Risk-based Decision Making for Deterioration Processes Using POMDP

Jannie S. Nielsen
Postdoc, Dept. of Civil Engineering, Aalborg University, Aalborg, Denmark

John D. Sørensen
Professor, Dept. of Civil Engineering, Aalborg University, Aalborg, Denmark

ABSTRACT: This paper proposes a method for risk-based decision making for maintenance of deteriorating components, based on the partially observable Markov decision process (POMDP). Unlike most methods, the decision polices do not need to be stationary and can vary according to seasons and near the end of the lifetime. The approach is demonstrated through two examples, and the total expected costs are similar to those of another efficient method.

1. INTRODUCTION

For deterioration processes, maintenance can theoretically be optimally planned using risk-based methods. Finding the optimal decisions involves solving a pre-posterior decision problem with a large number of decisions, and the decision problem can, in principle, be solved using either the normal or extensive analysis method (Raiffa and Schlaifer, 1961). Because the number of branches in a traditional decision tree increases exponentially with the number of time steps, it is generally not possible to calculate the expected costs for all branches.

Previously, decision trees for maintenance planning have been solved approximately by finding the optimal stationary decision rules (Straub, 2004) (Nielsen, 2013). For maintenance of offshore structures, inspections and repairs can only be performed during periods with relatively low wind speeds and wave heights, and the probability of inspections and repairs being possible within a time period will depend on the season. Therefore, the decision maker could benefit from having time variant decision policies that follow the seasons. Also near the end of the lifetime, decision policies will be different, as preventive repairs should not be made close to the end of lifetime.

In this paper, an approach for solving these decision problems is considered, where the decision policies do not need to be stationary. Here the continuous damage size is discretized, and the deterioration processes are modeled using dynamic Bayesian networks. Hereby, the approaches used for partially observable Markov decision processes (POMDP) can be used. The advantage of this approach is that the computation time is only linear with the number of time steps. The optimal decision policies are found sequentially from the last decision for, in principle, all possible belief states. In reality, the belief state is represented by a vector with a sum equal to one, and the number of different belief states is infinite.

In practice, the optimal decisions and expected costs can be found for a number of grid points, and the results from the approximated belief state closest to the real one is used. The method has previously been applied for a case where the damage state can take three values: no damage, damage, and failure (Nielsen and Sørensen, 2012), and the belief state could, therefore, simply be expressed by the probability of being in the damaged state, as it was assumed to be known whether or not failure had occurred. The grid points for that case could simply be chosen as evenly distributed probabilities of damage between zero and one, and the accuracy was determined by the number of grid points.

In this paper, the approach is extended to deterioration processes with more than three damage states. The approach is demonstrated for a maintenance problem for offshore wind turbines.

2. Bayesian pre-posterior decision problem

Figure 1 shows the decision tree considered in this paper. Information from condition monitoring and inspections is included, and decisions are made on inspections and repairs. Traditionally this type of decision problems is solved by using stationary decision rules. The decision on repair depends directly on the most recent inspection outcome, and inspections are scheduled equidistant or when the probability of failure given no detection at previous inspections exceeds a threshold. In both cases, the inspections are scheduled from the beginning, and extra information obtained during the lifetime is not considered.

If condition monitoring is available, they can be included using this approach, by setting a threshold for the monitoring outcome for when inspections are made. However, only the most recent monitoring outcome is considered in this case, and in case of uncertainties on the monitoring outcome, a better and more informed decision could be made by including the history of monitoring outcomes. To do so, Bayesian updating needs to be done during the lifetime to estimate the current probability distribution for the damage size. For this, discrete dynamic Bayesian networks (DBN) can be applied, as they enable computationally efficient Bayesian updating (Straub, 2009). Then a threshold for the probability of failure can be used as decision rule for when to make inspections, simulations can be applied to determine expected costs and the optimal value of the decision rules, and within the simulations a discrete Bayesian network can be applied for updating of probabilities. This method is computationally expensive, as time-consuming simulations are needed to find the optimal decision rule, and the decision rules are stationary. If no time-invariant uncertainties are present, an alternative approach is to use a method that exploits the Markovian assumption of independence between the future and past given the present.

