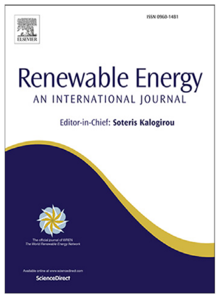
Journal Pre-proof



An autonomous decision-making agent for offshore wind turbine blades under leading edge erosion

Javier Contreras, Athanasios Kolios

PII: S0960-1481(24)00590-1
DOI: https://doi.org/10.1016/j.renene.2024.120525
Reference: RENE 120525

To appear in: *Renewable Energy*

Received date : 4 January 2024
Revised date : 15 April 2024
Accepted date : 18 April 2024

Please cite this article as: J. Contreras and A. Kolios, An autonomous decision-making agent for offshore wind turbine blades under leading edge erosion, *Renewable Energy* (2024), doi: https://doi.org/10.1016/j.renene.2024.120525.

Title Page

An autonomous decision-making agent for offshore wind turbine blades under leading edge erosion

Javier Contreras[a,*], Athanasios Kolios[b]

[a] *Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 16 Richmond St, Glasgow – G1 1XQ, Scotland, UK*

[b] *Department of Wind and Energy Systems Structural Integrity and Loads Assesment, Technical University of Denmark, Risø Campus Frederiksborgvej 399, Roskilde 4000, Denmark*

Abstract

The increasing pressure of offshore wind developments is leading to projects being located in areas with more difficult access and greater weather barriers. As these constraints increase, O&M costs also grow in importance. Therefore, the current scenario requires a careful planning to avoid unnecessary costly maintenance decisions or unexpected failures. To overcome the problem of increasing O&M costs and difficult access, this manuscript presents an autonomous decision-making Reinforcement Learning (RL) agent to improve O&M planning for the Leading Edge Erosion (LEE) problem. The method developed in this work makes use of a linear degradation model to account for the damage progression dynamics and site-specific weather models. The RL-based agent proposed in this manuscript is able to reduce expected O&M costs in the range of 12-21% when compared with condition-based policies.

List of Abbreviations

ANN	Artificial Neural Network
CFD	Computational Fluid Dynamics
CfD	Contracts for Difference
CTV	Crew Transfer Vessel
DQN	Deep Q Networks
HLV	Heavy Lift Vessel
LEE	Leading Edge Erosion
MIP	Mixed Integer Programming
MDP	Markov Decision Process
NLP	Non-linear programming

Highlights

- A decision-support framework for LEE maintenance is presented.
- Weather, material and repair success uncertainties are considered in the framework.
- An autonomous RL-based agent is trained considering site-specific conditions.
- The agent improves expected O&M costs against condition-based baseline policies.

An autonomous decision-making agent for offshore wind turbine blades under leading edge erosion

Javier Contreras[a,*], Athanasios Kolios[b]

[a] *Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 16 Richmond St, Glasgow – G1 1XQ, Scotland, UK*

[b] *Department of Wind and Energy Systems Structural Integrity and Loads Assesment, Technical University of Denmark, Risø Campus Frederiksborgvej 399, Roskilde 4000, Denmark*

Abstract

The increasing pressure of offshore wind developments is leading to projects being located in areas with more difficult access and greater weather barriers. As these constraints increase, O&M costs also grow in importance. Therefore, the current scenario requires a careful planning to avoid unnecessary costly maintenance decisions or unexpected failures. To overcome the problem of increasing O&M costs and difficult access, this manuscript presents an autonomous decision-making Reinforcement Learning (RL) agent to improve O&M planning for the Leading Edge Erosion (LEE) problem. The method developed in this work makes use of a linear degradation model to account for the damage progression dynamics and site-specific weather models. The RL-based agent proposed in this manuscript is able to reduce expected O&M costs in the range of 12-21% when compared with condition-based policies.

List of Abbreviations

ANN Artificial Neural Network

CFD Computational Fluid Dynamics

CfD Contracts for Difference

CTV Crew Transfer Vessel

DQN Deep Q Networks

HLV Heavy Lift Vessel

LEE Leading Edge Erosion

MIP Mixed Integer Programming

MDP Markov Decision Process

NLP Non-linear programming

OWTs Offshore Wind Turbines

O&M Operation and maintenance

PN Petri Nets

RL Reinforcement learning

WARER Whirling Arm Rain Erosion test Rig

Keywords

Leading edge erosion; Wind turbine blade O&M; Blade erosion degradation; Wind turbine O&M optimisation.

Highlights

- A decision-support framework for LEE maintenance is presented.
- Weather, material and repair success uncertainties are considered in the framework.
- An autonomous RL-based agent is trained considering site-specific conditions.
- The agent improves expected O&M costs against condition-based baseline policies.

---

1. Introduction

The rise of renewable energies and the challenging carbon-emission reduction goals set for the upcoming years have driven the exploration of offshore energy opportunities. In this context, offshore wind turbines (OWTs) are one of the most promising offshore energy sources. With the knowledge and expertise gained from the bottom-fixed sites, the development of floating wind technologies unlocked a large range of potential sites. Despite the knowledge of OWTs being much more premature than that of onshore ones (64.3 GW vs 841.9 GW capacity installed worldwide) the potential benefits of its large-scale deployment, such as the potential to install larger turbines or the reduction of the environmental impact of wind farms are propelling its growth. According to the Global Wind Energy Report 2023 produced by the Global Wind Energy Council (GWEC), the wind energy market is expected to grow by 15% on average per year and the compound annual growth rate of offshore wind reach 32% in the next five years.

Despite the promising outlook for the offshore wind industry, several issues still need to be addressed to make this technology as competitive as its onshore counterpart. The O&M costs of OWTs are estimated to account for 25-30% of the total lifecycle costs [1]. Offshore maintenance activities are estimated to be five to ten times more expensive than those performed onshore [2, 3]. When combined with the required weather

*Corresponding author
Email addresses: javier.contreras-lopez@strath.ac.uk (Javier Contreras), atko@dtu.dk (Athanasios Kolios)

windows for maintenance activities, this can result in O&M costs that are double those of onshore turbines [4]. The combination of accessibility challenges and the lower reliability of large rotor turbines offshore turbines [5] creates a challenging scenario leading the operators to use preventive or reactive maintenance resulting in unnecessary O&M costs [6].

Given the challenges of maintenance planning, the use of decision support tools is vital for offshore wind farm operators. Many efforts have recently been made to develop different tools to optimise one or many of the different existing maintenance methods: routine inspections, corrective maintenance, preventive maintenance, condition-based maintenance, predictive maintenance or opportunistic maintenance. Several different approaches have been used. These include methods such as Mixed Integer Programming (MIP), Non-linear Programming (NLP), stochastic models, Markov models, Petri Nets (PN) models, analytical models, fuzzy models, intelligent algorithmic models, and multi-objective models, to name a few. Regardless of the method used, scholars have targeted different levels for optimisation, ranging from individual components such as the tower, foundation, or drivetrain, to the entire turbine or wind farm. The objectives for optimisation include O&M costs, logistics costs, availability, reliability, and environmental impact. Some of the most recent publications are summarised here. Saleh et al. [7] proposed a PN model combined with RL for the O&M of wind turbines. Elusakin et al. [8] developed a stochastic PN model for O&M planning of floating offshore wind turbines. Yan and Dunnet [9] studied the maintenance of OWTs under the PN paradigm and considering periodic maintenance, condition-based maintenance and reactive maintenance policies. Ge et al. [10] designed a maintenance planning optimisation algorithm based on MIP to minimise power generation losses on maintenance activities. Li et al. [11] proposed a multi-objective maintenance strategy optimisation framework at wind-farm level considering uncertainty in the maintenance performance. In [12], Schouten et al. introduce a single-component model for maintenance optimisation under time-varying costs that is applicable to offshore wind turbine maintenance. Aafif et al. [13] provides an optimal preventive maintenance strategy for a wind turbine gearbox based on its temperature. In [14], Yong and Qirong propose an optimisation maintenance scheme for the maintenance missions considering the time windows based on a hybrid ant colony algorithm. In [15], Zou and Kolios propose a framework to improve maintenance decision-making by quantifying the value of information of condition monitoring.

The modelling of the O&M of OWTs at turbine level or wind farm level requires a deep knowledge about the failure modes of the components that carry the highest weights in the maintenance activities. Damage is usually discretised in states and its progression represented with a probabilistic description of the transition between them. The calibration of these require the possession of considerable amounts of failure and maintenance data of the same or similar equipment in sites with similar weather conditions to provide good results. Alternatively, the use of detailed models, can provide with a numerical testing environment to obtain synthetic data. Higher level models require more computationally affordable damage descriptions that can mimic the real behaviour of damage degradation. Being the rotor one of the most critical components

[16, 17] and LEE one of the failure modes carrying the higher criticality [17–20], its O&M planning requires a careful analysis. The unattended evolution of LEE can have aerodynamic, environmental and structural implications increasing in importance and finally being able to produce the catastrophic failure of the blade. Lifetime assessments of erosion protection systems can be found in the literature, such as the works performed by Hasager et al. [21, 22] and [23]. In [21], the lifetime assessment of leading edge protection systems of Vestas V52 turbines for sites in the Danish Seas was performed, finding expected lifetimes between 2 and 13 years. Also, in [22], for sites in the North and Baltic Sea, the expected lifetime of coatings was in the range of 1 to 25 years.There have been many efforts to estimate the life of protective coatings but, to the best knowledge of the authors, there are no studies focusing on the predictive maintenance of this failure mode. Under this high uncertainty in coating lifetime and weather effects, there is a need for a decision support tool to improve the decision-making capability of wind farm operators. The potential benefits of its application increase with its application in harsher environment. In this sense, the current study presents a novel autonomous decision-making RL agent to optimise OWT LEE O&M costs. The uncertainties in weather scenarios, maintenance performance and LEE protective coating behaviour are considered in this paper. The proposed agent, once trained, can provide an action suggestion at any stage of the turbine service life. Also, the proposed agent can be retrained once real operation data becomes available improving its accuracy an providing further O&M cost reduction.

The remainder of this paper is structured as follows: Section 2 presents the methodology used for the optimisation of the O&M planning. Section 3 provides the assumptions and considerations of the O&M model used in this study. Section 4 presents two case studies to evaluate the performance of the proposed decision-support agent. Section 5 offers a discussion about the benefits and limitations of the framework presented as well as some follow-up opportunities. Finally, Section 6 summarises the conclusions of the application of the proposed methodology.

2. Methodology

This section delineates the methodology employed in this study, which is divided into two subsections. The first subsection elucidates the computational framework for LEE degradation and turbine operation simulation, while the second one delves into the decision-making framework for the optimisation of O&M costs.

2.1. Computational framework

This subsection provides a description of the *environment* and the computational framework that defines the dynamics of the degradation of the system.

LEE is a degradation phenomenon that affects wind turbine blades in several aspects (acoustic, aerodynamic and structural). The relations between the parameters affecting this problem is shown in Figure 1.

