Modeling Wind Turbine Operational States from SCADA Data Using Hidden Markov Models

Fernando de Sá[1], Fabricio L. e Silva[2], Rodrigo Toso[3], Hélio H. de Souza[2], Douglas Cardoso[4], Ângela Ferreira[5], Kele Belloze[2], and Diego Brandão[2(✉)]

[1] Universidade Federal Fluminense, Niterói, Brazil
fernandosa@id.uff.br
[2] Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rio de Janeiro, Brazil
{fabricio.silva,helio.souza,kele.belloze,diego.brandao}@cefet-rj.br
[3] Microsoft, Redmond, USA
rfran@microsoft.com
[4] Universidade de Coimbra, Coimbra, Portugal
docardoso@uc.pt
[5] Research Center in Digitalization and Intelligent Robotics (CeDRI), SusTEC, Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
apf@ipb.pt

Abstract. People need to act quickly to stop climate change. Wind energy is a big step toward making energy systems that are better for the environment. But it's hard to keep wind turbines working well because they are very complicated systems that have to deal with bad weather and mechanical problems. A Supervisory Control and Data Acquisition (SCADA) system is used by wind farms to keep an eye on the turbines' health in real time. Even with this kind of system, it's hard to find important patterns across time, especially when it comes to how turbines break down. This is because SCADA data doesn't have a lot of time resolution and there aren't enough failure and non-failure classes. This article looks at how to use Hidden Markov Models (HMMs) to model and categorize the operational states of wind turbines using SCADA data from a real wind farm. The proposed method regards standard operation and various pre-failure conditions as concealed states. This method lets the model find temporal dependencies by using sequences from past data. We use a maximum likelihood-based method to pick features so that the model can figure out which SCADA variables are most important for each turbine component being tested. The proposed methodology is assessed across six operational states, encompassing normal operation and the pre-failure conditions of the chosen components. The macro F1 score ranges from 0.78 to 0.89, with an average of 0.83, showing that the method used correctly identifies temporal patterns related to component degradation.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16762, pp. 215–230, 2027.
https://doi.org/10.1007/978-3-032-30524-4_14

Keywords: Wind Energy · Hidden Markov Models · Fault Detection · Predictive Maintenance

1 Introduction

Wind and solar energy are two of the most important components of the global transition to less polluting energy systems. Their growth helps reduce greenhouse gas emissions and promotes the diversification of the energy mix. Thus, it is clear that they are aligned with global sustainability goals, such as the UN Sustainable Development Goals, especially SDG 7 (Affordable and Clean Energy) and SDG 13 (Climate Action) [3,7].

In the last few years, there have been a lot more wind farms, especially. But it's still hard to make sure that wind turbines work right. These turbines are subject to a wide range of mechanical and environmental conditions, which quickly wear down parts like gears, generators, and bearings. These parts are very important. If they break, they can cause big service interruptions and high maintenance costs. This is why wind farms need to be watched in real time [12].

As mentioned earlier, modern wind farms use Supervisory Control and Data Acquisition (SCADA) systems to continuously collect large volumes of operational data. Although they are often used to monitor turbine performance and provide information to detect faults in data and improve predictive maintenance, they still have significant limitations, primarily due to low temporal resolution, noise, and a large imbalance between normal operation and fault events.

In recent years, various machine learning and statistical approaches have been proposed to address these problems, including support vector machines, regression-based models, and deep learning methods [13,24]. Some of these approaches have shown promising results; however, they ultimately require large labeled datasets, are affected by class imbalance, and, in most cases, hinder interpretability, resulting in models with high accuracy but that do not allow for a true understanding of the problem, particularly in complex systems such as wind turbines.

Hidden Markov Models (HMMs) are a classical statistical technique that provides an interesting probability-based framework for modeling system activities as transitions between hidden states. They were designed to model sequential processes and temporal dependencies in time series data. So, HMMs make it possible to model how the parts of a wind turbine wear down over time under different operating conditions [19,25]. This quality makes them a good way to find patterns in how wind turbines work over time.

There have been a number of studies that looked into using HMMs to keep an eye on wind turbines by looking at both vibration signals and operational data [21,30]. But most of the studies that are already out there use high-frequency vibration measurements or simple experimental setups, which makes them less useful in real-world situations where only SCADA data is available. Conversely, as previously noted, utilizing SCADA data introduces further challenges, especially owing to reduced sampling rates and the challenges in identifying subtle degradation patterns.

