



Stochastics and Statistics

Partially Observable Markov Decision Processes incorporating epistemic uncertainties

R. Faddoul^{a,*}, W. Raphael^{a,*}, A.-H. Soubra^b, A. Chateauneuf^c^a Civil and Environmental Engineering Department ESIB, Saint-Joseph University, Mar Roukos, PoB 11-154, Riad El Solh, Beirut, Lebanon^b Institut de Recherche en Génie Civil et Mécanique, BP 152, 44603 St-Nazaire Cedex, France^c LaMI, Polytech'Clermont-Ferrand, BP 206, 63174 Aubière Cedex, France

ARTICLE INFO

Article history:

Received 10 April 2013

Accepted 19 August 2014

Available online 22 September 2014

Keywords:

Uncertainty modeling

Markov process

Dynamic programming

Epistemic uncertainty

Decision analysis

Maintenance

ABSTRACT

The use of Markov Decision Processes for Inspection Maintenance and Rehabilitation of civil engineering structures relies on the use of several transition matrices related to the stochastic degradation process, maintenance actions and imperfect inspections. Point estimators for these matrices are usually used and they are evaluated using statistical inference methods and/or expert evaluation methods. Thus, considerable epistemic uncertainty often veils the true values of these matrices. Our contribution through this paper is threefold. First, we present a methodology for incorporating epistemic uncertainties in dynamic programming algorithms used to solve finite horizon Markov Decision Processes (which may be partially observable). Second, we propose a methodology based on the use of Dirichlet distributions which answers, in our sense, much of the controversy found in the literature about estimating Markov transition matrices. Third, we show how the complexity resulting from the use of Monte-Carlo simulations for the transition matrices can be greatly overcome in the framework of dynamic programming. The proposed model is applied to concrete bridge under degradation, in order to provide the optimal strategy for inspection and maintenance. The influence of epistemic uncertainties on the optimal solution is underlined through sensitivity analysis regarding the input data.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Most countries worldwide face the problem of ageing civil engineering infrastructures. This is especially the case in the developed countries where investment focus is increasingly shifting from extending the infrastructure to the inspection, maintenance and rehabilitation (IM&R) of the existing assets. For example, government statistics in the United States show that the proportion of public non-capital spending for infrastructure increased from 39 percent in 1960 to 57 percent in 1994 (CBO, 1999). Although mathematical optimization and operational research methods started to be gradually applied for maintenance optimization since the early 1940s, it is until the last two decades that maintenance optimization became a topic of highest priority for civil engineering infrastructure managers. Several researches emphasized on the usefulness of the characterization of the uncertainties (pertaining to reliability analysis and to risk-based decision problems in engineering) as epistemic or aleatory (Der Kiureghian & Ditlevsen, 2009; Hofer, 1996). The decision process regarding the

optimal maintenance policy is often very sensitive to the input statistical parameters (e.g. mean values, standard deviations, Markov transition matrices) of the probabilistic models (Avrachenkov, Filar, & Haviv, 2001; Cranahan, 1988). However, these parameters are estimated by classical inference methods where the available data is often very limited (DeStefano & Grivas, 1998; Madanat, Mishalani, & Wan Ibrahim, 1995; Mishalani & Madanat, 2002). This lack of information leads to second-order uncertainties due to the inevitable difference between the observed sample from which inference is made and the real population. This second-order statistical uncertainty, shrouding the real values of the parameters, could be usually quantified by several means, such as confidence intervals, Bayesian probability distributions and fuzzy sets (Corotis, 2009). Keeping in mind that exact and tractable modeling of reality is a far-fetch if not an impossible goal to achieve, it remains that a desired attribute of a mathematical model is to be able to use, as best as possible, the available information. In fact, a mathematical optimization model using only a portion of the relevant information will produce biased optimal solutions. Parametric uncertainty is one type of information that is seldom taken into account by most of the existing models for IM&R optimization. Therefore, point estimators for the transition matrices in Markov Decision Processes (MDP) are simply used instead of exploiting the additional information that can be available. If, for example, a probability

* Corresponding author. Tel.: +961 3 731450.

E-mail addresses: faddoul_rafi@yahoo.com (R. Faddoul), wassim.raaphael@fi.usj.edu.lb, wassim012@hotmail.com (W. Raphael), Abad.Soubra@univ-nantes.fr (A.-H. Soubra), alaa.chateauneuf@cust.univ-bpclermont.fr (A. Chateauneuf).

density function is available for the elements of MDP transition matrices, then the use of the whole information contained in such a distribution will lead to more realistic life cycle cost estimation and therefore to better IM&R decisions. It will also be shown that the expected costs usually increase with full uncertainty consideration. Hence, neglecting the epistemic uncertainty leads in general to misleading low life cycle costs.

In the next two sections, a literature review on the epistemic uncertainties in the transition matrices and on the dynamic programming of finite horizon MDPs or POMDPs (Partially Observable Markov Decision Processes) are presented. Next, we describe the parameters uncertainties in MDPs. Then, we propose a methodology for incorporating the epistemic uncertainty in MDPs. The paper ends with the presentation of a numerical example that considers IM&R optimization management of a highway concrete bridge deck.