This phenomenon is caused by fatigue degradation through a repeated number of impacts of airborne particles (rain, insects and other airborne particles) onto the outermost layers of the blade. The dynamics of this process are affected by a number of parameters such as the impact energy, coating material parameters and weather conditions. The present computational framework provides a method to consider uncertainty in the abovementioned parameters. This computational framework is presented in [24] and depicted in Figure 2. To account for the uncertainty in climatic, material and aerodynamic parameters, the techniques described below can be used.

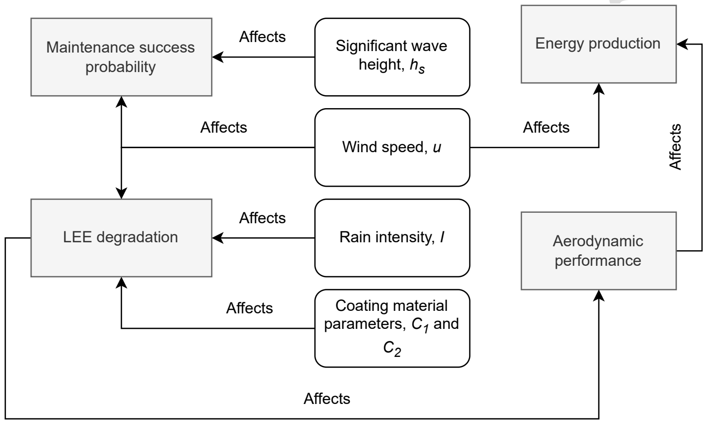


Figure 1: Relations between parameters

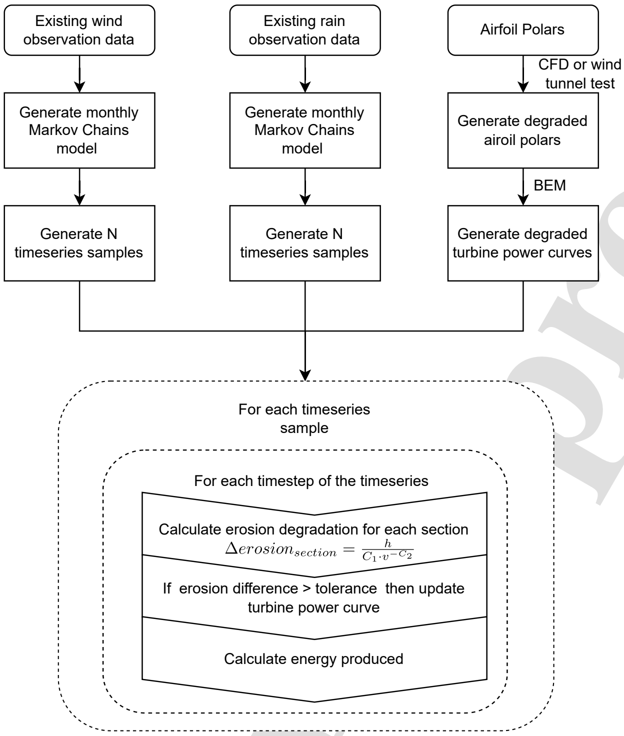


Figure 2: LEE calculation framework. Source: [24]

First, synthetic weather data needs to be generated for the location of the turbine. Rain intensity, wind speed and significant wave height time series should be generated in order to compute damage degradation and maintenance success rates. Depending on the availability of data for the project's location, various approaches can be considered. If a considerable amount of observations is not available, data can then be obtained from the ERA5 reanalysis data [25]. A Markov chains model [26] can then be used to generate synthetic wind series as shown in [24]. Significant wave height is an important parameter to account for the limitations in the logistics for offshore wind turbine maintenance activities. The generation of significant wave height series should be dependent on wind speed. Different approaches can be used to achieve this conditioned on data availability. In this case, an Artificial Neural Network (ANN) was used to mimic the significant wave height, h_s, patterns registered by the FINO1 platform. The parameters of the neural network used are the significant wave height of the two previous time steps, the wind speed of the current and two previous time steps and the calendar month (to account for seasonality). The proposed ANN is composed of a hidden layer of 4 neurons using the sigmoid activation function and an output layer with the significant wave height value.

LEE is known to cause effects on the aerodynamic performance of wind turbine blades [18, 27–29], resulting in reduced lift and increased drag forces. These effects lead to a decrease in the power generated by the turbine. The estimated annual energy production losses can range from 1.5% to 10%, depending on the turbine's characteristics and site-specific climatic conditions [27, 30–33]. Estimating changes in aerodynamic performance is a non-trivial task, often requiring the application of 2D and 3D Computational Fluid Dynamics (CFD) numerical models due to the limited availability of observational data [29, 34–36]. Once the blade's performance at various levels of LEE degradation is determined, the degraded power curves of the turbine can be constructed. These curves are used to assess the energy losses of the turbine. The energy produced at each time step, ΔE_i, is calculated using Equation 1, where $P(u, d)$ represents the power obtained from the degraded power curves, and Δt is the computational time step. Energy losses due to LEE degradation are then estimated as the difference between the pristine and degraded power curves.

$$\Delta E_i = P(u, d) \cdot \Delta t \tag{1}$$

Considering the high uncertainty in the behaviour of various coating materials is essential because the agent needs to account for uncertain degradation dynamics. To address this, the proposed method leverages the inherent uncertainty found in the Whirling Arm Rain Erosion test Rig (WARER) results, as shown in Figure 3. In these tests, leading edge protection coatings are subjected to water droplet impacts until they reach their final degradation. By analysing the evolution of the coating's degradation, the accumulated volume of water impacting the blade, and the velocity of the section being tested, curves showing the coating's failure can be obtained, as illustrated in Figure 3. The curve fitting used in this case follows Equation 2.

$$H = C_1 \cdot v(r)^{-C_2} \tag{2}$$

being H, the accumulated rain impingement to erosion failure and C_1, C_2 material parameters calibrated using experimental WARER test data for a specific protection system and $v(r)$ the local rotor speed. For this study, damage evolution is assumed to be linear, as assumed in other relevant works related to LEE in the literature [20] and following the experimental behaviour reported by [37], and damage accumulation calculated using the Palmgren-Miner rule, Equation 3. In this work, damage has been accumulated using average 10-min wind speed and rain data. The study of the influence of more granular data has not been part of this study, but the authors believe that the granularity used in this study can be considered representative for the lifetime analysis of the turbines.

$$\Delta d = \frac{h_i}{C_1 \cdot v(r)^{-C_2}} \tag{3}$$

with h_i being the accumulated rain impingement during time-step i.

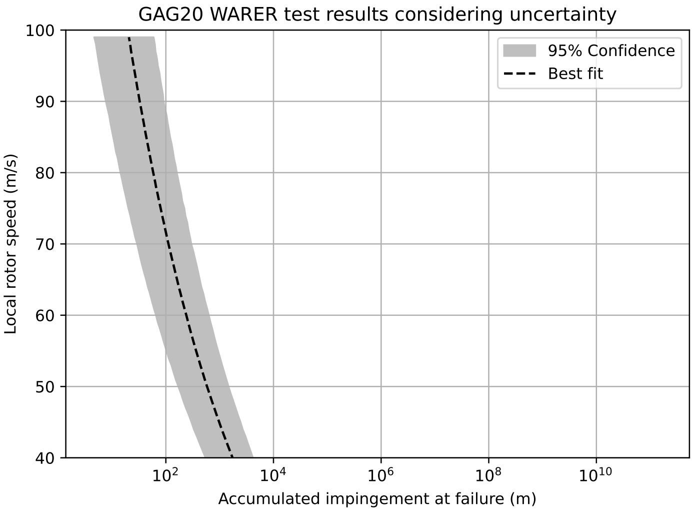


Figure 3: Accumulated impingement at failure for the GAG20 coating

In the literature, the evolution of LEE damage is typically described using a five-stage framework, which is described in Table 1. In this study, a continuous damage parameter, denoted as d, is defined within the interval [0,1], allowing for the representation of the damage severity across these stages. Figure 4 illustrates the mapping of these stages to the damage levels within the [0,1] range, providing a clear visualisation of how different damage severity levels are associated with the stages of LEE damage evolution.

Table 1: LEE stages

Stage	Description
0	Incubation stage
1	Formation of minor pits
2	Growth of pits
3	Partial removal of topcoat
4	Total removal of topcoat and initiation of delaminations

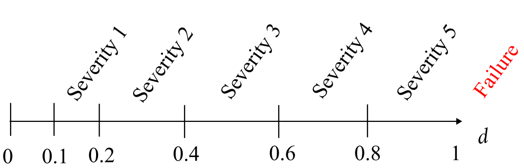


Figure 4: Damage, d, assigned to different damage severity categories

The proposed framework operates on two different timescales: one for computational modelling (computational time step) and another for decision-making (decision time step). In this study, the computational time step is set to 10 minutes, while the decision time-step is 1 calendar month. To mimic real-world conditions, the agent operates without prior knowledge of the model but relies on observations. The agent's state representation at each time step includes the parameters presented in Table 2.

Table 2: RL Agent state parameters

Parameter	Description	Range
Time from last maintenance, t_{lm}	Represents the last time a successful maintenance was performed	≥ 0
Time until decommissioning, t_{td}	Remaining time of the life of the turbine	$[0, 300]$
Estimated maximum damage, D_{max}	Maximum level of damage of the turbine as estimated through the model and updated through inspection data when available	$[0, 1]$
Current calendar month	Calendar month	$[1, 12]$
Average annual erosion rate, a_d	Prognostic feature for the agent representing the average annual erosion rate expected the turbine given the information available	≥ 0

At each decision step, the RL agent is presented with three possible actions: continue operating normally with no maintenance activities, attempt inspection, and attempt repair. The variable D_{max} is updated at each decision time step using the average annual erosion rate, unless new maintenance information is acquired. When new maintenance data, denoted as D_{ins}, becomes available, D_{max} is updated using the equation below:

$$D_{max} \rightarrow \frac{D_{max} + D_{ins}}{2} \tag{4}$$

The average annual erosion rate is initially set at 0.3, representing the average rate for the coating and the specific study site. Whenever new inspection data becomes available, a_d is updated using a weighted average, where the weights are proportional to the time between inspections. Greater weight is assigned to inspection data collected over longer intervals.

