Managing offshore wind turbines through Markov decision processes and dynamic Bayesian networks

Morato, P. G.; Papakonstantinou, K. G.; Andriotis, C.P.; Rigo, Philippe

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the 13th International Conference on Structural Safety & Reliability (ICOSSAR)

ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

Managing offshore wind turbines through Markov decision processes and dynamic Bayesian networks

P. G. Morato[a], K.G. Papakonstantinou[b], C.P. Andriotis[c], and P. Rigo[a]

[a]*ANAST, Dept.of ArGEnCo, University of Liege, Belgium, E-mail: pgmorato, ph.rigo@uliege.be.*
[b]*Dept.of Civil & Environmental Engineering, The Pennsylvania State University, USA, E-mail: kpapakon@psu.edu.*
[c]*Faculty of Architecture & the Built Environment, TU Delft, The Netherlands, E-mail: c.andriotis@tudelft.nl.*

ABSTRACT: Efficient planning of inspection and maintenance (I&M) actions in civil and maritime environments is of paramount importance to balance management costs against failure risk caused by deteriorating mechanisms. Determining I&M policies for such cases constitutes a complex sequential decision-making optimization problem under uncertainty. Addressing this complexity, Partially Observable Markov Decision Processes (POMDPs) provide a principled mathematical methodology for stochastic optimal control, in which the optimal actions are prescribed as a function of the entire, dynamically updated, state probability distribution. As shown in this paper, by integrating Dynamic Bayesian Networks (DBNs) with POMDPs, advanced algorithmic schemes of probabilistic inference and decision optimization under uncertainty can be uniquely combined into an efficient planning platform. To demonstrate the capabilities of the proposed approach, POMDP and heuristic-based I&M policies are compared, with emphasis on an offshore wind substructure subject to fatigue deterioration. Results verify that POMDP solutions offer substantially reduced costs compared to their counterparts, even in traditional problem settings.

1 INTRODUCTION

Civil and maritime infrastructures are exposed to deterioration mechanisms, such as fatigue or corrosion, and are thus at risk of structural failure. Deterioration models are, nonetheless, intrinsically uncertain, characterized by uncertainties customarily reaching coefficients of variation in the order of 25-30% (DNV 2015). In-service inspection and maintenance planning, i.e., collecting information through inspections and undertaking maintenance actions when needed, becomes therefore of paramount importance to optimally manage such systems throughout their lifetime. To this end, inspection and maintenance (I&M) planning aims to identify a strategy able to optimally balance structural failure risk against inspection and maintenance efforts and cost. Finding an optimal I&M policy demands, however, in most practical cases, the solution of a complex sequential decision-making problem under uncertainty.

Originally targeted for management of oil and gas platforms, risk-based inspection planning approaches simplify the I&M decision problem by evaluating only a predefined subset of heuristic rules out of all possible policies, thus alleviating the computational complexity (Faber 2002). Modern risk-based I&M planning methods evaluate the set of prescribed heuristic rules in a simulation environment, conducting Bayesian inference via dynamic Bayesian networks (Luque & Straub 2019). Heuristic-based policies are however compromised by the limited number of explored and evaluated policies out of an immense policy space.

In contrast, partially observable Markov decision processes (POMDPs) constitute a principled mathematical framework for sequential decision-making under uncertainty, in which the policy is defined as a function of a sufficient statistic, i.e., the dynamically updated history of actions and observations. Recent works on POMDPs for infrastructure management can be found in (Papakonstantinou & Shinozuka 2014a, b, Memarzadeh et al. 2015). With the advent of point-based solvers in this class of applications (Papakonstantinou et al. 2018), POMDP-based policies can be efficiently traced for medium-to-large problems of deteriorating structures and structural components (Papakonstantinou & Shinozuka 2014b, Morato et al. 2022).

In this paper, we adopt the methodology proposed in our earlier work (Morato et al. 2022), integrating dynamic Bayesian networks (DBNs) into the underlying structure of partially observable Markov decision processes, and we apply it for optimally managing an offshore wind structural detail subject to fatigue de-

terioration. Formulation schemes are described for encoding non-stationary stochastic deterioration processes in models parametrized by the relevant random variables or in terms of the deterioration rate. Both parametric and deterioration rate POMDP models are built based on the fatigue deterioration mechanism experienced by the offshore structural detail and defined according to offshore wind industrial standards (DNV 2015, DNV 2016). POMDP and heuristic-based policies are then computed and thoroughly compared for typical I&M and lifetime extension planning settings, and results verify that POMDP solutions offer substantially reduced costs in all the explored settings.

2 JOINT DBN-POMDP FRAMEWORK

2.1 *Stochastic structural deterioration modeling and mitigation via DBNs and heuristic decision rules*

The evolution of the stochastic deterioration process experienced by a structural component can be quantified in terms of a group of influencing random variables. DBNs encode the relationship amongst the involved random variables through conditional structures, enabling efficient inference, e.g., updating the deterioration process based on inspection outcomes. In most cases, the involved random variables are continuous and must be properly discretized in order to guarantee exact inference (Straub 2009).