2.1. *Markov decision model for deterioration processes*

A traditional Markov decision problem uses the fact that the optimal decision at a given time only depends on the current state of the component, not the history of damage development. If the component health is directly observed at every time step, the optimal decision for each time step can be found for all possible damage states sequentially from the last decision. Dynamic programming can be applied, such that the expected costs found for later time steps are used when computing the optimal decisions for earlier time steps (Dasgupta et al., 2006).

If the component health is not directly observed, but instead observed through an indicator, the problem is a partially observable Markov decision process (POMDP). Here, the optimal decision at each time step only depends on the current belief state for the component health. In other words, it depends only on the current probability distribution for the damage size, as it summarizes the prediction from the model and all past observations. In principle, the optimal decision can then be found for all possible probability distributions for the damage size for each time step. However, in reality, there are infinitely many possible probability distributions, so an approximation needs to be made. This can be done by finding the expected costs and optimal decisions for a number of grid points, and then interpolate between these grid points, when expected costs for other points are needed. To use the approach for deterioration processes with (discretized) continuous damage sizes, grid points need to be selected and a method to interpolate between grid points needs to be developed.

In order to make a grid, the probability distribution for the damage size is approximated by a 2-parameter Weibull distribution with scale parameter a and shape parameter b, with cumulative distribution function:

$$F_X(x) = 1 - \exp\left(-\left(\frac{x}{a}\right)^b\right) \qquad (1)$$

The Weibull distribution is discretized and truncated before the failed state, as it is assumed to be known if failure has occurred. The calculation grid

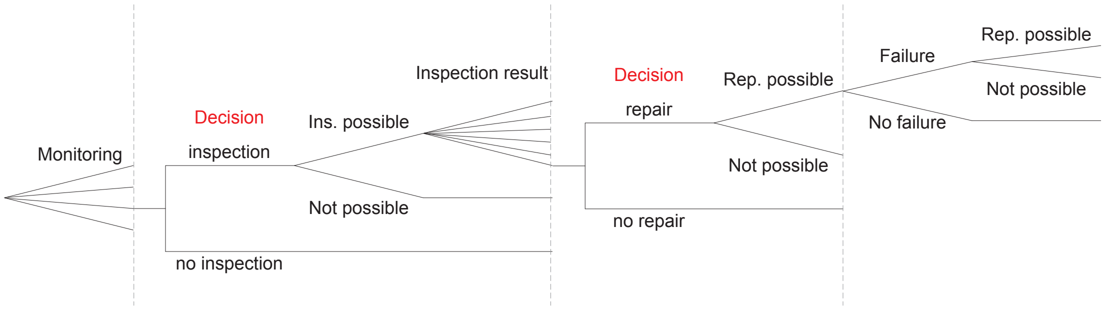

Figure 1: Decision tree for decisions on inspections and repairs and damage indications from condition monitoring and inspections. At the dashed lines all branches continues as the ones illustrated. Repeated for each time step in the model.

is made by doing this for a range of values of a and b.

There are two approximations regarding the use of this approach. First, there is the variation between the true probability distribution and the discretized Weibull distribution closest to the true distribution. Second, there is the approximation introduced by interpolation between grid points for the a and b parameters of the Weibull distribution.

In order to find the Weibull distribution closest to the true distribution, a selection criterion needs to be set up. Possible choices include least square estimates based on the distretized probability mass function and least square estimates based the discretized cumulative distribution function. The latter has been used here, as differences in probabilities for nearby damage states are less critical than for damage states far from each other.

A nonlinear optimizer can then be applied for estimating the optimal values of a and b for any distribution. Thereafter, multidimensional linear or cubic interpolation can be used. For this application, the probability distribution for all grid points are known, and an alternative interpolation method is to calculate the sum of the squares of the errors for all the distributions and select the distribution with the lowest value. With this approach, the nearest distribution is chosen and as such no interpolation is performed. For the same number of grid points, this method is less accurate, but it is much faster, as the time-consuming nonlinear optimization is not needed. Therefore, a denser grid can be used for this method with same computation time, and it has been used for the examples in this paper.