Thus, this study evaluates the use of Hidden Markov Models for the multi-class classification of wind turbine operational states using real SCADA data. The proposed approach models normal operation and multiple pre-failure conditions as hidden states and aims to identify temporal degradation patterns from sequential operational data.

The main contributions of this work are as follows. First, we evaluate the use of Hidden Markov Models for the multi-class classification of wind turbine operational states using SCADA data. Second, we employ a feature selection strategy based on maximum likelihood estimation to identify variables associated with component degradation. Third, we present an experimental evaluation of the proposed approach that accounts for multiple pre-failure processes affecting key turbine components.

The remainder of this paper is organized as follows. Section 2 introduces the theoretical foundations of Hidden Markov Models, including their main algorithms and the feature selection approach adopted in this work. Section 3 reviews related studies on wind turbine fault detection and condition monitoring. Section 4 describes the proposed methodology. Section 5 presents and discusses the experimental results. Finally, Sect. 6 concludes the paper.

2 Theoretical Background

Hidden Markov Models (HMMs) are probabilistic models widely used to represent sequential processes in which the underlying system states are not directly observable but can be inferred from observed data [6,19]. In this framework, the system evolves over a finite set of hidden states, generating observations according to state-dependent probability distributions, which makes HMMs suitable for modeling time-series data and dynamic systems.

Formally, an HMM is defined by the parameter set $\lambda = (A, B, \pi)$, where A denotes the state transition probability matrix, B represents the emission probability distributions associated with each hidden state, and π is the initial state probability vector.

The transition probability matrix $A = \{a_{ij}\}$ defines the probability of transitioning from state S_i to state S_j:

$$a_{ij} = P(q_{t+1} = S_j \mid q_t = S_i) \tag{1}$$

where q_t denotes the hidden state at time t. The emission distribution B defines the probability of observing o_t given the current hidden state.

Figure 1 illustrates the structure of a Hidden Markov Model. The hidden states, denoted by q_1, q_2, q_3 (in yellow), represent latent system conditions and evolve over time according to the transition probabilities a_{ij}. Each state generates observable variables o_1, o_2, o_3 (in blue) according to emission probability distributions $b_i(o_t)$, which specify the likelihood of an observation given the current state.

Three fundamental computational problems are typically associated with HMMs [19]: (i) the *evaluation problem*, which consists of computing the likelihood of an observation sequence given a model; (ii) the *decoding problem*, which

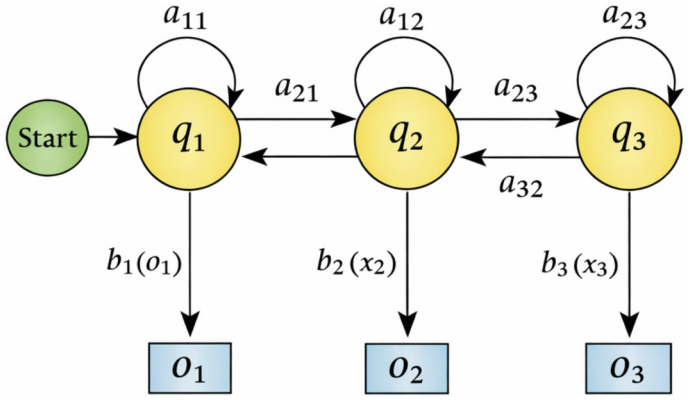

Fig. 1. Conceptual representation of a Hidden Markov Model showing hidden states, transition probabilities, and observation variables.

aims to determine the most probable sequence of hidden states; and (iii) the *learning problem*, which involves estimating the model parameters that maximize the likelihood of the observed data.

2.1 Algorithmic Formulation

The likelihood of an observation sequence $O = (o_1, o_2, \ldots, o_T)$ is computed using the Forward algorithm [14]. The forward variable $\alpha_t(i)$ represents the joint probability of observing the partial sequence up to time t and being in state S_i:

$$\alpha_t(i) = P(o_1, o_2, \ldots, o_t, q_t = S_i \mid \lambda) \tag{2}$$

It can be computed recursively as:

$$\alpha_{t+1}(j) = \left(\sum_{i=1}^{N} \alpha_t(i) a_{ij} \right) b_j(o_{t+1}), \tag{3}$$

where N denotes the number of hidden states.

