*Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)*
*Edited by* Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
©2023 ESREL2023 Organizers. *Published by* Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P555-cd

# Condition-based maintenance optimization of a large-scale system with a POMDP formulation: evaluation of a heursitic policy

Matthieu Roux

*LGI EA 2606, Univ. Paris-Saclay, CentraleSupelec, France. E-mail: matthieu.roux@centralesupelec.fr*

Yi-Ping Fang

*LGI EA 2606, Univ. Paris-Saclay, CentraleSupelec, France. E-mail: yiping.fang@centralesupelec.fr*

Anne Barros

*LGI EA 2606, Univ. Paris-Saclay, CentraleSupelec, France. E-mail: anne.barros@centralesupelec.fr*

This study investigates the question of the evaluation of a heuristic condition-based maintenance policy applied to a distributed multi-unit system. In particular, the system is composed of many units which function and degrade independently. We call it distributed since the system's total output is the sum of the individual output of each unit, where the failure of one unit has no impact on the functioning state of other units (i.e., no series/parallel structure). Taking into account the large-scale nature of the problem, which consists of a large number of units, is crucial because a good maintenance policy should coordinate the decisions at the scale of the system. Said differently, maintenance decisions cannot be taken independently unit-by-unit for the following reasons. First, the maintenance resource is limited and should be wisely allocated across the system. Second, as deploying on-site a maintenance crew is expensive, a good maintenance policy should also try to limit the number of deployments and group maintenance interventions. A condition-based maintenance policy relies on condition monitoring information, which we assume to be imperfect. We model this imperfection by assuming that remote sensors inaccurately estimate the true degradation state of the units. The decision-maker should then choose, at each time step, whether a maintenance operation or an inspection must be performed based on the information one has collected. We formulate the problem as a partially observable Markov decision process (POMDP). However, due to the curse of dimensionality, it cannot be solved via well-known approximate dynamic techniques. We then propose a heuristic algorithm based on a decomposition of the problem. The contribution of this work is to propose a framework to evaluate and validate this algorithm. We first validate the approach on a realistic-sized instance and show that the obtained policy has the expected properties (in terms of structure or value of information for different qualities of condition monitoring). Second, we validate the design of our procedure by showing that in a variety of scenarios, our heuristic performs better than its simpler (and more naive) alternatives.

*Keywords*: condition-based maintenance, partially observable Markov decision process (POMDP), large-scale multi-unit system, heuristic policy validation

## 1. Introduction

A large variety of industrial systems are often organized as a fleet, composed of many independent units that function and degrade over time. Together, those items contribute to the whole fleet's global service level. For example, wind turbines in large offshore wind farms, trains, or distributed data centers are complex systems organized as fleets. However, as they degrade over time, those items will eventually fail if they are not adequately maintained, affecting the fleet's total revenue or service level. Thus, it is essential to implement a maintenance policy consisting of deciding on the inspection of the material and the maintenance operations. For a couple of decades, condition-based maintenance (CBM) strategies have proven to be effective, providing better solutions with respect to the overall cost or availability of fleet units. Essentially, this family of maintenance policies leverages the information collected via condition monitoring (remote sensors or human inspections) to infer the degradation state of the monitored item and schedule the next maintenance operation as wisely as possible. In such context, the general state of mind is often to balance two concurrent goals: i) prevent unexpected failures of degraded items by scheduling preventive maintenance interventions, while at the same time ii) save costs by scheduling as few interventions as necessary and use the full lifetime of the items (Alaswad (2017)).

This work is inspired by Yildirim (2017), which focuses on optimizing the CBM policy of turbines in an offshore wind farm. In particular, the possibility to group several maintenance interventions to save costs (e.g., setup cost, deployment cost), which is a form of opportunistic maintenance, is a crucial element to optimize since deploying a maintenance crew on-site is very expensive and requires a specific material. The authors propose a heuristic mixed-integer linear program (MILP) to optimize the maintenance decisions with a rolling horizon approach, where condition monitoring coming from remote sensors is used at each new re-optimization. In this study, we propose to go one step further and explore a model where the condition monitoring is imperfect, meaning that the decision-maker does not have access to the exact degradation states when making maintenance decisions. In reality, such an approach is motivated by the fact that many remote sensors only partially capture the data required to infer the degradation state. Moreover, with the development of monitoring technologies, many companies are now considering implementing CBM policies. What we observed from our interactions is that they are often concerned by one of the following questions: 1) What if the sensors are not as accurate as guaranteed by the manufacturer? How could failing sensors affect the performance of the overall CBM policy? 2) What are the costs and benefits of implementing CBM? If several levels of condition monitoring systems (i.e., characterized by their monitoring accuracy) are available at different costs, how can we choose the ”best” one? As those questions are related, we propose investigating a CBM optimization framework to manage a fleet system with imperfect monitoring in this article. This study then allows us to compute the value of information (VoI) of a given monitoring technology, defined in Memarzadeh (2016) as ”the maximum cost a decision-maker is willing to pay for getting this information” and compare different alternatives.