In each time step, interpolation has been performed at two points. One after the condition monitoring outcome is obtained, and one after corrective repair. The outcome of the calculations are the decision policies for inspections and repairs for each time step. For inspections, the decision policies are given as function of the a and b values corresponding to the nearest Weibull distribution to the probability distribution updated after condition monitoring. For preventive repairs, the decision policies are given as function of the inspection outcome as well as the a and b values corresponding to the nearest Weibull distribution. Updating the probability distribution for the damage size due to deterioration and observations is performed using a discrete Bayesian network approach.

3. Example 1

The method is illustrated using a damage model with 10 damage states of equal size and constant transition probability. The lower interval boundaries are $0, 1, 2, \ldots, 9$, and the last state is the failed state. This corresponds to linear damage growth. The lifetime is 20 years, and the mean time to failure is 20 years. It is assumed that the damage size cannot skip any state. Initially, the damage size is assumed to be in the first state with probability equal to one.

3.1. Model

The computation is run for a lifetime equal to 20 years, and the step length is one month. This gives 240 time steps in total. In the beginning of each time step, results from an online condition monitoring system is obtained. The outcome depends on the damage size in the same way as for PoD (probability of detection) curves commonly used in risk-based inspection planning. However, here more outcomes are possible. The damage size causing each monitoring outcome is assumed lognormal distributed with parameters given in table 1, and figure 2 shows the probability of obtaining each outcome as function of damage size. After the monitoring outcome is obtained, a decision can be made to make an inspection. Here a similar model is used, see table 2 and figure 3. After the inspection outcome is obtained, or if no inspection is made, a decision can be made to make a preventive repair. Then, the deterioration model is used to update the damage size, and if failure happens during the time step, a corrective repair is made.

State	Description	Mean	COV
1	no alarm	-	-
2	low alarm	2.0	1.0
3	high alarm	5.0	1.0
4	failure	9.0	0.0

Table 1: Mean and coefficient of variation (COV) for the damage sizes causing each monitoring outcome.

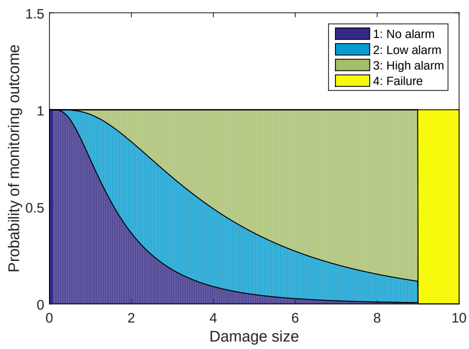

Figure 2: Probability of each monitoring outcome as function of damage size.

The costs are set relative to the costs of an inspection, such that the expected costs of an inspection is one, the expected costs of a preventive repair is 20, the expected costs of failure is 500, and the expected costs of lost production per time step is 100.

State	Description	Mean	COV
1	no detection	-	-
2	mild damage	2.0	1.0
3	some damage	4.0	0.8
4	significant damage	6.0	0.6
5	severe damage	8.0	0.4
6	failure	9.0	0.0

Table 2: Mean and coefficient of variation (COV) for the damage sizes causing each inspection outcome.

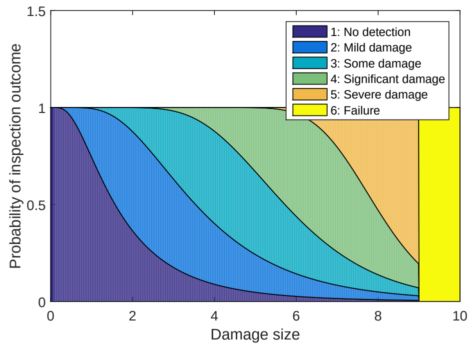

Figure 3: Probability of each inspection outcome as function of damage size.

