessment of each WT in the group.
3. **Unique training** consist in characterizing the performance based on the whole set of WTs and assessing the performance of each WT.

The methodology has been implemented through the three strategies in order evaluate which one provides most accurate results. Note that, even if the size of the training data sets employed for performance characterization varies, the explanatory variables comprising V are the same.

3.3. Approach for anomaly detection in WTs

Despite the different strategies being used for the training of the models, the process for anomaly detection remains the same. The training records were normalized to avoid the effect of the difference among magnitudes when PCA was applied. The amount of principal components kept was set in two ($M = 2$) based on the results obtained with this configuration of the parameter in one of the WTs. Other advantages related to the ease to represent results were also considered. The operating regimes were defined by the K-means method, using the Hartigan and Wong [46] algorithm and the euclidean distance. The elbow method showed that a number of $K = 5$ operating regimes

Table 2
Description of the quantitative variables composing the data.

Variable	Min.	1stQu.	Median	Mean	3rdQu.	Max.
w	−69.00	0.00	284.00	683.20	1027.00	3060.00
$\vec{x}_1$	0.00	2906.00	5124.00	5636.00	7790.00	41 385.00
$\vec{x}_2$	0.00	4072.00	5810.00	6625.00	8457.00	41 385.00
$\vec{x}_3$	−195.60	33.00	1188.50	1027.50	1693.10	1803.80
$\vec{x}_4$	0.00	0.00	9390.00	8086.00	13 311.00	14 214.00
$\vec{x}_5$	0.00	127.70	226.60	204.90	282.70	360.00
$\vec{x}_6$	0.00	41.55	45.51	42.30	49.74	850.00
$\vec{x}_7$	0.00	44.28	47.87	45.42	52.38	850.00
$\vec{x}_8$	0.00	57.65	62.04	61.82	68.15	129.43
$\vec{x}_9$	0.00	57.63	62.19	61.95	68.41	129.48
$\vec{x}_{10}$	0.00	57.52	61.97	61.70	68.05	130.28
$\vec{x}_{11}$	0.00	40.06	46.03	43.63	51.14	67.45
$\vec{x}_{12}$	0.00	38.57	43.56	41.29	47.94	66.05
$\vec{x}_{13}$	0	43.19	51.47	48.28	57.54	586.86
$\vec{x}_{14}$	0	43.16	51.26	48.6	59.24	155.24
$\vec{x}_{15}$	0	41.87	49.14	46.43	56	584.34
$\vec{x}_{16}$	−128.8	41.72	48.19	45.29	53.49	784.61
$\vec{x}_{17}$	0	41.29	47.71	44.81	52.99	511.96
$\vec{x}_{18}$	−7.739	41.135	46.596	44.055	51.672	484.959
$\vec{x}_{19}$	−200.95	0	0	24.4	85.93	180.28
$\vec{x}_{20}$	−184.87	0	0	24.41	85.92	217.26
$\vec{x}_{21}$	−295.25	0	0	24.39	85.92	275.05

defined as $G = \{A, B, C, D, E\}$ was appropriate in all the WTs of the individual training, all the groups of the group training and the unique training.

Transition matrices were calculated for each training by analyzing the transition sequence among the defined operating regimes. The threshold to validate transitions was established in $t_1 = 0.1$, which means that records corresponding to transitions with less than such probability to occur will be considered not concluent.

