

Article

Use of Markov Decision Processes in the Evaluation of Corrective Maintenance Scheduling Policies for Offshore Wind Farms

Helene Seyr  and Michael Muskulus 

Department of Civil and Environmental Engineering, Norwegian University of Science and Technology NTNU, NO-7491 Trondheim, Norway

* Correspondence: helene.seyr@gmail.com; Tel.: +47-400-867-61

Received: 18 June 2019; Accepted: 2 August 2019; Published: 3 August 2019



Abstract: Optimization of the maintenance policies for offshore wind parks is an important step in lowering the costs of energy production from wind. The yield from wind energy production is expected to fall, which will increase the need to be cost efficient. In this article, the Markov decision process is presented and how it can be applied to evaluate different policies for corrective maintenance planning. In the case study, we show an alternative to the current state-of-the-art policy for corrective maintenance that will achieve a cost-reduction when energy production prices drop below the current levels. The presented method can be extended and applied to evaluate additional policies, with some examples provided.

Keywords: maintenance planning; maintenance strategy; maintenance; corrective maintenance; repair; offshore wind energy; maintenance scheduling; optimization; modeling

1. Introduction

Offshore wind energy is an established form of energy generation in Europe and is globally gaining interest in countries all around the world, especially in East Asia. However, in most places, electrical energy produced by offshore wind farms (OWFs) is still more expensive than other electricity generation methods. Many improvements have already been achieved for different factors influencing the cost of electricity generated from wind. The size of the wind turbine support structures has been optimized, by e.g., minimizing the use of expensive materials. Turbine efficiency has been improved by e.g., optimizing the shape and materials of turbine blades. The overall production of a wind farm can be improved by studying wake effects and optimizing the control of individual turbines. Within the offshore wind research community, the optimization of operation and maintenance has recently been gaining interest from researchers all around the world. One reason for this is the high share of operation and maintenance cost in the overall energy costs up to a third of the price of electricity produced is due to operations and maintenance [1]. Reducing these operation and maintenance costs will improve the total cost of energy production and help achieve cost competitiveness with other generation methods, such as onshore wind or solar energy. Different groups have developed simulators that model the operation of OWFs, as reviewed in [2]. With these simulators, the researchers are able to investigate different maintenance scheduling policies by comparing the simulation results of different policies. Existing models depend almost exclusively on Monte Carlo simulations, i.e., running a large number of simulations with the same inputs, in order to investigate uncertainties and variations in different inputs and variables (wave heights, wind speeds or failure occurrence) used. Additionally, different policies can only be evaluated manually, by implementing each strategy individually in the model. Some exceptions to the dependency on Monte Carlo simulations are some newer approaches,

using stochastic models [3] or genetic algorithms [4]. The influence of uncertainties on the optimal scheduling of corrective maintenance has been investigated for the repair time [5], and the weather forecast [3,6].

In this paper, we present a method that can be applied to compare different maintenance scheduling policies for an OWF at a given location (with known weather) for a specific failure type with a known repair time. In contrast to most of the existing tools and models, no simulations are required and expected values for different performance indicators [7], like downtime or production losses can be compared for the different policies. With the presented method, including uncertainties is straightforward—it can be included directly into the model as opposed to running Monte Carlo simulations with different parameters, as has been done by most of the existing models. Uncertainty in the sea state is included in the presented case study. Section 2 explains the details of the method used and gives the details about the mathematical structure. Implementation of the method is explained in Section 3. Section 4 presents a case study applying the presented method. Discussions and an outlook on alternative policies that could be evaluated with this framework are given in Section 5.

2. Methodology

2.1. Markov Decision Process

The method we present in this paper is based on a Markov decision process (MDP). A MDP is a stochastic control process that can be seen as an extension of a Markov chain, adding actions and rewards [8]. The MDP can be described as a 5-tuple: $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where \mathcal{S} is a set of states, \mathcal{A} a set of actions, \mathcal{P} the transition probabilities between states, given actions, \mathcal{R} a real-valued reward (or penalty) function that calculates the reward (or penalty) of any given state and γ a discount factor. As the name suggests, this process assumes the Markov property, therefore the effects of an action taken in a state only depend on that state and not the prior history of the process. An example of a Markov decision process is presented in Figure 1. In the present framework, the set of states \mathcal{S} includes a finite number of states—in the example in Figure 1, six states are shown. (Infinite sets of states are possible in the framework of Markov decision processes. For more information about the mathematical concept, please refer to the literature, e.g., [8].) Each state can be described by one or multiple properties. These can be e.g., a location (distance from some fixed point), reward given in the respective state, or, in the case of offshore wind farm maintenance, the status of the turbine, a sea state observation, or the time needed to complete a repair. Each state differs from all other states in at least one characteristic, so no duplicates exist. The actions in the MDP can be either deterministic or stochastic. Deterministic actions lead to a (fixed) new state that the process will continue in after the current state. A stochastic action specifies a probability distribution over the next states. The transition probabilities between states depend on the action undertaken in that state and specify the new state, subject to that action. Therefore, for each state and possible action in that state, there is at least one positive transition probability to another state. For each state and action, the transition probabilities sum to one. A deterministic action is a special case of a stochastic action, with exactly one positive transition probability equal to one. The example in Figure 1 includes two stochastic actions and the associated transition probabilities. The reward function is a real valued function, assigning a value to each state and action combination. When a negative value is assigned by the reward function, it is often called a penalty function instead. In the example in Figure 1, each of the six states has one of two reward values, namely 1 and 0.

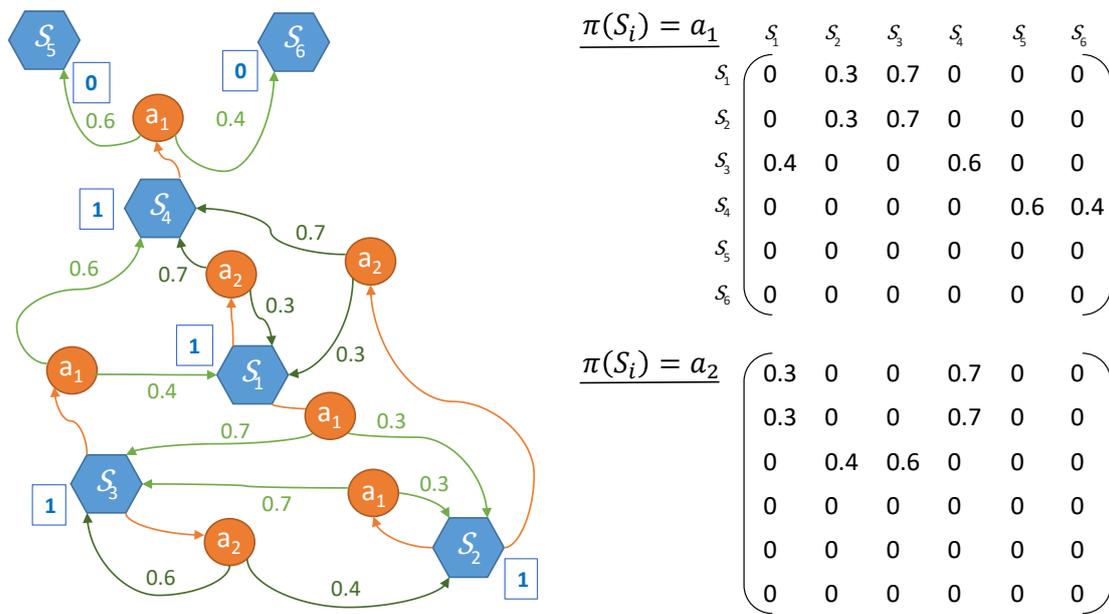


Figure 1. An example of a Markov decision process (left). The blue hexagons represent the states, with their rewards indicated in blue boxes next to the state. Orange circles indicate the two actions that can be taken in each state. Subject to the action, the transition probabilities are indicated with green arrows and the value displayed next to the arrow. Transition probabilities following action a_1 are shown in light green, while transition probabilities following action a_2 are shown in dark green. The policies “choose action a_1 in all states” (upper) and “choose action a_2 in all states” (lower) are presented by their transition matrices (right).