A parametric DBN structure encodes the probability P of deterioration d conditional on a set of random variables θ. In this case, probability of damage d_{t+1}, at time $t+1$, evolves conditional on the damage at the previous time, d_t, set of random variables θ and observations $o_0,...,o_t$:

$$P(d_{t+1},\theta_{t+1}|o_0,...,o_t) = \sum_{d_t}\sum_{\theta_t} P(d_{t+1},\theta_{t+1}|d_t,\theta_t)P(d_t,\theta_t|o_0,...,o_t) \quad (1)$$

After collecting an observation o_{t+1} with likelihood $P(o_{t+1}|d_{t+1})$, the deterioration process, conditional on all observations up to time $t+1$, can be updated through Bayesian inference:

$$P(d_{t+1},\theta_{t+1}|o_0,...,o_{t+1}) \propto P(o_{t+1}|d_{t+1})P(d_{t+1},\theta_{t+1}|o_0,...,o_t) \quad (2)$$

The deterioration process can be alternatively encoded in a deterioration rate DBN, tracing the damage evolution d as a function of the deterioration rate τ. In this case, damage d_t at deterioration rate τ_t, conditional on observations $o_0,...,o_t$, is quantified in one time step as:

$$P(d_{t+1},\tau_{t+1}|o_0,...,o_t) = \sum_{d_t}\sum_{\tau_t} P(d_{t+1},\tau_{t+1}|d_t,\tau_t)P(d_t,\tau_t|o_0,...,o_t) \quad (3)$$

Bayesian inference considering an observation o_{t+1}, with likelihood $P(o_{t+1}|d_{t+1})$, can then be performed as:

$$P(d_{t+1},\tau_{t+1}|o_0,...,o_{t+1}) \propto P(o_{t+1}|d_{t+1})P(d_{t+1},\tau_{t+1}|o_0,...,o_t) \quad (4)$$

In a structural reliability context, the probability of a failure event $P_{F,t}$, at time t, corresponds to the probability of being in a damage state $P(d_{F,t})$. Additionally, an annual risk performance measure can be computed as the failure probability between two successive years, i.e., $\Delta P_{F,t} = P_{F,t+1} - P_{F,t}$.

The risk of structural failure can be controlled through an I&M policy regulated by a set of predefined heuristic decision rules, e.g., equidistant inspections or planned maintenance interventions after an indication event. DBNs models, either parametric, deterioration rate based, or others, can be employed, in a simulation environment, to identify the most optimal heuristic from the complete set of evaluated decision rules. The total discounted reward $R_{T_i}^{(h)}$, resulting from a set of heuristic decision rules h, can be evaluated for each simulation as the sum of inspection C_i, repair C_r, decommissioning C_d, and failure C_f costs, discounted by the factor γ:

$$R_{T_i}^{(h)} = \sum_{t=t_0}^{t_N} \gamma^t \left[C_i(t) + C_r(t) + C_d(t) + \Delta P_F(t) C_f\right] \quad (5)$$

The total expected utility $\mathbf{E}[R_T(h)]$ can then be computed through a Monte Carlo simulation of n_{ep} episodes (policy realizations):

$$\mathbf{E}[R_T(h)] = \frac{\sum_{i=1}^{n_{ep}} \left[R_{T_i}(h)\right]}{n_{ep}} \quad (6)$$

ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

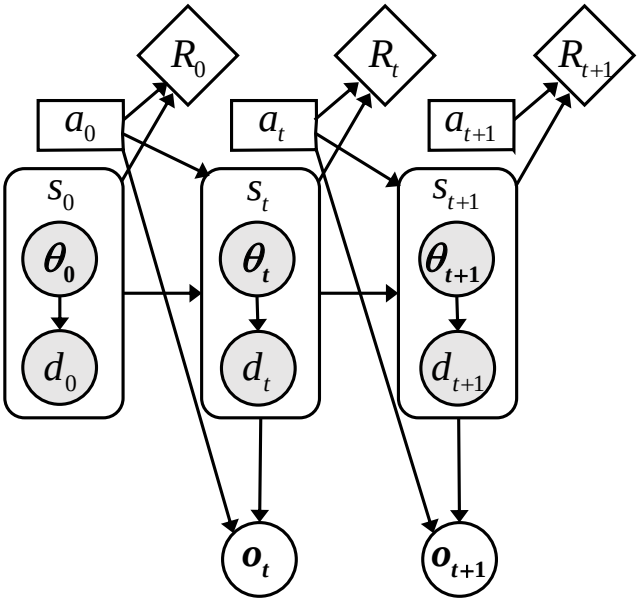

Figure 1. Dynamic decision network of a POMDP built based on a parametric DBN model (Morato et al. 2022).

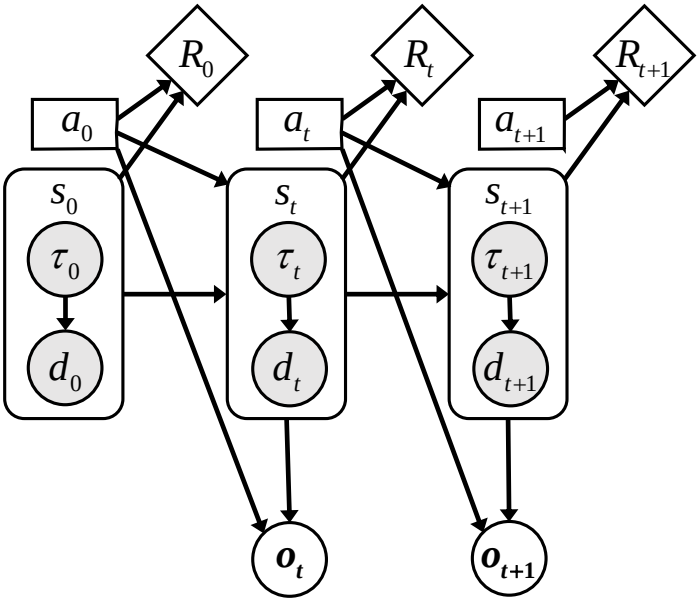

Figure 2. Dynamic decision network of a POMDP built based on a deterioration rate DBN model (Morato et al. 2022).