To do so, we propose to model the problem as a partially observable Markov decision process (POMDP). It is an interesting framework that allows modeling a sequential decision-making problem based on some partial knowledge about the system’s state. For us, it consists of finding the maintenance policy (scheduling inspections and maintenance operations based on imperfect knowledge about the degradation state of each item in the fleet) that minimizes the total expected cost. POMDPs were first introduced by Sondik (1971), but much progress have been made in the solving algorithms in the last 20 years, explaining why they are applied more and more to industrial use cases, e.g., Papakonstantinou (2014a), Papakonstantinou (2014b). Solving POMDPs is a complex problem (see Kaelbling (1998)), and exact algorithms, like Cassandra (2013), quickly suffer from the curse of dimensionality and history. However, recent advances in approximate dynamic programming algorithms made it possible to solve with good precision much larger problems (in terms of size of the finite state, action, and observation spaces); among those state-of-the-art solvers, we can cite, for example, SARSOP (Kurniawati (2008)) or PERSEUS (Spaan (2005)).

Nevertheless, even the most advanced approximate algorithms cannot handle more than tens of thousands of states in a reasonable amount of time. For our applications, we would like to be capable of handling fleets of around a hundred units. Considering that each item can be in five degradation states, this would still result in a huge state space of size $5^{100} \sim 10^{35}$. The same reasoning would also apply to the action and observation spaces. Moreover, in order to make cost-effective maintenance decisions, leveraging, for example, the opportunities offered by opportunistic maintenance, maintenance decisions cannot be taken independently unit-by-unit. This economic dependency, further reinforced by a limited maintenance resource, forces us to consider maintenance policies where interventions have some degree of coordination. To address this issue, we propose a heuristic policy with the advantage of efficiently scaling up to large-scale systems. The term ”heuristic” means that we have no mathematical guarantee that the result will be close to the optimal solution of the original problem, i.e., here the optimal maintenance policy. The proposed approach consists of a hybrid optimization framework mixing dynamic programming (DP) and integer linear programming (ILP). In this article, we will, in particular, focus on validating such an approach. We will show that, even if we do not have any guarantee of performance, various numerical tests still indicate that it is an interesting approach providing valuable results within reasonable computation time.

## 2. Problem description

### 2.1. *Large-scale fleet system*

We start first by introducing in more detail what we call a fleet system. In the rest of the paper, we consider that a fleet is an industrial system composed of many independent items. Each item (or unit) may be viewed as a multi-component sub-system, but this is outside the scope of our study. The most important characteristic is that each item functions and degrades independently of other items in the fleet. This is a reasonable assumption for many systems, such as trains or wind turbines. Indeed, external factors may simultaneously affect the degradation level of several items (e.g., weather conditions, load, and state of the rail infrastructure). However, here we will neglect this second-order aspect (which could be considered in later extensions of this work).

Items progressively degrade following a particular degradation process and will eventually fail if no preventive maintenance operation has been carried out in time. A failed item is a problem for the management of the fleet since it corresponds to lost profit (modeled as an opportunity cost) or, more generally, a degradation of the level of service (too many failed trains in a fleet are problematic because it forces the company to cancel some journeys and refund affected passengers). As explained in the introduction, the management challenge of such a fleet comes from two main factors: 1) the combinatorial nature of the problem, which grows exponentially with the number of items in the fleet, and 2) the fact that maintenance decisions need to be coordinated at the scale of the whole fleet.

### 2.2. *Modeling the problem as a POMDP*

In a previous paper (Roux (2022)), we modeled the maintenance problem of a single-item system as a POMDP. The goal was to study the impact of different condition monitoring accuracies on the optimized maintenance policy and its associated value. This work is a natural extension where we apply the POMDP model to a larger system composed of many units. Therefore, the notation is very similar, and we invite the reader to have a look at this previous paper for more details on the 1-item POMDP.