The thresholds $\vec{Y}_{thr}$ have been set based on the values for the KPI of each operating regime. Let $\vec{W}_k$ be the vector of length E_k of unique values of the power measures corresponding to the kth operating regime $\vec{W}_k = (w_{k,1}, w_{k,2}, \dots, w_{k,E_k})$. Another vector $\vec{f}_k$ stands for the frequencies of each of the unique values of the kth regime $\vec{f}_k = (f_{k,1}, f_{k,2}, \dots, f_{k,E_k})$. Therefore, a new vector of percentages of power of the group provided by each unique value $\vec{h}_k = (h_{k,1}, h_{k,2}, \dots, h_{k,E_k})$ can be calculated as shown in Eq. (15). Finally, the unique values are sorted increasingly and the cumulative percentage of power provided by the first q values is calculated by Eq. (16), from which the vector of cumulative percentage of power provided to the group $\vec{H}_k = (H_{k,1}, H_{k,2}, \dots, H_{k,E_k})$ is obtained. In the case study, the unique value of power corresponding to the $H_{k,q} = 0.02$ is set as threshold $y_{thr,k}$, which means that the records assigned to the jth operating regime corresponding to values that do not achieve this threshold will be classified as critical. Prior to this process, the distribution of unique values of power of each group is analyzed, and outliers are rejected by statistical criteria in an iterative process. The results provided by the methodology with $H_{k,q} = 0.02$, $H_{k,q} = 0.03$, $H_{k,q} = 0.05$ and $H_{k,q} = 0.1$ were compared, and $H_{k,q} = 0.02$ gave the best results attending to the amount of alarms detected in each WT and the capacity to detect differences among WTs with less overfitting than other thresholds.

$$h_{k,e} = \frac{w_{k,e} f_{k,e}}{\sum_{e=1}^{E_k} w_{k,e} f_{k,e}} \quad \forall\, k \in \{1, 2, \dots, K\} \tag{15}$$

$$H_{k,q} = \frac{\sum_{e=1}^{q} w_{k,e} f_{k,e}}{\sum_{e=1}^{E_k} w_{k,e} f_{k,e}} \quad \forall\, k \in \{1, 2, \dots, K\} \tag{16}$$

The performance assessment is carried out individually although different strategies are used for training. In the case study, 59 840 records were used as new records. From each new record x_{J+1}, the measures corresponding to the subset of variables V was selected and normalized taking into account the centers an deviations of the performance characterization phase. Then the values are transformed to data in the principal components PC1 and PC2 to obtain c_{J+1}, by multiplying the normalized values of the new register by the eigenvectors corresponding to each principal component, as may be seen in Eq. (11). Then, the operating regime whose centroid is the nearest to those measures is assigned to the new record as set by Eq. (17).

$$g_j = \min_{k \in (1,K)} \|c_j - \vec{\theta}_k\| = \min_{k \in (1,K)} \sqrt{\sum_{m=1}^{M} \left(c_{j,m} - \theta_{k,m}\right)^2} \tag{17}$$

As it may be seen in Fig. 1, once the cluster is assigned for the new record, if the likelihood to change from $G_i = g_{(j-1)}$ to $G_l = g_j$ exceeds the threshold $t_1 = 0.1$, the transition is accepted. Finally, the value obtained for the power of the new record y_j is compared to the minimum threshold for the assigned group $y_{thr\,G_j}$. The records with power below their corresponding thresholds will be considered critical, and the remaining ones validated.

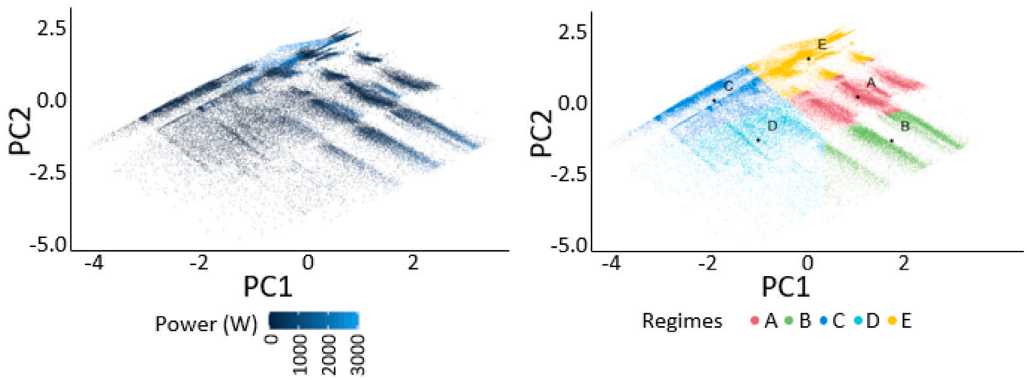
3.4. Results

The amount of alarms and their duration are evaluated for the three strategies in order to assess which one performs best. These indicators depend on the definition of thresholds for each regime, which depend on the definition of operating regimes. Two types of scenarios are obtained regarding performance characterization processes. In some of them, the percentage of variance captured by the principal components is higher than in others, which directly affects the distribution of records on the principal components.