2.2 *Optimal I&M planning through partially observable MDPs*

Dynamic Bayesian networks (DBNs) can be integrated into the underlying structure of partially observable Markov decision processes for optimal inspection and maintenance (I&M) planning, as proposed in (Morato et al. 2022). A POMDP is a 7-tuple $\langle S, A, O, T, Z, R, \gamma \rangle$ controlled stochastic process in which the decision maker (intelligent agent) interacts in a stochastic environment. For a more complete overview of POMDP theoretical foundations and detailed formulations, the reader is directed to (Papakonstantinou & Shinozuka 2014a, b).

The state space S of a POMDP based on a parametric or deterioration rate DBN model is defined as the joint space of $d \times \tau$ or $d \times \theta$, respectively. Figures 1 and 2 represent the dynamic decision network corresponding to POMDPs based on parametric and deterioration rate DBN models. The condition state due to the deterioration process is probabilistically tracked by its belief state $\mathbf{b}(s) \equiv P(s)$ or probability distribution over states, and the POMDP dynamics consist, therefore, of an agent taking an action a_t, at time step t, transferring the state $s_t \in S$ to state $s_{t+1} \in S$, according to the transition model $T \equiv P(s_{t+1}|s_t)$. If a maintenance action is not planned, the deterioration process evolves naturally; in this case, the action do-nothing $a_{DN} \in A$ is linked with a transition model T_{DN} defined as $P(d_{t+1}, \theta_{t+1}|d_t, \theta_t)$ or $P(d_{t+1}, \tau_{t+1}|d_t, \tau_t)$, and equivalent to the transition model formulated in Equations 1 and 3. A perfect repair maintenance action $a_{PR} \in A$ transfers, instead, the belief $\mathbf{b_t}$ at time step t to its initial belief state $\mathbf{b_0}$:

$$\mathbf{P}(s'|s, a_{PR}) = \begin{pmatrix} b_0(s_0) & b_0(s_1) & \cdots & b_0(s_{|S|}) \\ b_0(s_0) & b_0(s_1) & \cdots & b_0(s_{|S|}) \\ \vdots & \vdots & \ddots & \vdots \\ b_0(s_0) & b_0(s_1) & \cdots & b_0(s_{|S|}) \end{pmatrix} \quad (7)$$

The quality of an inspection technique can be quantified through an observation model Z, defined as the probability of collecting an observation $o \in O$ at state $s \in S$. If inspections provide binary indication outcomes, i.e., either observing detection o_D or no-detection o_{ND}, the observation model Z_I can be often deduced as $P(o|s) = PoD(s)$, from Probability of Detection curves, PoD, corresponding to the inspection type. If no inspection is conducted, the observation model Z_{NI} assumes that observation $o_0 \in O$ is collected independently of the state $P(o_0|s_{t+1}) = 1$, thus leaving the belief state unaffected.

The total discounted reward, or sum of discounted rewards, is denoted in POMDP terminology as value function V_T. In a partially observable environment, the estimated reward collected after taking an action a at belief state $\mathbf{b}$ is the average of rewards associated

to action a and states $s \in S$:

$$R(\mathbf{b},a) = \sum_{s \in S} b(s)R(s,a) \tag{8}$$

In an I&M framework, both maintenance and inspection actions should be determined and can be combined into maintenance-inspection decision groups. For instance, two maintenance actions: do-nothing (DN) and repair (PR), combined with two inspection decisions: no-inspection (NI) and visual inspection (VI), result in four action groups: DN-NI, DN-VI, PR-NI and PR-VI. Costs are then assigned to each of these combinations. Considering the do-nothing & no-inspection action (DN-NI), the reward $R_F(s,a_{DN-NI})$ corresponds simply to the risk of structural failure, assigning a failure cost C_f to the failure states $S_F \subseteq S$ (Section 2.1). One can also specify the failure risk only as a function of the initial state $s \in S$, leading to a faster computation with a point-based solver, and by defining $\bar{R}(s,a_{DN-NI})$ equal to the failure cost C_f if $s \in S_F$, and, otherwise, equal to 0:

$$R_F(s,a_{DN-NI}) = \sum_{s' \in S_F} \left\{P(s'|s,a_{DN-NI})C_f\right\} - \bar{R}(s,a_{DN-NI}) \tag{9}$$

If an action also features inspections a_{DN-I}, then, an inspection cost C_i is added to each state, along with failure risk:

$$R_O(s,a_{DN-I}) = R_F(s,a_{DN-NI}) + C_i \tag{10}$$

Similarly, a repair cost C_r is included, for all states $s \in S$, if a repair and no-inspection action a_{PR-NI} is undertaken:

$$R_R(s,a_{PR-NI}) = C_r \tag{11}$$

By extension, the cost associated with a decommissioning and no-inspection action a_{DEC-NI} is defined by adding a cost C_d to all states $s \in S$:

$$R_{DEC}(s,a_{DEC-NI}) = C_d \tag{12}$$

In the previous rewards definitions, costs are considered as negative rewards. One can also define positive rewards if, for instance, the infrastructure remains operative and thus yielding a positive income.

For most practical applications, the POMDPs state space is high-dimensional and the problem can be computationally intractable if solved by exact value iteration or grid-based approaches. State-of-the-art point-based POMDP solvers (Smith & Simmons 2006, Kurniawati et al. 2008) are, however, capable of scaling solutions to spaces of realistic dimensions, as demonstrated in (Papakonstantinou et al. 2018). Point-based solvers restrict the computation of Bellman backups to only a subset of reachable belief points, thus significantly improving computational efficiency. The value function $V_T(\mathbf{b})$ is parameterized by a set Γ of hyperplanes (α-vectors), each of them associated with an action a; and the optimal policy π^* corresponds to the α-vector that maximizes the value function $V_T(\mathbf{b})$:

$$V_T(\mathbf{b}) = \max_{\alpha \in \Gamma} \sum_{s \in S} \alpha(s)b(s) \tag{13}$$

State-of-the-art point-based solvers are developed for solving infinite horizon POMDPs, yet, in many cases, the decision maker deals with finite horizon policies, e.g., 20 years lifetime. The state space can then be augmented, following the approach in (Papakonstantinou & Shinozuka 2014a) to transform finite horizon POMDPs into infinite horizon ones.