The set of items is noted as $I$. For the sake of simplicity, we assume that all items are identical (e.g., same degradation process, maintenance cost, contribution to the fleet's performance), but this is not a restrictive assumption, and this could easily be generalized. The degradation state of item $i \in I$ at time $t$ is noted $s_{i,t} \in \hat{\mathcal{S}} = \{S_1, S_2, S_3, S_4, F\}$, where $F$ is the unique failure state.

At each time step, one action $a_{i,t} \in \hat{\mathcal{A}} = \{\mathbf{NA}, \mathbf{PM}, \mathbf{CM}, \mathbf{I}\}$ must be selected by the decision-maker for each item:

- **NA** corresponds to the default action of doing nothing
- **PM** is a preventive maintenance operation, which should be performed on a degraded yet still functioning item to put it back to the as-good-as-new state $S_1$
- **CM** is a corrective maintenance operation to repair a failed item
- **I** corresponds to a perfect inspection of the item; the decision-maker can pay to improve the knowledge about the item state and perfectly access to $s_{i,t}$, which complements the imperfect remote sensors information.

A cost is associated with each individual action, corresponding to the cost of performing preventive/corrective maintenance or an inspection ($c_{PM}, c_{CM}, c_I$). Regarding the remote condition monitoring system, we assume that each item is equipped with a sensor providing imperfect monitoring data. Periodically, every $K$ time steps (called the observation period), an observation $o_{i,t}$ can be used to estimate the state $s_{i,t}$ of the item. Eventually, we assume that in that model, failures are self-announcing, meaning that if an item fails, it will be immediately known by the decision-maker without any delay or ambiguity.

For the entire fleet, states are modeled by the finite set $\mathcal{S}$ containing all the combinations of individual degradation states of the items; the same applies to action and observation space, leaving us with the generic notation $s_t \in \mathcal{S}$, $a_t \in \mathcal{A}$ and $o_t \in \mathcal{O}$ as state, action, and observation at time $t$ for the whole fleet.

We formulate the problem as the following optimization program:

$$\min_{\pi\in\Pi} \mathbb{E}\Big[\sum_{t=0}^{+\infty} \gamma^t \cdot c(s_t, a_t)\Big] \qquad (1)$$

where $a_t = \pi(b_t)$, with $b_t$ being the belief about the system state at time $t$; $b_t[s]$ represents the probability that the system is in state $s \in \mathcal{S}$ given all the information the decision-maker has accumulated so far. The cost $c(s_t, a_t)$ when action $a_t$ is selected is the sum of the individual costs resulting from the maintenance operations on the items (i.e., cost of preventive or corrective maintenance, cost of inspection or opportunity cost from a failed item) plus a deployment cost that is paid when we have to deploy a maintenance crew to perform at least one intervention on the fleet (i.e., paid as soon as some other action than do noting is selected). The discount factor $\gamma$ enables us to sum the costs over an infinite time horizon. Finally, a resource constraint limits the number of simultaneous interventions that can be scheduled at a given time step, limiting in practice the actions to a subset of feasible actions $\tilde{\mathcal{A}} \subset \mathcal{A}$.

## 3. Heuristic solving approach

Because applying traditional POMDP solving algorithms is impossible for such a large problem, we adopt a heuristic approach. The idea is to use the Q-values computed from solving the much simpler 1-item problem as a proxy to design a meaningful cost function for an ILP.

### 3.1. *Modified 1-item POMDP*

We first consider the same problem as previously explained but now assume it has only 1 item. In that situation, the resource constraint becomes non-limiting and thus can be ignored. If we apply a point-based value iteration algorithm (e.g., SARSOP, Kurniawati (2008) or PERSEUS, Spaan (2005)), we can compute by approximate dynamic programming the Q-values $Q^*(s, a)$. They give us the approximate, but with good precision, expected cost corresponding to taking action $a$ when the item is in state $s$, and then, following the optimal 1-item maintenance policy.

The only thing we need to decide is how to determine the allocated fraction of the deployment cost, which is by nature a global cost (i.e., difficult to attribute to a particular unit in case several interventions have been opportunistically grouped together). Here, with only 1 item in the system, it means that we must pay the deployment cost each time we intervene on the item, whereas in reality, in a multi-item system, a cost-effective maintenance strategy would try to group maintenance interventions in order to save deployment costs. In fact, this cost is essential because its value strongly impacts how likely we are to schedule inspections: the more expensive the deployment cost, the less likely we are willing to schedule inspections, because they are more costly and thus less worthwile.