In addition to the Markov decision process that describe how the system works, our setup contains a set of policies Π . A policy $\pi \in \Pi$ is a mapping from \mathcal{S} to \mathcal{A} , and can be understood as a decision makers rule for choosing one of the possible actions $a \in \mathcal{A}$ in each state. In order to follow a policy, one must (a) determine the current state s , (b) determine the action to be executed in that state $a = \pi(s)$, (c) determine the new state s' and continue, alternating (b) and (c). The goal of using a MDP is of course to find an optimal (or at least better than existing) maintenance strategy. In the framework of the MDP, this is done by finding an optimal policy π . In order to evaluate a policy (and ultimately finding the optimal policy), it is necessary to determine expectation of the total reward gained by following it (in order to optimize it). Intuitively, one could try to sum all rewards obtained in the MDP when following the policy, but this can quickly become overwhelming. (Typically, summing all rewards will yield an infinite sum, namely for all MDPs with either infinite state space or for MPDs with infinite horizon. For more information about these cases, refer to e.g., [8].) The solution is to use an objective function to map the sequence of rewards to (single, real) utility values. Options to obtain an objective function are (1) setting a finite horizon, (2) using discounting to favour earlier rewards over later rewards and (3) averaging the reward rate in the limit.

Instead of optimizing the policy, in some cases, it might be desirable to compare different policies with each other. When combining an MDP with a fixed policy that chooses exactly one action for each state, the result is a Markov chain. This is because all of the actions are defined by the policy and one is left with the transition probabilities between states. One example of a resulting Markov chain is visualized in Figure 2. In this Markov chain, the value of each state S_i can be calculated based on the reward $\mathcal{R}(S_i)$ of that state and based on the values of the states that can be reached. It is calculated as

$$V(S_i) = \mathcal{R}(S_i) + \sum_j P_{ij} V(S_j), \tag{1}$$

where P_{ij} is the transition probability between state S_i and state S_j from \mathcal{P} . The equations in (1) are known as Bellman equations, named after Richard Bellman. We can solve the linear equation system (LES) defined by the transition probabilities and reward function to find the values $V(S_i)$ for each state. When comparing two policies, one can look up the value of a specific state one is interested in, usually a ‘starting’ point. In the case of OWF maintenance, this could e.g., be a state in which a failure occurs and the value could then be representative of the time it takes for this failure to be corrected, with a penalty incurred for each step taken without resolving the failure. A case study comparing different policies is presented in Section 4.

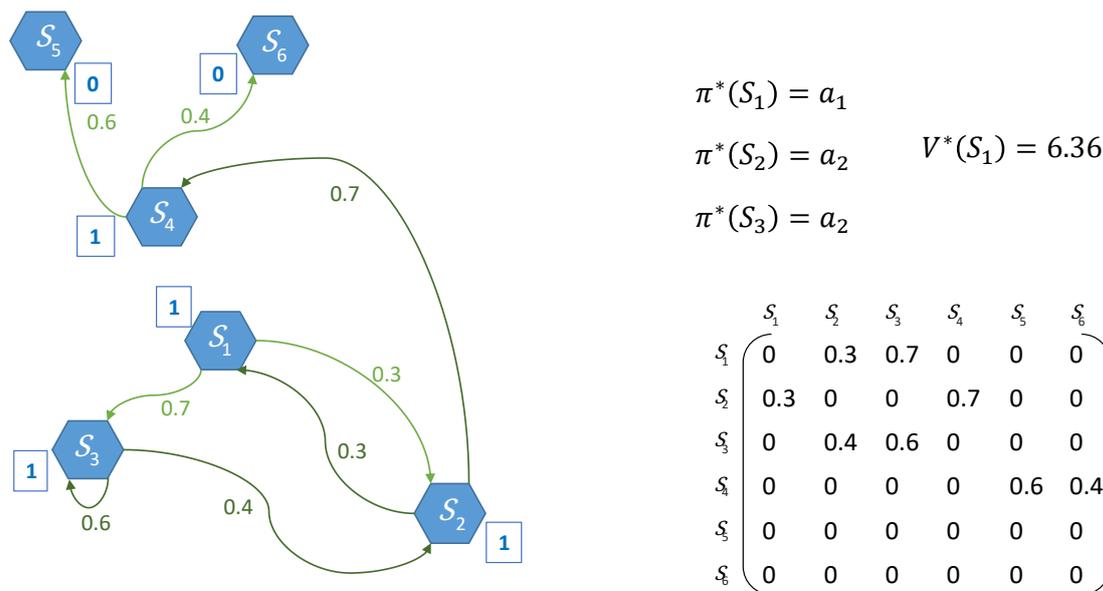


Figure 2. An example of a Markov chain (MC), as the result of selecting one policy in the setup of Figure 1 (left). The policy displayed on the right-hand side is the optimal policy for this MDP, when starting in S_1 . The transition probabilities for the MC are shown in the matrix (right).

In the example shown in Figure 1, a possible policy would be to always choose action a_1 . The corresponding Markov chain is presented on the right-hand side of the figure in the form of its transition probabilities. The rewards (presented in the figure in blue next to the states) are $\mathcal{R}(S_1) = \mathcal{R}(S_2) = \mathcal{R}(S_3) = \mathcal{R}(S_4) = 1$, and $\mathcal{R}(S_5) = \mathcal{R}(S_6) = 0$. In order to calculate the value of each of the states, we solve the equation system defined by the Bellman Equation (1):

$$\begin{pmatrix} V(S_1) \\ V(S_2) \\ V(S_3) \\ V(S_4) \\ V(S_5) \\ V(S_6) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0.3 & 0.7 & 0 & 0 & 0 \\ 0.4 & 0 & 0 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} V(S_1) \\ V(S_2) \\ V(S_3) \\ V(S_4) \\ V(S_5) \\ V(S_6) \end{pmatrix},$$

$$\begin{pmatrix} -1 & 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & -0.7 & 0.7 & 0 & 0 & 0 \\ 0.4 & 0 & -1 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} V(S_1) \\ V(S_2) \\ V(S_3) \\ V(S_4) \\ V(S_5) \\ V(S_6) \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

The values of the states are then

$$\begin{pmatrix} V(S_1) \\ V(S_2) \\ V(S_3) \\ V(S_4) \\ V(S_5) \\ V(S_6) \end{pmatrix} = \begin{pmatrix} 5.05 \\ 5.05 \\ 3.62 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

If one is interested in comparing the value of a specific state under two different policies, the calculation is repeated for that policy and the values compared. It is also possible to find the *optimal* policy without comparing the values based on the resulting Markov chains. In order to find the optimal policy, we define the optimal value function V^* by the recursive set of equations

$$V^*(S_i) = \mathcal{R}(S_i) + \max_{a \in \mathcal{A}} \left[\sum_j P(S_j | S_i, a) V^*(S_j) \right], \quad (2)$$

so the optimal value of a state S_i is the reward in the state, plus the maximum over all actions we could take in the state. This is a the generalized form of the Bellman Equation (1) for policies. In the example shown in Figure 1, the possible actions are a_1 and a_2 . The idea behind this maximum is that in every state we aim to choose the action that maximizes the value of the future. The optimal value function V^* can be found by e.g., value iteration. When V^* is known, the optimal policy π^* can be found by picking the action that maximizes the expected optimal value:

$$\pi^*(S_i) = \operatorname{argmax}_{a \in \mathcal{A}} \left[\sum_j P(S_j | S_i, a) V^*(S_j) \right]. \quad (3)$$

In the example shown in Figure 1, the optimal policy is to conduct action a_1 in states S_1 and S_4 , and a_2 in states S_2 and S_3 . Figure 2 shows the Markov chain corresponding to the optimal policy.