2.2. Decision-making framework

The decision-making process is executed by an agent trained using Reinforcement Learning. In this context, the agent is trained by interacting with the *environment*, receiving rewards and penalties to maximise a *reward* signal R. The problem is framed as a Markov Decision Process (MDP). Using this formalism, at each time step t, the agent receives some representation of the environment's *state*, $S_t \in \mathcal{S}$, and selects an *action*, $A_t \in \mathcal{A}(s)$. In the subsequent time step, the agent receives a numerical *reward*, $R_{t+1} \in \mathcal{R} \subset \mathbb{R}$ and receives the representation of the new state of the environment, S_{t+1}. In an MDP, the dynamics of the *environment* (S_t, R_t) are entirely characterised by the dynamics function $p(S, A)$ that depends only on the immediately preceding state and action (S_{t-1}, A_{t-1}).

$$p(s', r \mid s, a) \doteq \Pr\{S_t = s', R_t = r \mid S_{t-1} = s, A_{t-1} = a\} \tag{5}$$

Therefore, the interaction between agent and environment in a finite MDP gives rise to a trajectory $\{S_0, A_0\}, \{R_1, S_1, A_1\}, \ldots, \{R_T, S_T, A_T\}$ being T the termination state. The flexibility of the MDP framework makes it ideal for modelling O&M tasks, including the one addressed in this work. The final goal of the agent in RL is the maximisation of the cumulative sum of rewards, referred to as *return* G_t, following an action:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=t+1}^{T} \gamma^{k-t-1} R_k \tag{6}$$

being $\gamma \in [0, 1]$ the discount factor used in continuous task problems, where $T = \infty$, to avoid the potential issue of G_t approaching infinity. For finite episodic tasks, γ shall be taken as 1 to avoid suboptimal solutions in the optimisation of G. However, reducing γ can aid in stabilising the training process and encourage riskier decision-making [38]. By doing this, the agent increases the importance of the rewards and shorter time horizons and can have a target with a lower variance. This can be of great importance in high uncertainty scenarios such as the problem analysed in this work. To assess the preference for different actions in a given state, the agent utilises *value functions* or *action-value functions*. The *action-value function* of a state s under a policy π, denoted $q_\pi(s, a)$, is defined as follows:

$$q_\pi(s, a) \doteq \mathbb{E}_\pi[G_t \mid S_t = s, A_t = a] \tag{7}$$

The optimal value function $q^*(s, a)$ provides the maximum values in all states and can be determined by solving the Bellman equation:

$$q^*(s, a) = \mathbb{E}[R(s, a) + \gamma \sum_{s'} P(s' \mid s, a) \max_{a'} q^*(s', a')] \tag{8}$$

the optimal policy π^* is then constructed by following:

$$\pi^*(s) = \arg\max_a q^*(s,a) \tag{9}$$

To achieve the optimal policy, one of the strategies is to make use of the ε-*greedy* policy, which can be expressed as follows:

$$A_t = \begin{cases} \arg\max_a q^*(s,a) & \text{with probability } 1-\varepsilon \\ A \in \mathcal{A}(S_t) & \text{with probability } \varepsilon \end{cases} \tag{10}$$

where the agent balances the exploration, $\arg\max_a q^*(s,a)$ with the exploration, *random* action, by utilising the exploration rate, $\varepsilon \in [0,1]$. Typical approaches consider a decaying exploration rate over time to explore more intensively the state space frequented by the best-known policy to the agent. In this case, the update rule for the exploration is as follows:

$$\varepsilon_i = \varepsilon_0 + (\varepsilon_f - \varepsilon_0) \cdot \frac{\min(i,f)}{f} \tag{11}$$

where i is the step, ε_0 the initial learning rate, and ε_f the final learning rate. The values used were 0.6, 0.03 and 10^5 for ε_0 , ε_f and f, respectively.

Given the nature of the problem at hand, Temporal Difference (TD) learning methods, in which the values are updated online based on the difference of temporally successive estimates, can be beneficial. In this case, the method chosen to solve the problem is *Q-learning* [39]. *Q-learning* is an off-policy TD method used to find the action-value function of the states to find the optimal or nearly optimal policy. To address this problem, Deep Q Networks (DQN) are used for function approximation. The value $q(s,a)$ is approximated as $\hat{q}_\pi(s,a,w) \approx q_\pi(s,a)$, where w represents the set of weights for the DQN. This approach was chosen to improve the generalisation of the agent and better approach different regions of the state space given the continuous value of the damage state and the large state-action space of the problem. To use this method, two separate networks need to be kept, one called the online or behaviour network with weights w, which is the one being updated every step, and the target network, which shares architecture with the first but has a different weight vector w^- that is updated less frequently. In the agent's design, the weight vector update frequency C is set to 10^4 steps (months). The adoption of this approach, along with the use of the experience replay buffer M, help break the correlation of the sequence and stabilise the training of the agent. Throughout the learning process, Q-learning updates are applied to minibatches extracted from the experience replay, following the equation below:

$$w_{t+1} \leftarrow w_t + \alpha \frac{1}{N} \sum_{i=1}^{N} [R_i + \gamma \arg\max_{a_i} \hat{q}(s_i', a_i'; w_t^-) - \hat{q}(s_i, a_i; w_t)] \cdot \nabla w \cdot \hat{q}(s_i, a_i; w_t) \tag{12}$$

where the subindex i is used to denote the sample in the batch, t is the time index at which the weights are updated and ∇w the gradient of the weights. Here, α represents the learning rate, and N is the number of samples in the minibatch. The chosen size of the minibatch for the RL agent solving the LEE degradation O&M optimisation problem is 128. The weights learnt by the agent approximate the optimal state-action function $q^*(s,a)$ regardless of the followed policy. Then, the agent can approximate the optimal policy π^* by choosing the action with the greatest state-action value:

$$\widehat{\pi^*} = \arg\max_a \hat{q}^*(s,a;w) \approx \pi^* \tag{13}$$

The experiences from the replay buffer are not sampled uniformly but by a priority, P, assigned on its importance, using what is termed as prioritised replay buffer [40]. When stored in the replay buffer, each experience is assigned a priority based on its TD-error, creating what is termed a *prioritised replay buffer* [40]. These priorities are then used to calculate a probability distribution for sampling, which has been calculated as:

$$p_k = \frac{P(k)^\alpha}{\sum_{j=1}^{N} P(j)^\alpha} \tag{14}$$

With α as a parameter emphasising higher probabilities, p_k as the sampling probability of experience k, and N as the size of the experience replay buffer, sampling weights, denoted as w_s, are used to compensate for the bias introduced by the sampling probability distribution. These weights are calculated using the following expression:

$$w_{s_k} = \left(\frac{1}{N} \cdot \frac{1}{P(k)}\right)^\beta \tag{15}$$

During the training of the agent, the loss calculated for each experience is weighted by w_s to increase the importance of experiences with higher priorities. In this case, values of 0.6 and 0.4 were used for the parameters α and β, respectively.

The Deep Neural Network used is a fully connected network composed of three hidden layers with 300, 600, and 150 units, respectively, and it employs the $ReLU$ activation function. The output layer provides the state-action value, $\hat{q}(s,a;w)$, for each of the actions available for the agent. The activation function for the output layer is linear, allowing the network to provide negative q-values, as expected for the rewards of the environment. The optimisation algorithm chosen for training the network is ADAM [41], using a fixed learning rate, α, of 0.0001. The reward function defined for this problem is shown in Equation 16. The reward is composed of 3 terms, the aerodynamic losses, C_{aero}, the maintenance costs, C_{om}, and the downtime costs, C_{dt}. C_{aero} is computed as the difference in production between the original and the eroded power curves of the turbine, C_{om} using the costs provided in Table 3 and C_{dt} as energy lost during downtime. Maintenance costs are obtained following the procedure depicted in Figure 5. This function is defined to produce rewards ≤ 0 for which a zero initialisation of Q-values will encourage exploration. The algorithm outlining the training of the RL agent is depicted in Figure 6 and outlined in Algorithm 1.

$$R_i = C_{aero} - C_{om} - C_{dt} \tag{16}$$

Table 3: Repair costs per damage severity - 3 blades. Data obtained from [42] and [43]. m_b, m_a and m_e are the booking, access and execution costs, respectively.

Damage severity	m_b(£)	m_a(£)	m_e(£)
0 (Inspection)	1,600	1,000	3,200
1	2,000	1,000	4,000
2	2,000	1,000	4,000
3	3,000	1,000	6,000
4	5,000	1,000	36,000
5	0	250,000	3,500,000
6	0	250,000	5,000,000

Table 4: Weather repair constraints.

Damage category	Logistic requirements	Duration (h)	Max. significant wave height (m)	Max 10-min avg. wind speed (m/s)
1: LE discoloration, paint or bugs	CTV, rope access	6	1.5	11
2: Coat/paint damage, surface: Missing less than 10 cm^2	CTV, rope access	15	1.5	11
3: Coat/paint damage, surface: Missing more than 10 cm^2 Damaged leading edge protection Damaged leading edge tape LE erosion, down to laminate	CTV, rope access	18	1.5	11
4: LE erosion, down to laminate and first layer laminate	CTV, crawler platform	40	1.5	12
5: LE erosion, through laminate / Open LE	HLV, blade dissassembly	72	1.8	10
6: LE erosion, blade failure	HLV, blade disassembly	72	1.8	10

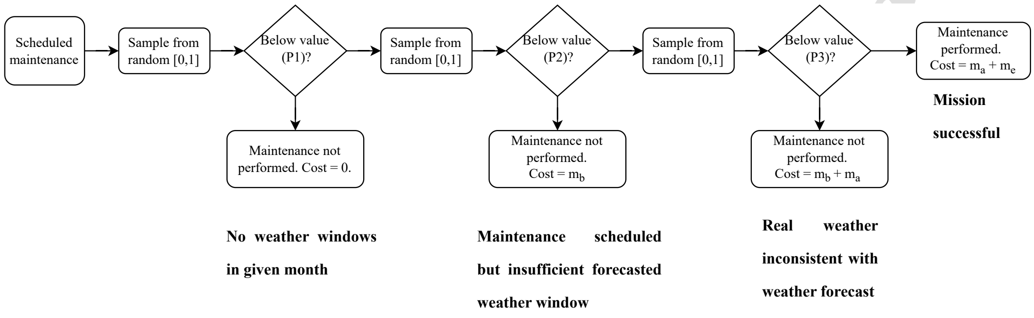


Figure 5: Repair modelling.