For state decoding, the Viterbi algorithm is used to determine the most probable sequence of hidden states [5]. The Viterbi variable $\delta_t(i)$ is defined as:

$$\delta_t(i) = \max_{q_1, \ldots, q_{t-1}} P(q_1, \ldots, q_t = S_i, o_1, \ldots, o_t \mid \lambda) \tag{4}$$

and is computed recursively as:

$$\delta_{t+1}(j) = \max_i \left[\delta_t(i) a_{ij} \right] b_j(o_{t+1}). \tag{5}$$

Model parameters are estimated using the Baum–Welch algorithm, a special case of the Expectation–Maximization (EM) framework, which iteratively updates the model parameters to maximize the likelihood of the observed data [28].

2.2 Feature Selection

In this work, the likelihood computed by the Forward algorithm is used as a criterion for evaluating the relevance of individual SCADA variables. For each variable, an observation sequence is constructed and its likelihood $P(O \mid \lambda)$ is computed under the corresponding HMM.

This likelihood tells us how well the model captures the time-related structure of the variable. Variables that produce elevated likelihood values are regarded as more informative for depicting degradation patterns and are consequently chosen for model development.

This strategy adheres to the maximum likelihood principle [22], which is fundamental to parameter estimation and model evaluation in HMMs.

HMMs have been extensively utilized in the analysis of sequential data for tasks including speech recognition, activity recognition, and signal processing [15,18,23]. When it comes to monitoring wind turbines, they offer a natural way to use SCADA time-series data to model changes between different operational states, such as normal operation and degradation states.

This study utilizes Hidden Markov Models to represent the temporal progression of operational states in wind turbine components through SCADA measurements. The next section explains how the proposed model was built and trained.

3 Related Work

In recent years, a significant amount of research has been conducted on how to use data-driven methods, leveraging SCADA data and sensor readings, to identify problems and monitor the integrity of wind turbines. The growing volume of operational data from wind farms has led to increased use of machine learning, along in conjunction with probabilistic models, in developing models able of detecting anomalous operational patterns and supporting predictive maintenance strategies. This section is a narrative review of the literature on using Hidden Markov Models (HMMs) to find problems and figure out what state a wind turbine is in, with a focus on SCADA-based methods.

3.1 Search Methodology

We used the Elicit platform [1] for literature search. This tool does a semantic search on big academic platforms like OpenAlex and Semantic Scholar [2]. Elicit goes beyond keyword searches to do a contextualized retrieval of studies. This improves search coverage and gets rid of biases that can happen in traditional searches. To this end, we formulated the subsequent research question:

> "How effectively can Hidden Markov Models applied to SCADA data classify normal and pre-failure operational states in wind turbine components?"

[1] elicit.com.

The search returned an initial set of 500 candidate papers. These studies were screened based on their abstracts using the following inclusion criteria:

- Application of Hidden Markov Models as the principal method;
- Using SCADA data;
- Focus on wind turbine components or systems;
- Classification of operational states (normal, anomaly, or pre-failure);
- Empirical validation with quantitative performance metrics;
- Use of real-world data (excluding purely simulated studies);
- Publication in peer-reviewed venues;

After the search results from the Elicit platform were returned, 16 studies were selected for in-depth analysis. Subsequently, a manual review was conducted to confirm the selection and ensure compliance with the inclusion criteria.

At this point, two studies were identified in a different language (a Korean version of another paper already included and another in Portuguese), and another was excluded because it didn't use Hidden Markov Models.

Consequently, the final collection comprised 13 distinct studies included in this review. Figure 2 shows a summary of the whole process of choosing literature.