3 DETERIORATION ENVIRONMENT

A monopile foundation, dominant in most installed offshore wind turbines, is an assembly of rolled plates welded transversely and forming a hollow steel pipe. A transverse butt weld is therefore deemed to be a representative structural detail in this case. The fatigue deterioration of the joint is modeled, following DNV-GL design standards (DNV 2016), by a cumulative fatigue damage law. A limit state $g_{SN}(t)$ is then formulated based on a cumulative damage Miner's rule, over time t:

$$g_{SN}(t) = \Delta - \nu t \left[\frac{q^{m_1}}{C_{1,SN}} \gamma_1 \left\{ 1 + \frac{m_1}{h}; \left(\frac{S_1}{q}\right)^h \right\} + \frac{q^{m_2}}{C_{2,SN}} \gamma_2 \left\{ 1 + \frac{m_2}{h}; \left(\frac{S_1}{q}\right)^h \right\} \right] \tag{14}$$

where $C_{1,SN}$, $C_{2,SN}$, m_1, and m_2 are material parameters corresponding to a 'D' category bi-linear SN

ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

Table 1. Random variables and deterministic parameters for modeling the fatigue deterioration.

Parameter	Distribution	Mean	Std
Miner's cumulative damage model			
$C_{1,SN}$*	Normal	12.564	0.2
$C_{2,SN}$*	Normal	16.006	0.2
q** (MPa)	Trunc. Normal	10.209	2.55
q***	Trunc. normal	8.834	2.21
Δ	Lognormal	1	0.3
h	Deterministic	0.8	-
ν (cycles/s)	Deterministic	0.16	-
m_1	Deterministic	3	-
m_2	Deterministic	5	-
Fracture mechanics model			
lnC_{FM}**	Normal	-26.432	0.126
lnC_{FM}***	Normal	-26.501	0.131
d_0 (mm)	Exponential	0.11	0.11
Y	Lognormal	1	0.1
d_c (mm)	Deterministic	20	-
m (mm)	Deterministic	3	-

*Fully correlated.
**Inspection and maintenance planning application.
***Lifetime extension planning application.

curve with S_1 stress range value at the knee; the expected stress range is parameterized by Weibull factors q and h; ν represents the cycle rate; and Δ corresponds to the fatigue limit. Note that γ_1 and γ_2 stand for lower and upper incomplete gamma functions, respectively. Assuming the structure is designed to the limit, the loading scale factor q is back calculated considering a fatigue design factor of one for the I&M planning setting, and a fatigue design factor of two for the lifetime extension planning scenario. Table 1 lists all relevant parameters.

Since inspections cannot reveal the accumulated fatigue damage computed through Miner's rule, fracture mechanics models are normally utilized instead for in-service inspection and maintenance planning, as in this case the crack size belief state can be updated based on collected crack observations. In this sense, a probabilistic fracture mechanics model is calibrated with the objective of achieving the same structural reliability level computed previously by the cumulative fatigue damage law (Eq. 14). The crack growth is described here with a Paris' law model,

Table 2. Description of the discretization scheme applied to DBN-POMDP deterioration rate and parametric models.

Variable	Interval boundaries
Deterioration rate model	
S_d	$[0, d_0 : (d_c - d_0)/(\lvert S_d\rvert - 2) : d_c, \infty]$
S_τ	$[0 : 1 : 20]$
S_τ*	$[0 : 1 : 60]$
Parametric model	
S_d	$0, \exp\left\{\ln(10^{-2}) : \frac{\ln(d_c) - \ln(10^{-2})}{\lvert S_d\rvert - 2} : \ln(d_c)\right\}, \infty$
S_K	$0, \exp\left\{\ln(10^{-4}) : \frac{\ln(2) - \ln(10^{-4})}{\lvert S_K\rvert - 2} : \ln(2)\right\}, \infty$

*Lifetime extension planning setting.

originally introduced in (Ditlevsen & Madsen 1996):

$$d_{t+1} = \left[d_t^{\frac{2-m}{2}} + \frac{2-m}{2} C_{FM} (Y \pi^{0.5} S_e)^m n \right]^{\frac{2}{2-m}} \quad (15)$$

where the crack depth is modeled by d, with crack growth parameters C_{FM}, and m, n cycles per time step, geometric factor Y, and the same loading as for the damage cumulative law, described by the expected stress range $S_e = q\Gamma(1 + 1/h)$ through the parameters q and h. Table 1 lists all relevant parameters and the fatigue limit state is formulated as $g_{FM}(t) = d_c - d(t)$. Assuming a through-thickness failure, the critical crack size d_c corresponds to the plate thickness.

We then translate the proposed probabilistic fracture mechanics model into both a deterioration rate-based and a parametric dynamic Bayesian network, alleviating computational complexity in the latter by combining the time-invariant parameters C_{FM}, Y and q into the chance node $K \equiv C(Y\pi^{0.5}S_e)^m n$. The state transition models $p(d_{t+1}, \tau_{t+1} | d_t, \tau_t)$ and $p(d_{t+1}, K_{t+1} | d_t, K_t)$ are constructed through sequential Monte Carlo simulations, relying on Equation 15, and are discretized according to the scheme shown in Table 2. The observation quality $p(o|d)$ is modeled depending on the inspection type and will be explained on the respective case studies. An accurate enough discretization is achieved by including 60 crack states $|d|$, and resulting in a root mean square error of $2.4 \cdot 10^{-3}$ when comparing the reliability index with a Monte Carlo simulation featuring two eddy current inspections at years 8 and 16.