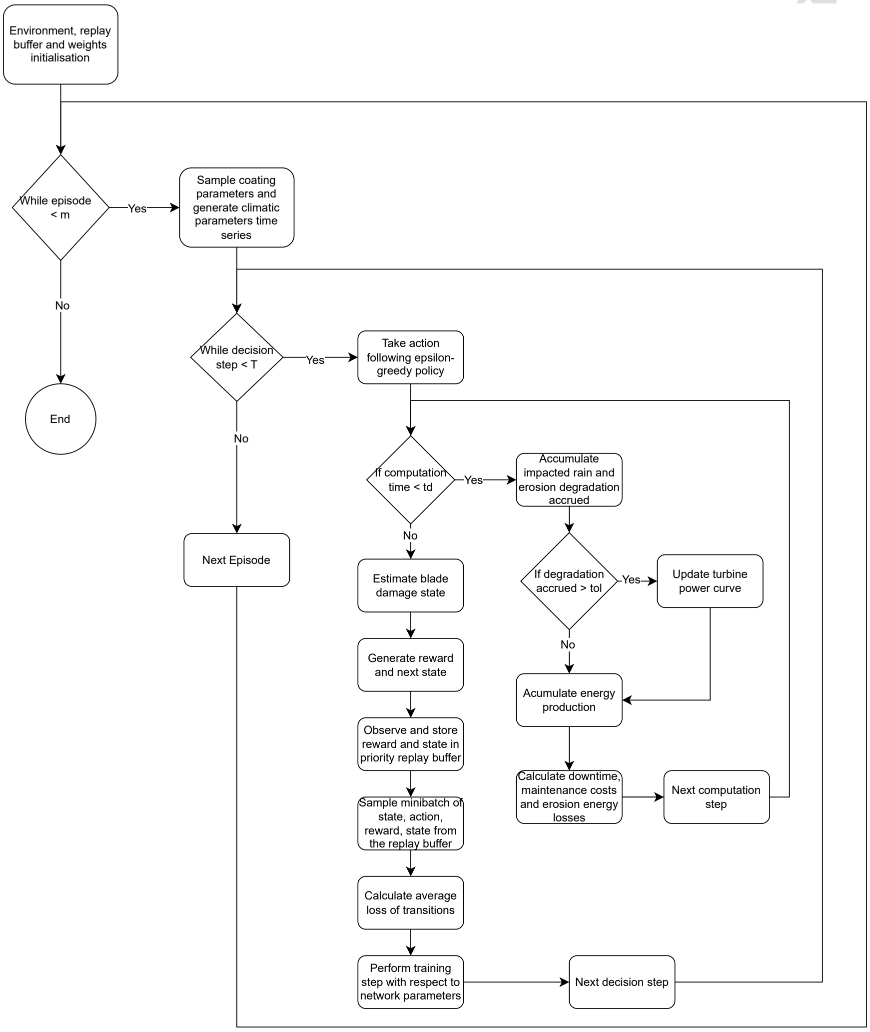


Figure 6: Leading edge erosion RL algorithm flowchart

Algorithm 1 Deep Q-learning for wind turbine blade LEE O&M optimisation with experience replay buffer

Initialise priority replay buffer M to capacity N
Initialise action-value function $\hat{q}$ with random weights w
Initialise action-value function $\hat{q}$ for target network with weights $w^{-} = w$
Environment initialisation ▷ Wind turbine definition, Blade degradation power curves, weather data, maintenance success probabilities
Generate k transitions to pre-fill M using a random policy
for episode $= 1$, m **do**
    Reset environment, $s = s_0$
    Generate random material coating parameters C_1, C_2 and weather scenarios $I(t), u(t)$.
    for decision step $t_d = 1$,T **do**
        With probability ε select a random action a_t
        otherwise select $a_{t_d} = \arg\max_a \hat{q}^*(s, a; w)$
        execute action a_{t_d}
        while computation time $t_c \leq t_d$ **do**
            Accumulate impacted rain.
            Calculate real erosion degradation accrued, Δd.
            if $\Delta d \geq tol$ **then**
                Update turbine power curve due to erosion degradation.
            Accumulate energy production, ΔE.
            Calculate downtime and maintenance costs C_{dt} and C_{om}.
            Calculate erosion energy losses.
        Estimate blade damage state, D_{max}
        Generate reward r_{t_d+1} and next state s_{t_d+1}
        Observe r_{t_d+1} and s_{t_d+1}
        Store transition $(s_{t_d}, a_{t_d}, r_{t_d+1}, s_{t_d+1})$ in M
        Sample minibatch of transitions $(s_j, a_j, r_{j+1}, s_{j+1})$ from M
        Calculate average loss of transitions
        Perform training step with respect to network parameters w
        Every C steps reset $w^{-} = w$

3. O&M considerations

For the O&M simulations, the following assumptions were considered:

- Only the O&M costs related to blade damage due to LEE are considered in this study; no other failure modes are taken into account.
- Turbine operation is assumed to commence at the beginning of January.
- Imperfect repairs are considered, where the true damage state of each calculation point, denoted as d, is set to a value drawn from a Gaussian distribution with $d \sim \mathcal{N}(\mu, \sigma^2)$, where $\mu = 0.05$ and $\sigma = 0.001$, and truncated at the interval $[0, 1]$.
- Imperfect inspections are also considered, with inspected damage denoted as D_{ins}. Inspected damage follows a Gaussian distribution with parameters $\mu = d$ and $\sigma = 0.1$, truncated within the interval $[0, 1]$.
- If any real damage calculation point on the blade reaches $d = 1$, the turbine will be preventively stopped until it undergoes repair or replacement. This study assumes that when the blade reaches this degradation level, alarms from other systems such as SCADA will trigger the preventive shutdown.
- An energy cost of 50 GBP/MWh is considered, in line with the Contracts for Difference (CfD) strike price signed for CfD4 in the UK in 2022.
- Probabilistic definitions of repair success are discredited by month to mimic real O&M scheduling. The associated cost of a repair is a function of the damage and the month when the repair is attempted.
- For condition-based maintenance strategies, referred to as AC, repairs are attempted upon reaching an estimated damage, D, above a specified damage threshold.
- Energy production losses resulting from reduced aerodynamic performance of the blade due to erosion are considered, following the calculation framework outlined in [24] and summarised in this study.
- Energy production losses stemming from downtime and preventive stops are also taken into account.
- Maintenance costs are as specified in Table 3.
- Inspections are mandated for all maintenance strategies during the early operation phase of the turbine, specifically during months 3 to 6, to ensure more stable results that closely resemble real-life operations.
- For this study, inspection costs used were assumed as a deterministic values. Notwithstanding, the proposed framework can accommodate a probabilistic description of inspection costs for different damage levels or inspection techniques.

This study models the maintenance success rate for a maintenance mission in three sequential steps as shown in Figure 5. First, it considers the probability of a given month to have wind and significant wave height values below the threshold, denoted as P_1. Second, it evaluates the probability of the forecasted weather complying with a required weather window, known as P_2. Finally, it assesses the probability of the actual weather matching the weather predictions, labelled as P_3. The weather constraints for different maintenance methods and the required weather windows are provided in Tables 4 and 3. Synthetic weather data generation techniques, as described earlier, are used to obtain values for P_1 and P_2. In the absence of data, real weather is assumed to deviate from the forecast with increasing uncertainty. For the calculation of P_3 values, a Gaussian distribution is employed, centred on the forecast value, with a standard deviation increasing by 4% daily.

4. Case studies

To assess the effectiveness of the proposed framework, two case studies were conducted. Both cases share the same location and turbine model but differ in terms of maintenance success probabilities. In Case 2, there is a lower maintenance success rate and a more pronounced seasonal influence, resulting in a higher difference between the success rates during spring-summer and autumn-winter months. These probabilities are presented in Appendix A.

For these cases, the O&M costs related to leading-edge erosion were analysed under condition-based maintenance policies, *AC*, and the policies generated by the RL agents. Two AC policies were selected as baselines for comparison with the performance of RL agents: *AC 0.4* and *AC 0.3*. These *AC* policies initiate maintenance when the blade reaches 0.4 and 0.3 D_{max}, respectively. The results are analysed and compared in terms of several aspects, including the average estimated damage throughout the turbine's lifetime, the estimated damage when maintenance is attempted, the evolution of the frequency of maintenance activities over time, the average time between maintenance actions in relation to the estimated annual damage rate, repair frequency per calendar month and the percentage of O&M actions taken. Finally, a thorough cost analysis based on a number of cost metrics is shown to compare the analysed policies.

Both case studies are situated at the FINO1 platform, located 45 km off the coast of Germany. The 5MW NREL offshore wind turbine serves as the model for simulating these scenarios, with its characteristics detailed in Table 5.

Table 5: 5 MW NREL Turbine data. Data extracted from [44]

Property	Value
Rated power	5 MW
Control	Variable speed, collective pitch
Drivetrain	High speed, multiple-stage gearbox
Rotor diameter	126 m
Hub height	90 m
Cut-In / Rated / Cut-out wind speed	3 m/s / 11.4 m/s / 25 m/s
Cut-in / Rated rotor speed	6.9 rpm, 12.1 rpm
Rated tip speed	80 m/s

For these case studies, a training period of 10^6 months was employed for training the RL agents. Simulations were conducted using γ values ranging from 0.95 to 1 at intervals of 0.01. The two best-performing agents are compared with condition-based maintenance strategies featuring damage repair thresholds of 0.3 and 0.4.

Both condition-based and RL maintenance strategies utilised updates in the estimated maximum damage, D_{max}, and the average annual erosion rate, a_d, to estimate the blade's condition. To evaluate the results of the various O&M strategies, 5,000 simulations spanning 25 years each were performed.

When assessing risk, the expected O&M cost value must be supplemented with additional metrics. Therefore, the policies will be compared based on the following metrics: Conditional Value at Risk ($CVaR_{0.95}$), which represents the average of values above the 95th percentile; the median; the expected cost (mean); and Value at Risk ($VaR_{0.95}$).

4.1. Case study 1

In this case, the maintenance success probabilities shown in were derived from direct simulation considering the weather repair constraints shown in Tables A.1 to A.3 and the synthetic weather generation explained under section 2.1.

In this subsection, we present and analyse the results of Case Study 1. Figure 7 provides a summary of the most relevant aspects of the different policies. Figure 7a illustrates the evolution of the average maximum blade LEE damage over time. At the start of the operation, the 90^{th} percentile damage approaches the damage threshold of AC strategies. The periodic waviness in the data series is attributed to the seasonality of maintenance success probability, with a period of 12 months, and the distinct strategies employed for maintenance scheduling. AC strategies exhibit a more regular damage pattern compared to RL strategies.

It's worth noting that RL strategies tend to utilise most of the LEE's lifespan before decommissioning. This tendency is more pronounced in the case of the *RL CS1* $\gamma = 0.98$ RL agent.

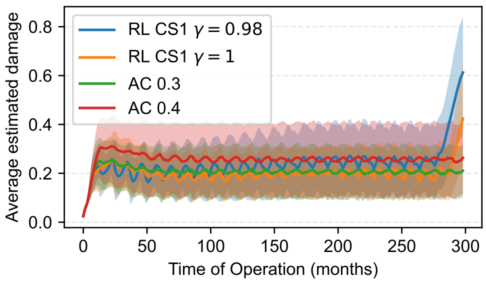
Figure 7b displays the distribution of D_{max} for the maintenance attempts of the different strategies. *AC* strategies follow an exponential-like distribution with peaks at their respective damage thresholds (0.3 and 0.4), which decrease with the success of maintenance activities. Conversely, the RL agents employ different strategies. *RL CS1* $\gamma = 1$ demonstrates a Gaussian distribution with a mean of 0.3, while *RL CS1* $\gamma = 0.98$ shows a wider Gaussian-like distribution with a mean around 0.35.