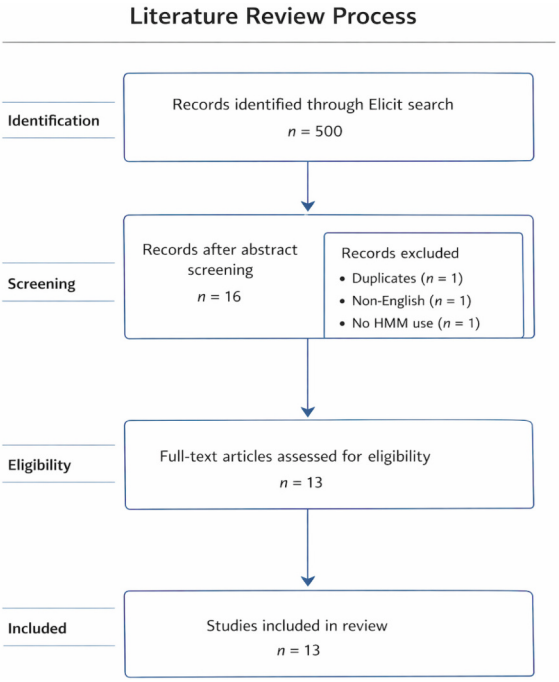

Fig. 2. Overview of the literature review process, including Elicit-based screening and subsequent manual validation

3.2 More About Included Studies

The selected studies examine various components of wind turbines, with a focus on bearings, gearboxes, and generators. The kind of data used is a big difference. Numerous studies use high-frequency sensor data, particularly vibration signals, which provide richer temporal information and enable more comprehensive characterization of mechanical dynamics. For instance, Shin et al. [21], and Zhao et al. [30] used HMMs on vibration-based measurements and obtained very good classification results. Another use of HMMs was presented by Jang et al. [9], who investigated their use to monitor rotating machinery operating under varying conditions.

Several studies use traditional HMM formulations to model degradation processes [11], a variant using an adaptive approach is presented in [10]. In this adaptive approach, the authors demonstrate its ability to make predictions at multiple stages. To demonstrate that a hybrid probabilistic approach can efficiently handle complex system dynamics, the authors Sasaki [20] present the HMM with probabilistic principal component analysis (PPCA).

Another line of research is based on the analysis of SCADA data, aiming to model real-world operational situations, but faces additional challenges such as lower temporal resolution (e.g., 5 or 10-min intervals), higher noise levels, and severe class imbalance, as demonstrated by Windmann et al. [26] and Poyraz et al. [17]. These authors examined anomaly detection using SCADA data in manufacturing environments, demonstrating the practical effectiveness of these methods. Another proposal, still in this context, was made by Yang et al. [29]. The authors propose a dynamic HMM with real-time updates to improve adaptability to changes in environmental conditions, thus advancing this line of research.

Furthermore, some studies indicate that results in SCADA-based environments are often more limited. Parekh et al. [16] report that standard HMMs may fail under severe class imbalance, highlighting the importance of data preprocessing and modeling choices. In this context, Chesterman et al. [4] highlight the diversity of modeling strategies and the challenges associated with assessing the normal behavior of operational wind farms.

Other works like Herp et al. (2017) [8] examine state transition dependencies using a Bayesian framework based on Hidden Markov Models (HMMs), whereas Leahy et al. (2016) employ Support Vector Machines (SVM) as a non-HMM benchmark for SCADA-based fault detection, providing a valuable point of reference for comparison.

The performance varies substantially across studies, someone achieve F-measures above 0.9 when using high-frequency data and carefully engineered preprocessing frameworks [27]. In contrast, studies based on SCADA data often report lower performance or increased sensitivity to noise and imbalance [Omitted for review].

Another important point is how to define operational states. There are several ways to define pre-failure conditions, including thresholds, historical maintenance records, and anomaly detection. The time frame for pre-failure states is often not clearly defined, making it difficult to compare studies.

As presented, several extensions of HMMs have been proposed to address the limitations of standard HMM formulations. They consist of degradation-aware models, adaptive HMMs with dynamic parameter modifications, and composite methodologies that integrate y probabilistic mechanisms. Table 1 summarizes the studies included in this review.