Furthermore, the probability that an inspection or repair is not possible within a time interval is included, and it can vary according to seasons. For each season, each lasting three months, a probability of the actions not being possible is defined. In general, there are stricter weather requirements for more complicated actions, so the probability that corrective repairs are not possible is larger than for preventive repairs. And if preventive repairs are not possible, neither are corrective. Therefore, the probability that corrective repairs are not possible during a time step is provided conditioned that preventive repairs are possible. Similarly, the probability that preventive repairs are not possible is provided conditioned that inspections are possible.

The calculations are performed both for the case without and with seasons. Without seasons, it is assumed that inspections and repairs can always be made during the time step for which they are planned. When seasons are included, inspections are still always assumed to be possible. Preventive repairs are always possible during the summer half, but in the winter half there is a probability of 0.1

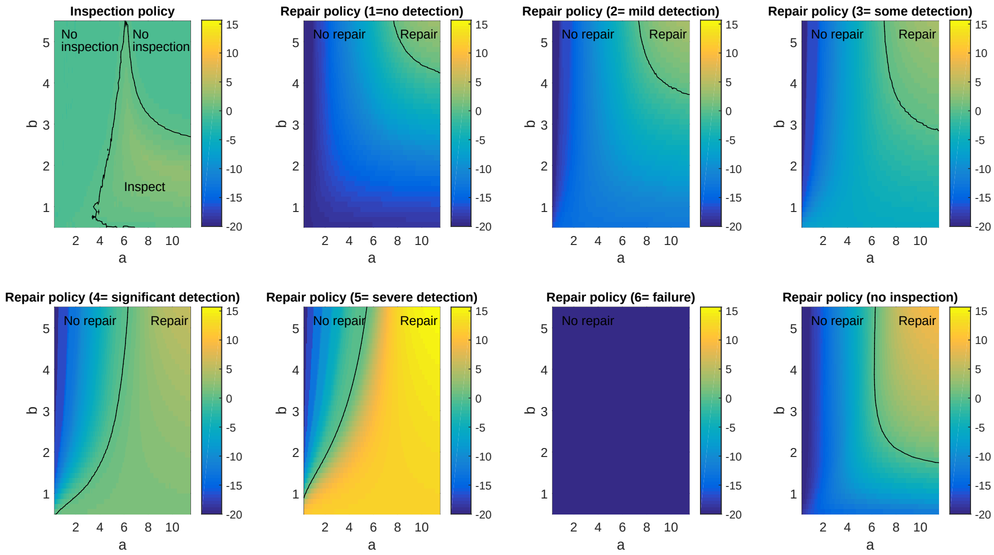

Figure 4: Decision policies 10 years into the lifetime for Example 1. The color shows the difference in costs for making the inspection/repair or not doing it. Positive values indicate that the optimal decision is to inspect/repair, and the black lines divide the regions with different optimal decisions. The repair policy is shown for each inspection outcome and for no inspection.

that it cannot be made during a time step. For corrective repairs, the probability that it cannot be done during a time step given that a preventive repair can be made is 0.05 during summer half and 0.4 during the winter half.

The range and spacing for the parameters a and b need to be chosen based on a tradeoff between accuracy and computation time. The probability distribution for the damage size was found for each time step for various combinations of observations, and the corresponding range of a and b values was found. Values of a are chosen in the range from 0.25 to 11.5 with 0.05 distance between values. For b the values are in the range from 0.5 to 5.5 with distance 0.125. Additionally, the expected costs are found for the failed case. This gives 9267 probability distributions in total for the damage size. The computation time per time step was around 2 minutes, and the total computation time for 240 time steps was around 8 hours on an Intel Core i7 processor using parallel computing in Matlab.