The frequency of attempted repair activities over the turbine's service life is presented in Figure 7c. *AC* strategies maintain a constant maintenance rate throughout the turbine's life, whereas *RL* strategies tend to accumulate more maintenance activities at the beginning of their service life and reduce them as decommissioning approaches. This trend is more pronounced in the *RL CS1* $\gamma = 0.98$ policy but is also evident in the *RL CS1* $\gamma = 1$ policy.

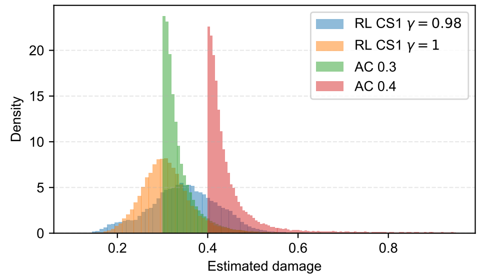
Figure 7d illustrates the average time between maintenance actions for the different policies analysed. *AC 0.4* shows the longest time between maintenance actions for all values of a_d. As expected, the time between maintenance actions for this policy is greater than *AC 0.3*. RL agents adopt different approaches, with *RL CS1* $\gamma = 0.98$ being closer to *AC 0.3*, while *RL CS1* $\gamma = 1$ follows a safer strategy for $a_d \leq 0.4$.

Figure 7e provides insights into the planning of maintenance activities by calendar month. It's important to note that this graph displays all maintenance attempts, not just the successful ones. *AC* policies show a curve with lower values in the months of April to October, with similar shapes and values. This is because maintenance success probabilities are higher during those months, reducing the need for maintenance actions in the coming months. In contrast, *RL* policies exhibit a different behavior, with a significant increase in maintenance planning intentions for the period from October to February. RL agents have learned the benefit of anticipating maintenance, as failure to do so would lead to an increase in the blade's damage state and higher maintenance costs. RL policies adopt a more conservative approach in this regard compared to *AC* policies.

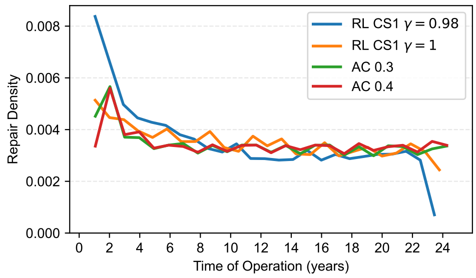
Finally, Figure 7f presents the percentage of different actions taken. Given that *AC 0.4* has a higher damage repair threshold, it's expected that the 'operate' action is more frequent (85.29% of months) compared to *AC 0.3* (82.77%). The fixed inspections scheduled for all policies remain at 1.34%, with RL agents showing a marginal increase in the use of inspections (2.00% and 2.22% for *RL CS1* $\gamma = 1$ and *RL CS1* $\gamma = 0.98$, respectively). *RL CS1* $\gamma = 1$ employs the highest maintenance intention rate (16.98%), while *RL CS1* $\gamma = 0.98$ adopts a rate of 14.98%, falling between *AC 0.3* and *AC 0.4*.



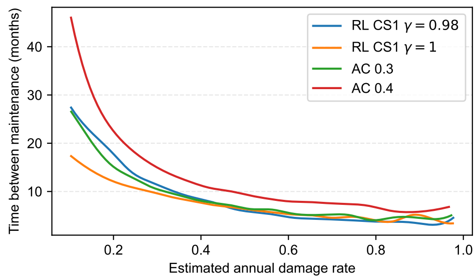
(a) Evolution of maximum estimated damage, D_{max} vs time of operation. The shadowed regions represent the 10-90 percentile band



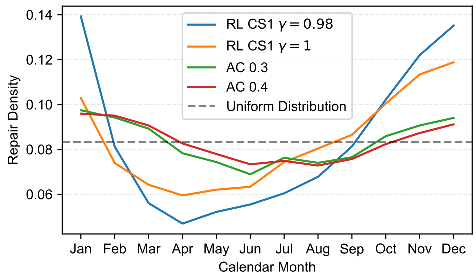
(b) Distribution of attempted repairs by estimated damage. *AC* policies have a fixed threshold while *RL* policies are left free to modify it to optimise expected O&M costs



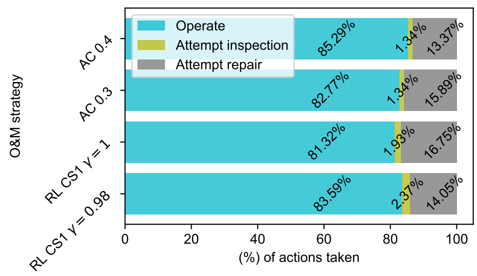
(c) Evolution of attempted repairs over the years of operation. *AC* policies have a fixed policy with forced inspections at the beginning of operation



(d) Average time between maintenance actions vs estimated annual damage rate, a_d



(e) Repair attempt frequency over calendar months. The dashed line represents the uniform maintenance planning distribution



(f) O&M actions taken by the policies

Figure 7: Case study 1 O&M policy analysis

Figures 8 and 9 display the distribution of the O&M cost for the evaluated O&M maintenance policies, while Table 6 presents various cost metrics compared to the baseline policy *AC 0.3*. Concerning cost distribution, *AC 0.4* exhibits a higher number of values at the lower end of the cost spectrum. This can be attributed to the fixed policy of *AC 0.4*, which entails some risk to the blade's condition but proves effective for scenarios involving slow damage growth.

In contrast, both *AC 0.3* and *RL CS1* $\gamma = 0.98$ show similar cost distributions, with a slight advantage in median values observed for *RL CS1* $\gamma = 0.98$. On the other hand, *RL CS1* $\gamma = 1$ outperforms in terms of the average, $CVaR_{0.95}$, and $VaR_{0.95}$ values. It presents reductions of 21.4%, 46.1%, and 8.7%, respectively, when compared to the *AC 0.3* policy, along with a marginal increase in the median value (8.5%).

RL CS1 $\gamma = 0.98$ closely resembles the behavior of *AC 0.3* by achieving a 5% reduction in median cost, with slight increases observed in the $CVaR_{0.95}$ and $VaR_{0.95}$ values. Conversely, *AC 0.4* displays a 12.1% reduction in the median value but experiences significant increases in the remaining metrics.

Table 6: Cost metrics for Case study 1

Label	Median	Average	$CVaR_{0.95}$	$VaR_{0.95}$
RL CS1 $\gamma = 0.98$	95.0%	100.7%	102.4%	105.4%
RL CS1 $\gamma = 1$	108.5%	78.6%	53.9%	91.3%
AC 0.3	100.0%	100.0%	100.0%	100.0%
AC 0.4	87.9%	192.0%	273.2%	432.8%

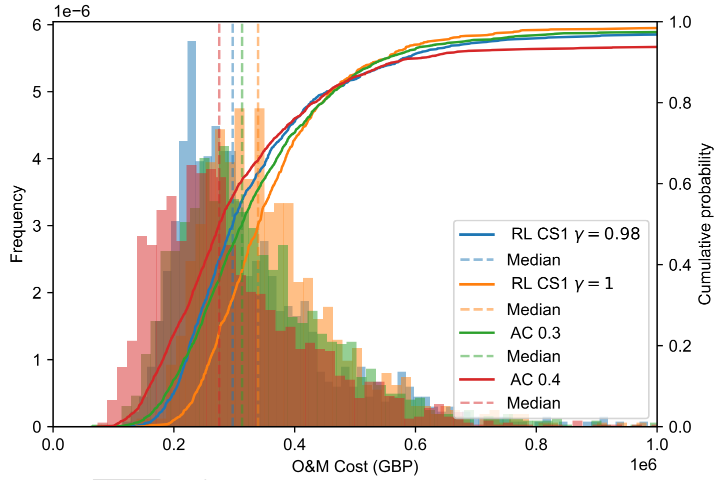
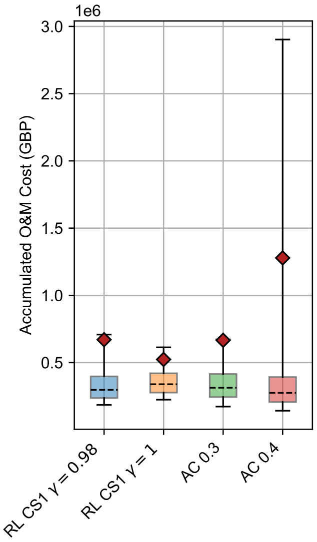
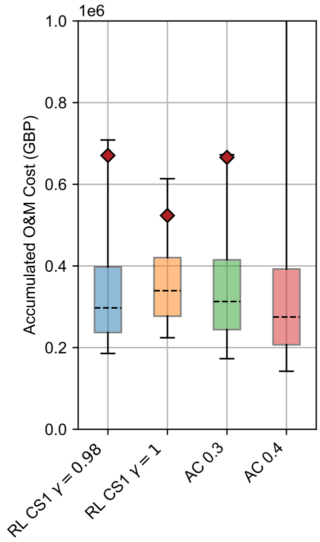


Figure 8: O&M cost distribution of the CS1 policies analysed. The dashed lines represent the median of the distribution. The right axis shows the cumulative probability of the distribution



(a) Cost distribution of CS1 O&M policies



(b) Cost distribution of CS1 O&M policies (zoomed in)

Figure 9: Cost distribution of CS1 O&M policies. The minimum and maximum values of the whiskers repesent P5 and P95, respectively and the red marker the average cost. The right plot is a zoomed in version of the one on the left.

4.2. Case study 2

In this case, the same location, turbine and costs are assumed with the main difference being the maintenance success probabilities which have been modified to show a greater seasonal influence and a lower maintenance success rate to assess the robustness of the proposed agent under more difficult conditions. The probabilities used are shown in Tables A.4 to A.6.

In this subsection, we present and analyse the results of case study 2. Figure 10 summarises the most relevant aspects of the different policies. Figure 10a illustrates the evolution of the average maximum blade LEE damage over the turbine's operational period. During the turbine's operation, the 90th percentile damage increases above the thresholds of the *AC* strategies, reaching 0.47 for *AC 0.4* and 0.39 for *AC 0.3*. The wavy pattern in the data series is attributed to the seasonality of maintenance success probability, exhibiting a periodic behaviour with a 12-month cycle, and the distinct strategies in maintenance scheduling. While *AC* strategies demonstrate a similar regularity in the damage pattern compared to RL strategies, the last 50 months of operation show a noticeable difference. RL strategies tend to make more extensive use of the blade's leading-edge erosion resistance before decommissioning. This trend is managed differently by *RL CS2* $\gamma = 0.98$ and *RL CS2* $\gamma = 0.99$, with the agent having $\gamma = 0.98$ progressively reducing the average damage.