Table 1. Summary of the 13 studies included in the literature review

Study	Data Type	Component	HMM Variant	Notes
Zhao et al. (2009) [30]	Vibration	Bearings	HMM + fuzzy quantization	Robust to outliers
Jang et al. (2012) [9]	Vibration	Gearbox	Standard HMM	Rotating machinery monitoring
Windmann et al. (2015) [26]	SCADA	General	Standard HMM	Industrial dataset
Leahy et al. (2016) [13]	SCADA	Multi-component	Non-HMM (SVM)	Included for comparison
Herp et al. (2017) [8]	SCADA	General	HMM-related (Bayesian)	State transition modeling
Sasaki et al. (2017) [20]	Sensor (strain)	Blades	HMM + PPCA	Hybrid probabilistic model
Shin et al. (2018) [21]	Vibration	Bearing, Gearbox, Generator	Standard HMM	High performance (F1 up to 0.918)
Poyraz et al. (2018) [17]	SCADA	General	Standard HMM	Power curve-based detection
Jiang et al. (2023) [10]	Sensor	Bearings	Adaptive degradation HMM	Multi-step prediction
Chesterman et al. (2023) [4]	SCADA	General	Not specified	Comparative review/implementation
Jiang et al. (2024) [11]	Sensor	Bearings	Degradation HMM	Reliability modeling
Parekh et al. (2025) [16]	SCADA	Multi-component	Standard HMM	Failed under class imbalance
Yang et al. (2025) [29]	SCADA	Multi-component	Dynamic HMM	Real-time adaptation

3.3 Discussion

The reviewed literature indicates that Hidden Markov Models provide a suitable probability structure for modeling time relationships in wind turbine operational data. Their ability to represent sequential transitions between latent states makes them particularly appropriate for capturing degradation processes and intermediate operating conditions.

However, the effectiveness of HMM-based approaches varies substantially with data features. Studies based on high-frequency sensor data generally show higher performance, as these signals provide detailed temporal information that supports the identification of degradation patterns. In contrast, SCADA-based approaches face additional challenges, including lower temporal resolution, higher noise levels, and pronounced class imbalance, which can significantly affect model performance.

Furthermore, several studies show that standard HMM formulations may be insufficient when applied directly to raw SCADA data. As a result, extensions such as degradation-aware models, adaptive HMMs, and hybrid probability-based approaches have been proposed to improve robustness plus predictive effectiveness. Nevertheless, these approaches often introduce additional complexity and may rely on assumptions that are not always compatible with realistic contexts.

Another limitation seen across the literature concerns the definition and labeling of pre-failure states. Different studies use heterogeneous criteria, such as threshold-based methods. In particular, the study stresses scenarios in which

only low-frequency operational data are available and where preprocessing, along with feature selection, is central to identifying degradation patterns.

In this context, the present work assesses the use of standard Hidden Markov Models for multiclass classification on SCADA data under realistic conditions. In particular, the study stresses scenarios in which only low-frequency operational data are available and where preprocessing along with feature selection are central to identifying degradation patterns.

4 Methodology

This study investigates the use of Hidden Markov Models (HMMs) for the multi-class classification of operational states in wind turbine components, using SCADA data, with a particular focus on modeling pre-failure states. Figure 3 describes the methodology used with four main steps: data acquisition, time-series construction and feature selection, HMM modeling, and performance evaluation.

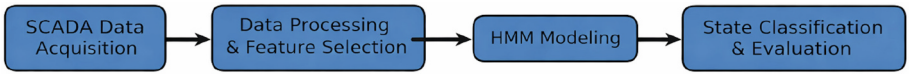

Fig. 3. Overview of the proposed methodology.

4.1 Dataset

The experiments used a publicly available wind turbine dataset [1], comprising SCADA measurements from turbines operating in a wind farm. The dataset comprises 79 operational variables and a failure log that records past faults affecting different parts of the turbine. The dataset exhibits common features of real SCADA data, including noise, missing values, and an imbalance between normal and failure conditions.

4.2 Data Processing and Temporal Sequence Construction

The SCADA data are modeled as multivariate time series, in which each observation comprises multiple operational variables recorded simultaneously.

To capture temporal behavior, the time series were segmented into subsequences of 2000 observations. Considering the SCADA sampling interval, each subsequence corresponds to a fixed temporal window preceding a failure event.

Subsequences representing normal operation were extracted from periods without recorded failures. Pre-failure sequences were defined as the observations immediately preceding each failure event.

Each sequence τ_i is associated with one of the following operational states:

- S_1: normal operation
- S_2: gearbox pre-failure
- S_3: generator pre-failure
- S_4: generator bearing pre-failure
- S_5: hydraulic group pre-failure
- S_6: transformer pre-failure

In this context, pre-failure states are interpreted as transitional stages associated with the progression of anomalies toward failure.

Initially, each subsequence is represented by a matrix of size 2000×79, corresponding to the number of time steps and measured variables.

4.3 Feature Selection

Due to the high dimensionality and potential redundancy of SCADA variables, a feature selection step was performed.