2. Epistemic uncertainties in the transition matrices

Several studies have been suggested in the literature since the early 1950s in order to take into account the effect of uncertainties in the transition matrices. Bellman (1961) proposed an “adaptive control” of Markov chains by recognizing that the problem can be transformed into an equivalent dynamic programming problem with a completely known transition law. The state space of the equivalent problem is the Cartesian product of the original state space by the set of all probability distributions on the parameter set. This approach has the advantage of being adaptive but it suffers, as pointed out by Bellman himself, from the “dimensionality curse”; i.e. the problem becomes quickly intractable as the state space grows exponentially with higher dimensions. A customized version of this approach in the field of IM&R of civil engineering structures was proposed by Durango and Madanat (2008). One can say that most of the research tackling the issue of parametric uncertainty in sequential decision optimization problems focuses on investigating robust optimization techniques (Givan, Leach, & Dean, 2000; Iyengar, 2005; Nilim & El Ghaoui, 2005). In the same framework, ambiguity is investigated in the context of infinite horizon MDP with finite state and action spaces where it is modeled by constraining the transition probability matrix to lie in a pre-specified polytope (Bagnell, Ng, & Schneider, 2001; Satia & Lave, 1973; White & Eldeib, 1994). Such an approach (robust optimization) considers the optimization problem as a game between the decision maker and the nature which is supposed to be malevolent. In other terms, the worst case scenario is assumed, leading to policies which are deemed too conservative by several authors. For this reason, several studies attempted to avoid this drawback. Kuhn and Madanat (2006) used the Hurwicz criterion (Revelle, Whitlatch, & Wright, 1997) which allows the decision maker to set his or her own ‘optimism level’, as a number between 0 and 1. This number is used to balance the decisions by interpolating between the best case and the worst case scenarios. Delage and Mannor (2009) proposed a chance constrained MDP with a set of percentile criteria that represent the trade-off between the optimistic and the pessimistic views. However, as for the models of Kuhn and Madanat, a so-called risk of the policy level must be chosen at the discretion of the decision maker. Moreover, these approaches do not use all of available information about the epistemic uncertainty.

The authors of the present paper believe that for problems which do not include hard constraints (i.e. constraints which must be met compulsory otherwise the solution found is deemed not feasible), the whole information available about the uncertainty can be taken into account by the optimization model. A particular example of optimization problems, where the constraints are soft, is the management optimization of infrastructure facilities which is subject to limited budget constraints. In this case, constraining the expected direct cost that is to be paid by the manager during each time period, to be less than or equal to the imposed budgets limits is deemed to be sufficient.

In this paper, we present an extension of dynamic programming algorithms used to solve finite horizon MDPs or Partially Observable MDPs (POMDPs) in order to take into account second-order epistemic uncertainties. The whole information available about the uncertainty is taken into account by the optimization model. In addition, we propose a methodology based on the use of Dirichlet distributions which avoid, in our sense, much of the controversy dominating the literature about the different methods used to estimate the Markov transition matrices (maximum likelihood estimation, regression using state expectation, regression using state distribution, etc.). We illustrate the methodology by applying it to a Generalized Partially Observable Markov Decision Process (GPOMDP) (Faddoul, Raphael, & Chateaufneuf, 2009). A literature review on POMDPs and GPOMDPs is provided in the following section. Although the proposed methodology for including epistemic uncertainties in the decision process is motivated by IM&R optimization, the formulation is quite general and can be relevant to any MDP.

3. Dynamic programming for finite horizon POMDPs

The methodology presented in this paper uses dynamic programming to solve finite horizon MDPs. This technique has been used extensively for maintenance optimization (Wang, 2002). However, a legitimate question arises about the suitability of the Markov property to describe the state evolution of civil engineering structures. Conflicting arguments are presented in literature concerning this subject. Although Neves, Frangopol, and Cruz (2006) have criticized the Markov property assumption for civil engineering structures, Orcesi and Cremona (2010) have demonstrated that the homogeneous Markov assumption was justified in the case of the French national bridge stock. Whatever the case may be, a technique proposed by Robelin and Madanat (2007) allows the dynamic programming to take into account the possible non-Markovian effects of actions and/or deterioration processes.

A dynamic programming model is said to be deterministic if the application of an action or a degradation process at the beginning of stage n will result in a precisely known state of the system at the beginning of stage $n + 1$. On the other hand, in probabilistic dynamic programming, the application of an action or a degradation process at the beginning of stage n will lead to a probabilistic distribution of the state of the system at the beginning of the next stage (Pham & Wang, 1996).

In a maintenance model, large uncertainties are associated with the results of any maintenance action and/or degradation process, due to two main factors:

- 1 imperfectness of maintenance actions;
- 2 stochastic nature of the degradation process between the application time of the maintenance action and the beginning of the next stage.

Due to the abovementioned uncertainties, probabilistic dynamic programming models were used in the majority of maintenance optimization problems.

However, classical probabilistic dynamic programming models in maintenance assume perfect inspections at the beginning of each time period. This type of models suffers from two main drawbacks; namely:

- 1 The inspections are assumed to be perfect and this is rarely the case;
- 2 The optimization of inspection planning (which is vital because inspections have usually significant costs) is not possible.

The use and development of POMDPs was essentially driven by the goal of dealing with the first one of these two shortcomings of classical probabilistic MDPs.