#### 3.1.1. *Zero deployment cost*

The first naive idea would be to neglect the deployment cost and compute the Q-values for a 1-item problem with no deployment cost at all. By doing so, we underestimate the cost of maintenance interventions on the unit, but it has the advantage of being easy to compute.

#### 3.1.2. *Full deployment cost*

The second simplistic idea consists of the opposite of allocating the full deployment cost in the computation of the Q-values. In that case, it is clear that we tend to overestimate the cost of maintenance interventions, but it is still straightforward to compute.

#### 3.1.3. *Proposed method*

The proposed method we will test in the next section is an intermediate approach. It consists of modeling the interaction between a unit $i$ and the rest of the fleet. Here, the entity "the rest of the fleet" is not finely modeled; we only consider its aggregated behavior. In the end, it is all that matters because this is all we need to know to schedule interventions for the unit $i$: if too much resource is used by the rest of the fleet, the considered item cannot schedule any intervention; but similarly, if the rest of the fleet uses some resource, unit $i$ and the rest of the fleet can now "share" the deployment cost, thus encouraging group interventions to save deployment cost.

This model enables us to compute the Q-values of a 1-item POMDP which is supposed to approximate the interaction between a particular unit and the rest of the fleet. Heuristically, the underlying idea is that it leads to modeling more precisely the maintenance resource limitation and leveraging the opportunistic maintenance.

### 3.2. *An ILP to schedule operations at the scale of the fleet*

Now that we have presented the three alternatives for the modified 1-item POMDP, we show how we intend to use it to produce a maintenance policy at the fleet level. To do so, we formulate a sequential optimization framework where we solve an ILP at the beginning of each observation epoch, i.e., every $K$ time steps. We use this optimization program to schedule the interventions to be performed for the next $K$ time steps. Such a choice of a re-optimization period is convenient because it corresponds to the observation period, meaning that we can produce a maintenance schedule that leverages the monitoring information. The way of proceeding is the following:

(1) At the beginning of an observation epoch, we collect the imperfect observation $o_t$ and update the belief state $b_t$;
(2) The ILP is solved, providing a maintenance schedule for the next $K$ time steps;
(3) We execute the maintenance schedule for the whole observation epoch, then start the process again.

Let:

- $x_t \in \{0,1\}$ be a binary variable indicating whether we should schedule a deployment at time $t$;
- $I_w \subset I$ be the subset of items in a working state at time $t = 0$;
- $I_f \subset I$ be the subset of items being failed at time $t = 0$;
- $z^I_{i,t} \in \{0,1\}$ be a binary variable indicating whether we should schedule an inspection of item $i$ at time $t$;
- $z^{PM}_{i,t} \in \{0,1\}$ be a binary variable indicating whether we should schedule a preventive maintenance on item $i$ at time $t$;
- $z^{CM}_{i,t} \in \{0,1\}$ be a binary variable indicating whether we should schedule a corrective maintenance on item $i$ at time $t$;
- $z^{NA}_i \in \{0,1\}$ be a binary variable indicating whether we should not schedule any intervention during the next $K$ time steps on item $i$;
- $b_i$ be the belief state of item $i$ at time $t = 0$ after taking into account the observation;
- $\tilde{Q}^I(b,t)$ (resp. $\tilde{Q}^{PM}(b,t)$) be an estimation of all the expected future costs associated to an item initially described by the belief $b$ if we schedule an inspection (resp. preventive maintenance) at time $t$ and then follow the optimal policy;
- $\tilde{Q}^{CM}(t)$ be an estimation of all the expected future costs associated to an item that is initially failed if we schedule corrective maintenance at time $t$ and then follow the optimal policy;
- $\tilde{Q}^{NA}(b)$ be an estimation of all the expected future costs associated to an item initially described by the belief $b$ if we schedule to do nothing on the next $K$ time steps and then follow the optimal policy;
- $r_a$ be the resource required for intervention $a \in \hat{\mathcal{A}}$;
- $R$ be the total maintenance resource available at each time step.