Each variable was evaluated individually using the likelihood produced by the HMM. For each variable, sequences corresponding to normal operation and the respective pre-failure condition were used to estimate model parameters and compute the log-likelihood.

The likelihood is used as a proxy to assess how well each variable captures the statistical structure of the degradation process. Variables with higher likelihood values were selected to compose the final feature subset.

After this step, each subsequence is represented by a matrix of size 2000×5.

4.4 HMM Modeling

The *hmmlearn* library was used to implement the HMM model in Python, with multivariate Gaussian emission distributions and diagonal covariance matrices.

The model was configured with six hidden states, corresponding to the previously defined operational conditions $(S_1, ..., S_6)$. Model parameters were estimated using the Baum–Welch algorithm, and state decoding was performed using the Viterbi algorithm.

To reflect the progressive nature of degradation processes, a left-right topology was adopted, restricting transitions to the same state or to the next state.

4.5 Model Evaluation

The dataset was divided into training and testing subsets using a 70–30 split, resulting in eight test sequences τ_i.

Model performance was assessed using confusion matrices and the macro-averaged F1-score, which provides a balanced evaluation across all classes, particularly in the presence of class imbalance.

For each class c, precision and recall are defined as:

$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c} \tag{6}$$

$$\text{Recall}_c = \frac{TP_c}{TP_c + FN_c} \tag{7}$$

where TP_c, FP_c, and FN_c denote the number of true positives, false positives, and false negatives for class c, respectively.

The F1-score for each class is computed as:

$$F1_c = 2 \cdot \frac{\text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \tag{8}$$

The macro-averaged F1-score is then obtained as the arithmetic mean across all classes:

$$F1_{\text{macro}} = \frac{1}{C} \sum_{c=1}^{C} F1_c \tag{9}$$

where C is the number of classes.

This metric was selected due to its sensitivity to class imbalance, as it assigns equal weight to each class regardless of its frequency.

The adopted split does not explicitly enforce temporal separation between training and testing sequences, which may influence generalization performance.

5 Results and Discussion

This section presents the results, in which the first part analyzes the feature selection process. The second part presents the classification performance on the test samples using HMM.

The feature selection procedure identified one representative variable for each monitored component based on the maximum likelihood criterion described in the methodology. Table 2 summarizes the selected features.

Table 2. Selected features for each component.

Component	Selected Feature
Gearbox	Average ambient temperature [°C]
Generator	Minimum average generator rotation speed [rpm]
Generator Bearing	Average absolute wind direction [m/s]
Hydraulic Group	Average controller-center temperature [°C]
Transformer	Active generator power [Wh]

The selected variables appear to capture the relationships with the monitored components. From this selection, some characteristics can be interpreted based on what occurs in the physical operation of the turbine; an example is the relationship between the generator's rotational speed and the active power,

while others capture less clear operating conditions that may influence component degradation. This suggests that the probability-based selection procedure can identify variables that encode relevant temporal patterns, even when the relationship with the component is not explicitly causal.

Based on historical failure events, the time series were divided into subsequences that formed training and test subsets, in a 70%–30% ratio, resulting in eight test samples τ_i. Each sample contains subsequences associated with the six operational states described previously.

The classification performance was evaluated using confusion matrices and macro-averaged F1-scores. Table 3 shows the results for the test sample τ_1. We obtained a model that achieved a macro F1-score of 0.85.

Table 3. Confusion matrix for test sample τ_1.

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	1938	62	0	0	0	0
S_2	0	2000	0	0	0	0
S_3	0	705	1295	0	0	0
S_4	0	0	0	1076	924	0
S_5	0	0	0	0	1972	28
S_6	0	0	0	0	0	2000

The model achieved perfect classification for gearbox pre-failure (S_2) and transformer pre-failure (S_6), indicating that the selected features provide strong discriminatory signals for these components. In contrast, higher confusion was observed between generator-related states (S_3 and S_4). This behavior suggests that the degradation patterns of these components exhibit similar temporal signatures within the selected feature space, making them harder to distinguish.

For the second test sample τ_2, the model achieved a macro F1-score of 0.89, as presented in Table 4.

Table 4. Confusion matrix for test sample τ_2.