As it might be seen in Fig. 2, in the cases with lower percentage of variance, the records with higher power are not split from others (see Fig. 2(a)), and they are classified in the same regimes (see Fig. 2(b)). Therefore, wider and more similar distributions are obtained for the power values of the regimes in these scenarios, and the defined thresholds are lower and more similar among different regimes on each training of the model as shown in Table 3.

In the opposite scenario shown in Fig. 3, a higher percentage of variance is captured by the principal components, leading to a better segregation of records with different power output (see Fig. 3(a)) and larger power differences among different operating regimes (see Fig. 3(b)). As the power distributions differ more for the distinct operating regimes in these scenarios, so they do the thresholds obtained for each regime (see Table 3).

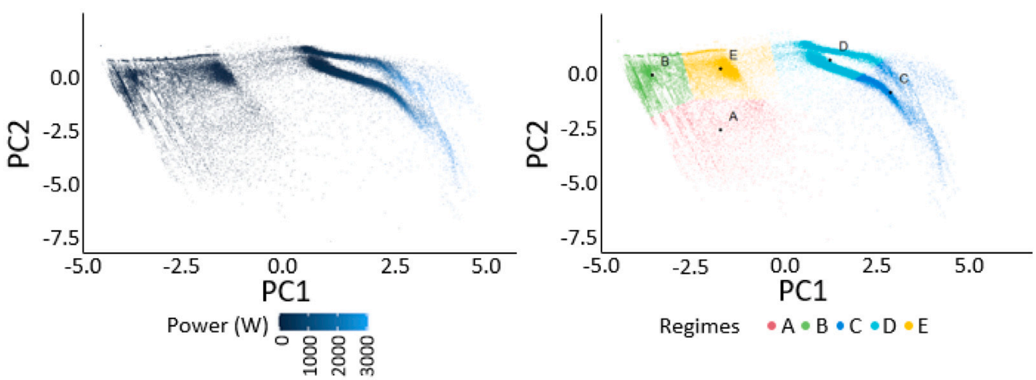
As it may be seen in Table 3, this occurred for the three strategies. Besides, the transition sequence analysis proved that it is more likely to switch operating regimes in the cases with less percentage of variance captured.



(a) Distribution of records in principal components.

(b) Definition of operating regimes.

Fig. 2. Scenario with low percentage of variance.



(a) Distribution of records in principal components.

(b) Definition of operating regimes.

Fig. 3. Scenario with high percentage of variance.

Table 3
Percentage of variance captured by principal components and thresholds for each training.

	% of variance	$y_{thr,1}$ regime A	$y_{thr,2}$ regime B	$y_{thr,3}$ regime C	$y_{thr,4}$ regime D	$y_{thr,5}$ regime E
WT 1	34%	70 W	404 W	0 W	0 W	169 W
WT 2	60%	1449 W	0 W	0 W	458 W	59 W
WT 3	56%	0 W	0 W	715 W	81 W	0 W
WT 4	58%	0 W	1224 W	0 W	99 W	0 W
WT 5	34%	342 W	0 W	220 W	0 W	69 W
WT 6	34%	0 W	0 W	191 W	407 W	80 W
WT 7	35%	0 W	207 W	0 W	78 W	370 W
WT 8	60%	70 W	571 W	921 W	0 W	0 W
Group 1	59%	0 W	0 W	81 W	765 W	0 W
Group 2	33%	0 W	170 W	100 W	0 W	368 W
Unique	39%	234 W	0 W	118 W	0 W	76 W