2.2. Probabilistic Weather Input

In the modeling of offshore wind farm maintenance, the uncertainty in the local weather, and more specifically the wave height, is often included. In the given framework, the local weather can be included into the model by adding a sea state (wave height) property in the definition of the states. Some of the existing maintenance approaches [2] use Markov chains to model the weather. Similar to these approaches, transition probabilities between different wave height bins can be used in the MDP. Given some source of historical weather data, the wave height data are sorted into so-called “bins”-categories summarizing wave heights in a given interval. The size of these bins should be adjusted based on the application, for offshore wind farm maintenance 0.4 m is a useful interval step [5]. Then, the probabilities to transition from one bin to all other bins are calculated based on the number of occurrences of transitions between these bins in the given data source. In order to be able to investigate seasonality, one can calculate separate matrices for e.g., each month of the year.

2.3. Repair Time Modelling

Another factor that influences the decision-making in offshore wind farm maintenance is the time it takes to bring a wind turbine component that has failed back to an operational state. Throughout this article, we will use the term “repair time”. This repair time is the cumulative amount of time spent during maintenance actions, and we assume a fixed repair time, without uncertainty. The repair time should not be confused with the time between a failure and its resolution, which we will refer to as “downtime”. The repair time can be included into the MDP as a parameter to the states. During maintenance, the MDP will move from states with a high remaining repair time, to states with

a lower remaining repair time, until a state with no remaining repair time is reached, where the process will stop.

2.4. Calculation of Production Loss

We want to use the MDP to evaluate different policies for offshore wind farm maintenance. One aspect to compare is the production loss of a turbine or wind farm under a given policy. The production loss can only be estimated, as explained in [7], as we cannot measure the absence of production. Therefore, a method to estimate the production loss is needed. Given information about the wind speed from e.g., measurements and a knowledge about how the power production depends on the wind speed, it is straightforward to find an estimate of the production losses. One could include the wind speed as a parameter into the states of the MDP as was done with the wave height. As we do not need to use the wind speed as a decision criterion for the policies, we use a matrix with conditional probabilities of wind speed values given a wave height, similar to how it has been presented in [5]. Given some source of weather data, the wind speeds are first sorted into bins—a bin size of 1 m/s is sufficient for production loss calculations. For each state in the MDP with a given wave height parameter, the conditional probability for each wind bin is populated in a matrix that can later be used to look up these values. The expected production loss for each state in the MDP can be calculated based on these relative probabilities and a power curve for the turbine type of interest. A power curve can be either obtained directly from the manufacturer or a linearized power curve can be used, based on the turbine model. If one does not have information about an actual turbine model, reference turbines like [9] or [10] can be used. In order to obtain the (expected) production loss, the production values for each discrete wind speed bin, as obtained from the power curve, are weighted (multiplied) with the (conditional) probability from the matrix. The sum of these weighted production values is then the production loss for the state. For a state S_i , given n discrete wind speed steps with conditional probabilities $P(u_k|S_i)$ for $k = 1 \dots n$, and a power curve $p(\cdot)$ the expectation of the production loss $L(S_i)$ is

$$E(L_{S_i}) = \sum_{k=1}^n p(u_k)P(u_k|S_i). \quad (4)$$

3. Implementation

In order to use the presented method to evaluate different maintenance policies, it is necessary to implement it in a programming language. The resulting program can then be used to evaluate well-known policies and compare them to alternative options. In our analysis, the implementation was conducted in Python 3.

In order to define the MDP, we define the states, actions, policies and the reward function. The set of states \mathcal{S} can be generated, by defining the composition of a state and then generating a list of possible states. A state could e.g., be a tuple of several parameters $S_i = (p_1(i), p_2(i), p_3(i))$, where each of the parameters can take different values (e.g., $p_1(i) \in \{0, 0.4, 0.8, 1.2, \dots, 10.0, 10.4\}$ a wave height, $p_2(i) \in \{5, 4, 3, 2, 1, 0\}$ the number of remaining repair hours, and $p_3(i) \in \{\text{'at shore'}, \text{'offshore'}\}$ the vessel location). The different actions $a \in \mathcal{A}$, will have different outcomes depending on the state. Possible actions for an implementation for the offshore wind maintenance planning could be “go out to the wind farm”, “repair the turbine”, and “return to shore”. As described above, each policy $\pi \in \Pi$ is a set of rules, defining which action should be taken in which state. It can be implemented as a set of conditional expressions ensuring that only transitions between states which correspond to the actions defined by the policy are possible. When investigating multiple policies with similar rules, a high level policy can be implemented first, and the characteristic parameter changed for each individual policy in the evaluation. The reward function is a function assigning a real value to each state. It is also possible for the reward value to be dependent on the action taken to reach the state. Implementation of

this reward function highly depends on the structure of the states; in most cases, it will be a function depending on one or more parameters of the state.

In order to calculate the value of a policy, the first step is to define the equation system resulting from plugging the policy into the MDP, thereby forming a Markov chain. The LES has been observed to follow some rules and the matrix defining it can be produced following these steps:

1. Find the number of states n , and find a mapping of the states, assigning each of them a natural number, effectively applying an order to the states.
2. Create the $(n \times n)$ matrix containing the transition probabilities for the investigated policy.
3. Calculate the entries of the reward vector, where the i -th entry corresponds to $\mathcal{R}(S_i)$.
4. The matrix \mathcal{P} and vector \mathcal{R} define an equation system

$$(\mathcal{P} - I_n)V = -\mathcal{R},$$

which can be solved using a linear algebra routine in e.g., Matlab or Python. Depending on the structure of the matrix and vector, different algorithms might be used to achieve fast computation.

5. In order to investigate different properties of a policy, the same matrix is used in a LES combined with different reward functions for each property.

4. Case Study

4.1. MDP Definition

In this case study, the states $S \in \mathcal{S}$ of the MDP are tuples of the form $S = (\text{location}, \text{wave height}, \text{repair time left}, \text{steps waited})$, where ‘location’ can take on either of the values ‘port’ or ‘turbine’. The significant wave height (‘wave height’) takes values in steps of 0.4 m between 0 m and 10.4 m. The ‘repair time’ starts off with an initial value, specific to the turbine component that is investigated. The values for repair time are taken from [11], the most recent source for offshore wind turbine failure and repair data. Different components and types of repairs have been investigated in this case study, each with a distinct mean time to repair and worker requirement. For the example of the major blade repair, with 21 h mean time to repair, the values for the ‘repair time’ range from 0 to 21 h in steps of 1 h. For other components and repair times, the values have a different range. The steps are, however, set to 1 h, for all repair types and components investigated. This results in a different number of states for different types of repair. The ‘steps waited’ also take steps of 1, starting at 0 and ranging up to 3 depending on the maintenance policy. A summary of the parameters for the states is shown in Table 1. The set of actions $\mathcal{A} = \{\text{stay}, \text{wait}, \text{reset wait time}, \text{go out}, \text{repair}, \text{return}\}$, where the actions ‘wait’ and ‘reset wait time’ are only used in some of the policies. How the actions are used in the different policies is detailed below in Section 4.4, a summary of the possible actions is provided in Table 2. The transition probabilities between states \mathcal{P} depend on the transition probabilities of the significant wave height values. These probabilities are calculated based on the weather data from FINO 1 [12]. More details on how the probabilities are calculated are given in Section 4.2. The reward function \mathcal{R} is used to evaluate different aspects of the maintenance policies. To evaluate the influence of the policy-change on the expected downtime of the turbine, a penalty is used for the steps it takes to end up in a repaired state. To calculate the expected production losses, the reward function \mathcal{R} represents a penalty of the production losses. These are calculated based on the correlation of wind speeds and wave height and a linearized power curve for the NREL 5 MW turbine [9]. The details of this calculation are presented below in Section 4.3. Discounting is not used in this case study and hence the discount factor set to $\gamma = 1$. To evaluate a maintenance policy, we investigate the value of the initial states. These states are those in which the failure occurred and hence the repair has not started. As we assume cumulative repairability (i.e., when a repair has to be interrupted, progress is kept and the repair can be continued at a later stage), these are all states with the initial repair time values. Since the failure can occur at any wave height, multiple states with this repair value exist. These are weighted with their probability of occurrence and the values summed before reporting.