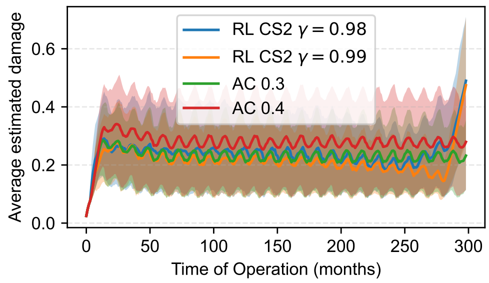
Figure 10b presents the distribution of D_{max} for the maintenance attempts of various strategies. *AC* strategies exhibit an exponential-like distribution with peaks at their respective damage thresholds (0.3 and 0.4), which decrease with the success of maintenance activities. In contrast, the RL agents adopt different strategies. *RL CS2* $\gamma = 0.99$ showcases a Gaussian-shaped distribution with a mean around 0.3, while *RL CS2* $\gamma = 0.98$ displays a more skewed distribution, peaking around 0.4.

The frequency of attempted repair activities over the turbine's service life is shown in Figure 10c. *AC* strategies maintain a consistent maintenance rate throughout the turbine's life, whereas RL strategies aim to reduce repair activities as the turbine approaches the end of its operational life. Both RL strategies exhibit a peak in maintenance activities during the final years, with year 23 for *RL CS2* $\gamma = 0.99$ and years 20-21 for *RL CS2* $\gamma = 0.98$.

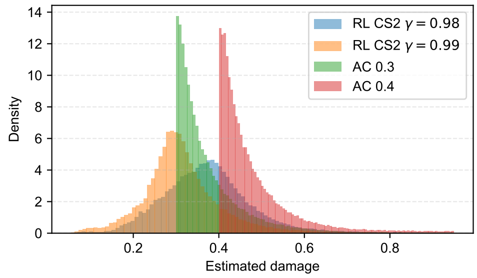
Figure 10d displays the average time between maintenance actions for the different policies analysed. *AC 0.4* shows the longest intervals between maintenance actions for all a_d. As expected, the time between maintenance actions in this policy is greater than that of *AC 0.3*. RL agents follow distinct policies, with *RL CS2* $\gamma = 0.98$ resembling the approach of *AC 0.3*, while *RL CS2* $\gamma = 0.99$ adopts a more cautious strategy for $a_d \leq 0.4$. However, *RL CS2* $\gamma = 0.99$ appears to face generalisation issues for $0.8 \leq a_d \leq 1.0$.

Figure 10e provides insight into maintenance planning by calendar month. Notably, this graph illustrates all maintenance attempts, not just the successful ones. *AC* policies and *RL CS2* $\gamma = 0.99$ exhibit a similar curve with lower values during the months from April to October. This behaviour aligns with higher maintenance success probabilities in those months, reducing the need for maintenance actions in the coming months. In contrast, the *RL CS2* $\gamma = 0.98$ policy deviates from this pattern, displaying a pronounced increase in maintenance planning from October to December.

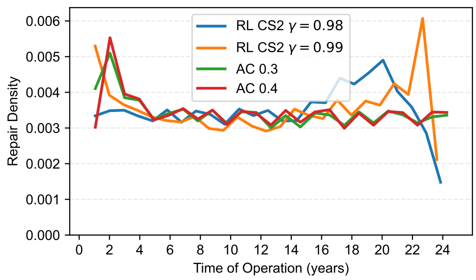
Lastly, Figure 10f presents the percentage of different actions taken. As *AC 0.4* has a higher damage repair threshold, it is unsurprising that the "operate" action is more prevalent (77.37% of months) compared to *AC 0.3* (73.23%). Fixed inspections are scheduled for all policies at a rate of 1.34%, with RL agents demonstrating an increase in inspection usage, particularly *RL CS2* $\gamma = 0.98$ (13.3%) compared to *RL CS2* $\gamma = 0.99$ (5.61%). Furthermore, *RL CS2* $\gamma = 0.99$ exhibits a slightly higher repair intention rate than *AC 0.3* (25.95% vs. 25.44%), and *RL CS2* $\gamma = 0.98$ adopts a repair attempt rate of 23.12%, positioning it between *AC 0.3* and *AC 0.4*.



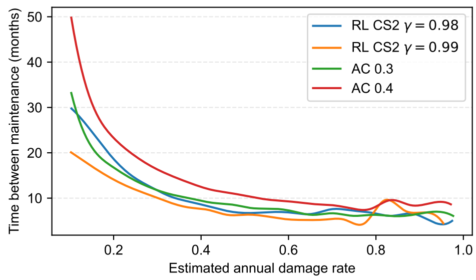
(a) Evolution of maximum estimated damage, D_{max} vs time of operation. The shadowed regions represent the 10-90 percentile band



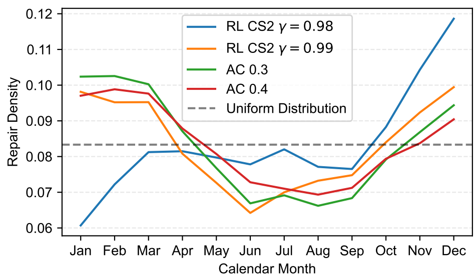
(b) Distribution of attempted repairs by estimated damage. *AC* policies have a fixed threshold while RL policies are left free to modify it to optimise expected O&M costs



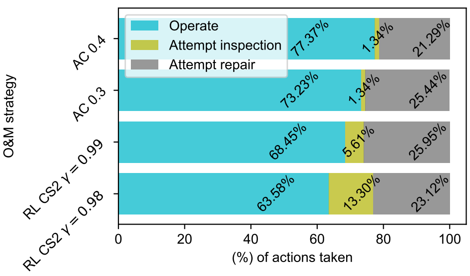
(c) Evolution of attempted repairs over the years of operation. *AC* policies have a fixed policy with forced inspections at the beginning of operation



(d) Average time between maintenance actions vs estimated annual damage rate, a_d



(e) Repair attempt frequency over calendar months. The dashed line represents the uniform maintenance planning distribution



(f) O&M actions taken by the policies

Figure 10: Case study 2 O&M policy analysis

Figures 11 and 12 display the distribution of O&M costs for the evaluated O&M maintenance policies, and Table 7 presents various cost metrics compared to the baseline policy, *AC 0.3*. Regarding cost distribution, *AC 0.4* has more values in the lower end, which can be attributed to the fixed policy of *AC 0.4* taking risks with the blade's condition and being successful for slowly growing damage cases. *AC 0.3* reaches a higher cumulative probability (0.91) at £1.5 million than *AC 0.4* (0.88), while higher values are achieved by RL policies, specifically *RL CS2* $\gamma = 0.98$ (0.948) and $\gamma = 0.99$ (0.95). In terms of cost metrics, *RL CS2* $\gamma = 0.98$ and $\gamma = 0.99$ outperform *AC 0.3*, with reductions in the range of 12-13%, 16-19%, and 73-78% for Average, $CVaR_{0.95}$, and $VaR_{0.95}$, respectively. They also exhibit a slight increase in the median value (11.5% and 6.2%, respectively). In contrast, *AC 0.4* shows a 6.2% reduction in the median value but experiences significant increases in the other metrics.

Table 7: Cost metrics for Case study 2

Label	Median	Average	$CVaR_{0.95}$	$VaR_{0.95}$
RL CS2 $\gamma = 0.98$	111.5%	86.8%	80.9%	26.9%
RL CS2 $\gamma = 0.99$	106.2%	87.3%	84.0%	21.7%
AC 0.3	100.0%	100.0%	100.0%	100.0%
AC 0.4	93.8%	159.4%	154.4%	308.5%

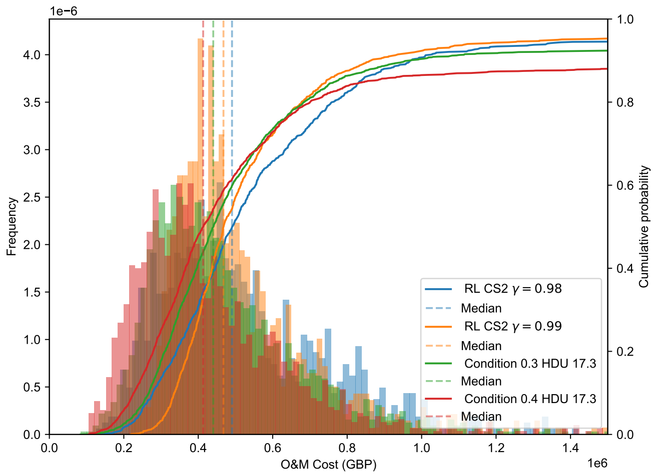
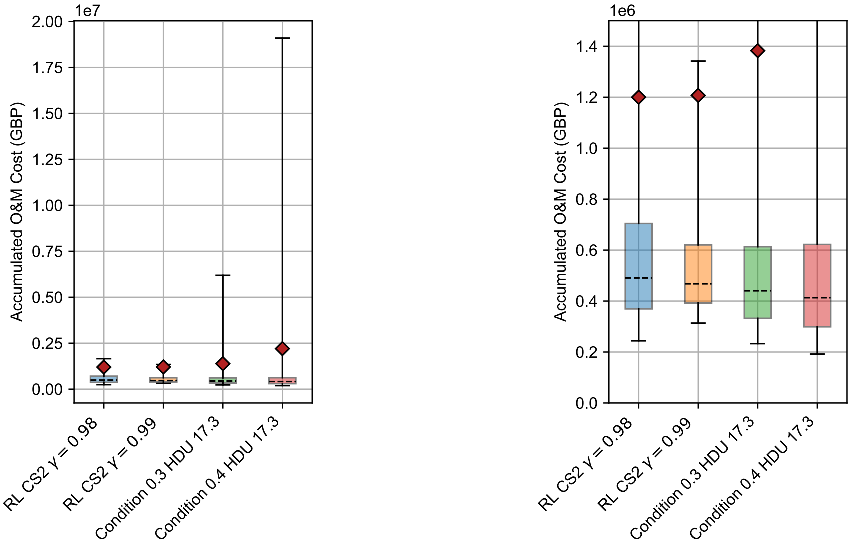


Figure 11: O&M cost distribution of the CS2 policies analysed. The dashed lines represent the median of the distribution. The right axis shows the cumulative probability of the distribution



(a) Cost distribution of CS2 O&M policies

(b) Cost distribution of CS2 O&M policies (zoomed in)

Figure 12: Cost distribution of CS2 O&M policies. The minimum and maximum values of the whiskers represent P5 and P95, respectively and the red marker the average cost. The right plot is a zoomed in version of the one on the left.