It is expected that the distributions for the power values for each operating regime will be different, since these operating regimes were set according to explanatory variables that condition the power output. Moreover, it is expected that these operating regimes are stable, because records were measured within ten-minute intervals in which it is not expected to obtain great differences regarding the set of explanatory variables. According to the obtained results, the percentage of variance captured by the principal components has more effect on the efficacy of the implementation of the methodology than the employed strategy. In the cases with low percentage of variance captured, the clusters obtained from the combination of PCA and K-means do not match the real operating regimes, while in the rest, a better definition of operating regimes was possible. In the case study, the 60% of the information of the subset of explanatory variables was enough to enable the operating regime identification. It is a remarkable result since several of the initial 23 variables were rejected and the dimensionality of the eight explanatory ones was reduced to only two principal components.

However, a series of limitations were identified regarding the case study. Some of them are related to the problem statement through the implementation of PCA and Markov chains. The PCA poses the hypothesis of lineality among variables, and the use of Markov chains assumes that the likelihood to operate in a certain regime depends only on the previous one. Besides, as homogeneous Markov chains were used, it is assumed that the likelihood of switching operating regimes remains constant over time. This is a high simplification of the problem, in which external and internal variables of WTs are involved. Some other limitations were related to the characteristics of the case study, which were mainly related to the lack of data and information. As only

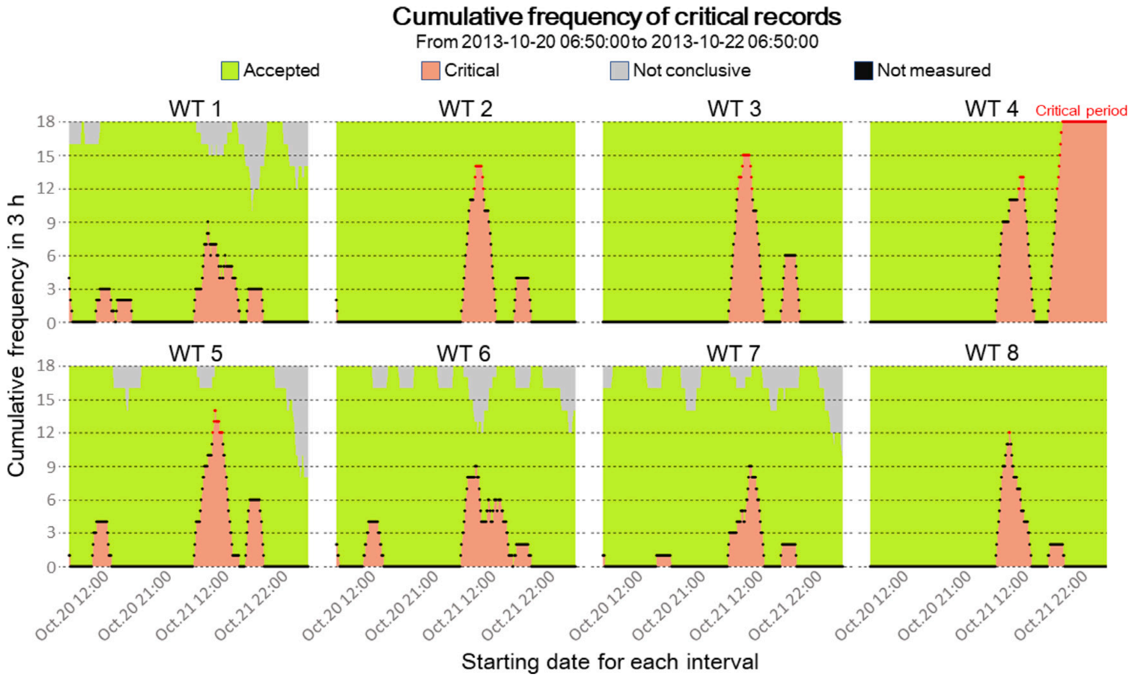


Fig. 4. Example of the proposed display.

operational data was available, labeling the data was no possible, and there was no option to validate the obtained results.