Table 1. The different parameters of the states and their possible values.

Parameter	Possible Values	Comment
Location	{port, turbine}	
Wave height [m]	{0, 0.4, 0.8, ..., 10, 10.4}	
Repair time [h]	{0, 1, 2, ..., max - 1, max}	maximum (<i>max</i>) depends on component
Steps waited	{0, 1, 2, 3}	depends on strategy

Table 2. The different actions and how they influence the parameters of the next state.

Action	Parameters for Next State
stay	
wait	steps waited +1
reset wait time	steps waited = 0
go out	location = turbine
repair	repair time -1
return	location = port

4.2. Weather Input

In this case study, the weather data used to calculate the transition probabilities between wave heights (and subsequently states) comes from the FINO 1 measurement campaign. The data from the FINO 1 measurements have some missing observations. Additionally, the wind speeds are provided in 10 min aggregated means while wave height measurements are provided for 30 min intervals, which is not convenient for the calculation of production losses (Section 4.3). The transition probabilities have therefore been calculated based on the interpolated time series also used in [13,14]. In order to calculate the transition probabilities, the significant wave height is categorized in steps of 0.4 m first. This means that all wave height observations between 0 m and 0.4 m will be collected in one so-called bin. The same is done for the values between 0.4 m and 0.8 m, and so on. We have chosen to calculate separate matrices with the transition probabilities for each month, by sorting the data beforehand. This has the advantage that we can investigate and observe how the season affects the optimal policy.

4.3. Calculation of Production Loss

As described above in Section 2.4, the expected production loss for each state is calculated based on probabilities of the wind speed given the sea state, and a power curve. The probabilities are based on data from the FINO 1 measurement campaign. The same 1 h-interpolated FINO 1 data [13] that was used to calculate the wave height transition probabilities was used. For each observation point, a wave height value and a wind speed value are known. The wave height and wind speed are then categorized. For the wind speed, the step size is 1 m/s, so each observation for wind speeds between 0 m/s and 1 m/s will be collected together. The same is done for wind speeds between 1 m/s and 2 m/s and so on. The wave heights are categorized as described in Section 4.2. Then, the conditional probabilities of these wind speeds subject to the wave height at the same point in time are gathered. For the calculation of the production loss, information about the power curve is needed in addition to the weather. In our case study, a linearized power curve is used for the NREL 5 MW turbine [9], as was also done in [5]. The linearized power curve is based on the cut-in and cut-off wind speed as well as the wind speed where the rated power (5 MW) is reached. When solving the MDP, the production loss values are used to calculate the reward of each state. The weighted values of the initial states are then summed and reported, as described above in Section 4.1. The loss of production is calculated in terms of electric power (kWh). If one is interested to compare this directly to the cost of maintenance, the energy needs to be valued in terms of money. This can be done by either using a (variable) electricity market price or a (fixed) feed-in-tariff.

4.4. Policies

This section presents the different maintenance policies that are investigated and compared in the case study. As described in Section 2, a single policy assigns an action $a \in \mathcal{A}$ to each state $S \in \mathcal{S}$. A summary of all policies, with different parameters is shown in Table 3. For each policy, the possible actions under this policy are listed.

Table 3. Names of the different policies investigated in this article, as well as the maximum number of steps that can be waited under this policy and the possible actions.

Policy Name	Max (Steps Waited)	Possible Actions
go-right-away	0	stay, go out, repair, return
wait-1-step	1	stay, wait, reset wait time, go out, repair, return
wait-2-steps	2	stay, wait, reset wait time, go out, repair, return
wait-3-steps	3	stay, wait, reset wait time, go out, repair, return
0.8 m-limit	0	stay, go out, repair, return
1.2 m-limit	0	stay, go out, repair, return
2.0 m-limit	0	stay, go out, repair, return
2.4 m-limit	0	stay, go out, repair, return
2.8 m-limit	0	stay, go out, repair, return

4.4.1. Go-Right-Away

In order to be able to conduct maintenance, a vessel has to be at the turbine and the wave height needs to be below a defined threshold of 1.6 m. This is a value, based on the often presented wave height limit of 1.5 m for vessel access [15], modified to fit the wave height resolution of the case study. In this strategy, as soon as the wave height is below the threshold of 1.6 m, the vessel is sent to the wind turbine. We assume a travel time of one step (1 h) in this case study, which might be short compared to some wind farms. However, since we are using weather data from FINO 1, which is next to the Alpha Ventus wind farm in the North Sea, we are already assuming a wind farm relatively close to shore which will have a shorter travel time. Once the vessel reaches the turbine, repair is conducted if the wave height is still below the threshold. As soon as the wave height crosses the threshold, the repair is interrupted and the vessel returns to port. We assume that the repair is cumulative, i.e., when the repair is interrupted, it can be continued at a later stage without any loss of progress. The return to port takes one step (1 h) again. As soon as the wave height crosses below the limit again, another access is made until the turbine is repaired. We do not take into account any restrictions to the working time of the maintenance crew or vessel crew, so it is possible to have one access and conduct the full repair without ever returning to port. This is a simplification that could be justified, if the boat has living quarters and enough personnel on board to rotate in shifts. Figure 3 shows a decision diagram for this policy. In every state of the Markov decision process, the diagram can be used to find the action that the policy prescribes for that state. In Figure 4, a minimal MDP is shown for this policy. Here, two steps of repair are required and two wave heights are considered, namely below and above the limit. The probability to stay below the limit is denoted as $P(-,-)$, the probability to change wave height from below the limit to above the limit is denoted as $P(-,+)$ and so on. Assuming the state is ('port', 'above limit', '2'), the first check is whether the repair time is greater than zero, which it is. The next check is whether the vessel is at the turbine, which it is not. Thus, the next inquiry is whether the wave height is below the limit, which it is not. The action is then 'stay'. The state will be the same in the next step with a probability of $P(+,+)$ and will change to ('port', 'below limit', '2') with a probability of $P(+,-)$. In this state, the action will be 'go out'.

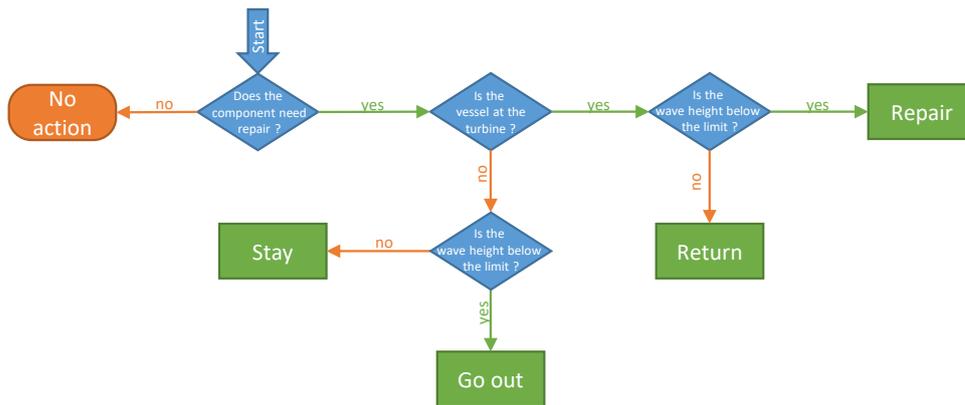


Figure 3. The decision tree for the original (go-right-away) policy. This assessment is conducted for each state and influences the transition probabilities in the MDP, by choosing an action for each state. An example of how the policy is applied can be seen in Figure 4.

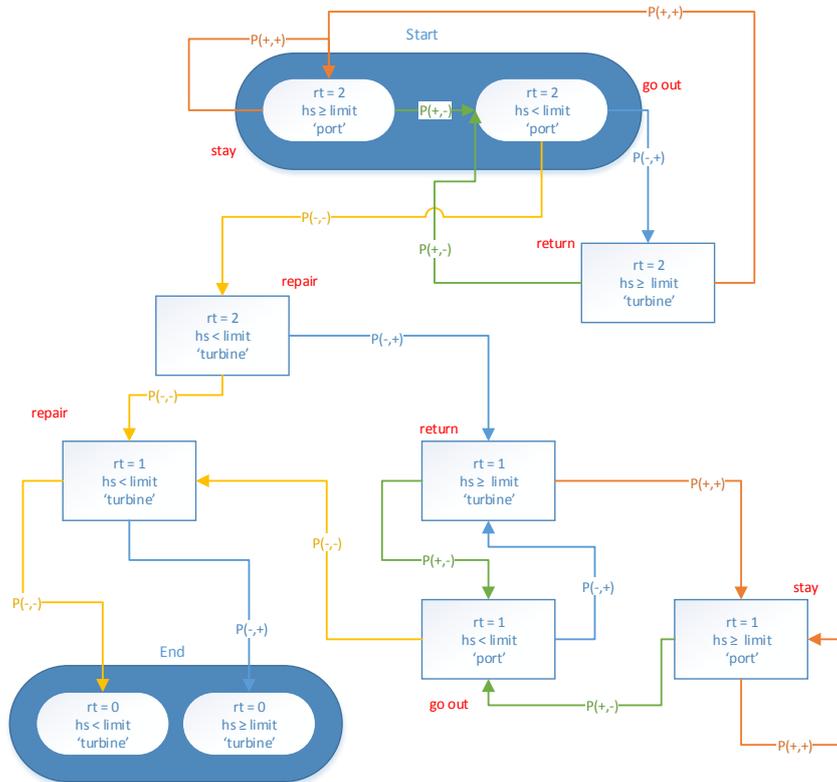


Figure 4. A minimal example of the process for the original (go-right-away) policy. Here, only two steps of repair are shown. The process starts in states with repair time (rt) equal to 2 h ($rt = 2$), which is then reduced to 1 h ($rt = 1$) and finally 0 h ($rt = 0$). The wave height (hs) is categorized as being below ($<$) or above (\geq) the threshold (limit), and transition probabilities are adjusted to accommodate this simplification. $P(+,-)$ is the probability to get from a wave height above threshold ($hs \geq \text{limit}$) to a wave height below threshold ($hs < \text{limit}$). With 'start', we mark the states in which a maintenance decision maker would start the decision of when to repair, i.e., the point in time when the failure occurs/is reported. The decision taken in each state is marked in red, next to the respective state. How the decision is made, based on the state and maintenance policy can be understood from Figure 3.

4.4.2. Wait-n-Steps

An alternative to accessing the wind farm as soon as the wave height is below the threshold is to wait a certain number of steps in good weather, before going out with the vessel to conduct maintenance. The intuition behind this policy is that, if the sea has been calm for several time-steps, it is more likely to stay calm (i.e., below the wave height limit) due to persistence. Waiting a certain amount of time in good weather assures that the observation below the limit was not just an outlier and one can avoid interrupting the maintenance operations. In the investigated policies, the number of waiting steps is fixed and independent of the observed wave height in the state. In our case study, we investigated wait-times of one step, two steps and three steps. Each step represents 1 h. The other aspects of the strategy remain as before. Again, the repair is assumed to be cumulative, so, if the repair is interrupted, progress is kept and it can be continued and completed at a later stage. The maintenance is aborted and the vessel returns to shore as soon as the wave height is above the threshold. The time it takes to access the turbine and return to port respectively is one step (1 h). The decision diagram for this policy is shown in Figure 5.

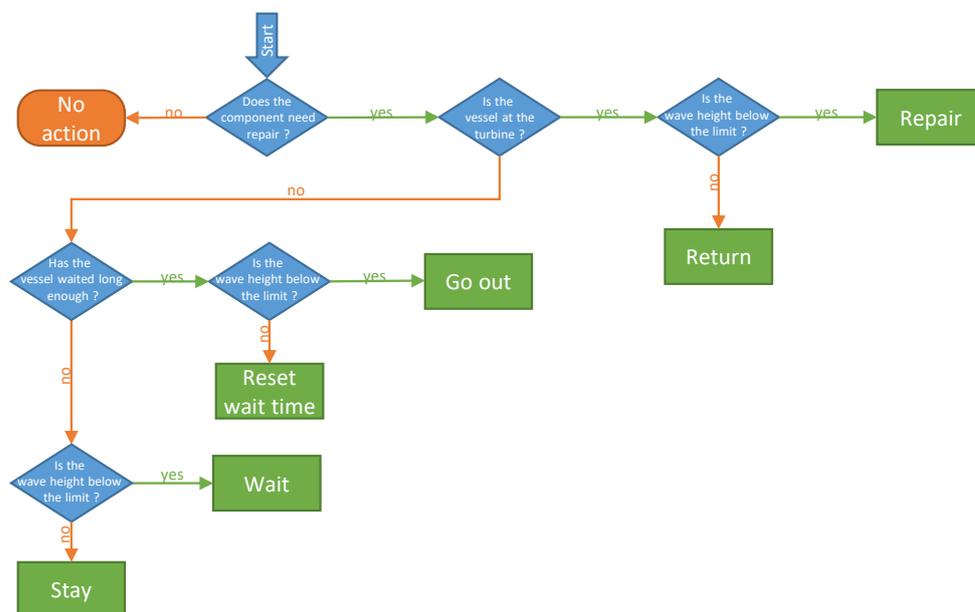


Figure 5. The decision tree for the wait-n-steps policy. First, the decision maker checks, whether a repair is necessary (repair time > 0). Depending on the location (at turbine), a wait-time check is conducted. This depends on the number of wait steps specified by the policy (1 h, 2 h, 3 h). Finally, the weather is checked and the correct action chosen for this state. This assessment is conducted for each state and influences the transition probabilities in the MDP, by choosing an action for each state.

4.4.3. Different-Limits

The third type of policy that is being investigated in this article has a second wave height threshold. One limit (new) is used for the decision of going out to the wind turbine and the other (original) threshold of 1.6 m is used for the decision to start and continue the repair. It is also used for triggering a possible return of the vessel to the harbour. We investigate both lower (stricter) and higher (laxer) wave height limits for access (new limits), specifically we investigate the limits 0.8 m, 1.2 m, 2 m, 2.4 m, and 2.8 m. The repair is again assumed to take a fixed amount of time and can be completed by accumulating enough maintenance (repair) actions. Again, as soon as the wave height is above the (original) wave height threshold, the repair is aborted and the vessel returns to shore. The decision tree for this policy is identical to the one of the go-right-away policy shown in Figure 3,

only that the weather check uses a different threshold in a port and at the turbine. The go-right-away policy is a special case of this strategy, where both limits are 1.6 m.

4.5. Repair Data Input

For the repair time values, data from [11] are used. They present the mean time to repair [h], number of workers needed and mean annual failure rates for 19 wind turbine components. For each component, three types of failures are distinguished, namely ‘major replacement’, ‘major repair’ and ‘minor repair’. Each of these have their own values, leading to a total of 57 different combinations of component and repair type, with specific repair time and worker requirements. We have investigated some selected turbine components and failure types, namely major gearbox replacement, major blade repair, and minor electrical repair. The repair time value is used to generate the possible states for the MDP, whereas the worker requirement is used for cost calculations. The values that have been used are summarized in Table 4.

Table 4. Turbine components that are investigated in the case study and their repair parameters.

Component	Cumulative Repair Time [h]	Average Number of Workers Needed
Gearbox, major replacement	231	17.2
Blade, major repair	21	3.3
Electrical, minor repair	5	2.2

4.6. Cost Data Input

The costs for vessel and workers are dependent on the number of accesses, the total operation time (travel time and working time combined), the vessel charter costs, the vessel hourly costs, the number of workers needed for the repair, and the worker hourly wages. In the case study, these costs are all set according to values from the literature, provided in Table 5. In order to value the production losses in terms of money, the market price for electricity or feed-in-tariffs can be used. Since the price of electricity varies a lot, both between seasons, time of the day and countries, not a single “correct” electricity price can be used to analyze the production losses. In order to show the variation in electricity prices and their influence on the optimal maintenance policy, we include an analysis of the corrective maintenance cost in the case study. In order to gain some insight into the electricity prices in Europe, we used [16,17].

Table 5. Input used to calculate the maintenance cost.

Hourly vessel	287.5 €	calculated from daily cost from [13]
Hourly worker	55.3 €	calculated from annual worker salary
Mobilisation vessel	1000 €	arbitrary (similar to number from [18])

4.7. Results

In this section, some aspects under which the different maintenance policies have been compared are presented. Some of these aspects are similar to the key performance indicators presented in [7].

4.7.1. Repair Actions

The number of repair actions with each policy can be used as a control in order to detect possible mistakes in the implementation of the strategy. All policies include cumulative repair, no degeneration and the work is not continued after the repair is completed. Therefore, the expected repair time calculated and returned by each maintenance policy should be equal to the repair time needed to bring the investigated component back to a state as-good-as-new. Due to memory and rounding errors, this differs insignificantly between policies, in the magnitude of 10^{-10} h in our study.

4.7.2. Downtime

The expected downtime of a maintenance action or repair can be used to evaluate a maintenance policy. The downtime of a turbine is defined as the time the turbine is in a non-operational state caused by either a fault, or by a maintenance action. With an increase in downtime, the time-based availability of the turbine is reduced, often leading to lost production and a lower energy-based availability [7]. The downtime will incur production losses and the decision maker is therefore most likely interested to reduce it. In order to calculate the expected downtime for a policy, the reward function of the MDP is modified such that every step the process takes (i.e., every transition from one state to the next) gets a penalty of 1, representing the time that is lost in this step. When the MDP is then solved, thus the value of each state is calculated, and the average of values of the starting states weighted by the probability of occurrence gives the expected downtime until the cause of downtime (in this case a failure) is resolved. The starting states are those states with a repair time equal to the expected repair time and can be understood as the time of occurrence of the failure. The turbine downtime is, unsurprisingly, higher for the more restrictive policies. For the ‘wait-n-steps’-policies, downtime is always higher than for the original ‘go-right-away’ strategy. For the ‘different-limits’ policies, those with a less restrictive limit are observed to have a slightly lower downtime than the original strategy. Due to the threshold for access being less restrictive, the vessel is more often at the turbine location. It can be avoided to “waste” one time step of calm weather for the access. This increases the likelihood of the vessel being already at the turbine location when the weather is calm enough to conduct a repair and therefore a faster resolution of the failure. For policies with a stricter limit than 1.6 m, the downtime increases, depending on the repair time and month, to up to three times the downtime of the original policy. Figure 6 shows the downtime for each policy and each month for the major gearbox replacement with a repair time of 231 h.

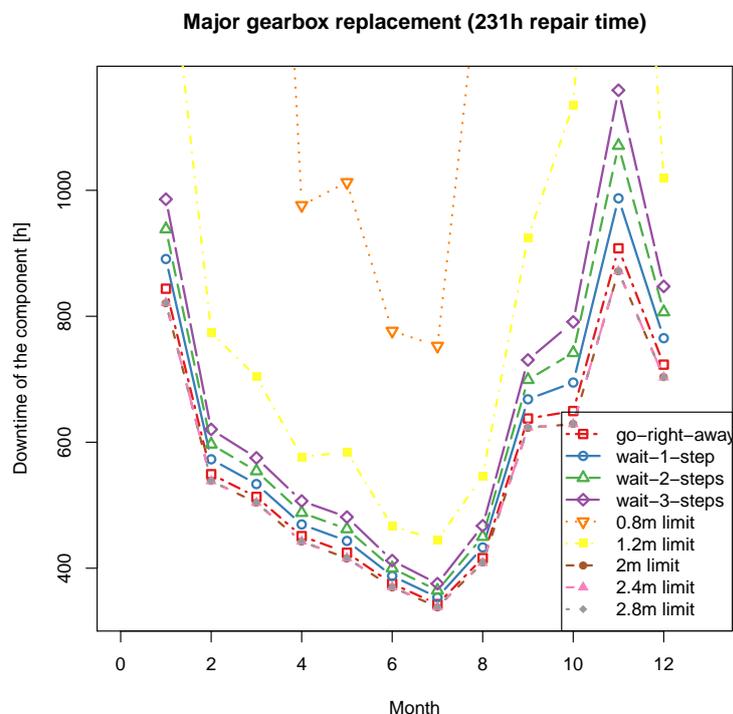


Figure 6. Downtime of the wind turbine due to a major gearbox replacement for different policies. Policies with a less restrictive wave height threshold for the vessel access have a lower downtime than more restrictive policies.

4.7.3. Production Losses

The second aspect that is used to evaluate a maintenance policy is the production lost due to the downtime of the turbine. As explained in Section 2.4, we calculate the production loss based on the wave height in each state. In the MDP, the expected production loss for each policy can be calculated, by using the lost production as ‘reward’ in the process. Then, the value of the starting states represents the production loss that can be expected by using the evaluated policy. Results for the production loss are shown in Figure 7, for a minor repair of the electrical system. It can be observed that the policies with a laxer wave hold threshold for vessel access have a slightly lower production loss than the original policy. The more restrictive policies on the other hand lead to an increase in lost production, up to more than three times the values of the original policy. For the calculation of the losses in terms of monetary value, different electricity prices have been used in this case study, based on data from Eurostat [16] for various countries. These results are shown combined with other maintenance costs below in Section 4.7.5.

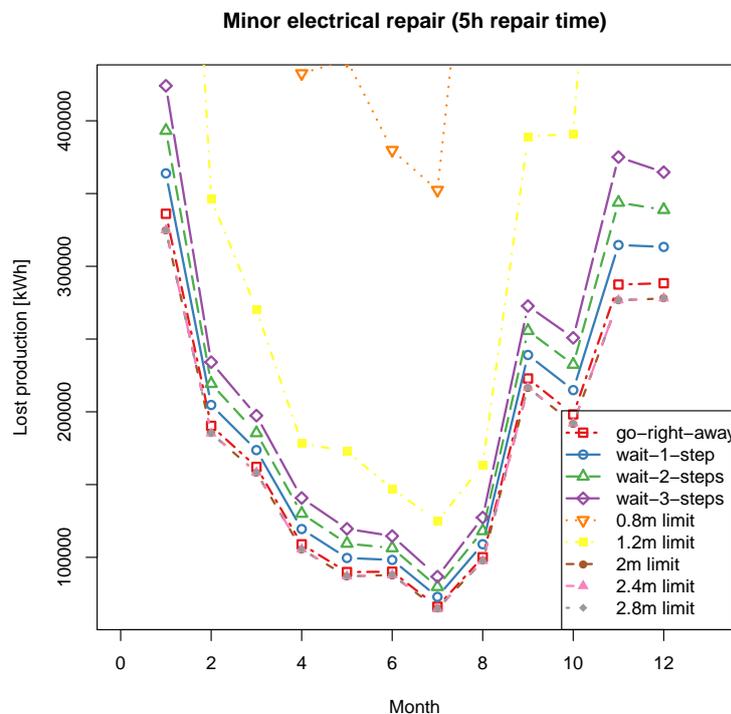


Figure 7. Lost production in kWh for a minor repair of the electrical system, with a repair time of 5 h. The policies with a higher (less restrictive) wave height threshold for vessel access show slightly lower losses in production than the original ‘go-right-away’-policy.

4.7.4. Number of Vessel Accesses and Returns

Another aspect that can be used to compare different maintenance policies is the number of vessel accesses. This number is of interest, since usually each vessel mobilization induces a fixed cost for the maintenance provider or wind farm operator. Hence, the decision maker is interested in keeping the total number of vessel mobilizations low, while still trying to conduct a repair as fast as possible. The number of vessel accesses for each policy can be monitored, again by modifying the reward function. The reward is set to 1 for each state in which the selected action is ‘go out’. Each time a vessel is sent from the port to the turbine, the reward will increase by one and after the process has finished, the expected number of vessel accesses can be calculated in the same way as the number of repair actions or downtime. As the MDP is stopped as soon as the repair is complete, the number of vessel returns will always be one less than the number of accesses, and can be calculated by following

the same logic as for the vessel accesses, switching the action from ‘go-out’ to ‘return’. The number of accesses needed before a completed repair implies vessel and worker costs. Figure 8 shows that the policies with a wave height threshold of 1.6 m (‘go-right-away’ and ‘wait-n-steps’) perform very similar in terms of number of vessel mobilizations. The policies with a more restrictive wave height threshold (0.8 m and 1.2 m) show fewer vessel mobilizations. The policies with a higher threshold for waves (2 m, 2.4 m, 2.8 m) show very high numbers of vessel mobilizations, up to 10-times the values of the ‘go-right-away’-policy. This is likely caused by the wave height limit for repairs, which remains at 1.6 m also for those policies. When the vessel goes out in harsher weather than is allowed during repairs, and this weather persists for longer than the travel time, the vessel has to return to port right away and no repair can be conducted.

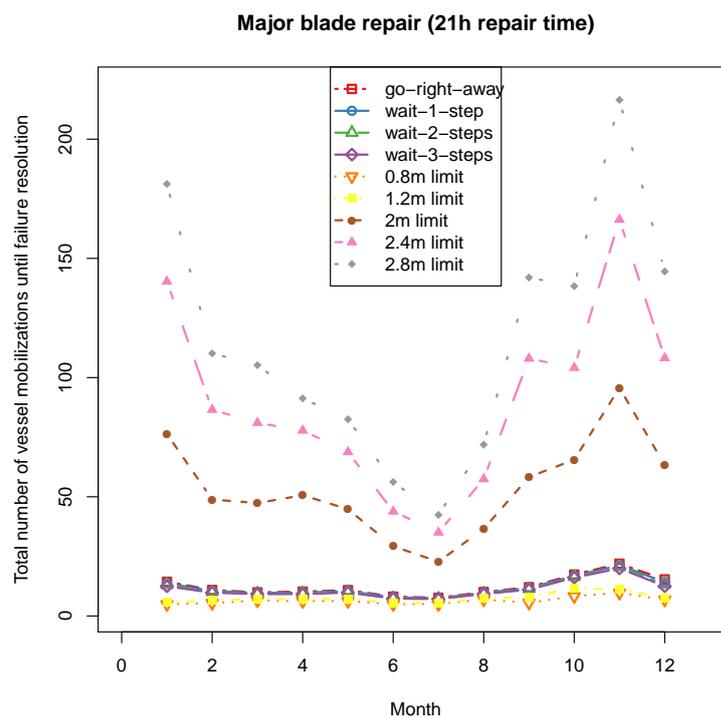


Figure 8. Total number of vessel accesses until the major blade repair is completed—for different policies.

4.7.5. Total Cost of Maintenance

The results for the total cost calculation are naturally the most complex, as they combine the cost calculations with the production losses. In the given framework with cumulative repair and no penalty for an unsuccessful repair attempt, one expects that the ‘go-right-away’ strategy will be the cheapest option, as this strategy leads to the fastest resolution of the failure. Our case study confirms this under the current electricity prices and assumed worker and vessel costs. Figure 9 shows the example of a major blade repair, and the total cost of maintenance for different policies. Should, however, the electricity price drop, and reach levels below 2.4 Euro-cent, the ‘wait-1’ strategy surpasses the original (go-right-away) strategy, as the avoidance of unnecessary vessel mobilizations will outbalance the losses due to turbine downtime. This can be seen from Table 6, for the major gearbox replacement for the month of June. As the production losses highly depend on the repair time and weather, no universal “cut-off” point between policies exists, but has to be investigated on an individual basis. Should the current trends of dropping yield for the electricity producer continue, we expect to see novel policies surpassing the cost-performance of the current state-of-the-art policy.

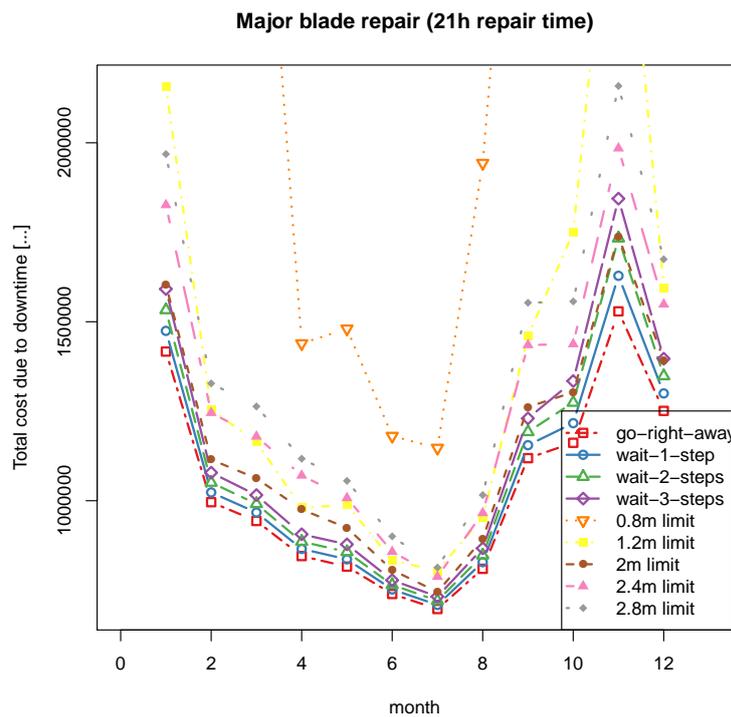


Figure 9. Results for the total costs for the major repair of a turbine blade. We assume an electricity price of 30.84 Euro-cent (Germany second half 2017 from [16]). The original strategy is the cheapest option independent of the season.

Table 6. Electricity price at which a break-even is reached between two policies in €. A value is calculated for each month of the year, for the major gearbox replacement, with a repair time of 231 h. This comparison is not complete and solely meant as an example to show that novel policies indeed become cheaper than the original policy for low enough electricity prices.

Month	simple = wait1	simple = wait2	simple = wait3	wait1 = wait2	wait1 = wait3	wait2 = wait3
1	0.0117	0.0114	0.0110	0.0110	0.0106	0.0103
2	0.0163	0.0153	0.0144	0.0143	0.0134	0.0125
3	0.0118	0.0108	0.0099	0.0098	0.0090	0.0082
4	0.0211	0.0192	0.0175	0.0173	0.0158	0.0142
5	0.0193	0.0178	0.0164	0.0163	0.0150	0.0137
6	0.0247	0.0228	0.0212	0.0210	0.0194	0.0178
7	0.0180	0.0168	0.0157	0.0156	0.0146	0.0136
8	0.0175	0.0158	0.0143	0.0141	0.0127	0.0113
9	0.0113	0.0104	0.0096	0.0095	0.0088	0.0081
10	0.0095	0.0086	0.0078	0.0077	0.0070	0.0063
11	0.0075	0.0070	0.0066	0.0066	0.0061	0.0057
12	0.0223	0.0209	0.0196	0.0195	0.0182	0.0169

5. Discussion

The most important takeaway from this paper should be the methodology that has been presented. The Markov decision process is a powerful tool and yet so versatile that it can be modified to fit a multitude of use cases. Uncertainties in different parameters can be included, by adding a parameter to the state, representing e.g., the probability of a successful repair, or the occurrence of a new failure.