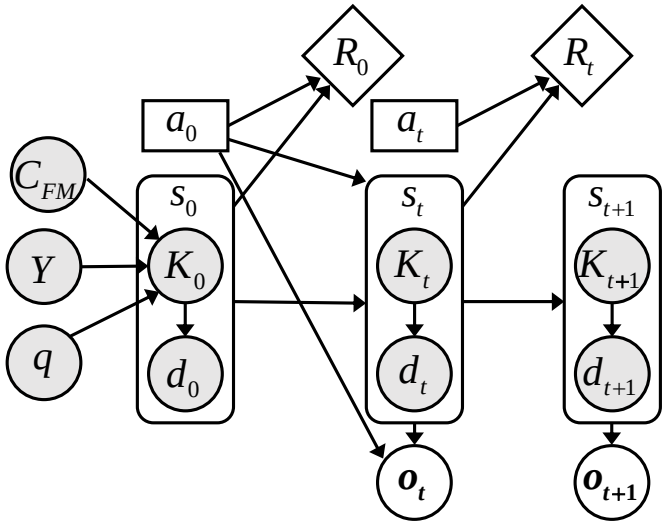

Figure 3. Parametric DBN-POMDP dynamic decision network designed for the numerical experiments.

The developed DBN structures serve as the backbone for the evaluation of all the heuristic decision rules explored in the numerical investigations, and can be directly integrated into the underlying structure of the POMDP models, as described in Section 2.1. The POMDP state space for the deterioration rate model contains the joint distribution of d and τ, as illustrated in Figure 2, summing up to a total of 1,260 states. For the finite horizon I&M planning setting examined in Section 4.1, the state space is augmented to 13,860 states, due to the fact that time needs to be also included in the state vector.

The joint distribution of d and K forms the parametric POMDP model state space, as illustrated in Figure 3, summing up in this case up to 4,000 states. When translated into a finite horizon model applicable to the I&M planning setting, the parametric model results to 156,000 states.

4 NUMERICAL EXPERIMENTS

4.1 *Inspection and maintenance planning*

In this first example, we explore a typical risk-based inspection planning setting with an assumed 20-year finite horizon, in which the inspection quality is modeled by probability of detection curves and with the possibility of planning perfect repair maintenance actions.

In this scenario, the transition model assigned to the action do-nothing a_{DN} is modeled according to the fatigue deterioration rate introduced in Section 3, and the perfect repair transition model a_{PR} transfers the current belief state $\mathbf{b_t}$ to its initial belief $\mathbf{b_0}$, as stated in Equation 7.

Table 3. Inspection quality.

Inspection technique	X_0	b
Eddy current (EC)	1.16	0.9
Ultrasonic testing (UT)	0.41	0.642
Visual inspection (VI)	83.03	1.079

Table 4. Definition of costs/rewards.

Failure	-1000 (money units)
Eddy current inspection	-1 (money units)
Ultrasonic inspection*	-1.5 (money units)
Visual inspection*	-0.5 (money units)
Perfect repair*	-100 (money units)
Production**	+5 (money units)
Replacement**	-100 (money units)
Decommissioning**	-20 (money units)
Discount factor	0.95 (-)

*Inspection and maintenance planning setting.
**Lifetime extension planning setting.

While only one inspection type is available in most traditional inspection and maintenance planning applications, three inspection techniques are possible here, namely, eddy current, ultrasonic testing and visual inspection. Table 3 lists the parameters corresponding to each inspection technique, following the probability of detection (PoD) formulation proposed by (DNV 2015):

$$PoD(a) = 1 - \frac{1}{1 + (a/X_0)^b} \tag{16}$$

Based on the proposed transition and observation models, we construct a finite horizon deterioration rate POMDP by combining the following action and observation decisions: do-nothing & no-inspection (DN-NI), do-nothing & eddy current inspection (DN-EC), do-nothing & ultrasonic inspection (DN-UT), do-nothing & visual inspection (DN-VI), and perfect-repair & no-inspection (PR). The costs for this case are listed in Table 4. Note that perfect-repair is not paired with any inspection type since an observation is generally expected to be suboptimal after the component is fully repaired.

Table 5. Comparison between POMDP and heuristic-based policies in I&M and lifetime extension settings.

Policy	$\mathbf{E}[R]$ (95% C.I.)	%SARSOP
I&M planning: 20 years finite horizon		
POMDP-SARSOP	-29.53	
POMDP-FRTDP	-29.53	0%
Heur. EQ-INS (EC)	-39.62 (0.47)	-34.2%
Heur. THR-INS (EC)	-38.97 (0.35)	-32.0%
Heur. THR-INS (UT)	-48.63 (0.45)	-64.7%
Heur. THR-INS (VI)	-71.94 (0.18)	-143.6%
Lifetime extension planning: Infinite horizon		
POMDP-SARSOP*	41.20	
POMDP-FRTDP*	41.04	<1%
POMDP-SARSOP**	41.11	<1%
POMDP-FRTDP**	40.51	<2%
Heur. EQ-INS/DEC	16.16 (0.37)	-60.8%
Heur. EQ-INS/REP	-2.15 (0.17)	-105.2%
Heur. THR-INS/DEC	15.96 (0.34)	-61.3%
Heur. THR-INS/REP	1.16 (0.60)	-97.2%

*Deterioration rate POMDP (Fig. 2).
**Parametric POMDP (Fig. 3).

The finite horizon POMDP model is then computed via SARSOP (Kurniawati et al. 2008) and FRTDP (Smith & Simmons 2006) point-based solvers. Furthermore, policies regulated by predefined heuristics are also evaluated. Heuristics include planning of (i) equidistant inspections (EQ-INS) or inspections upon exceedance of an annual failure probability threshold (THR-INS); and (ii) repairs automatically scheduled upon crack detection. Results from both POMDP and heuristic-based policies are reported in Table 5.