3.5. From model results to decision making

In the case study, the frequency of critical values is evaluated in a time window of $d = 18$ records, which corresponds to 3 h. The threshold for the cumulative frequency of critical values is set in $R_{thr} = 12$, which means that if 12 of the last 18 records are classified as critical, then a critical period is detected, and an alarm is activated.

In order to support the maintenance decision making, a display panel is proposed to visualize the information provided by the results of the methodology. For each WT, the results obtained for the instants within the last 48 h are shown in Fig. 4. The cumulative frequency of critical, accepted, not conclusive or not measured values for the three hours prior to those instants is shown. An alarm message is also shown when the threshold is met.

This display eases the decision making process through easy-to-interpret visualizations of the performance of the whole WT system during the last 48 h. Summarizing and simplifying the obtained results is a key factor which can help to introduce this type of analysis on industry, since it allows the reduction of reaction times for decision making.

4. Concluding remarks

In this paper, a methodology focused on WT performance assessment at system level is proposed, which contributes to the growing field of performance based maintenance by providing a new approach based on condition monitoring from the perspective of non binary failures. The methodology was implemented in a case study in which two years of real SCADA data about eight WTs were used aiming to characterize and assess the performance of the assets. Three different strategies were employed, and performance characterization was performed considering historical data for each turbine separately in the first one, groups of turbines in the second one and the complete set of turbines in the third one. However, the results showed that the characteristics of the data had more effect on the goodness of performance characterization than the followed strategy in the case study.

The selection of explanatory variables and the reduction of the dimensionality of the data bases was performed by means of PCA. Combining correlation analysis and PCA proved to be a simple but effective method while selecting explanatory variables. In the case study, the amount of variables introduced to the model reduced to one third of the original amount, which also impacts in the efficiency of the model in computational terms. This method also gave good results on the posterior reduction of the dimensionality, which proved to be a crucial part of the methodology, as better defined operating regimes were obtained when the ability of the principal components to explain the data variability increased. K-means was used for the identification of operating regimes, but the effectiveness of this method was conditioned by the previous step. Results showed that the performance of the PCA and K-means for the identification of operating regimes had more effect than the strategy used on the accuracy of the results. The implementation of the first module from a strict data based approach provides an alternative for the cases in which there is not enough information about the variables or there is less know-how about the application field.

The holistic point of view of the methodology and the combination of tools and methods employed in the case study eases the interpretation of what each result implies. The time window framework proposed for the second module provides an easy to interpret summary of all the obtained results, which, combined with the proposed display for the visualization of these results, eases the maintenance decision making. Since no in-deep knowledge about the model is required to interpret the results, the display allows bringing the methodology closer to industrial environments as a simple tool for WT performance monitoring.

The case study was generally limited by the information available for the research. The existence of alarm data or maintenance logs would have allowed the validation of the results by checking which of the identified anomalies really correspond to WT performance losses prior to failures. In such cases, it would have been possible to analyze how maintenance policies based on the proposed methodology could enable the anticipation of maintenance activities to these failures. Moreover, accessing more information about the WTs would give some insights into the causes of the differences among WTs, leading to the definition of new strategies based on industrial and empirical knowledge of the implementation of the methodology. However, despite the limitations caused by the lack of information, the implementation of the methodology in the case study proved its capability to identify operating regimes. Identifying operating regimes allows the definition of well studied performance based maintenance policies as it integrates information about the operational context of the WTs from a more realistic and data-based perspective.

From this point of view, this case study is a previous step to greater ones. Building maintenance policies based on the implementation of the methodology and analyzing their performance would be interesting, as well as comparing it with the results provided by other existing policies. To this aim, it is needed more information about the WTs, their historical information and the cost structures. Accessing to this information would also allow the definition of more consistent strategies and the validation of the results of the new case studies.

From the data analytic perspective, some other techniques and methods could be considered for the identification of operating regimes and transition validation in further case studies, in order to avoid the assumptions of the one presented here. The methods employed for the newer case studies depend on their specific characteristics and must give logical solutions to the different challenges faced during the implementation of the methodology.