In a POMDP (Eckels, 1968; Monahan, 1982), the state of the system at the beginning of each time period cannot be fully observed. Thus, the manager of the system must rely on a characterization of a partially observed state; i.e. a belief state which is usually described by probabilistic distributions. In a classical POMDP, the “belief state” of the system at the beginning of stage n is defined by the vector $v^n = [v_1^n, v_2^n, \dots, v_k^n]$ where the v_i^n are the probabilities associated with the different possible values θ_i^n of the exact state θ^n of the system at the beginning of stage n , i.e. $v^n = [\Pr(\theta_1^n), \Pr(\theta_2^n), \dots, \Pr(\theta_k^n)]$.

The effect of a maintenance action or a degradation process can be modeled (as in a probabilistic dynamic programming problem) by the transition matrices A_{an} and M respectively where the element a_{ij} ($i = 1, \dots, k; j = 1, \dots, k$) of matrix A_{an} represents the probability that the system evolves from the state θ_i^n to the state θ_j^n if we implement the maintenance action a at the beginning of stage n and where the element m_{ij} ($i = 1, \dots, k; j = 1, \dots, k$) of matrix M represents the probability that the system evolves from the state θ_i^n to the state θ_j^{n+1} as a result of the degradation process.

The belief state ${}^av^n$ of the system during stage n , after the implementation of a maintenance action, will be equal to the matrix product of the vector v^n by the maintenance transition matrix A_{an} ; i.e. ${}^av^n = v^n \times A_{an}$. Similarly, the belief state v^{n+1} of the system at the beginning of stage $n + 1$, that is, after the evolution of the system, due to the Markovian degradation process, will be equal to the matrix product of the vector ${}^av^n$ by the Markovian degradation process transition matrix M ; i.e. $v^{n+1} = {}^av^n \times M$. Thus, a maintenance action and/or a Markovian degradation process (having probabilistic consequences on the system degree of degradation) will result in a probability distribution, i.e. an exactly well defined belief state of the system.

Similarly to a deterministic dynamic programming problem, solving a finite time horizon POMDP by dynamic programming requires a backward calculation (beginning from the last time period) of the costs associated with each of the belief states v^n of stage n by choosing the action that minimizes the total cost $c(v^n)$ (Bellman, 1961). This cost is composed of the cost of that action $ca(a^n)$ and the discounted optimal cost $\alpha \times {}^*c(v^{n+1} | a, v^n)$ associated with the forecasted belief state of stage $n + 1$, knowing that we applied the action a and that we were in the belief state v^n at the beginning of stage n . i.e. the recursive relation will be:

$$c(v^n) = ca(a^n) + \alpha \times {}^*c(v^{n+1} | a, v^n) \quad (1)$$

One can note that Eq. (1) is similar to a deterministic dynamic programming recursive relation but where we replace the state θ^n by the belief state v^n .

Extensions and generalizations of the POMDPs described above were proposed by Corotis, Ellis, and Jiang (2005) and Faddoul et al. (2009) through the development of GPOMDPs. A GPOMDP allows for easy mathematical modeling of complex sequence of actions to be undertaken during each time period of the planning horizon. For example, a sequence of actions may consist of two inspection decisions followed by one maintenance decision. This type of modeling is relevant to maintenance problems, as a cheap inspection may be carried out first, in order to decide whether a more precise and costly inspection should be applied before taking the decision relative to the type of the repair technology.

4. Modeling of parameters uncertainties in MDPs

The transition matrices relative to degradation processes, maintenance actions and inspections, are usually estimated by using one, or a combination, of the following techniques (DeStefano & Grivas, 1998; Kallen, 2007): (i) expert evaluation based on past experience; (ii) expert evaluation based on analytical reasoning, taking into account the physical and chemical properties underlying the considered

evolution processes; (iii) statistical estimation based on historical observation samples of the system itself and/or similar systems; (iv) estimation based on stochastic mathematical models which might be fitted to available historical data, completed by physical knowledge to calibrate some of the model parameters.

Whatever the applied estimation methods, uncertainties will usually veil the exact and true values of the estimated matrices. These uncertainties can be typically expressed by: (i) fuzzy sets, (ii) mean value and variance of the estimated parameters, (iii) confidence intervals for a specified significance level, (iv) full probability distributions. In what follows, we present the probabilistic approaches that may be used to describe the epistemic uncertainties to be included in a MDP.

Consider the case of an estimation procedure, of a transition matrix T , based on N observations of the true state of the system at the beginning of stage $n + 1$, knowing that it was at an exactly defined state θ_i^n at the beginning of stage n . Such a situation corresponds to the case where the inspections used for the estimation of T are assumed to be performed at regular time intervals equal to the transition period of the Markov process. Moreover, it is assumed that these inspection techniques are perfect, i.e. they reveal the true state θ of the structure. The cases where the inspections used for the estimation of the transition matrices are imperfect and/or performed at varying time intervals, are treated in Appendices A.1 and A.2.

The rows of T can be considered as the vector parameters of multinomial distributions of dimension K (K being the number of states). i.e. the observed states of a sample of size N , that was in state θ_i^n at the beginning of stage n , will be distributed according to the following multinomial probability distribution:

$$f(x_1, \dots, x_K; N, p_{i1}, \dots, p_{iK}) = \frac{N!}{x_1! \times \dots \times x_K!} \times p_{i1}^{x_1} \times \dots \times p_{iK}^{x_K} \quad (2)$$

where x_i is the number of observations where the system state is θ_i^{n+1} at the beginning of stage $n + 1$, and p_{ij} is the element at row i and column j of the transition matrix.