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	2000	0	0	0	0	0
S_2	0	1954	46	0	0	0
S_3	0	0	2000	0	0	0
S_4	0	0	0	2000	0	0
S_5	0	0	0	861	1139	0
S_6	0	0	0	0	330	1670

The results for τ_2 show improved class separability compared to τ_1, although some confusion persists in states associated with hydraulic and transformer systems (S_5 and S_6). This variability across samples indicates that classification performance is influenced by the specific characteristics of each failure event and the associated temporal patterns.

The remaining test samples produced consistent results, as summarized in Table 5.

Table 5. Macro F1-score

Sample	Macro F1-score
τ_3	0.80
τ_4	0.78
τ_5	0.78
τ_6	0.87
τ_7	0.80
τ_8	0.84

Overall, the model achieved macro F1-scores ranging from 0.78 to 0.89, with an average of approximately 0.83 across the eight test samples. These results indicate that the HMM-based approach is able to capture temporal patterns associated with degradation processes in wind turbine components using SCADA data.

In general, the results obtained indicate that HMMs tend to achieve moderate results when using data from SCADA. These results corroborate what is indicated in the literature, where high-frequency data are better learned by HMMs, while medium or low-frequency data show less expressive results, especially as in the case of this work, where there is a significant imbalance between the classes.

Particularly, although some classes show higher misclassification rates, especially those associated with generator-related components. The results demonstrate that HMMs provide a coherent and interpretable probabilistic framework for modeling sequential operational behavior, making the technique quite interesting to support predictive maintenance strategies in wind power systems, especially in scenarios where only SCADA data is available.

6 Final Remarks and Future Work

The relevance of wind energy as a component in reducing greenhouse gas emissions and achieving climate goals has grown amidst the transition to sustainable energy systems. In this sense, maintaining wind turbines in efficient operation is of paramount importance to ensure an efficient energy supply.

This work evaluates the use of Hidden Markov Models (HMMs) for multiclass classification of operational states in wind turbine components of the Portuguese

company EDP, using SCADA data. The modeling used allows the representation of temporal degradation patterns in the turbines' operational data.

The results of experiments on a publicly accessible dataset indicate that the HMM-based method can distinguish between various operating conditions and identify temporal relationships in SCADA time series. The assessed test samples had macro F1 scores ranging from 0.78 to 0.89.

Finally, the data demonstrate that HMM provides a reliable, probability-based framework for emulating the sequential operational behavior of wind turbines in practical conditions, based on SCADA data. The findings also highlight the effects of class imbalance and low temporal resolution on classification performance.

Further work on this project will investigate integrated models that combine HMMs with other machine learning techniques and will extend the proposed method to include additional turbine components. In addition, it is believed that the robustness and predictive performance of the models can be improved by combining SCADA data with complementary monitoring sources, such as vibration measurements and meteorological data.

Acknowledgment. This work was supported by FCT - Fundação para a Ciência e Tecnologia, I.P., by projects: CeDRI, UID/05757/2025 (DOI: 10.54499/UID/05757/2025) and UID/PRR/05757/2025 (DOI: 10.54499/UID/PRR/05757/2025); SusTEC, LA/ P/0007/2020 (DOI: 10.54499/LA/ P/0007/2020). Specifically, the financial assistance of the FCT grant CETP/0003/2022 (DOI 10.54499/CETP/0003/2022) towards this research is hereby acknowledged. Additionally, the authors would like to thank the following Brazilian Agencies: FAPERJ, CAPES, and CNPq. DB thanks Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for the financial support through the grants E-26/210.798/2024 and E-26/204.262/2025.

Declaration of Generative AI in Scientific Writing. During the preparation of this work, the authors used Elicit to assist with the literature review and Grammarly for English-language refinement. Following the use of these tools, the authors reviewed and edited the manuscript as necessary and assume full responsibility for its content.

References

1. EDP Open Data. https://opendata.edp.com/pages/homepage/
2. Bernard, N., et al.: Using artificial intelligence for systematic review: the example of elicit. BMC Med. Res. Methodol. **25**(1), 75 (2025)
3. Büyüközkan, G., et al.: A novel renewable energy selection model for united nations' sustainable development goals. Energy **165**, 290–302 (2018)
4. Chesterman, X.: Overview of normal behavior modeling approaches for SCADA-based wind turbine monitoring. Renew. Energy (2023)
5. Churbanov, A., Winters-Hilt, S.: Implementing EM and Viterbi algorithms for hidden Markov model in linear memory. BMC Bioinform. **9**(1), 224 (2008)
6. Ghojogh, A., Sepanj, M.H., Ghojogh, B.: On the relation of state space models and hidden Markov models. arXiv preprint arXiv:2601.13357 (2026)
7. Global Wind Energy Council: Global wind report 2021 (2021)