This methodology is thought to be implemented in offshore wind farms, and the results obtained in the onshore case study of this paper gives hope that the methodology can also be implemented offshore cases, as onshore scenarios may be more challenging. Besides, it would be interesting to evaluate how this methodology can be implemented in other sectors. Market competitors could benefit from the application of performance based maintenance policies on systems with high energy consumption, obtaining more efficient processes. Besides, it might also be useful in processes in which many factors affect the performance of the systems, even if in such cases there is not much knowledge about which of the factors have more effect. However, it must be reasonable to make the assumption that those systems operate under different conditions which will characterize the operating regimes.

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.

Data availability

The data that has been used is confidential.

Acknowledgments

This study was partially funded by the Spanish Agencia Estatal de Investigación (AEI) — Ministerio de Ciencia e Innovación through the research project PID2021-122132OB-C22.

References

[1] A. Van Horenbeek, L. Pintelon, Development of a maintenance performance measurement framework-using the analytic network process (ANP) for maintenance performance indicator selection, Omega (United Kingdom) 42 (1) (2014) 33–46.
[2] K. Fraser, Facilities management: the strategic selection of a maintenance system, J. Facil. Manage. 12 (1) (2014) 18–37.
[3] M. Shafiee, J.D. Sørensen, Maintenance optimization and inspection planning of wind energy assets: Models, methods and strategies, Reliab. Eng. Syst. Saf. 192 (2019) 105993.
[4] IRENA, Future of Wind: Deployment, Investment, Technology, Grid Integration and Socio-Economic Aspects, International Renewable Energy Agency, 2019, pp. 1–88.
[5] J. Martin, D. Ramsey, The economics of wind energy, J. Appl. Corp. Finance 21 (2) (2009) 100–109.
[6] A. Mérigaud, J.V. Ringwood, Condition-based maintenance methods for marine renewable energy, Renew. Sustain. Energy Rev. 66 (2016) 53–78, http://dx.doi.org/10.1016/j.rser.2016.07.071, URL https://www.sciencedirect.com/science/article/pii/S1364032116304099.
[7] P. Sandborn, G. Haddad, A. Kashani-Pour, X. Lei, Development of a maintenance option model to optimize offshore wind farm sustainment, in: 32nd ASME Wind Energy Symposium, American Institute of Aeronautics and Astronautics, Inc., 2014, pp. 1–7.
[8] J. Izquierdo, A.C. Márquez, J. Uribetxebarria, A. Erguido, On the importance of assessing the operational context impact on maintenance management for life cycle cost of wind energy projects, Renew. Energy 153 (2020) 1100–1110, http://dx.doi.org/10.1016/j.renene.2020.02.048, URL https://www.sciencedirect.com/science/article/pii/S0960148120302354.
[9] F.P.G. Márquez, M. Papaelias, Chapter 3 - An overview of wind turbine maintenance management, in: Non-Destructive Testing and Condition Monitoring Techniques for Renewable Energy Industrial Assets, Butterworth-Heinemann, 2020, pp. 31–47.
[10] UNE EN 13306, E.C. for Standardization, Maintenance - maintenance terminology, in: UNE EN 13306, UNE, 2017, p. 98.
[11] R. Ahmad, S. Kamaruddin, An overview of time-based and condition-based maintenance in industrial application, Comput. Ind. Eng. 63 (1) (2012) 135–149.
[12] V.J. Jimenez, N. Bouhmala, A.H. Gausdal, Developing a predictive maintenance model for vessel machinery, J. Ocean Eng. Sci. 5 (4) (2020) 358–386.
[13] H.P. Bloch, F.K. Geitner, Practical Machinery Management for Process Plants. Volume 2: Machinery Failure Analysis and Troubleshooting, Exxon Chemical Co., 1983.