5. Discussion

The analysis of both case studies has led us to the conclusion that RL agents have been able to improve the target metric of the optimisation, which is the expected value of the O&M cost, within a certain range. In the case of CS1, with maintenance probabilities based on site-specific weather constraints, the reduction in expected (average) O&M costs was 21.4% when compared with the baseline AC 0.3 condition-based policy. Alongside the reduction in average costs, there was also a decrease in several relevant cost metrics related to risk-based decision-making, such as $CVaR_{0.95}$ and $VaR_{0.95}$ with values of 46.1% and 8.7%, respectively. The same trend was observed in CS2, an environment that has a greater uncertainty in the repair success derived from harsher climatic conditions, with reductions of 13.2%, 19.1% and 73.1% for the average, $CVaR_{0.95}$ and $VaR_{0.95}$ O&M costs. A considerable $VaR_{0.95}$ reduction is provided by the RL agent for CS2, highlighting the importance of predictive maintenance in cases of reduced maintenance accessibility of offshore assets. This expected cost reduction comes with an increase in the median cost, making the condition-based policies (AC) more cost-efficient in some cases. Additionally, γ values between 0.98 and 1.0 have proven to be the most effective in achieving this reduction. Overall, RL agents have successfully identified a cost advantage by reducing maintenance activities towards the end of the turbines' operational life. The use of inspections by RL agents has increased as maintenance success rates decreased; the inspection intention rate grew from 2.0% in CS1 to a range of 5-13% in CS2, explaining the importance of a reduced uncertainty of the damage state for low accessibility sites. Regarding maintenance planning by calendar month, RL agents did not provide a clear indication of a single planning strategy, which would require further investigation towards potential convergence issues.

The presented framework has proven to be effective in high-uncertainty scenarios, with the material parameters C_1 and C_2 having the greatest influence on the degradation dynamics. This information is valuable for the initial planning of the O&M of the turbine. To reduce the uncertainty in the degradation dynamics, the probabilistic description of the abovementioned parameters can be modified once real operation data becomes available to improve the performance of the agent. Unfortunately, the modification of the description of the stochastic variables requires the retraining of the agent, which can be time-consuming.

This framework can be used by operators at the early O&M design stage at the wind farm level. By analysing the behaviour of the best agents, important qualitative metrics can be extracted to define global policies such as the damage threshold for optimal maintenance scheduling for a particular failure mode if considered alone. If combined with additional components and failure modes, this framework can provide O&M policies at the wind turbine level. In this study, only the leading-edge erosion failure mode of the blade was considered. Nevertheless, it can be extended to accommodate different failure modes as long as a degradation function can be defined. This would require the inclusion of, at least, two parameters for the DQN per failure mode. One of the parameters would be the estimation of the state of the component and failure mode, and the other a prognostic parameter to improve the O&M planning of the agent. The selection of the failure modes to consider should be based on risk priority to provide efficiency to the framework.

In this study, material parameters C_1 and C_2 have been assumed to remain constant throughout the life of the turbine. It is important to note that there are many types of repair available (protection tapes, protective coatings, and epoxy or polyurethane fillers) the durability of which is not well known yet. An interesting opportunity to overcome this issue would be the inclusion of SHM in the turbine to provide timely inspection data. Moreover, this would reduce the cost of inspection data for low-accessibility sites, which has proven to be determinant for O&M for cost reduction. Also, there is potential for improvement in the quantification of uncertainty in the damage state and prognostic features of the agent. In the proposed definition of the RL agent, there is no quantification of the uncertainty about $D_{\max}$ and a_d made by the agent, which can be bypassed by the usage of the parameters t_{tm} and t_{td}. Another interesting direction of providing additional functionality to this framework would be the inclusion of opportunistic maintenance as an action for the agent. It would be interesting to explore the damage level at which opportunistic maintenance becomes attractive, as this is sometimes the case when unexpected failures of different components of the turbines occur.

6. Conclusion and further remarks

The proposed O&M blade LEE maintenance optimisation based on RL is able to produce an improvement in average costs in the range 12-21%, a reduction in risk of failure of the blades and reductions in $CVaR_{0.95}$ and $VaR_{0.95}$ O&M costs under this failure mode against condition-based policies. In contrast, condition-based policies can show lower median costs, and be more cost-effective in some low degradation cases. The proposed agent has highlighted the importance of a reduced uncertainty in the known condition of the blade when the opportunities for repair are fewer, with a growth from 2.0% (CS1) to 13.0% (CS2) in the scheduling of inspections. This framework has proven to be robust as to produce consistent improvements in different settings. Besides, the provided framework has the option to be re-trained with real data of different turbines of a site during operation to reduce the uncertainty in the material parameters and approximate better the degradation dynamics of this failure mode.

Notwithstanding, the high uncertainty underlying this problem sets a difficult scenario for decision-making in which the interpretability of the recommendations and the models used is key for practitioners to modify their current way of operating. Also, the need to incorporate the risk-critical failure modes to produce a common maintenance strategy calls for computationally efficient frameworks in which the logistics of the whole wind farm is considered and the opportunities for maintenance actions when not strictly required can be studied. In order to reduce the complexity of the models, a thorough understanding of the problem at hand is required, and this is why frameworks such as the proposed are required. Once there is a more profound knowledge about the dynamics of the failure mode and the relevance of different parameters modifying them, computationally efficient reduced-order models can be built for strategic wind farm decision-making. Techniques such as intelligent PN [7, 45] are promising for this last step in which the maintenance optimisation of assets in similar conditions can be jointly considered.

7. Acknowledgements

This study is part of the ENHAnCE project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 859957. Weather data was made available by the FINO (Forschungsplattformen in Nord-und Ostsee) initiative, which was funded by the German Federal Ministry of Economic Affairs and Energy (BMWi) on the basis of a decision by the German Bundestag, organised by the Projekttraeger Juelich (PTJ) and coordinated by the German Federal Maritime and Hydrographic Agency (BSH). CFD results were obtained using the ARCHIE-WeSt High Performance Computer (www.archie-west.ac.uk) based at the University of Strathclyde.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

[1] T. Stehly, P. Duffy, 2021 cost of wind energy review, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States) (2022).

[2] M. Hofmann, I. B. Sperstad, Nowicob–a tool for reducing the maintenance costs of offshore wind farms, Energy Procedia 35 (2013) 177–186.

[3] M. Shafiee, Maintenance logistics organization for offshore wind energy: Current progress and future perspectives, Renewable energy 77 (2015) 182–193.

[4] M. Arshad, B. C. O'Kelly, Offshore wind-turbine structures: a review, Proceedings of the Institution of Civil Engineers-Energy 166 (4) (2013) 139–152.

[5] D. Cevasco, S. Koukoura, A. Kolios, Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications, Renewable and Sustainable Energy Reviews 136 (2021) 110414.

[6] L. Wang, A. Kolios, X. Liu, D. Venetsanos, R. Cai, Reliability of offshore wind turbine support structures: A state-of-the-art review, Renewable and Sustainable Energy Reviews 161 (2022) 112250.

[7] A. Saleh, R. Remenyte-Prescott, D. Prescott, M. Chiachío, Intelligent and adaptive asset management model for railway sections using the ipn method, Reliability Engineering & System Safety (2023) 109687.

[8] T. Elusakin, M. Shafiee, T. Adedipe, F. Dinmohammadi, A stochastic petri net model for o&m planning of floating offshore wind turbines, Energies 14 (4) (2021) 1134.

[9] R. Yan, S. Dunnett, Improving the strategy of maintaining offshore wind turbines through petri net modelling, Applied Sciences 11 (2) (2021) 574.

[10] X. Ge, Q. Chen, Y. Fu, C. Chung, Y. Mi, Optimization of maintenance scheduling for offshore wind turbines considering the wake effect of arbitrary wind direction, Electric Power Systems Research 184 (2020) 106298.

[11] M. Li, X. Jiang, J. Carroll, R. R. Negenborn, A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty, Applied Energy 321 (2022) 119284.

[12] T. N. Schouten, R. Dekker, M. Hekimoğlu, A. S. Eruguz, Maintenance optimization for a single wind turbine component under time-varying costs, European Journal of Operational Research 300 (3) (2022) 979–991.

[13] Y. Aafif, A. Chelbi, L. Mifdal, S. Dellagi, I. Majdouline, Optimal preventive maintenance strategies for a wind turbine gearbox, Energy Reports 8 (2022) 803–814.

[14] Y. Wang, Q. Deng, Optimization of maintenance scheme for offshore wind turbines considering time windows based on hybrid ant colony algorithm, Ocean Engineering 263 (2022) 112357.

[15] G. Zou, A. Kolios, Quantifying the value of negative inspection outcomes in fatigue maintenance planning: Cost reduction, risk mitigation and reliability growth, Reliability Engineering & System Safety 226 (2022) 108668.

[16] J. Walgern, K. Fischer, P. Hentschel, A. Kolios, Reliability of electrical and hydraulic pitch systems in wind turbines based on field-data analysis, Energy Reports 9 (2023) 3273–3281.

[17] J. C. Lopez, A. Kolios, Risk-based maintenance strategy selection for wind turbine composite blades, Energy Reports 8 (2022) 5541–5561.

[18] R. Herring, K. Dyer, F. Martin, C. Ward, The increasing importance of leading edge erosion and a review of existing protection solutions, Renewable and Sustainable Energy Reviews 115 (2019) 109382.

[19] M. H. Keegan, D. Nash, M. Stack, On erosion issues associated with the leading edge of wind turbine blades, Journal of Physics D: Applied Physics 46 (38) (2013) 383001.

[20] C. Hasager, L. Mishnaevsky Jr, C. Bak, J. I. Bech, S. Fæster, N. F.-J. Johansen, How can we combat leading-edge erosion on wind turbine blades?, Danmarks Tekniske Universitet, Institut for Vindenergi, Risø Campus 4000 (2021).

[21] C. Hasager, F. Vejen, J. Bech, W. Skrzypiński, A.-M. Tilg, M. Nielsen, Assessment of the rain and wind climate with focus on wind turbine blade leading edge erosion rate and expected lifetime in danish seas, Renewable Energy 149 (2020) 91–102.

[22] C. B. Hasager, F. Vejen, W. R. Skrzypiński, A.-M. Tilg, Rain erosion load and its effect on leading-edge lifetime and potential of erosion-safe mode at wind turbines in the north sea and baltic sea, Energies 14 (7) (2021) 1959.

[23] J. I. Bech, C. B. Hasager, C. Bak, Extending the life of wind turbine blade leading edges by reducing the tip speed during extreme precipitation events, Wind Energy Science 3 (2) (2018) 729–748.

[24] J. C. López, A. Kolios, L. Wang, M. Chiachio, A wind turbine blade leading edge rain erosion computational framework, Renewable Energy 203 (2023) 131–141.

[25] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al., The era5 global reanalysis, Quarterly Journal of the Royal Meteorological Society 146 (730) (2020) 1999–2049.

[26] J. R. Norris, J. R. Norris, Markov chains, no. 2 in 1, Cambridge university press, 1998.

[27] A. Sareen, C. A. Sapre, M. S. Selig, Effects of leading edge erosion on wind turbine blade performance, Wind Energy 17 (10) (2014) 1531–1542.

[28] R. Kyle, F. Wang, B. Forbes, The effect of a leading edge erosion shield on the aerodynamic performance of a wind turbine blade, Wind Energy 23 (4) (2020) 953–966.