4.2 *Lifetime extension planning*

In this second example, we consider a lifetime extension planning setting, in which the decision maker opts between doing-nothing, replacing, or decommissioning the structure. Suppose an offshore wind turbine on operation for 16 years without planned inspections or repairs up to that point. The initial belief state for this problem thus corresponds to the state of the structure, $\mathbf{b_{16}} = \mathrm{T}^{16}\,\mathbf{b_0}$, at year 16.

Both deterioration rate and parametric infinite horizon POMDP models are implemented and solved through point-based solvers, by combining the following actions and observations decisions: do-nothing & no-inspection (DN-NI), do-nothing & eddy current inspection (DN-I), replacement & no-inspection (REP), and decommissioning & no-inspection (DEC). The do-nothing action, both including and excluding inspections, is modeled exactly as in the I&M planning setting, and a replacement is assumed as a perfect repair. The decommissioning action transfers the current belief state $\mathbf{b_t}$ to an absorbing state s_{dec}, in which no further rewards can be collected, i.e., the structure is no longer in operation (N-OP). In this infinite horizon setting, the transition model for the deterioration rate POMDP is still specified considering 60 deterioration rates (Table 2), whereas the crack transitions according to the last deterioration rate for those deterioration rates beyond 60. Table 4 lists all the costs/rewards considered for this experiment. Note that a positive reward is now collected every time the structure is operative.

Table 5 reports the results for both POMDP and heuristic-based policies. Heuristic rules consist, in this setting, in planning equidistant inspections (EQ-INS) or inspecting after reaching an annual failure probability threshold (THR-INS); and either a replacement (REP) or a decommissioning (DEC) action is ordered after a crack detection occurs.

5 RESULTS AND DISCUSSION

POMDP-based policies outperform heuristic-based policies in all the explored settings, resulting in a total expected reward benefit ranging from 32% to 144%. Table 5 reports the results corresponding to both I&M planning and lifetime extension planning investigations. For each policy, either POMDP or heuristic-based, Table 5 lists the expected total rewards $\mathbf{E}[R]$ along with the 95% confidence intervals (95% C.I.), and the relative difference in expected total rewards between each policy and SARSOP (%SARSOP).

In terms of POMDP-based policies, the difference between SARSOP and FRTDP point-based solvers is less than 1% in all the experiments. While SARSOP solver quickly reduces the lower bound within seconds of computational time, the FRTDP solver is able to reduce the upper bound faster, leading to convergence within the allocated computational time for the finite horizon I&M planning setting. Figure 4 shows the evolution of the expected total reward for each solver over its computational time.

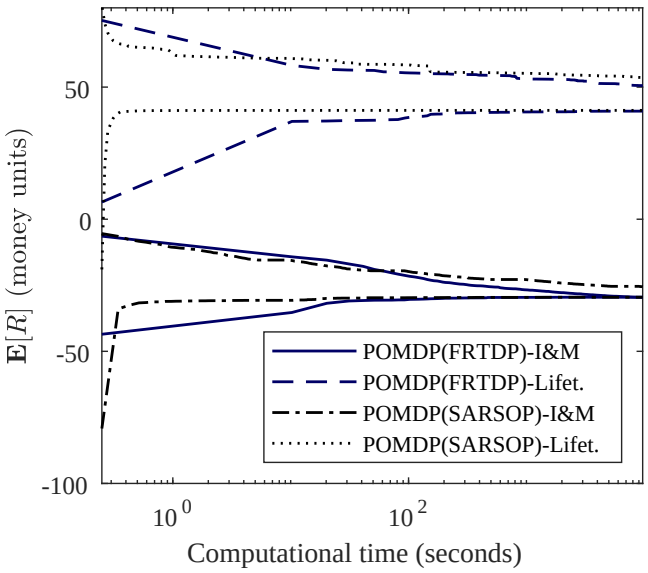

Figure 4. Evolution of expected total rewards over computational time for each POMDP point-based solver.

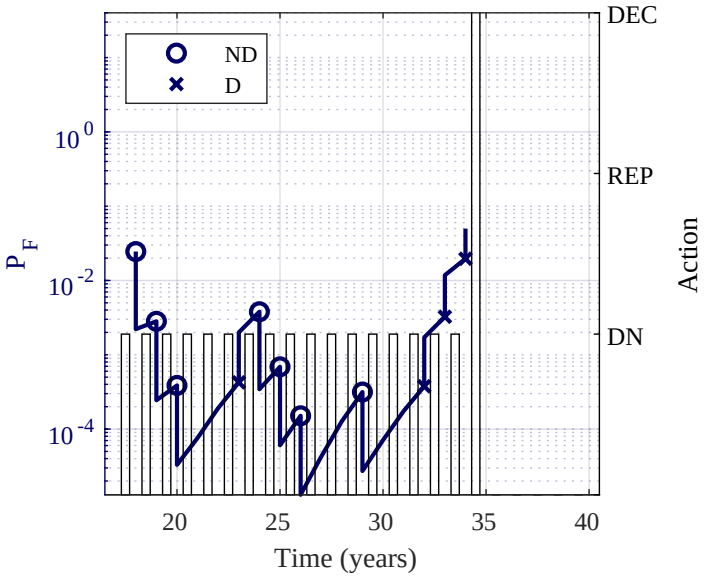

Figure 5. Realization of a POMDP policy in the lifetime extension planning application.