[14] O. Merkt, On the use of predictive models for improving the quality of industrial maintenance: An analytical literature review of maintenance strategies, in: Federated Conference on Computer Science and Information Systems, FedCSIS 2019, Vol. 18, Polish Information Processing Society – as since 2011, 2019, pp. 693–704.
[15] M. Tahan, E. Tsoutsanis, M. Muhammad, Z.A. Abdul Karim, Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review, Appl. Energy 198 (2017) 122–144.
[16] U. Leturiondo, O. Salgado, L. Ciani, D. Galar, M. Catelani, Architecture for hybrid modelling and its application to diagnosis and prognosis with missing data, Measurement 108 (2017) 152–162, http://dx.doi.org/10.1016/j.measurement.2017.02.003, URL https://www.sciencedirect.com/science/article/pii/S0263224117300908.
[17] M. Morshedizadeh, M. Kordestani, R. Carriveau, D.S.-K. Ting, M. Saif, Power production prediction of wind turbines using a fusion of MLP and ANFIS networks, IET Renew. Power Gener. 12 (9) (2018) 1025–1033.
[18] E. Lapira, D. Brisset, H. Davari Ardakani, D. Siegel, J. Lee, Wind turbine performance assessment using multi-regime modeling approach, Renew. Energy 45 (2012) 86–95.
[19] E. Gonzalez, E.M. Nanos, H. Seyr, L. Valldecabres, N.Y. Yürüşen, U. Smolka, M. Muskulus, J.J. Melero, Key performance indicators for wind farm operation and maintenance, Energy Procedia 137 (2017) 559–570.
[20] E. Gonzalez, B. Stephen, D. Infield, J.J. Melero, On the use of high-frequency SCADA data for improved wind turbine performance monitoring, J. Phys. Conf. Ser. 926 (1) (2017) 12009.
[21] A. Kusiak, Z. Zhang, A. Verma, Prediction, operations, and condition monitoring in wind energy, Energy 60 (2013) 1–12.
[22] D. Sandoval, U. Leturiondo, F. Pozo, Y. Vidal, Low-speed bearing fault diagnosis based on permutation and spectral entropy measures, Appl. Sci. 10 (13) (2020) 4666, http://dx.doi.org/10.3390/app10134666, URL https://www.mdpi.com/2076-3417/10/13/4666.
[23] P. Bangalore, L.B. Tjernberg, An artificial neural network approach for early fault detection of gearbox bearings, IEEE Trans. Smart Grid 6 (2) (2015) 980–987, http://dx.doi.org/10.1109/TSG.2014.2386305.
[24] Z. Gao, X. Liu, An overview on fault diagnosis, prognosis and resilient control for wind turbine systems, Processes 9 (2) (2021) 300, http://dx.doi.org/10.3390/pr9020300, URL https://www.mdpi.com/2227-9717/9/2/300.
[25] X. Liu, Z. Gao, M.Z.Q. Chen, Takagi–Sugeno Fuzzy model based fault estimation and signal compensation with application to wind turbines, IEEE Trans. Ind. Electron. 64 (7) (2017) 5678–5689, http://dx.doi.org/10.1109/TIE.2017.2677327.
[26] P. Tamilselvan, P. Wang, S. Sheng, J. Twomey, A two-stage diagnosis framework for wind turbine gearbox condition monitoring, Int. J. Progn. Health Manag. (2013) 21–31.
[27] J. Xiang, S. Watson, Y. Liu, Smart monitoring of wind turbines using neural networks, in: Sustainability in Energy and Buildings, Springer, 2009, pp. 1–8.
[28] R. Rahimilarki, Z. Gao, A. Zhang, R. Binns, Robust neural network fault estimation approach for nonlinear dynamic systems with applications to wind turbine systems, IEEE Trans. Ind. Inform. 15 (12) (2019) 6302–6312, http://dx.doi.org/10.1109/TII.2019.2893845.
[29] M.L. Wymore, J.E. Van Dam, H. Ceylan, D. Qiao, A survey of health monitoring systems for wind turbines, Renew. Sustain. Energy Rev. 52 (2015) 976–990.
[30] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, G. Nenadic, Machine learning methods for wind turbine condition monitoring : A review, Renew. Energy 133 (2019) 620–635.
[31] K. Leahy, C. Gallagher, P. O'Donovan, D.T.J. O'Sullivan, Issues with data quality for wind turbine condition monitoring and reliability analyses, Energies 12 (2) (2019) 201.