[29] L. Mishnaevsky Jr, C. B. Hasager, C. Bak, A.-M. Tilg, J. I. Bech, S. D. Rad, S. Fæster, Leading edge erosion of wind turbine blades: Understanding, prevention and protection, Renewable Energy 169 (2021) 953–969.

[30] A. Castorrini, A. Corsini, F. Rispoli, P. Venturini, K. Takizawa, T. E. Tezduyar, Computational analysis of performance deterioration of a wind turbine blade strip subjected to environmental erosion, Computational Mechanics 64 (4) (2019) 1133–1153.

[31] F. Papi, G. Ferrara, A. Bianchini, Practical considerations on wind turbine blade leading edge erosion modelling and its impact on performance and loads, in: Journal of Physics: Conference Series, Vol. 1618, IOP Publishing, 2020, p. 052005.

[32] M. Schramm, H. Rahimi, B. Stoevesandt, K. Tangager, The influence of eroded blades on wind turbine performance using numerical simulations, Energies 10 (9) (2017) 1420.

[33] D. Eisenberg, S. Laustsen, J. Stege, Wind turbine blade coating leading edge rain erosion model: Development and validation, Wind Energy 21 (10) (2018) 942–951.

[34] D. C. Maniaci, E. B. White, B. Wilcox, C. M. Langel, C. van Dam, J. A. Paquette, Experimental measurement and cfd model development of thick wind turbine airfoils with leading edge erosion, in: Journal of Physics: Conference Series, Vol. 753, IOP Publishing, 2016, p. 022013.

[35] M. Carraro, F. De Vanna, F. Zweiri, E. Benini, A. Heidari, H. Hadavinia, Cfd modeling of wind turbine blades with eroded leading edge, Fluids 7 (9) (2022) 302.

[36] F. Papi, L. Cappugi, S. Perez-Becker, A. Bianchini, Numerical modeling of the effects of leading-edge erosion and trailing-edge damage on wind turbine loads and performance, Journal of Engineering for Gas Turbines and Power 142 (11) (2020) 111005.

[37] G. S. Springer, Erosion by liquid impact, Osti (1 1976).

[38] V. François-Lavet, R. Fonteneau, D. Ernst, How to discount deep reinforcement learning: Towards new dynamic strategies, arXiv preprint arXiv:1512.02011 (2015).

[39] C. J. C. H. Watkins, Learning from delayed rewards (1989).

[40] T. Schaul, J. Quan, I. Antonoglou, D. Silver, Prioritized experience replay, arXiv preprint arXiv:1511.05952 (2015).

[41] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[42] J. S. Nielsen, D. Tcherniak, M. D. Ulriksen, A case study on risk-based maintenance of wind turbine blades with structural health monitoring, Structure and Infrastructure Engineering 17 (3) (2021) 302–318.

[43] Y. Yi, J. D. Sørensen, Reduction of operation and maintenance cost for wind turbine blades–reliability model (2019).

[44] J. Jonkman, S. Butterfield, W. Musial, G. Scott, Definition of a 5-mw reference wind turbine for offshore system development, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States) (2009).

[45] A. Saleh, M. Chiachío, J. F. Salas, A. Kolios, Self-adaptive optimized maintenance of offshore wind turbines by intelligent petri nets, Reliability Engineering & System Safety 231 (2023) 109013.

Appendix A. Repair success probabilities

Table A.1: CS1 P_1 probabilities. The first row represents the damage severity

	0 (Inspection)	1	2	3	4	5	6
Jan	0.6614	0.6614	0.6614	0.6614	0.6614	0.3665	0.3665
Feb	0.7075	0.7075	0.7075	0.7075	0.7075	0.4052	0.4052
Mar	0.7194	0.7194	0.7194	0.7194	0.7194	0.4138	0.4138
Apr	0.8004	0.8004	0.8004	0.8004	0.8004	0.4807	0.4807
May	0.8138	0.8138	0.8138	0.8138	0.8138	0.4812	0.4812
Jun	0.8533	0.8533	0.8533	0.8533	0.8533	0.5326	0.5326
Jul	0.8663	0.8663	0.8663	0.8663	0.8663	0.5356	0.5356
Aug	0.8388	0.8388	0.8388	0.8388	0.8388	0.5083	0.5083
Sep	0.7908	0.7908	0.7908	0.7908	0.7908	0.4722	0.4722
Oct	0.7169	0.7169	0.7169	0.7169	0.7169	0.3162	0.3162
Nov	0.6880	0.6880	0.6880	0.6880	0.6880	0.3813	0.3813
Dec	0.6605	0.6605	0.6605	0.6605	0.6605	0.3841	0.3841

Table A.2: CS1 P_2 probabilities. The first row represents the damage severity

	0 (Inspection)	1	2	3	4	5	6
Jan	0.8444	0.7615	0.7243	0.6891	0.4624	0.1000	0.1000
Feb	0.8653	0.7925	0.7595	0.7281	0.5264	0.1000	0.1000
Mar	0.8832	0.8186	0.7892	0.7611	0.5715	0.1000	0.1000
Apr	0.9071	0.8544	0.8298	0.8062	0.6418	0.1000	0.1000
May	0.9070	0.8556	0.8317	0.8088	0.6483	0.1000	0.1000
Jun	0.9191	0.8728	0.8514	0.8307	0.6846	0.1000	0.1000
Jul	0.9221	0.8772	0.8514	0.8356	0.6921	0.1000	0.1000
Aug	0.8945	0.8369	0.8103	0.7849	0.6118	0.1000	0.1000
Sep	0.8912	0.8314	0.8037	0.7772	0.5964	0.1000	0.1000
Oct	0.8442	0.7597	0.7216	0.6856	0.4571	0.1000	0.1000
Nov	0.8303	0.7409	0.7006	0.6624	0.4264	0.1000	0.1000
Dec	0.8412	0.7576	0.7198	0.6840	0.4567	0.1000	0.1000

Table A.3: CS1 P_3 probabilities. The first row represents the damage severity

	0 (Inspection)	1	2	3	4	5	6
Jan	0.9614	0.9414	0.9309	0.9191	0.8066	0.1000	0.1000
Feb	0.9613	0.9409	0.9302	0.9196	0.8124	0.3930	0.3779
Mar	0.9680	0.9510	0.9417	0.9321	0.8387	0.1000	0.1000
Apr	0.9703	0.9538	0.9449	0.9352	0.8432	0.4560	0.4560
May	0.9708	0.9550	0.9463	0.9374	0.8502	0.4124	0.4124
Jun	0.9666	0.9481	0.9383	0.9281	0.8320	0.2432	0.2571
Jul	0.9751	0.9606	0.9383	0.9446	0.8645	0.3236	0.2991
Aug	0.9689	0.9521	0.9433	0.9342	0.8447	0.6747	0.6898
Sep	0.9703	0.9545	0.9459	0.9369	0.8510	0.2917	0.2917
Oct	0.9590	0.9353	0.9223	0.9095	0.7857	0.1000	0.1000
Nov	0.9630	0.9425	0.9316	0.9199	0.8057	0.1000	0.1000
Dec	0.9690	0.9534	0.9447	0.9359	0.8492	0.1000	0.1000

Table A.4: CS2 P_1 probabilities. The first row represents the damage severity

	0 (Inspection)	1	2	3	4	5	6
Jan	0.4374	0.4374	0.4374	0.4374	0.6614	0.3665	0.3665
Feb	0.5006	0.5006	0.5006	0.5006	0.7075	0.4052	0.4052
Mar	0.5175	0.5175	0.5175	0.5175	0.7194	0.4138	0.4138
Apr	0.6406	0.6406	0.6406	0.6406	0.8004	0.4807	0.4807
May	0.6622	0.6622	0.6622	0.6622	0.8138	0.4812	0.4812
Jun	0.7282	0.7282	0.7282	0.7282	0.8533	0.5326	0.5326
Jul	0.7504	0.7504	0.7504	0.7504	0.8663	0.5356	0.5356
Aug	0.7036	0.7036	0.7036	0.7036	0.8388	0.5083	0.5083
Sep	0.6253	0.6253	0.6253	0.6253	0.7908	0.4722	0.4722
Oct	0.5140	0.5140	0.5140	0.5140	0.7169	0.3162	0.3162
Nov	0.4733	0.4733	0.4733	0.4733	0.6880	0.3813	0.3813
Dec	0.4362	0.4362	0.4362	0.4362	0.6605	0.3841	0.3841

Table A.5: CS2 P_2 probabilities. The first row represents the damage severity

	0 (Inspection)	1	2	3	4	5	6
Jan	0.7130	0.5799	0.5246	0.4748	0.4624	0.1000	0.1000
Feb	0.7488	0.6280	0.5768	0.5302	0.5264	0.1000	0.1000
Mar	0.7800	0.6701	0.6228	0.5793	0.5715	0.1000	0.1000
Apr	0.8229	0.7300	0.6885	0.6500	0.6418	0.1000	0.1000
May	0.8226	0.7320	0.6917	0.6541	0.6483	0.1000	0.1000
Jun	0.8448	0.7618	0.7248	0.6900	0.6846	0.1000	0.1000
Jul	0.8502	0.7694	0.4790	0.6983	0.6921	0.1000	0.1000
Aug	0.8001	0.7004	0.6566	0.6160	0.6118	0.1000	0.1000
Sep	0.7943	0.6912	0.6459	0.6041	0.5964	0.1000	0.1000
Oct	0.7126	0.5771	0.5207	0.4700	0.4571	0.1000	0.1000
Nov	0.6894	0.5489	0.4908	0.4388	0.4264	0.1000	0.1000
Dec	0.7077	0.5740	0.5181	0.4678	0.4567	0.1000	0.1000

Table A.6: CS2 P_3 probabilities. The first row represents the damage severity

	0 (Inspection)	1	2	3	4	5	6
Jan	0.9243	0.8862	0.8666	0.8447	0.8066	0.1000	0.1000
Feb	0.9241	0.8853	0.8653	0.8457	0.8124	0.3930	0.3779
Mar	0.9370	0.9044	0.8868	0.8688	0.8387	0.0940	0.0880
Apr	0.9415	0.9097	0.8928	0.8746	0.8432	0.4560	0.4560
May	0.9425	0.9120	0.8955	0.8787	0.8502	0.4124	0.4124
Jun	0.9343	0.8989	0.8804	0.8614	0.8320	0.2432	0.2571
Jul	0.9508	0.9228	0.7474	0.8923	0.8645	0.3236	0.2991
Aug	0.9388	0.9065	0.8898	0.8727	0.8447	0.6747	0.6898
Sep	0.9415	0.9111	0.8947	0.8778	0.8510	0.2917	0.2917
Oct	0.9197	0.8748	0.8506	0.8272	0.7857	0.1000	0.1000
Nov	0.9274	0.8883	0.8679	0.8462	0.8057	0.1000	0.1000
Dec	0.9390	0.9090	0.8925	0.8759	0.8492	0.1000	0.1000

Declaration of interests

☒ 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.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: