OMAE2022-79353

STOCHASTIC WEATHER WINDOW ANALYSIS IN OPERATIONS AND MAINTENANCE PLANNING POLICIES FOR OFFSHORE FLOATING MULTI-PURPOSE PLATFORMS

Taemin Heo, Ding Peng Liu, Lance Manuel

Dept. of Civil, Architectural & Environmental Eng.
The University of Texas at Austin
Austin, Texas 78712, USA

ABSTRACT

In an emerging "blue economy," large multipurpose floating platforms in the open ocean could possibly support a range of activities including energy generation, aquaculture, seabed mining, transport, tourism, and sea-based laboratories. A Markov Decision Process (MDP) framework is developed to deal with operations and maintenance issues that are inevitable; challenges arise from the complex stochastic weather conditions that need to be accounted for. Using data as well as contrasting synthetic simulations of relevant weather variables, we demonstrate the robustness/versatility of the MDP model.

1 INTRODUCTION

Offshore floating multipurpose platforms (see, for example, Fig. 1) are becoming increasingly common in the "blue economy," with their significant benefits derived from shared use of infrastructure, resources, and services. They can combine, for instance, energy generation such as from wind and waves, aquaculture, leisure, and transport functions. Satisfactory performance of these multiple functions requires that we must deal with a wide-ranging set of operations and maintenance (O&M) activities that can take different amounts of time and might require different levels of calmness in weather conditions. In 2015, Zanuttigh et al. [1] examined 12 multipurpose platforms schemes for aquaculture, wind and wave energy to identify their benefits, costs, impacts on environment and risks. Special issues such as the frequent feeding for aquaculture systems and fouling prob-

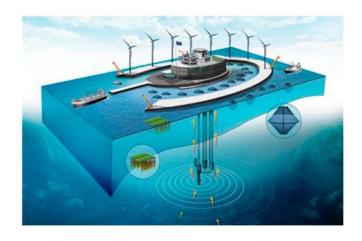


FIGURE 1. Multipurpose Platform Concept [3].

lems for wave energy converters lead to greater transportation cost while the frequency of O&M tasks for offshore wind turbines is much lower. A proper strategy of O&M planning can lead to cost reduction for O&M so as to make it comparable to that od single-purpose platforms [2].

Although there are many schedule planning strategies for O&M at offshore wind farms [4], few studies focus on multipurpose platforms. Since O&M tasks must take place nearly on a daily basis, their management plays a significant role in multipurpose platform systems [5]. Decision processes are needed in addition to mere task scheduling. Recently, a Markov Decision

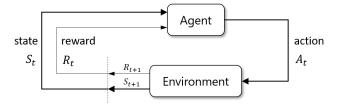


FIGURE 2. The agent-environment interaction.

Process (MDP) approach was introduced for such O&M optimization [6]. The method involves a stochastic weather window analysis that operators can employ to decide upon scheduling of work activities. Note that, in the ocean, wind, waves, and other weather-related variables influence the so-called Beaufort scale navigability, the selection of service vessels, crew personnel, etc. The MDP approach addresses balancing of loss of revenue and unsuccessfully completed O&M activities in overall cost minimization. The method was illustrated using observed metocean data at a planned site but was not assessed for robustness in performance for general metocean conditions. The present study aims to fill that gap using synthetic weather time series. Distinct weather patterns are simulated so as to represent: (1) highly variable seas (HVS); (2) very stormy seas (VSS); and (3) generally calm seas (GCS). The proposed approach seeks optimal policies in each of these distinct weather patterns to assess robustness. We compare the optimized policy for the observed metocean data set with policies for the synthetic data sets; robustness of the method is thus assessed.

2 METHODS

2.1 Theoretical Background

Figure 2 shows the agent-environment interaction that describes the general framework for the sequential decision-making problem. Distinct interface elements can define reinforcement learning including MDP, Multi-Armed Bandit (MAB), or Partially Observable Markov Decision Process (POMDP). The decision maker, also called the agent, seeks to maximize rewards by taking optimized actions.

For MDPs, actions influence immediate rewards and subsequent states through future rewards [7]. They involve state transitions and state-dependent delayed rewards, whereas bandits only explore best actions, regardless of state. POMDP further generalizes MDP by assuming hidden observations of the environment. We use MDP since weather conditions are fully observable in our problem.

There are three classes of methods for solving MDPs: dynamic programming (DP), Monte Carlo (MC) methods, and temporal-difference (TD) learning. MC methods are simple but not well suited for step-by-step incremental computation. TD

methods are a hybrid of MC and DP; they are model-free and fully incremental but complex to analyze. This study approximates a complete model of the environment from metocean data. DP methods are well established and best for a complete MDP. For these reasons, we use one type of DP—value iteration—to determine the optimized policy (see [8]). Value iteration is described in the following.

2.2 MDP Framework for O&M optimization

An MDP is a discrete-time stochastic control process defined by a 4-tuple comprising state (S), action (A), transition probability (P), and cost or reward (R) (see Figs. 2 and 3). To define one for this study, we partition the weather conditions into two groups: favorable and bad. Favorable conditions represent a calm sea state providing an operable environment. Figure 4 shows an illustrative situation; an operable environment is defined as those times when the Beaufort scale is below 5. Weather conditions are converted into a sequence of binary F (favorable) or B (bad) states for this work.

In such a binary sequence, the state space, S, is represented by a 2-tuple (x,i), where x represents the sustained duration of favorable conditions, indicating how long favorable weather has been continuing, while i represents downtime or how much time has passed since an O&M issue or need emerged.

The operator can take one of two actions: "Stay" or "Go". A "Stay" action means that the operator (or work crew) will stay at the port and wait for weather conditions to improve, while a "Go" action means the operator will set out for the needed work activity. Regardless of which action is taken by the operator, the weather stochastically changes its condition. Following a "Stay" action, the state (x,i) will transition to (x+1,i+1) or (0, i + 1), where the former occurs when favorable conditions continue, while the latter occurs when weather conditions turn bad. Following a "Go" action, the state (x, i) will transition to (x+1,0) or (0,i+1), where the former occurs when favorable conditions continue and the operator can complete the needed O&M activity, while the latter occurs when the operator fails to complete the O&M activity due to an inoperable environment in the middle of the work. To clarify these definitions, we provide an example.

Example. Assume that BFFF is a binary sequence that has been observed when an O&M need has emerged. At this point, x = 3 and i = 4. If the next update of the weather condition is favorable, the sequence becomes BFFFF but if the next weather condition is bad, the sequence becomes BFFFB. We must choose an action without knowing what the next weather condition will be. Thus, we have four possible outcomes from the action. For action "Stay", (3,4) can transition to (4,5) if the next condition turns out to be F or (0,5) if it turns out to be F. For action "Go", (3,4) can transition to (4,0) if the next

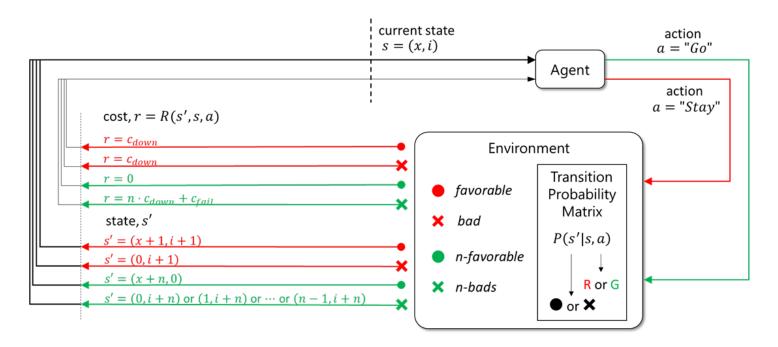


FIGURE 3. Markov decision process framework for offshore multi-purpose platform operation and maintenance optimization.

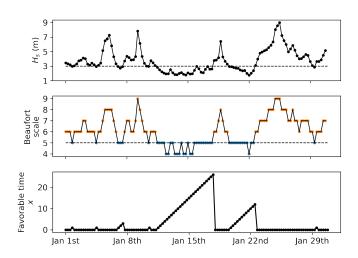


FIGURE 4. Example time series of significant wave height, H_s , along with corresponding Beaufort scale values color-coded as *favorable* (blue) and *bad* (orange). Sustained intervals of favorable conditions are also indicated.

condition is F or (0,5) if B.

Note that the intentionally simple example presented above applies to an O&M activity that requires 1 time step to complete. For a more general activity that requires n time steps to com-

plete, we need to monitor n-step transitions to fully define the outcome of a "Go" action. There will be a total of 2^n possible sequences then (we had only 2^1 possible sequences in the example above). Among these possible sequences, the O&M activity can be successfully executed in only one case that must offer n consecutive favorable conditions. Let "n-favorable" represent that single successful case while "n-bads" represent all the other cases. For the action "Stay", the operator will always observe the weather for 1 time step. Then, the outcomes of "Stay" are the same as for the n=1 case. For "Go", a state (x,i) transitions to (x+n,0) if n-favorable is realized but transitions to (0,i+1) or (1,i+1) or \ldots or (n-1,i+1) if n-bads is realized. To illustrate such a general O&M activity case, we provide an example for n=2.

Example. Return to the binary sequence BFFF considered before and its current state, (3,4). For an n=2 O&M activity, we need to monitor 2-step transitions. Then, an n-favorable situation occurs with BFFFFF while all others (BFFFBB, BFFFFB, BFFFFB) lead to an n-bads situation. For action "Stay", the outcomes are the same as n=1 case; (3,4) transitions to (4,5) or to (0,5). For action "Go", (3,4) transitions to (5,0) if an n-favorable situation results, to (0,6) for both BFFFBB and BFFFFBF; and to (1,6) for BFFFBF.

With the defined current state, s, and the next state, s', where $s, s' \in S$ and with action $a \in A$, a transition probability matrix P(s'|s,a) and cost function R(s',s,a) may be defined as shown

in Fig. 3. An example of the transition probability matrices with lexicographically ordered states is presented in Fig. 5. For visualization purposes, n is assumed to be 1 and the sustained time x with favorable conditions and the downtime i are limited to 2. As explained, a "Stay" action always increases downtime by 1. When the next weather condition is *favorable* (red circle), the sustained favorable time x also increases by 1 but if the next weather is bad (red x), the sustained favorable time goes 0. For the action "Go", the O&M activity succeeds if conditions are n-favorable (green circle). Thus, n-favorable makes downtime 0 and increases the sustained favorable time by n. For n-bads (green x), downtime increases by n and sustained favorable time goes to 0.

Note that the transition probability matrices need only two entries per row. These transition probabilities only depend on the sustained favorable times and are independent of downtime. Thus, the same entries repeat for states with the same current sustained favorable time but different downtimes. Figure 6 shows the repeating 1-step and *n*-step sustained favorable time transition probability matrices. Since the same entries repeat, we only need these sustained favorable time transition probabilities to construct full transition probability matrices. Furthermore, the *n*-step sustained favorable time transition probability matrix is the *n*-th power of the 1-step transition probability matrix. Consequently, we can estimate the 1-step sustained favorable time transition probabilities by counting transitions *favorable* and *bad* from preprocessed sustained favorable time series and normalizing them to get the complete transition probability matrices.

For the cost function R(s',s,a), we define c_{down} as the per unit time loss of revenue and c_{fail} as the unit failure cost. We assume that revenue loss starts being tallied right when an O&M issue arises and stops after successful O&M activity. We also assume that if the operator fails to complete the O&M activity, an incurred cost equal to $(n \cdot c_{down} + c_{fail})$ results, no matter when the failure to complete occurs since the desired action did not occur within the n time steps needed.

Finally, we use value iteration to arrive at the optimal policy by recursively updating value functions. The value function of a state, s, under a policy π (defining specific ways of acting), denoted $v_{\pi}(s)$, is the expected accumulated reward when starting in s and employing policy, π . We can easily achieve optimal policies once we have found the optimal value functions. The Bellman equation—a necessary condition for optimality associated with dynamic programming—separates the value function into two parts: an immediate reward and a discounted future value. Thus, the optimal value function can be written as:

$$v_{\pi_*}(s) = \max_{a \in A} \left[\sum_{s' \in S} P(s'|s, a) R(s, a, s') + \gamma \sum_{s' \in S} P(s'|s, a) v_{\pi^*}(s') \right].$$

Starting from the initial value function in which zero values are

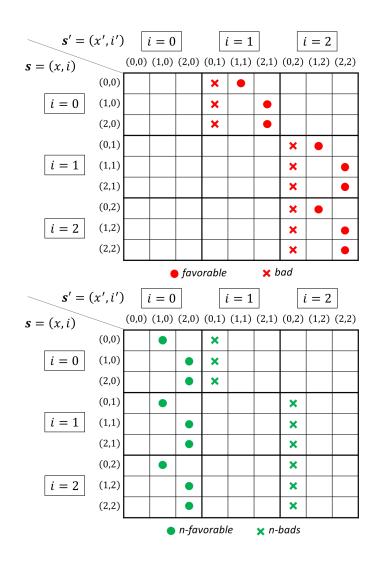


FIGURE 5. An example of transition probability matrices for (top) "Stay" and (bottom) "Go" actions.

assigned to all states, Algorithm 1 shows how the value iteration recursively updates the value function until it converges to the optimal value function.

2.3 Synthetic weather condition generation

To assess robustness in our MDP performance for general metocean conditions, we use a Markov chain (MC) simulation to generate synthetic weather conditions. Note that this new Markov chain is different from the one used for the MDP framework itself. This Markov chain consists of states and transition probabilities. To distinguish transition probability matrices involved in the MC-based synthetic weather condition generation from the transition probability matrices for MDP, let us denote them as TPM_{MC} and TPM_{MDP} . The state space of the Markov

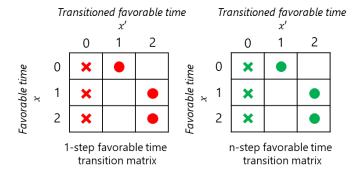


FIGURE 6. An example of repeating 1-step and *n*-step sustained favorable time transition probability matrices.

Algorithm 1 Value Iteration

Initialize
$$v_0(s) = 0$$

 $v_1(s) \leftarrow \max_a \sum_{s',r} P(s'|s,a)[R(s,a,s') + \gamma v_0(s')]$
 $k = 0$ and $\varepsilon =$ tolerance
while $|v_{k+1} - v_k| > \varepsilon$ **do**
 $k = k+1$
 $v_{k+1} \leftarrow \max_a \sum_{s',r} P(s'|s,a)[R(s,a,s') + \gamma v_k(s')]$
end while
 $\pi_*(s) = \operatorname{argmax}_a \sum_{s',r} P(s'|s,a)[R(s,a,s') + \gamma v_{k+1}(s')]$

chain are the values of x (sustained favorable time), which can transition to either 0 or x+1. Therefore, stochastic simulation of the Markov chain with TPM_{MC} generates time series of sustained favorable times that match weather patterns per the target TPM_{MC} .

We define three different TPM_{MC} targets corresponding to three distinct weather patterns: (1) highly variable seas (HVS); (2) very sustained seas (VSS); and (3) generally calm seas (GCS). The color-mapped 1-step sustained favorable time transition probabilities are presented in Fig. 7. Synthetic weather time series for the distinct conditions are presented in Fig. 8. Consistent with its name, HVS defines greatly fluctuating seas; the likelihood of bad weather conditions is always 50%. VSS defines seas with long sustained periods of either favorable or bad weather. For VSS, when x (sustained favorable time) is 0, the likelihood of continuing bad weather conditions is 90%. This means that with VSS, the weather tends to stay bad when the prior period weather was bad. On the other hand, for x > 1, i.e., when favorable weather conditions have lasted at least one period, the likelihood of favorable conditions continuing is 95%. Thus, when the sea enters favorable conditions, it tends to stay in those favorable conditions. As seen in Fig. 8, when the weather stays calm, the sustained favorable time steadily increases. The periods through the rise to the triangular peaks in the time series are periods of favorable weather. On the other hand, the sustained favorable weather time stays at zero if the weather is bad. Figure 8 clearly shows the targeted highly fluctuating behavior for HVS and the long sustained periods of relatively unchanging weather for VSS. Finally, GCS describes metocean conditions stochastically similar to those recorded at the planned Scottish site (SCT) [6]. The SCT transition probability matrix was computed by counting transitions from the actual metocean time series and normalizing them. Since the data set is of limited length, not all transitions are observed. The empty white cell in Fig. 7 is a result of this type of lack of observations. By design, GCS has the same target transition probabilities as SCT, and the empty (missing) value for SCT was interpolated, in defining GCS, from the adjacent transition probabilities. For VSS, GCS, and SCT, the sustained favorable time duration is limited to 50 time steps to avoid meaningless long dwell times of sustained favorable conditions.

Using the synthetically generated sustained favorable time series, we compute the transition matrices, TPM_{MDP} . The discussion concerning Fig. 5 and related explanation is what we need to count different transitions and normalize them to yield TPM_{MDP} . The procedure of analysis for given values of n, c_{down} , and c_{fail} is summarized below.

- 1. Choose a weather pattern and the corresponding TPM_{MC} .
- 2. Run stochastic simulations of sustained favorable times using Markov chain so as to yield 10^6 time steps starting from x = 0.
- 3. Count transitions *favorable* and *bad* from the generated sustained favorable time series and normalize them to obtain the 1-step sustained favorable time transition matrix.
- 4. Construct the complete TPM_{MDP} using the 1-step favorable time transition matrix.
- 5. Solve the MDP using value iteration to obtain the optimal policy.

3 O&M MODEL ROBUSTNESS ASSESSMENT

Note that in the examples presented, a single time step is equal to 6 hours. Two cases associated with different assumptions in loss of revenue due to downtime are discussed.

3.1 Case of constant downtime costs

The O&M issues for an energy generation sector are associated with constant (time-independent) downtime costs. From Fig. 3, we note that all costs are independent of downtime if c_{down} is constant. Hence, the optimal policy will also be independent of downtime. For simplicity, we select $c_{down} = -2$, and $c_{fail} = -1$ for this assessment; we are only really concerned with the ratio of c_{down} to c_{fail} when deciding on the optimal action for a given state. Our MDP framework, derives an optimal policy from the generated synthetic weather conditions without need for

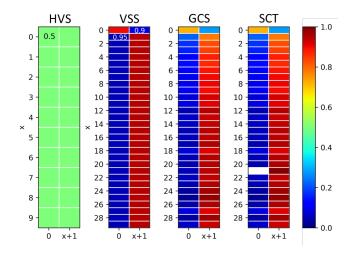


FIGURE 7. Transition probability matrices for the synthetic weather condition generation: (1) highly variable seas (HVS); (2) very sustained seas (VSS); (3) generally calm seas (GCS); (4) as-recorded seas at the Scottish Site (SCT).

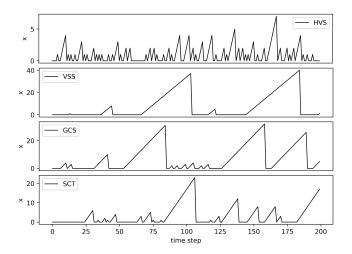


FIGURE 8. Three distinct synthetically generated weather conditions (HVS, VSS, and GCS) and observed seas at a planned Scottish site (SCT) Note that, for SCT, a time step is equivalent to 6 hours.

the actual transition probability matrices used to generate those data. Figure 9 compares computed optimal policies for n = 1 and n = 3 O&M activities.

Figure 9 shows that our MDP framework captures specific patterns for the different metocean conditions and produces optimal policies for O&M activities that require different planning time steps. For HVS, the policy indicates that an n=1 O&M activity is feasible for all sustained favorable times, but the n=3 O&M activity is risky to pursue due to the high variability of sea

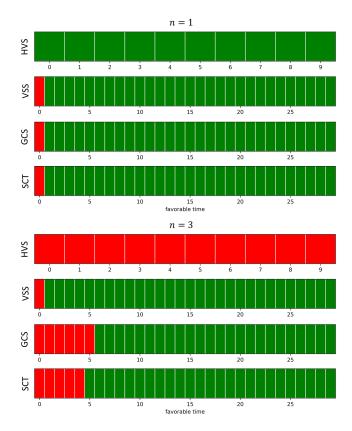


FIGURE 9. Optimal O&M policies for three synthetic metocean conditions and for the SCT case (Red: "Stay", Green: "Go") (top) n = 1 and (bottom) n = 3.

conditions. These results are rational and expected, given HVS's fluctuating weather pattern.

For VSS, the MDP suggests the same optimal policy for both the n=1 and n=3 O&M activities. Since VSS is associated with longer period of sustained favorable and bad conditions, our MDP framework suggests that both types of activities may be undertaken as long as the weather has entered the favorable period. Therefore, the optimal policy recommends "Stay" action if the favorable time is 0 and "Go" for $x \ge 1$.

The similar optimized policies derived for the GCS and SCT cases confirms the robustness of our method (GCS transition probabilities were synthetic targets selected to match observed SCT weather). Figure 8 shows that GCS and SCT conditions exhibit somewhat more weather variability than VSS. They have occasional long periods of favorable weather, but we also observe a few bumpy peaks (see Fig. 8) indicating fluctuating conditions. The optimized policies established in Fig. 9 are compatible with these weather patterns of GCS and SCT. This all leads to the same policy as VSS for the n=1 O&M activity, but we need to wait for longer times to start work for the n=3 activity. This is a

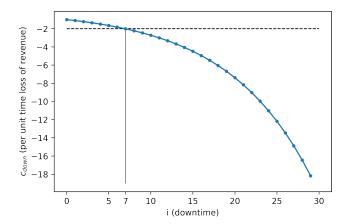


FIGURE 10. Time-dependent per unit time loss of revenue versus downtime

result of the sporadic fluctuating weather conditions of GCS and SCT. Due to this unpredictable weather variability of the GCS and SCT conditions, even though not as extreme as HVS, we need to wait longer than VSS to make sure the weather will be calm throughout the planned activity.

The example case study for the constant (i.e., time-independent) downtime costs has shown how the proposed MDP framework learns different weather patterns and optimizes policies for different types of O&M activities. The optimized policies conform to the expected weather patterns, and the MDP framework rationally identifies and schedules O&M activities that are acceptable to conduct and when are appropriate times for crews/vessels to depart for the work to be carried out, while taking into consideration overall cost and safety.

3.2 Case of time-dependent downtime costs

For O&M issues related to some activities, such as fish farming in an aquaculture sector, the per-unit-time loss in revenue can vary with time associated with delay in addressing the needed problem or service. We consider a case where such losses in revenue are assumed to increase exponentially with downtime. If c_{down} is a function of downtime, the derived optimal policy will also depend on the downtime. We assume that $c_{down}(i) = -e^{0.1 \cdot i}$; all other parameters are the same as in the previous case. Figure 10 shows the assumed $c_{down}(i)$ variation with downtime. For the constant c_{down} case considered earlier, we assumed $c_{down} = -2$. It is seen that $c_{down}(i)$ is closest to -2 when downtime, i = 7. Although it is not possible to match the optimized policy of the previous case due to the costs with increasing downtime, we might expect a similar O&M policy around i = 7.

Figure 11 shows derived optimal policies for n = 1 and n = 3 O&M activities for all the different weather conditions. The HVS optimized policy for n = 1 activity suggests that we should

not start work when downtime is zero (which is obvious), and indicates that the required O&M activity is feasible for all other states. As with the constant c_{down} case, our MDP framework shows that the n=3 activity is not feasible in any states due to the high probability of failure expected from the fluctuating weather patterns associated with HVS.

Up to a downtime of 14, VSS's optimized policy for the n=1 activity is the same as for the previous constant c_{down} case. When the downtime exceeds 14, it is beneficial to begin the work activity, regardless of the current weather regime since $c_{down}(i)$ values and loss of revenue become too large. For the n=3 activity, the optimized policy suggests that staying and waiting for favorable weather is preferred when the downtime is less than 3. This is because we might expect long-lasting calm conditions when VSS enters a favorable weather regime, and its benefit is greater when compared with trading off of missed time steps and resulting loss of revenue. The optimized policy for n=3 also indicates that we should not risk performing work in a bad weather regime even if $c_{down}(i)$ is so high because failure is very likely to occur.

For the n = 1 activity, the GCS and SCT weather patterns lead to similar results as for VSS. For the n = 3 activity, the recommended waiting times vary depending on the time of sustained favorable weather, x. The derived policy reflects a complex weather transition probability situation at the actual site in Scotland.

To further justify the results presented, we highlight optimized policies around i=7, in Fig. 11, with a yellow box. As mentioned before, the derived policies do not exactly match those for the constant downtime case. Nevertheless, the results are fairly similar in the two cases. Note that Fig. 11 also shows analytical results for the three synthetic weather conditions (i.e., excluding SCT). These analytically derived policies resulted from using the assumed TPM_{MC} by considering it as 1-step sustained favorable time transition matrix in the TPM_{MDP} solution. Except for a few states, the simulation-based optimized policies are almost identical to these analytical answers. This establishes the the claim of sample size-based adequacy and robustness of the MDP framework.

In summary, the case study involving time-dependent downtime costs shows that the proposed MDP framework realistically captures state transition patterns from the input data and offers different but expected optimal policies for the distinct metocean conditions.

4 CONCLUSIONS

The robustness of the proposed MDP framework for planning O&M activities around offshore multi-purpose platforms has demonstrated and optimal policies have been derived for greatly contrasting synthetic metocean conditions, one of which was selected to match actual site conditions at a Scottish off-

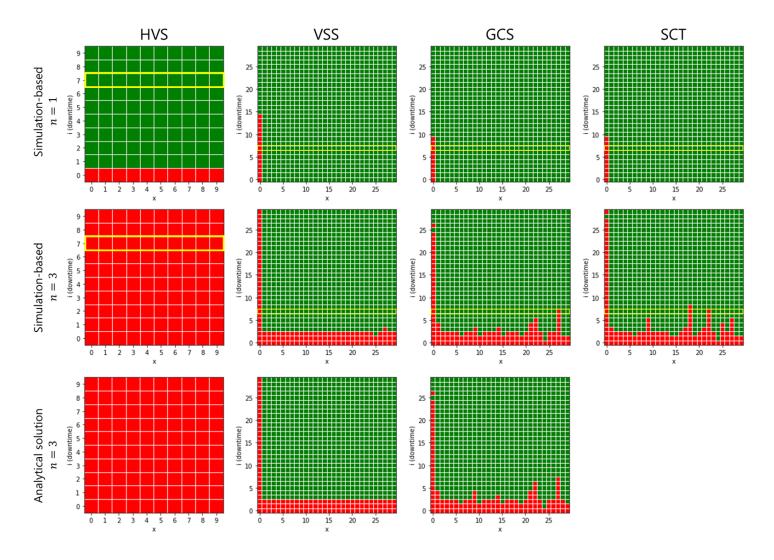


FIGURE 11. Optimal policies for three synthetic metocean conditions and for the SCT case (Red: "Stay", Green: "Go").

shore site. The activities addressed included shorter and longer work periods that are often needed. Other methods for sequential decision-making in reinforcement learning were briefly discussed. Two examples involving constant and time-dependent downtime costs and associated loss of revenue were evaluated for all the different metocean conditions. Derived optimized O&M plans are reasonable, given the specific weather patterns and planned activities, and may be employed in future blue economy projects involving various maritime activities.

REFERENCES

[1] Zanuttigh, B., Angelelli, E., Bellotti, G., Romano, A., Krontira, Y., Troianos, D., Suffredini, R., Franceschi, G., Cantù, M., Airoldi, L., et al., 2015. "Boosting blue growth in a mild

- sea: analysis of the synergies produced by a multi-purpose offshore installation in the Northern Adriatic, Italy". *Sustainability*, 7(6), pp. 6804–6853.
- [2] Foteinis, S., and Tsoutsos, T., 2017. "Strategies to improve sustainability and offset the initial high capital expenditure of wave energy converters (WECs)". *Renewable and Sustainable Energy Reviews*, **70**, pp. 775–785.
- [3] Sie, Y.-T., Château, P.-A., Chang, Y.-C., and Lu, S.-Y., 2018. "Stakeholders opinions on multi-use deep water offshore platform in Hsiao-Liu-Chiu, Taiwan". *International Journal of Environmental Research and Public Health*, 15(2), p. 281.
- [4] Seyr, H., and Muskulus, M., 2019. "Decision support models for operations and maintenance for offshore wind farms: A review". *Applied Sciences*, **9**(2), p. 278.

- [5] Röckmann, C., Lagerveld, S., and Stavenuiter, J., 2017. "Operation and maintenance costs of offshore wind farms and potential multi-use platforms in the dutch north sea". In *Aquaculture Perspective of Multi-Use Sites in the Open Ocean*. Springer, Cham, pp. 97–113.
- [6] Heo, T., Nguyen, P. T., Manuel, L., Collu, M., Abhinav, K., Xu, X., and Brizzi, G., 2020. "Operations and maintenance for multipurpose offshore platforms using statistical weather window analysis". In Global Oceans 2020: Singapore–US Gulf Coast, IEEE, pp. 1–7.
- [7] Sutton, R. S., and Barto, A. G., 2018. *Reinforcement Learning: An Introduction*. MIT press.
- [8] Papakonstantinou, K. G., and Shinozuka, M., 2014. "Planning structural inspection and maintenance policies via dynamic programming and markov processes. Part I: Theory". Reliability Engineering & System Safety, 130, pp. 202–213.