[32] P. Marti-Puig, A.M. Blanco, J.J. Cárdenas, J. Cusidó, J. Solé-Casals, Feature selection algorithms for wind turbine failure prediction, Energies 12 (3) (2019) 453.

[33] M. Schlechtingen, I.F. Santos, Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 2: Application examples, Appl. Soft Comput. 14 (2014) 447–460.

[34] N. Beganovic, D. Söffker, Structural health management utilization for lifetime prognosis and advanced control strategy deployment of wind turbines: An overview and outlook concerning actual methods, tools, and obtained results, Renew. Sustain. Energy Rev. 64 (2016) 68–83.

[35] K. Leahy, C. Gallagher, P. O'Donovan, K. Bruton, D.T.J. O'Sullivan, A robust prescriptive framework and performance metric for diagnosing and predicting wind turbine faults based on SCADA and alarms data with case study, Energies 11 (7) (2018) 1738.

[36] J. Izquierdo, A. Crespo Márquez, J. Uribetxebarria, Dynamic artificial neural network-based reliability considering operational context of assets, Reliab. Eng. Syst. Saf. 188 (2019) 483–493, http://dx.doi.org/10.1016/j.ress.2019.03.054, URL https://www.sciencedirect.com/science/article/pii/S0951832018311906.

[37] H. Chen, H. Liu, X. Chu, Q. Liu, D. Xue, Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network, Renew. Energy 172 (2021) 829–840, http://dx.doi.org/10.1016/j.renene.2021.03.078, URL https://www.sciencedirect.com/science/article/pii/S0960148121004341.

[38] Z. Song, Z. Zhang, Y. Jiang, J. Zhu, Wind turbine health state monitoring based on a Bayesian data-driven approach, Renew. Energy 125 (2018) 172–181, http://dx.doi.org/10.1016/j.renene.2018.02.096, URL https://www.sciencedirect.com/science/article/pii/S0960148118302404.

[39] P. Korkos, M. Linjama, J. Kleemola, A. Lehtovaara, Data annotation and feature extraction in fault detection in a wind turbine hydraulic pitch system, Renew. Energy 185 (2022) 692–703, http://dx.doi.org/10.1016/j.renene.2021.12.047, URL https://www.sciencedirect.com/science/article/pii/S0960148121017675.

[40] M. Lydia, S.S. Kumar, A.I. Selvakumar, G.E. Prem Kumar, A comprehensive review on wind turbine power curve modeling techniques, Renew. Sustain. Energy Rev. 30 (2014) 452–460.

[41] Q. Huang, Y. Cui, B. Tjernberg, Lina Bertling, Pramod, Wind turbine health assessment framework based on power analysis using machine learning method, in: 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), IEEE, Bucharest, Romania, 2019, pp. 1–5.

[42] Q. Xu, S. Lu, Z. Zhai, C. Jiang, Adaptive fault detection in wind turbine via RF and CUSUM, IET Renew. Power Gener. 14 (10) (2020) 1789–1796.

[43] I.T. Jolliffe, Mathematical and statistical properties of population principal components, in: Principal Component Analysis, Springer New York, New York, NY, 2002, pp. 10–28, http://dx.doi.org/10.1007/0-387-22440-8_2.

[44] L. Morissette, S. Chartier, The k-means clustering technique: General considerations and implementation in Mathematica, Tutor. Quant. Methods Psychol. 9 (1) (2013) 15–24.

[45] C. McKinnon, A. Turnbull, S. Koukoura, J. Carroll, A. McDonald, Effect of time history on normal behaviour modelling using SCADA data to predict wind turbine failures, Energies 13 (18) (2020) 4745, http://dx.doi.org/10.3390/en13184745, URL https://www.mdpi.com/1996-1073/13/18/4745.

[46] J.A. Hartigan, M.A. Wong, AK-means clustering algorithm, J. R. Stat. Soc. Ser. C. Appl. Stat. 28 (1) (1979) 100–108.