Policies based on inspections planned via a predefined annual failure probability threshold (THR-INS) and policies with planned inspections at equidistant intervals (EQ-INS) tend to reach similar total rewards. Moreover, decision rules featuring eddy current (EC) inspections result in better policies, under the proposed cost model, than those employing ultrasonic testing (UT) and visual inspections (VI). Heuristic decision rule evaluations also indicate that undertaking a decommissioning action (INS/DEC) after observing a crack results in higher profits than replacing the structure (INS/REP).

One can deduce that the optimality of heuristic-based policies will thus be importantly influenced by the ability of exploring the appropriate space of decision rules. Selecting optimal heuristics is case dependent and can be achieved by experience or by probing a large set of decision rules. POMDP-based policies, on the other hand, offer a mapping from the current belief state (dynamically encoding the entire prior history of actions and observations) to the optimal action, and the sequence of optimal actions might be non-trivial in certain scenarios. Consider, for instance, a POMDP policy realization in the lifetime extension application, dictating a decommissioning action after three successive crack detection indications, as illustrated in Figure 5.

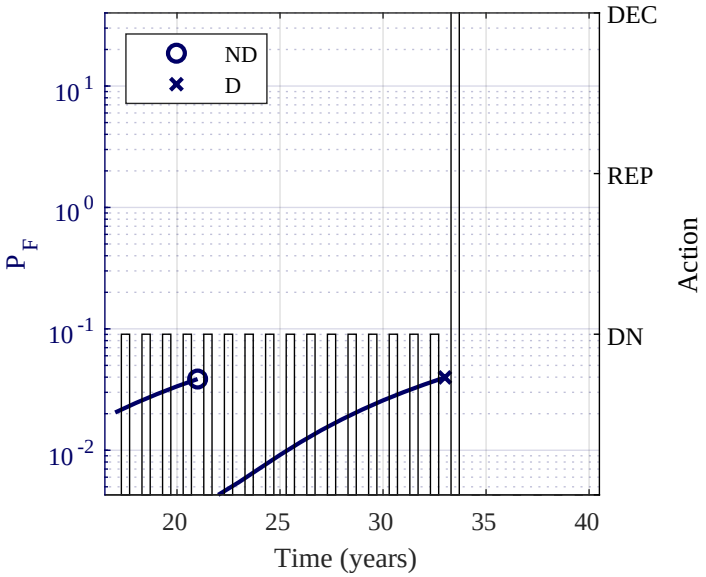

Figure 6. Realization of a heuristic policy in the lifetime extension planning application.

In contrast, the indication-based heuristic policy, for the same lifetime extension setting, assigns a decommissioning action after receiving a crack detection indication, as shown in Figure 6. Thus, the learned POMDP policy autonomously acknowledges that, in this particular application, the observation outcome might not be very accurate and more information should be collected before ordering a decommissioning action. If a detection is followed by a no-detection indication, a do-nothing action is then preferred.

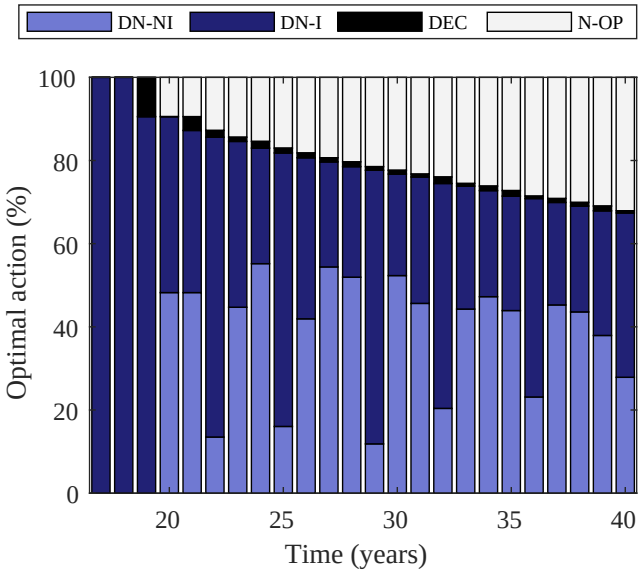

Figure 7. Histogram of POMDP optimal actions assigned based on 10^5 policy evaluations in the lifetime extension planning setting.

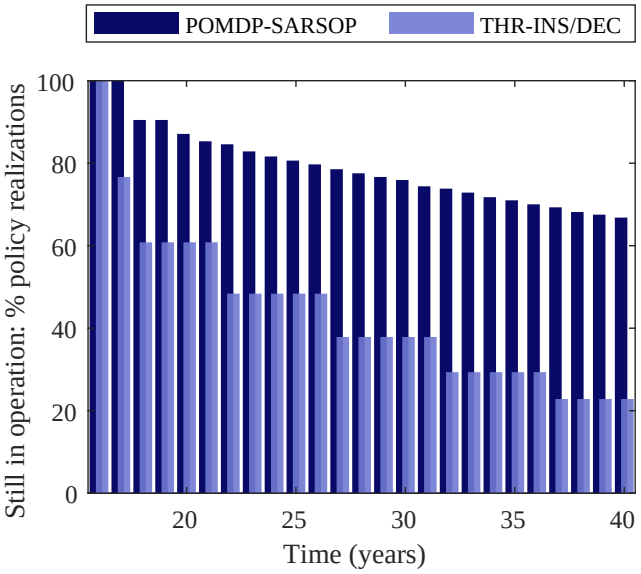

Figure 8. Comparison of POMDP and heuristic-based policies in terms of the policy realizations which indicate an active operation status of the structure over time.

As explained in Section 2.2, point-based POMDP policies are parameterized by a set of hyper-planes (α-vectors), each one of them associated to a corresponding optimal action. A frequency histogram derived from policy evaluations offers to the decision maker a summary of the actions taken over the investigated horizon. Figure 7 represents, for the lifetime extension application, the histogram of actions defined by the POMDP policy, collected through 10^5 policy realizations. Besides observing how POMDP adaptive policies provide optimal sequences of actions that can be hardly parametrized by a set of ad-hoc decision rules, Figure 7 also shows that the structure still remains in operation at year 40 for more than 67% of the evaluated policy realizations (N-OP stands for non-operational). It would be thus interesting to compare the behavior of POMDP and heuristic-based policies, not only in terms of expected reward (Table 5), but also examining the resulting expected life extension.

A further comparison is, therefore, displayed in Figure 8, indicating the percentage of realizations, for both POMDP and heuristic-based policies, for which the structure still remains in operation, i.e., a decommissioning action has not been assigned. The examined heuristic policy corresponds to the case 'THR-INS/DEC', performing inspections based on a predefined annual failure probability threshold and assigning a decommissioning action after a detection outcome.

Interestingly, the resulting heuristic policy resembles the behavior of a periodic strategy, assigning decommissioning actions every four to five years, providing also some further insights on the small observed differences between 'THR-INS/DEC' and 'EQ-INS/DEC' in Table 5. As seen in Figure 8, in terms of lifetime extension, more than 67% of the POMDP policy realizations characterize a structure that is in operation at year 40, whereas less than 25% of the cases based on the heuristic-based policy end up with an structure still in operation at year 40.

6 CONCLUSIONS

This paper examines the efficiency of integrating Dynamic Bayesian Networks (DBNs) and Partially Observable Markov Decision Processes (POMDPs) in a joint algorithmic context for optimal Inspection and Maintenance (I&M) planning. Time-invariant parameters and finite horizon settings can be implemented within this framework by simply augmenting the POMDP state space, generating high-dimensional sparse matrices, which can be efficiently solved by state-of-the-art point-based POMDP solvers.

The application of the methodology to the case of an offshore wind structural detail subject to a non-stationary fatigue deterioration process and modeled according to common offshore wind industrial stan-

dards verifies the computational efficiency and practical relevance of the proposed approach. The results show that POMDP-based policies outperform traditional heuristic-based policies in all tested settings. While in the analyzed inspection and maintenance planning setting, POMDP adaptive policies provide optimal total cost and actions by combining three potential inspection techniques and one maintenance action, in the lifetime extension setting, POMDP policies not only yield benefits in terms of expected reward, but also maintain the structure in operation for a longer time span compared to the analyzed heuristic strategies.

Further efforts are also currently under way in utilizing the presented framework and approximating optimal POMDP policies through deep reinforcement learning approaches with or without stochastic constraints (Andriotis & Papakonstantinou 2019, Andriotis & Papakonstantinou 2021) for settings featuring very high-dimensional state, action and observation spaces.

ACKNOWLEDGEMENTS

Dr. Morato gratefully acknowledges the support received by the National Fund for Scientific Research in Belgium F.R.I.A. - F.N.R.S. Dr. Papakonstantinou would further like to acknowledge that this material is also based upon work supported by the U.S. National Science Foundation under Grant No. 1751941 and Dr. Andriotis also acknowledges the support of the TU Delft AI Labs program.

REFERENCES

Andriotis, C.P. & Papakonstantinou, K.G., 2019. Managing engineering systems with large state and action spaces through deep reinforcement learning. Reliability Engineering & System Safety, 191, p.106483.

Andriotis, C.P. & Papakonstantinou, K.G., 2021. Deep reinforcement learning driven inspection and maintenance planning under incomplete information and constraints. Reliability Engineering & System Safety, 212, p.107551.

Ditlevsen, O. & Madsen, H.O. 1996. Structural Reliability Methods. New York: John Wiley & Sons.

DNV, G., 2015. DNVGL-RP C210. Probabilistic methods for planning of inspection for fatigue cracks in offshore structures.

DNV, G., 2016. DNVGL-RP C203: Fatigue design of offshore steel structures.

Faber, M.H., 2002. Risk-based inspection: The framework. Structural engineering international, 12(3), pp.186-195.

Kurniawati, H., Hsu, D., & Lee, W. S., 2008. SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces. In Proceedings of Robotics: Science and Systems IV, Zurich, Switzerland, 2008. doi:10.15607/RSS.2008.IV.009.

Luque, J. & Straub, D., 2019. Risk-based optimal inspection strategies for structural systems using dynamic Bayesian networks. Structural Safety, 76, pp.68-80.

Memarzadeh, M., Pozzi, M. & Zico K.J., 2015. Optimal planning and learning in uncertain environments for the management of wind farms. Journal of Computing in Civil Engineering, 29(5), p.04014076.

Morato, P.G., Papakonstantinou, K.G., Andriotis, C.P., Nielsen, J.S., & Rigo, P., 2022. Optimal inspection and maintenance planning for deteriorating structural components through dynamic Bayesian networks and Markov decision processes. Structural Safety, 94, pp.102140.

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

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

Papakonstantinou, K.G., Andriotis, C.P. & Shinozuka, M., 2018. POMDP and MOMDP solutions for structural life-cycle cost minimization under partial and mixed observability. Structure and Infrastructure Engineering, 14(7), pp.869-882.

Smith, T., & Simmons, R., 2006. Focused real-time dynamic programming for MDPs: Squeezing more out of a heuristic. In Proceedings of the AAAI Conference on Artificial Intelligence, pp. 1227-1232.

Straub, D., 2004. Generic approaches to risk based inspection planning for steel structures. [Doctoral thesis, Swiss Federal Institute of Technology in Zürich (ETH), Zürich, Switzerland].

Straub, D., 2009. Stochastic modeling of deterioration processes through dynamic Bayesian networks. Journal of Engineering Mechanics, 135(10), pp.1089-1099.