3.2. Results

The outcome of the computations is a set of decision policies for each time step in the model. Figure 4 shows an example for year 10, for the case without seasons. Not all policies are relevant for all values of a and b. For example, if an inspection should not be made, it does not matter what the optimal repair decision is for each inspection outcome. Therefore, the policies can be summarized in a single figure as shown in figure 5. As seasons are not included, policies for adjacent time steps are very similar. However, near the end of the lifetime (20 years), the policies will change. Figure 6 shows decision policies for year 18. As expected, damages should be larger before they are repaired compared to year 10.

When seasons are included in the model, the decision policies will generally vary during the year. Figure 7 shows the decision policies for inspections for all months in year 10. During the summer half, inspections should be made at damages lower than in the winter, such that repairs are less likely to be made during the winter.

To validate the efficiency of the found decision policies, simulations are run where the found policies are applied each time a decision is made. For comparison, simulations are also run for time-invariant decision policies, where the inspections

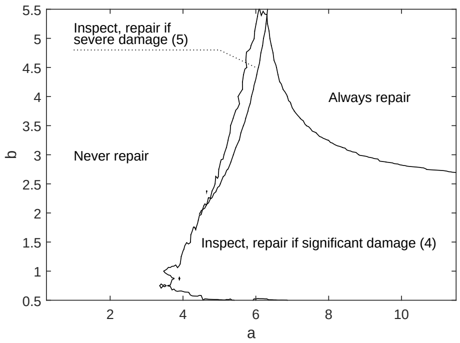

Figure 5: Summarized decision policies 10 years into the lifetime for Example 1.

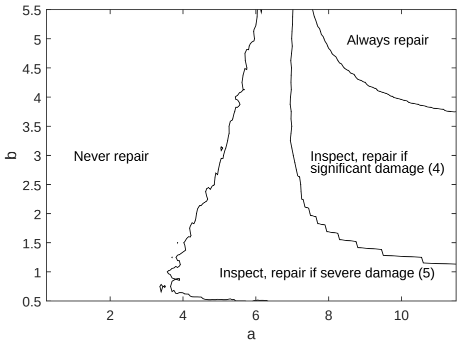

Figure 6: Summarized decision policies 18 years into the lifetime for Example 1.

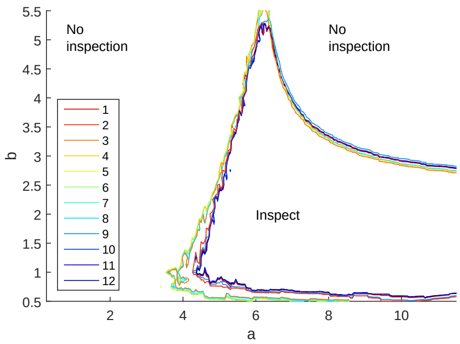

Figure 7: Decision policies for inspections for all months in year 10 for Example 1. The first three and last three months are considered winter months.

are made when the probability of failure during the following time step, which is updated using the monitoring outcome, is above a threshold value, and repairs are made when the inspection result is above a threshold value. Both threshold values are optimized using simulations. For both types of decision rules, Bayesian updating is performed during simulations using a DBN approach, and 100,000 simulations are made for each case. Figure 8 shows the expected costs for both cases. The two methods are almost equally good, but the threshold approach gives slightly lower costs compared to the POMDP policies, both for the case with and without seasons. For comparison, the expected costs are 347 when only corrective maintenance is used.

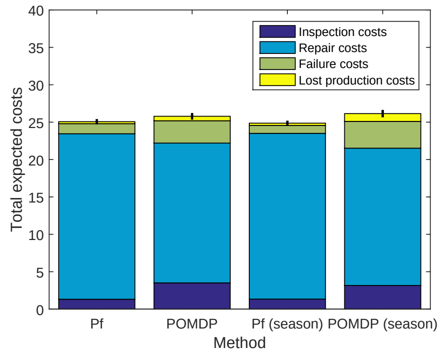

Figure 8: Expected costs for POMDP and threshold approach (Pf) both with and without seasons for Example 1. The vertical black lines show the 95% confidence intervals for the total costs.