The results presented in the case study Section 4.1 show that the Markov decision process is a valid approach to assess different maintenance policies for offshore wind farms. It has shown that, depending on the circumstances, the current state-of-the-art maintenance policy is indeed optimal. We have further shown that, with an electricity price below 3 Euro-cent, the ‘wait-1-step’-policy becomes better than the

original strategy in the given framework. This is assuming a crew transfer vessel with the presented values for maintenance costs can be used for the given repair and weather probabilities based on FINO 1 [12].

According to Fraunhofer ISE [17], the wind specific electricity prices in Germany are currently between 8 and 14 Euro-cent, while consumer end prices for electricity were at 31.23 Euro-cent in Germany in the second half of 2018 according to Eurostat [16]. This shows that only a small fraction of the end-consumer price is paid to the wind farm operator, roughly between 26–46% of the consumer end price goes to the energy producer in Germany. Fraunhofer ISE [17] predict the prices in Germany to further drop to around 5 to 11 Euro-cent by 2030.

When applying these percentages to other European countries, like Lithuania with an electricity price of 0.1097 Euro-cent in the second half of 2018 [16], a yield of 3–5 Euro-cent becomes realistic. This is without the prediction of a drop in the share of the percentage of the electricity price that goes to the producer. Factoring that into the previous calculation, a yield between 1.8–0.3 Euro-cent for Lithuania in 2030 can be predicted. Therefore, wind farm operators might soon be interested to look beyond the state-of-the-art strategy and investigate other policies.

Another aspect to consider is the limitation of this study concerning different vessel types. A gearbox replacement usually requires a lifting vessel with a crane, which generally have higher mobilization and hourly hire rates than the ones investigated in the current framework. In reality, the maintenance policy of waiting for a persistently calm sea might therefore already be economically viable in some cases for the current electricity prices.

A similar argument can be observed for wind farms that are far offshore, with longer travel times. The example presented here was based on FINO 1 [12] data, a measurement mast close to the Alpha Ventus wind farm very close to the coast. With an increasing travel time, the cost of a failed maintenance attempt (an unnecessary vessel mobilization) increases and it is expected that another policy than ‘go-right-away’ will be economically better and possibly already for current electricity prices.

The Markov decision process can be used to study and compare many different maintenance policies that have not been discussed here. It is also very straightforward to use the same process for wind farms with a longer travel time, other site-specific weather conditions, turbine types or cost numbers. Some examples for investigations in the future include:

- Maintenance policies taking into account work-time restrictions or shift lengths.
- Policies including multiple vessels.
- Policies including different vessel types.
- Investigations of multiple failures or turbines.
- A framework that takes into account incomplete repair actions or loss of repair progress in case of interrupted maintenance.
- Wind farms further from shore, with a longer travel time and harsher weather.

6. Materials and Methods

Wind and wave data from the FINO 1 project are provided by the Bundesministerium für Wirtschaft und Energie (BMWi), Federal Ministry for Economic Affairs and Energy and the Projektträger Jülich, project executing organization (PTJ). They can be downloaded from <http://fino.bsh.de/> by users from Europe, for research purposes.

The implementation of the method, as used for the case study, is freely available from <https://github.com/helenese/MDP>, licensed under Creative Commons Attribution—NonCommercial 4.0 International (CC BY-NC 4.0).

7. Conclusions

In this article the Markov decision process (MDP) has been presented as a useful method for offshore wind farm maintenance modeling. The method can be adapted to fit many use cases and uncertainties can be included without relying on Monte Carlo simulations. The case study has

validated the use of this concept and further indicates that under a hypothetically, lower electricity price alternative policies for the scheduling of repair will become more efficient than the current state of the art.

Author Contributions: H.S. and M.M. contributed to the research idea, H.S. conducted the analysis, M.M. supervised the analysis, H.S. wrote the paper, and M.M. reviewed the paper.

Funding: Part of the work leading to this publication was financed by the AWESOME project (awesome-h2020.eu), which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant No. 642108. Most of the work of the first author was completed without funding after the funding period was over.

Acknowledgments: The authors would like to acknowledge the two anonymous reviewers, whose comments and suggestions helped improve the quality of this paper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Feng, Y.; Tavner, P.; Long, H. Early experience with UK round 1 offshore wind farms. *Energy* **2010**, *163*, 167–181. [CrossRef]
2. Seyr, H.; Muskulus, M. Decision Support Models for Operations and Maintenance for Offshore Wind Farms: A Review. *Appl. Sci.* **2019**, *9*, 278. [CrossRef]
3. Gintautas, T.; Sørensen, J.D. Improved Methodology of Weather Window Prediction for Offshore Operations Based on Probabilities of Operation Failure. *J. Mar. Sci. Eng.* **2017**, *5*, 20. [CrossRef]
4. Stock-Williams, C.; Swamy, S.K. Automated daily maintenance planning for offshore wind farms. *Renew. Energy* **2018**. [CrossRef]
5. Seyr, H.; Muskulus, M. Value of information of repair times for offshore wind farm maintenance planning. *J. Physics: Conf. Ser.* **2016**, *753*, 092009. [CrossRef]
6. Seyr, H.; Muskulus, M. How Does Accuracy of Weather Forecasts Influence the Maintenance Cost in Offshore Wind Farms? In Proceedings of the 27th International Ocean and Polar Engineering Conference, International Society of Offshore and Polar Engineers, San Francisco, CA, USA, 25–30 June 2017.
7. Gonzalez, E.; Nanos, E.M.; Seyr, H.; Valldecabres, L.; Yürüßen, N.Y.; Smolka, U.; Muskulus, M.; Melero, J.J. Key Performance Indicators for Wind Farm Operation and Maintenance. *Energy Procedia* **2017**, *137*, 559–570. [CrossRef]
8. Puterman, M.L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
9. Jonkman, J.; Butterfield, S.; Musial, W.; Scott, G. *Definition of a 5-MW Reference Wind Turbine for Offshore System Development*; National Renewable Energy Laboratory: Golden, CO, USA, 2009. [CrossRef]
10. Bak, C.; Zahle, F.; Bitsche, R.; Kim, T.; Yde, A.; Henriksen, L.C.; Hansen, M.H.; Blasques, J.P.A.A.; Gaunaa, M.; Natarajan, A. The DTU 10-MW reference wind turbine. In *Danish Wind Power Research; Sound/Visual Production*, Trinity: Fredericia, Denmark, 2013.
11. Carroll, J.; McDonald, A.; McMillan, D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind. Energy* **2015**, *19*, 1107–1119.
12. Bundesamt für Seeschifffahrt und Hydrographie (BSH). FINO Datenbank. Available online: <http://fino.bsh.de> (accessed on 23 January 2018)
13. Dinwoodie, I.; Endrerud, O.E.V.; Hofmann, M.; Martin, R.; Sperstad, I.B. Reference Cases for Verification of Operation and Maintenance Simulation Models for Offshore Wind Farms. *Wind. Eng.* **2015**. [CrossRef]
14. Dornhelm, E.; Seyr, H.; Muskulus, M. Vindby—A Serious Offshore Wind Farm Design Game. *Energies* **2019**, *12*, 1499. [CrossRef]
15. van Bussel, G.; Bierbooms, W. The DOWEC Offshore Reference Windfarm: Analysis of Transportation for Operation and Maintenance. *Wind. Eng.* **2003**, *27*, 381–391. [CrossRef]
16. Eurostat, the Statistical Office of the European Union. Energy Statistics-Electricity Prices for Domestic and Industrial Consumers, Price Components. Available online: <https://ec.europa.eu/eurostat/web/energy/data/database> (accessed on 4 June 2019).

17. Kost, C.; Shammugam, S.; Jülch, V.; Nguyen, H.T.; Schlegl, T. *Stromgestehungskosten Erneuerbare Energien*; Fraunhofer-Institut für Solare Energiesysteme ISE: Freiburg, Germany, 2014. (In German)
18. Martin, R.; Lazakis, I.; Barbouchi, S.; Johanning, L. Sensitivity analysis of offshore wind farm operation and maintenance cost and availability. *Renew. Energy* **2016**, *85*, 1226–1236. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).