The ILP to solve at each observation epoch is the following:

$$
\begin{aligned}
\min_{x,z} \sum_{t=0}^{K-1} & \left( c_{deploy} \cdot x_t + u_t^I + u_t^{PM} + u_t^{CM} \right) \\
& + u^{NA} \\
\text{s.t.} \quad & u_t^I = \sum_{i \in I_w} \tilde{Q}^I(b_i, t) \cdot z_{i,t}^I \\
& u_t^{PM} = \sum_{i \in I_w} \tilde{Q}^{PM}(b_i, t) \cdot z_{i,t}^{PM} \\
& u_t^{CM} = \sum_{i \in I_f} \tilde{Q}^{CM}(t) \cdot z_{i,t}^{CM} \\
& u^{NA} = \sum_{i \in I_w} \tilde{Q}^{NA}(b_i) \cdot z_i^{NA} \\
& \sum_{t=0}^{K-1} \left( z_{i,t}^I + z_{i,t}^{PM} \right) + z_i^{NA} = 1, \quad \forall i \in I_w \\
& \sum_{t=0}^{K-1} z_{i,t}^{CM} = 1, \quad \forall i \in I_f \\
& z_{i,t}^a \leq x_t, \quad \forall i \in I, \forall a \in \{I, PM, CM\} \\
& \sum_{a,i} z_{i,t}^a \cdot r_a \leq R, \quad \forall t
\end{aligned}
\tag{2}
$$

## 4. Numerical results

We performed some numerical analysis to test the validity of our heuristic procedure. We tested our method on a fleet composed of 50 items, and the first promising result is that our approach can compute and simulate an inspection and maintenance policy in a reasonable amount of time (whereas the traditional approach using approximate dynamic programming is clearly intractable). In this section, we describe the different tests we performed and describe to what extent they can, at least partially, contribute to validate the proposed heuristic method.

### 4.1. *Validation of the iterative procedure*

The method we propose relies on an iterative procedure. As computations go, we can access better estimations of the aggregated behavior of the ”rest of the fleet”, thus enabling us to further improve the computation of our Q-values for the modified 1-item POMDP. We then expect our procedure to converge towards a low-cost solution (maybe not the global optimum, but at least some ’local’ minimum). In Fig. 1, we observe that the method computes better maintenance policies after a few iterations. Moreover, this is true for different scenarios of monitoring performance (i.e., the quality of the imperfect monitoring observations used at each observation epoch). From this observation, we can conclude that our modified 1-item POMDP is able to capture (maybe partially) the interaction between a given unit and the rest of the fleet and that a better model of the aggregated behavior of the rest of the fleet leads to a better solution, in terms of total cost.

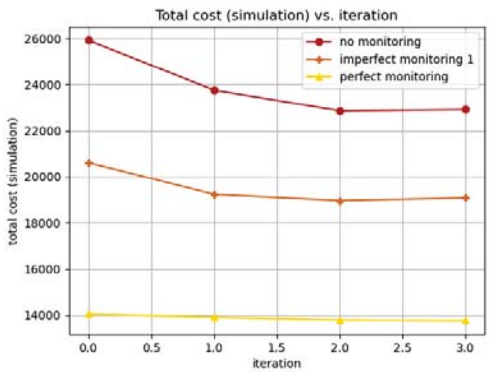


Fig. 1. Convergence towards a low-cost solution after a few iterations.

### 4.2. *Value of information*

When computing the average cost for different monitoring performances, we observe that our method can capture a value of information (VoI), as defined in Memarzadeh (2016). We observe in Fig. 2 that the better the condition monitoring system, the lower the overall maintenance cost. Even if we still have no guarantee of performance, this is another indication that our approach is consistent and produces relevant and sound policies.

### 4.3. *Comparison with simpler alternatives*

We compare our proposed method, which relies on the modified 1-item POMDP, with the two

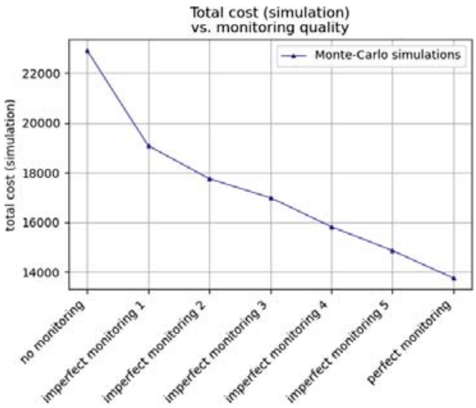


Fig. 2. Value of information.