4. EXAMPLE 2

For this example, the damage model is based on a fracture mechanical model. The damage size (crack length) a can be found based on the damage size in the previous time step using the following expression (Ditlevsen and Madsen, 2007):

$$a_t = \left((1 - \frac{m}{2}) C \Delta S^m \pi^{m/2} \Delta n + a_{t-1}^{1-m/2} \right)^{(1-m/2)^{-1}} \tag{2}$$

Where ΔS is the stress range, Δn is the number of stress cycles, and m and C are empirical model parameters. For this example, the time steps is one month, ΔS is assumed normal distributed with mean 60 and standard deviation 10, Δn is deterministic 10^6, and m is deterministic with value 3.5. The initial value a_0 is assumed exponential distributed with mean value 0.2. The value of C is found by calibration using Crude Monte Carlo simulations to give same mean time to failure as in Example 1. A value of $C = e^{-33.5}$ was found using 100,000 simulations.

4.1. Model

Next, a DBN model for the damage development was made following (Straub, 2009). As the damage model is exponential, the intervals for the damage size have an exponentially increasing size. To find the transition matrix, Monte Carlo simulations were used. A good accuracy could be obtained using 80 intervals for the damage size, but that would result in long computation time (more than one week) for the Markov model. Instead 30 intervals were used, even though it gave an overestimation of the probability of failure. This was corrected by decreasing all probabilities below the diagonal by a constant factor, and increasing the probabilities on the diagonal to keep a total probability of one for each interval. The factor was chosen such that the probability of failure after 20 years was equal to the value found from the original model using Monte Carlo simulations. Figure 9 shows the probability of failure as function of time for three cases: Monte Carlo simulations including 95% confidence intervals, DBN model with 30 states, and the edited DBN model with 30 states.

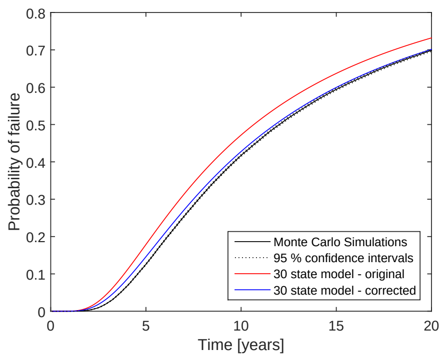

Figure 9: Probability of failure as function of time for the original model, DBN model with 30 states, and corrected model with 30 states for Example 2.

The range of *a* and *b* values was found in a similar way as in Example 1, and the range for *a* was 0.05 to 10 with a step length 0.05, and for *b* the range was from 2.5 to 8 with a step length 0.125. In total, 9001 probability distributions including the distribution for a failed component. Each time step has a computation time of around 9 minutes, giving a total computation time of 36 hours.

4.2. Results

The decision policies were found for the case with seasons, and figure 10 shows the summarized decision policies for year 10. Generally, inspections should not be made, but repairs should be made at distributions with lower scale parameters compared to the linear model in Example 1. To validate the efficiency of the decision policies, simulations have been run as in Example 1 and the total expected costs are compared to the threshold approach in figure 11. The two methods give almost the same total expected costs. For comparison, the expected costs for corrective maintenance only is 682.

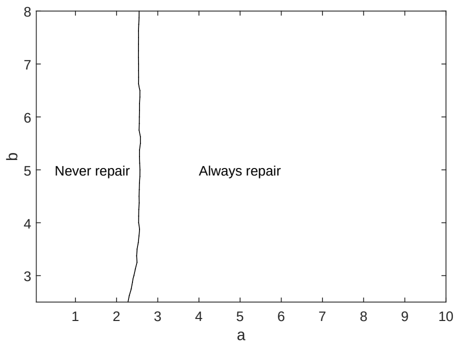

Figure 10: Summarized decision policies 10 years into the lifetime for Example 2.