8. Herp, J.: Dependency in state transitions of wind turbines. IEEE Trans. Industr. Electron. (2017)
9. Jang, M., Lee, J.M., Hwang, Y., et al.: Condition monitoring of rotating machine with a change in speed using hidden Markov model. J. Korean Soc. Noise Vibr. Eng. **22**(5), 413–420 (2012). https://doi.org/10.5050/KSNVE.2012.22.5.413
10. Jiang, C., Chen, Q., Lei, B.: A self-adaptive multi-step degradation hidden Markov model for operational reliability prediction of wind turbine bearings. In: 2023 5th International Conference on System Reliability and Safety Engineering (SRSE), pp. 19–25. IEEE (2023)
11. Jiang, C., Chen, Q., Lei, B.: A novel operational reliability and state prediction method based on degradation hidden Markov model with random threshold. Qual. Reliab. Eng. Int. **41**(2), 652–671 (2025)
12. Kobbacy, K.A.H., Murthy, D.P.: Complex System Maintenance Handbook. Springer (2008)
13. Leahy, K., Hu, R., Konstantakopoulos, I.C., et al.: Diagnosing wind turbine faults using machine learning techniques applied to operational data. In: International Conference on Prognostics and Health Management (ICPHM) (2016). https://doi.org/10.1109/ICPHM.2016.7542860
14. Li, T., Ma, J.: On theoretical justification of the forward-backward algorithm for the variational learning of Bayesian hidden Markov models. IET Signal Process. **16**(6), 674–679 (2022)
15. Mor, B., Garhwal, S., Kumar, A.: A systematic review of hidden Markov models and their applications. Arch. Comput. Methods Eng. **28**(3), 1429–1448 (2021)
16. Parekh, R., Kiran, R.: Risk analysis of wind turbine using machine learning and ai. In: Proceedings of the International Conference on Industrial Engineering and Operations Management (2025)
17. Poyraz, O.: Anomaly detection on wind turbines. In: SIU (2018)
18. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE **77**(2), 257–286 (2002)
19. Rabiner, L.R., Juang, B.H.: An introduction to hidden Markov models. IEEE ASSP Mag. **3**(1), 4–16 (1986)
20. Sasaki, R.: Health monitoring using hmm and PPCA. In: PHM Asia-Pacific (2017)
21. Shin, S.H.: Development of a fault monitoring technique for wind turbines using a hidden Markov model. In: Proceedings of Italian National Conference on Sensors (2018)
22. Sreevalsan-Nair, J.: Maximum likelihood. In: Encyclopedia of Mathematical Geosciences, pp. 851–853. Springer, Cham (2023)
23. Sulayman, M., et al.: User activity recognition using hidden Markov models. Sensors **19**(17), 3771 (2019)
24. Wakayama, T., et al.: Deep learning-based anomaly detection for wind turbine bearings. Mech. Syst. Signal Process. **210**, 111123 (2024)
25. Wang, L., Zhao, W., Liu, B., Li, Y.F.: Reliability assessment for partially monitored systems based on degradation hidden Markov models with time-varying parameters. IEEE Trans. Reliab. (2025)
26. Windmann, S.: Evaluation of model-based condition monitoring systems in industrial application cases. In: Proceedings of the International Conference on Machine Learning for Cyber Physical Systems (2015)
27. Wong, P.K., Zhong, J.H., Yang, Z.X., Vong, C.M.: A new framework for intelligent simultaneous-fault diagnosis of rotating machinery using pairwise-coupled sparse bayesian extreme learning committee machine. Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci. **231**(6), 1146–1161 (2017)

28. Yang, F., et al.: Statistical and computational guarantees for the Baum-Welch algorithm. J. Mach. Learn. Res. **18**(125), 1–53 (2017)
29. Yang, L.: Dynamic hidden Markov model for offshore wind farm meteorological risk prediction. Renew. Energy (2025)
30. Zhao, H., Gao, R.: Wind turbine bearing fault diagnosis based on hidden Markov model. Mech. Syst. Signal Process. **23**(6), 1893–1903 (2009)