Table 1. Comparison between our proposed method and the potential simpler alternatives; total cost evaluated via simulations for different settings of monitoring quality.

| Monitoring quality | Proposed method (cost) | Full deploy. | Zero deploy. |
|---|---|---|---|
| no monitoring | 22,920 | +10.3% | +13.9% |
| monitoring 1 | 19,090 | +8.0% | +23.3% |
| monitoring 2 | 17,760 | +7.6% | +26.6% |
| monitoring 3 | 17,000 | +6.9% | +29.5% |
| monitoring 4 | 15,830 | +4.7% | +35.2% |
| monitoring 5 | 14,870 | +3.4% | +41.4% |
| perfect | 13,760 | +2.1% | +55.8% |

much simpler alternatives evoked in sections 3.1.1 and 3.1.2 (i.e., zero and full deployment cost). In Table 1, we reported the total cost obtained via simulation for different scenarios of monitoring accuracy, defined by the conditional probabilities $\mathbb{P}(o_t|s_t)$ that give the probability of receiving observation $o_t$ given that the item is in degradation state $s_t$. It shows that our method systematically finds a maintenance policy of lower cost compared to the alternatives (and the gap is quite significant). We can conclude that the model we proposed for the interaction between a given unit and the rest of the fleet is relevant to consider in the sense that, since it leads to better solutions, it is a reasonable increase in the model complexity provided its better performance.

## 5. Conclusion

From this work, we can conclude that validating a heuristic maintenance policy is not straightforward. Here, we studied a situation where we assumed we had enough data to perform accurate optimization and simulations. This assumption enabled us to use simulations to compare different policies, but this may not always be possible (e.g., for relatively new material with little historical data). However, without a tractable exact or approximate (with bounds) algorithm to compute the optimal solution, evaluating the performance of a given heuristic optimization framework is challenging.

In this paper, we showed that, even if a thorough validation relying for example on a gap to the optimal solution is not possible, there are still some properties that we can exploit and search for in the heuristic policy to increase our confidence in the model. We showed using a numerical analysis that our maintenance optimization method could capture the value of information of different condition monitoring systems. Moreover, we validated the iterative feature of the proposed procedure by showing that iteration after iteration, the computed policy converges towards a low-cost solution (which may, however, only be a local minimum). Eventually, we illustrated on many different monitoring settings that the proposed method performs better than the more naive alternatives we first mentioned, which constitutes an additional element for validating our heuristic framework.

This work could be continued and improved in at least a couple of future directions. First, one could compare the proposed method with the current strategies applied to such systems, which would constitute an interesting reference benchmark. Second, in future work, we would like to use existing POMDP solvers and run them on small enough instances, with only a few items so that it remains reasonably computable, to see the proximity of our heuristic policy with the optimal one.

**Acknowledgement**

This work is funded by the Chair on Risk and Resilience of Complex Systems (CentraleSupelec, EDF, Orange, SNCF).

**References**

Alaswad, S., and Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability engineering & system safety*, 157, 54–63.

Yildirim, M., Gebraeel, N. Z., and Sun, X. A. (2017). Integrated predictive analytics and optimization for opportunistic maintenance and operations in wind farms. *IEEE Transactions on power systems*, 32(6), 4319–4328.

Memarzadeh, M., and Pozzi, M. (2016). Value of information in sequential decision making: Component inspection, permanent monitoring and system-level scheduling. *Reliability Engineering & System Safety*, 154, 137–151.

Sondik, E. J. (1971). The optimal control of partially observable Markov processes. *Stanford University*.

Papakonstantinou, K. G., and Shinozuka, M. (2014a). Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part I: Theory. *Reliability Engineering & System Safety*, 130, 202–213.

Papakonstantinou, K. G., and Shinozuka, M. (2014b). Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation. *Reliability Engineering & System Safety*, 130, 214–224.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2), 99–134.

Cassandra, A. R., Littman, M. L., & Zhang, N. L. (2013). Incremental pruning: A simple, fast, exact method for partially observable Markov decision processes. arXiv preprint arXiv:1302.1525.

Kurniawati, H., Hsu, D., & Lee, W. S. (2008). SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces. In *Robotics: Science and systems*, 2008.

Spaan, M. T., & Vlassis, N. (2005). Perseus: Randomized point-based value iteration for POMDPs. *Journal of artificial intelligence research*, 24, 195–220.

Roux, M., Fang, Y., & Barros, A. (2022, August). Impact of imperfect monitoring on the optimal condition-based maintenance policy of a single-item system. In *32nd European Safety and Reliability Conference (ESREL 2022)* (pp. 658–664).