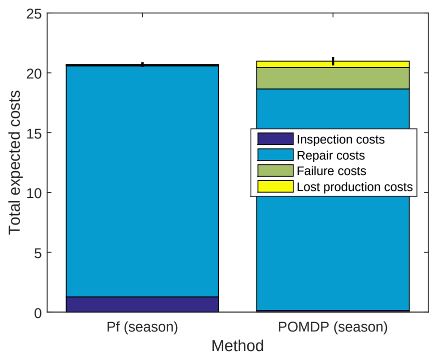

Figure 11: Expected costs for POMDP and threshold approach (Pf) for Example 2. The vertical black lines show the 95% confidence intervals for the total costs.

5. DISCUSSION

The examples show that the POMDP method is able to give almost as good results as the efficient simulation based threshold method. It was expected that it would give lower costs, especially, when seasons are included, as the POMDP method takes that into consideration and can have annual variations in the decision policies unlike the threshold method. Higher costs of lost production and lower accessabilities during the winter could possibly make the POMDP approach more beneficial.

For the threshold method, only the probability of failure during the following time step is considered, when decisions are made. It has a direct relationship with the expected failure costs, but not with the expected inspection outcomes and as such not the expected costs to preventive repairs. The POMDP method considers the entire probability distribution and, therefore, the relationships with both expected failure and repair costs. If the probability distribution could always be well approximated by a Weibull distribution and if proper interpolation was performed, the POMDP method should give the lowest costs. However, a 2-parameter Weibull distribution does not always give a perfect fit, especially when observations are included. A better fit could be obtained by introducing a lower bound using a 3-parameter Weibull distribution. If the computation time should still be limited, the number of a and b values should be reduced to keep the same total number of probability distributions.

The time-consuming part of the computation is to find the nearest grid point or, alternatively and even more time-consuming, to make a nonlinear fit to a Weibull distribution. Therefore, the number of times this is done will have a linear effect on the computation time. The number of times the interpolation is done is the product of the number of time steps, the number of grid points, and the number of branches for each time step. Additionally, the time spent on each interpolation depends on the number of damage states and the number of grid points.

A drawback of the POMDP model is the Markovian assumption, as time-invariant parameters are hard to include in the model. To do so, the grid points should be found for 'all possible' joint distributions for the damage size and a time-invariant model parameter. Even a relatively simple model with two parameters for each variable and a correlation would give a five-dimensional grid. As the computation time increases at least linear with the number of grid points, it will probably be too time-consuming. For the threshold approach, time-invariant parameters can easily be included in the model.

ACKNOWLEDGEMENTS

The work presented in this paper is part of the project "Reliability-based analysis applied for reduction of cost of energy for offshore wind turbines" supported by the Danish Council for Strategic Research, grant no. 2104-08-0014. The financial support is greatly appreciated.

6. REFERENCES

Dasgupta, S., Papadimitriou, C., and Vazirani, U. (2006). *Algorithms*. McGraw-Hill.

Ditlevsen, O. and Madsen, H. O. (2007). *Structural Reliability Methods*. Department of Mechanical Engineering, Technical University of Denmark.

Nielsen, J. S. (2013). "Risk-based operation and maintenance of offshore wind turbines." Ph.D. thesis, Department of Civil Engineering, Aalborg University.

Nielsen, J. S. and Sørensen, J. D. (2012). "Maintenance optimization for offshore wind turbines using POMDP in Reliability and Optimization of Structural Systems: Proceedings of the 16th working conference on the international federation of information processing (IFIP) working group 7.5 on reliability and optimization of structural systems. ed. A. Der Kiureghian, A. Hajian. American University of Armenia Press, Yerevan, Armenia, 2012. p. 175-182.

Raiffa, H. and Schlaifer, R. (1961). *Applied statistical decision theory*. Harvard University.

Straub, D. (2004). "Generic approaches to risk based inspection planning for steel structures." Ph.D. thesis, Swiss Federal Institute of Technology, ETH Zurich.

Straub, D. (2009). "Stochastic modeling of deterioration processes through dynamic bayesian networks." *Journal of Engineering Mechanics*, 135(10), 1089–1099.