The epistemic uncertainty related to the parameters p_{ij} , in a given row of the transition matrix, can be conveniently represented by the Dirichlet multivariate probability distribution since it is the Bayesian conjugate prior to the multinomial distribution.

Hence, the epistemic uncertainty related to the parameters p_{ij} , in a given row of the transition matrix, can be conveniently represented by the following Dirichlet multivariate probability distribution:

$$\text{Dir}(p_{i1}, \dots, p_{iK-1}; \alpha_1, \dots, \alpha_K) = \frac{\Gamma(\sum_{j=1}^K \alpha_j)}{\prod_{j=1}^K \Gamma(\alpha_j)} \prod_{j=1}^K p_{ij}^{\alpha_j-1} \text{ with parameters } \alpha_1, \dots, \alpha_K > 0 \quad (3)$$

where $\Gamma(\cdot)$ is the gamma function. As we have mentioned in Section 3, the belief state vector $v^{n+1} = [v_1^{n+1}, v_2^{n+1}, \dots, v_k^{n+1}]$ at the beginning of stage $n + 1$ is equal to the matrix product: $v^{n+1} = v^n \times T$, where T is a transition matrix, it comes:

$$v_i^{n+1} = \sum_{j=1}^k v_j^n \times p_{ji} \quad \forall i = 1, \dots, k \quad (4)$$

In other terms, the belief state vector v^{n+1} is a linear combination of the transition matrix rows which are random vectors having Dirichlet distributions. Thus, we have translated the epistemic uncertainty that was shrouding the estimates of the transition matrix elements into a probability distribution of the belief state v^{n+1} .

Several techniques are available for estimating the parameters of Dirichlet distributions from sample data (Minka, 2003; Wicker, Muller, Kalathur, & Poch, 2008). However, a simple and natural technique arises from the fact that the Dirichlet distribution is a conjugate prior to the multinomial distribution. If we take a prior $\text{Dir}(p_{i1}, \dots, p_{iK-1}; \alpha_1, \dots, \alpha_K)$ and a multinomial sample x_1, \dots, x_K , the posterior will be $\text{Dir}(p_{i1}, \dots, p_{iK-1}; \alpha_1 + x_1, \dots, \alpha_K + x_K)$. Hence,

the parameters of the prior can be thought of as pseudo-counts; i.e. the probability density function of a Dirichlet distribution returns the belief that the probabilities of K rival events are p_i given that each event has been observed $(\alpha_i - 1)$ times.

If, for example, we are given a sample of evolving structures which has resulted in a matrix S , where an element s_{ij} represents the number of structures which were in state θ_i^n and evolved to state θ_j^{n+1} , the Dirichlet parameters for each row of the transition matrix will simply be represented by a row of a matrix $\alpha_{i1}, \dots, \alpha_{iK}$ where an element $\alpha_{ij} = s_{ij} + 1$.

Even if an element s_{ij} of the sample matrix is equal to zero, the corresponding element p_{ij} of the estimated transition matrix may be, due to the statistical uncertainty, different from zero even though it will be usually relatively small. Hence, special attention needs to be paid for transition matrices where there is an impossibility that the structure evolves from some specified states to other specified states, as for example for the deterioration matrix which is usually an upper triangular matrix due to the fact that usually the performance of the structure cannot improve on its own without any maintenance action. For such matrices, the Dirichlet parameter vectors must only include the elements α_{ij} for which a transition is possible, and thus, they may have a lower dimension than the state space of the structure.

Even though, the Dirichlet distribution is typically chosen as the probability distribution for the rows of the transition matrices, one can choose any other multivariate distribution that might seem suitable for particular problems. When the available estimation data is given either in the form of estimates of the mean $\tilde{E}(p_{ij})$ and the variance $\tilde{\text{Var}}(p_{ij})$, or in the form of an estimated mean value $\tilde{E}(p_{ij})$ and a confidence interval I , a technique for the reparametrization of a Dirichlet distribution is presented in Droesbeke, Fine, and Saporta (2002, chap. 14).

Since the belief state vector v^{n+1} is equal to the matrix product $v^n \times T$, where T is a random transition matrix, then v^{n+1} will also be a random vector having a probability distribution.

5. Incorporating epistemic uncertainty in MDPs

A classical POMDP behaves like a classical deterministic MDP but over the belief states instead of the exact states. In such a model, the results of actions, degradations or inspections allow the probabilistic outcomes in the original state space to have exact outcomes in the belief state space (i.e. space of probability distributions). This is because the outcomes are belief states which are exactly defined and expressed by probability vectors. If we extend the classical POMDP model to the case of a probabilistic POMDP, i.e. making the POMDP behave like a probabilistic MDP but over the belief states rather than the exact states, then any result of action, degradation or inspection will be a probabilistic distribution of the belief states. It is in this second order probabilistic distribution that the epistemic uncertainty will be integrated.

In order to calculate the optimal expected cost $*c(v^n)$ associated with the belief state v^n , the dynamic programming recursive procedure will call, at the beginning of stage n , for the minimum expected costs $*c(v^{n+1}|v^n)$ calculated with the updated belief states at the beginning of stage $n + 1$. To take into account the uncertainties in estimating the matrices of actions, degradations and inspections, a probabilistic POMDP must evaluate the expected value $E[*c(v^{n+1}|v^n)]$ (rather than $*c(v^{n+1}|v^n)$ which evaluates the optimal expected cost of exactly updated belief states), knowing that the vector v^{n+1} has a probability distribution over the belief state space at the beginning of stage $n + 1$.

In our approach, the belief state vector $v^{n+1}|v^n$ is considered as a random vector. It is a function of the transition matrix rows which are

random vectors having a known distribution (e.g. Dirichlet distribution). Consequently, the expected cost associated with the probability distribution of $v^{n+1}|v^n$ will be:

$$E[*c(v^{n+1}|v^n)] = \int_{\text{statespace}} *c(v^{n+1}|v^n) \times f(v^{n+1}|v^n) \quad (5)$$

where $f(\cdot)$ is the probability distribution of the belief state $v^{n+1}|v^n$ resulting from epistemic uncertainties, and $E[\cdot]$ is the mathematical expectation operator.

The minimum expected cost $*c(v^{n+1}|v^n)$ in Eq. (5) is calculated recursively using a dynamic programming algorithm. As it can be seen from Eq. (5), the uncertainties in the transition matrices are accounted for by taking the expectation of the minimum expected costs $*c(v^{n+1}|v^n)$ over the probability density function $f(v^{n+1}|v^n)$ of the belief state $v^{n+1}|v^n$.

By using Eq. (5), we are applying the principle of separating the epistemic and the aleatory uncertainties (Hofer, 1996) related to the state of the system during any of the future periods where the epistemic and the aleatory uncertainties are represented respectively by $f(\cdot)$ and $(v^{n+1}|v^n)$. It should be emphasized here that once the system has reached the time period n all the uncertainties related to the state of the system are epistemic uncertainties since the system has one exact true state which is unknown to the manager (Hofer, 1996).

A convenient approach which could be used for computing Eq. (5) is the Monte Carlo simulation technique. A crucial concern to deal with is the issue of additional computational complexity that Monte Carlo simulations would bring to the problem. Notice however that the dynamic programming framework alleviates drastically this difficulty as follows:

If we assume that the deterioration transition matrices M_j are particular occurrences of a random variable \mathbf{M} , and if one samples the transition matrix SN_{\max} times for each stage, then the total number of sampled states in stage i will be equal to SN_{\max} power i ($(\text{SN}_{\max})^i$). If for example one samples the transition matrices for each stage three times, then the number of states sampled at the 20th stage will be 3^{20} .

We present in Fig. 1 the proposed algorithm for the case where the deterioration transition matrix is uncertain (the cases where the action and/or inspection matrices are uncertain are straightforward). The gray elements of the flowchart correspond to the added Monte Carlo simulations to the basic dynamic programming algorithm. In the upper part of the flowchart of Fig. 1, we have four nested loops:

- (i) In the outer loop, the stage index i is decremented from the end of the planning horizon ($i = N$) to time period ($i = 1$).
- (ii) The second nested loop considers all the K possible belief states v_k^i of the i th stage.
- (iii) The third nested loop considers all the a_{\max} alternative actions. In this loop, we minimize the cost $cs(v_k^i \times A_a) + ca(a) + *cf$ where $*cf$ is the average optimal future cost; $ca(a)$ is the cost of the maintenance action a and $cs(v_k^i \times A_a)$ is the immediate users cost due to the application of action a .
- (iv) The inner loop corresponds to the sampling of the uncertain transition matrix, the prescribed sampling number being SN_{\max} . At the end of this loop, we calculate the average of the obtained optimal future costs $*c(v_k^i \times A_a \times M)$.

Since there is only one belief state in the first stage, a specific sampling loop is performed for that initial belief state in the lower part of the algorithm.

As for the updating of the epistemic uncertainties, we adopt the open-loop feedback control approach in the sense that, after each stage, in the case of the use of Dirichlet matrices, the pseudo-counts of these matrices are updated according to the results of the

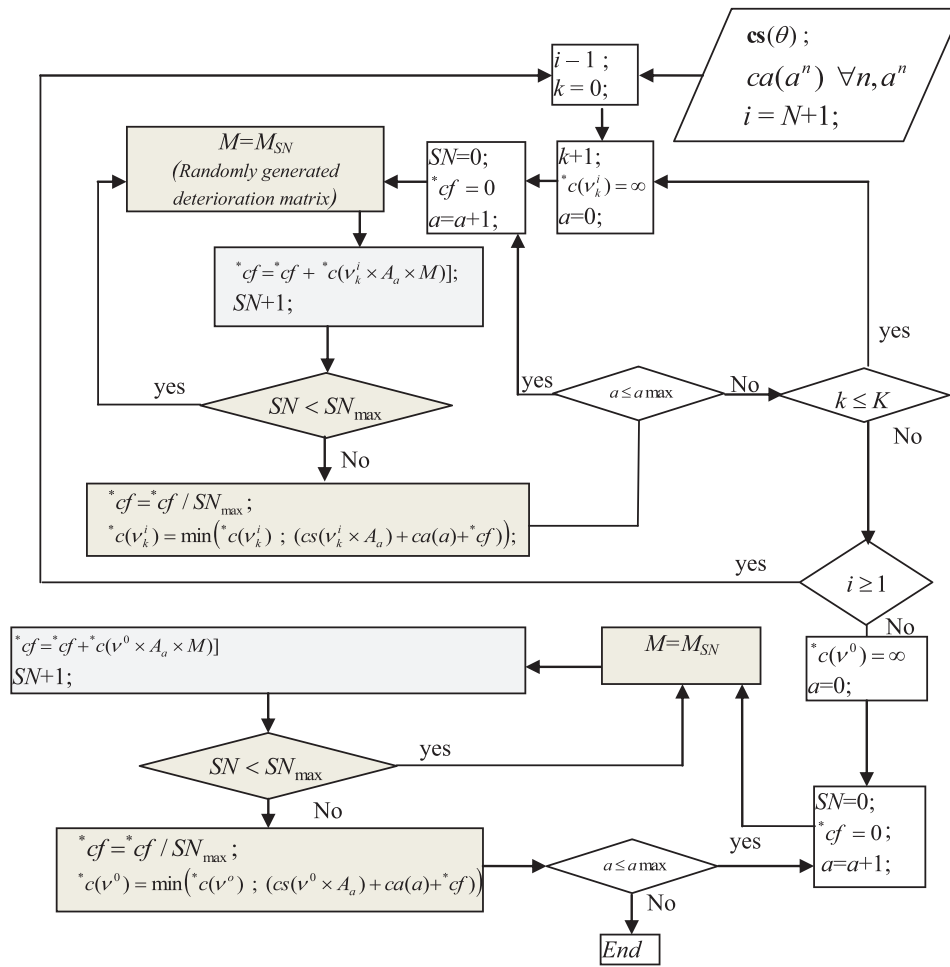


Fig. 1. Dynamic programming algorithm for uncertain deterioration matrix.

applied inspections to our structure and others of the same type as well.

As for the sampling of the transition matrices, we believe that two cases are relevant:

- A stationary uncertainty model:* In this modeling the same set of sampled matrices is used for every stage and state.
- A time varying uncertainty model:* In such a modeling, a new set of sampled matrices is used for each stage.

The first procedure corresponds to the case where there is only one true but unknown transition matrix for all the periods. The second procedure corresponds to the case where the epistemic uncertainty emanates from the influence of an unknown variable which affects the evolution process and which is subject to variation with time; i.e. there is a single known probability distribution from which a different transition matrix is true for each period. The methodology proposed in this paper can be used for both cases.

As mentioned above, the underlying principle of the algorithm described in this paragraph is also valid for the transition matrices of the actions and the uncertainties matrices of the inspections in a GPOMDP framework. It should be emphasized here that a probabilistic MDP solved by dynamic programming can be considered as a special GPOMDP for which the results of the actions are uncertain but for which we assume a single perfect inspection to be applied for free at the beginning of all time periods. Therefore, the approach for including epistemic uncertainty presented in this section can be easily extended to probabilistic MDPs.

6. Numerical application

Consider a highway concrete bridge whose deck is made of a single concrete slab; the herein data are for illustrative purposes only. We suppose that the performance of the concrete deck is described using five states as may be shown from Table 1. In this table, the costs $cs(a\theta_i^n)$ are generally the sum of the costs incurred by the user of the bridge due to malfunctioning of the deck along with the risk of failure which is expressed in monetary units and calculated as the product of the failure probability of the deck by the costs generated by such a failure.

The stochastic deterioration of the deck is modeled by a Markov chain with state space $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ having the following transition matrix:

	θ_1^{n+1}	θ_2^{n+1}	θ_3^{n+1}	θ_4^{n+1}	θ_5^{n+1}
$a\theta_1^n$	0.571	0.285	0.144	0	0
$a\theta_2^n$	0	0.571	0.285	0.144	0
$a\theta_3^n$	0	0	0.571	0.285	0.144
$a\theta_4^n$	0	0	0	0.714	0.286
$a\theta_5^n$	0	0	0	0	1

For this structure, four imperfect inspection techniques (i_0, i_1, i_2, i_3) are available, where i_0 means that no inspection is performed (i.e. we entirely rely on the prediction of the degradation model) and its cost is nil. The cost of inspection technique i_1 is 20 units and the cost of

Table 1
Costs due to deck condition.

Condition of the deck		Costs incurred during each stage due to the deck condition, $cs(\theta_i^n)$
θ_1	Very good condition	200 units
θ_2	Good condition	600 units
θ_3	Fairly good condition	1250 units
θ_4	Poor condition	2000 units
θ_5	Very poor condition	3500 units

Table 2
Conditional probability distribution of inspections results.

$P(r_k \theta_k^n)$	Inspection technique	r_1	r_2	r_3	r_4	r_5
θ_1^n	i_1	0.444	0.333	0.111	0.112	0.
	i_2	0.888	0.112	0.	0.	0.
	i_3	0.555	0.223	0.111	0.111	0.
θ_2^n	i_1	0.222	0.444	0.222	0.112	0.
	i_2	0.1	0.8	0.1	0.	0.
	i_3	0.223	0.555	0.111	0.111	0.
θ_3^n	i_1	0.1	0.2	0.4	0.2	0.1
	i_2	0.	0.111	0.777	0.112	0.
	i_3	0.	0.112	0.777	0.111	0.
θ_4^n	i_1	0.	0.112	0.222	0.444	0.222
	i_2	0.	0.111	0.111	0.555	0.223
	i_3	0.	0.	0.1	0.8	0.1
θ_5^n	i_1	0.	0.112	0.111	0.333	0.444
	i_2	0.	0.111	0.111	0.223	0.555
	i_3	0.	0.	0.	0.112	0.888

inspection techniques i_2 and i_3 is 40 units each. For each inspection technique we assume that there are k possible results r_k ($k = 1, \dots, 5$). The uncertainties associated with the results of the inspection techniques i_1 , i_2 and i_3 are expressed by the probability distributions shown in Table 2.

We assume that only four imperfect maintenance actions (a_0 , a_1 , a_2 , a_3) can be employed, where a_0 means that no maintenance is performed and its cost is nil. The cost of action a_1 (preventive maintenance action) is 800 units, the cost of action a_2 (corrective maintenance action) is 800 and the cost of action a_3 (replacement of the deck) is 3000 units. The probabilities associated with the consequences of the maintenance actions are expressed by the following matrices:

		$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
$A_0 =$	θ_1^n	1	0	0	0	0
	θ_2^n	0	1	0	0	0
	θ_3^n	0	0	1	0	0
	θ_4^n	0	0	0	1	0
	θ_5^n	0	0	0	0	1

Transition matrix for action a_0

		$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
$A_1 =$	θ_1^n	1	0	0	0	0
	θ_2^n	0.666	0.334	0	0	0
	θ_3^n	0.4	0.4	0.2	0	0
	θ_4^n	0	0.167	0.333	0.333	0.167
	θ_5^n	0	0	0.285	0.285	0.43

Transition matrix for action a_1

		$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
$A_2 =$	θ_1^n	0.9	0.1	0	0	0
	θ_2^n	0.142	0.716	0.142	0	0
	θ_3^n	0.1	0.3	0.5	0.1	0
	θ_4^n	0.4	0.3	0.2	0.1	0
	θ_5^n	0.2	0.4	0.2	0.1	0.1

Transition matrix for action a_2

		$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
$A_3 =$	θ_1^n	1	0	0	0	0
	θ_2^n	1	0	0	0	0
	θ_3^n	1	0	0	0	0
	θ_4^n	1	0	0	0	0
	θ_5^n	1	0	0	0	0

Transition matrix for action a_3

The element a_{ij} in each of these matrices corresponds to the probability that the deck, which is initially in the state θ_i^n (before the maintenance action), will be after the application of the maintenance action in the state $a\theta_j^n$ (the superscript “a” means that the state into consideration occurs immediately after the maintenance action before any degradation can take place).

We suppose that the initial belief state of the deck is $v = [0.2 \ 0.2 \ 0.3 \ 0.2 \ 0.1]$. The planning horizon is set to 14 years. The length of each time period is taken to be 2 years long and the discount rate is: $\alpha = 0.049$.

We should mention that in this example we adopted a discretization step of 1/15 in the $[0, 1]$ interval of the probability space describing the belief states. Thus, since the original state space has five possible outcomes, the total number of belief states in each stage is 650,630. For each belief state and for each uncertain transition matrix, Monte Carlo sampling was done by randomly generating transition matrices using Dirichlet distributions.

The results were computed using specialized GUI software for GPOMDP that we have developed. It can be noted (see Figs. 2 and 3) that the added complexity (that arises from considering the epistemic uncertainty) to the original dynamic programming algorithm is modest owing to the fast convergence of the solution when one increases the sampling size. For a sampling number greater than 20 the coefficient of variation of the estimated minimum expected cost is smaller than 0.0006 with a quasi-constant value of this cost (8024 units).

The time required for a sampling number equal to 20 (for which the COV was equal to 0.0006) was about 3 minutes on a core 2 duo Intel computer, when considering the uncertainty of the deterioration transition matrix.

6.1. Validation of the proposed methodology using a simulation analysis

Firstly, a simulation analysis is conducted in order to validate the expected costs calculated by our methodology against those obtained by simulation under the assumption that the transition matrices are uncertain. In a second stage, the simulated costs resulting from the use of the IM&R strategies prescribed by our model (i.e. by considering the epistemic uncertainties) and the ones resulting from the IM&R strategies prescribed by the classical POMDPs (i.e. without considering the epistemic uncertainties) are presented and discussed. The aim is to show the great interest of considering the epistemic uncertainties in the analysis. For this purpose, two sets of 100 simulations each (reducing thus the standard deviation by $1/\sqrt{100}$) were executed: In the first set, the IM&R strategy that is prescribed by the epistemic

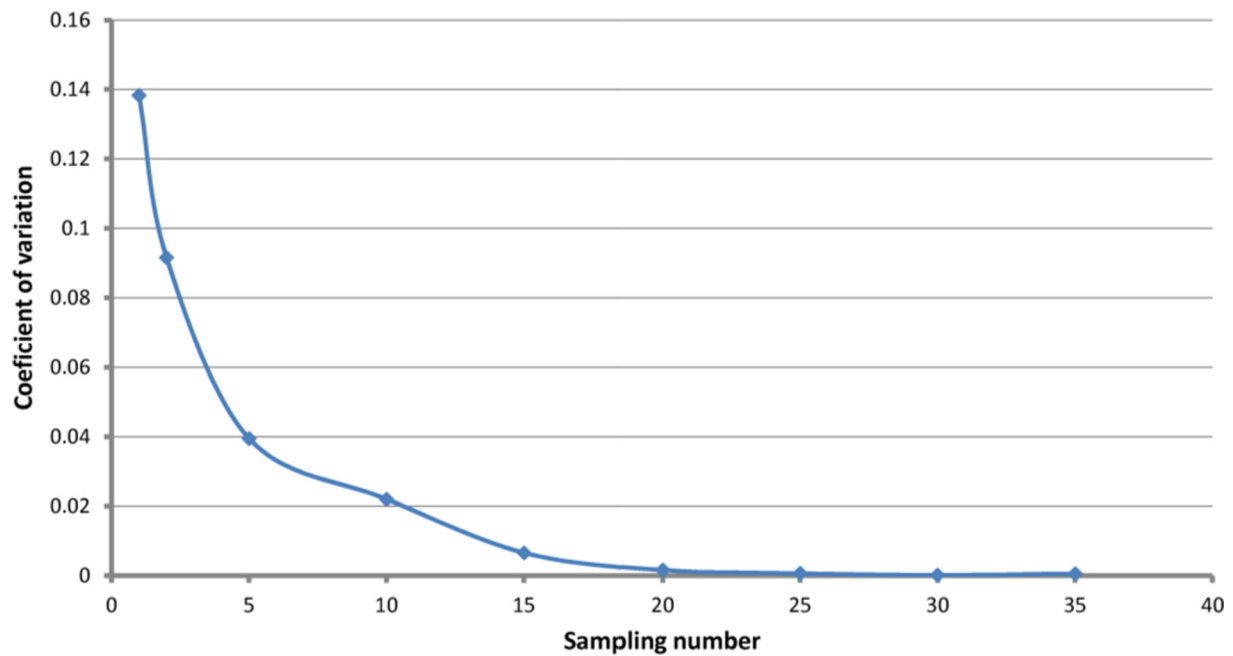


Fig. 2. Decrease of the coefficient of variation as the Monte Carlo sampling increases.

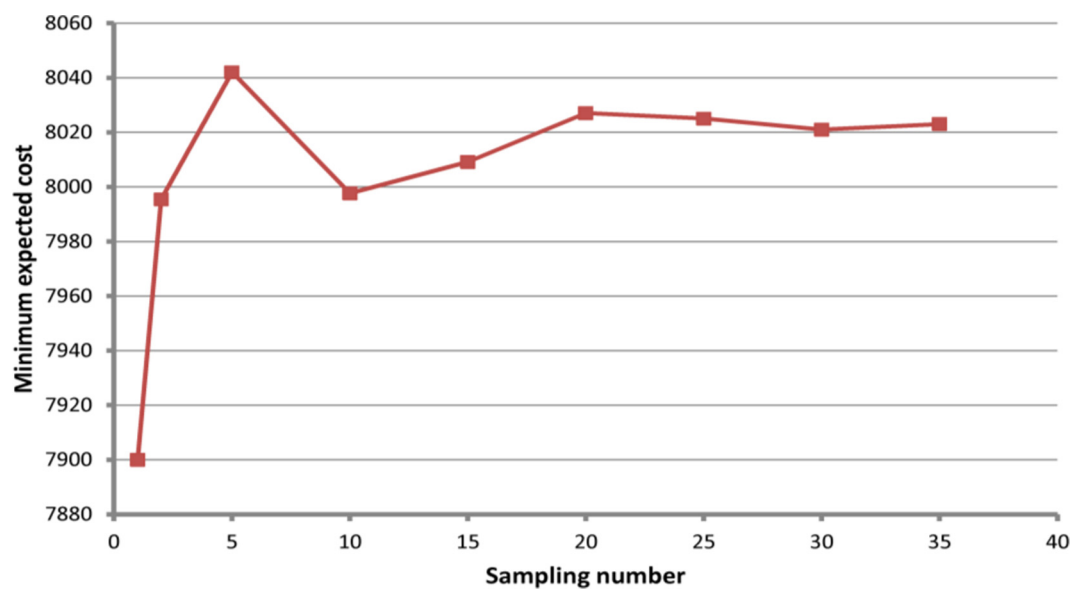


Fig. 3. Convergence of the solution as the Monte Carlo sampling increases.

uncertainty approach presented in this paper was applied; while, in the second set, the IM&R strategy that is prescribed by classical POMDP calculations (i.e. without epistemic uncertainty) was applied. An inspection-action scheme where the degradation process is subject to epistemic uncertainty (with an available sample size $n = 49$) was considered in the analysis. At the beginning of each time period, an adjusted degradation transition matrix was randomly generated from the Dirichlet matrix. Then, using the generated matrix, we randomly generate the evolution of the state of the deck by sampling the multinomial distributions which are the rows of the generated matrix. Using Bayes equation, the results of the inspections were also randomly generated based on the belief state of the deck and the uncertainty characterization of the inspection techniques (Table 2).

Although such a simulation is adequate for evaluating specified IM&R strategies, it is not feasible to use it in searching for opti-

mal strategies. This is due to the daunting computational complexity which may arise from such a brute force searching algorithm.

6.1.1. Validation in terms of the computed costs

By applying the IM&R strategy obtained using the approach presented in this paper (which takes into account the epistemic uncertainties), the simulation analysis has led to an average simulated cost of 6560 units and a standard deviation of 700 units. This cost is to be compared to the calculated expected cost (6380) which was obtained using the methodology presented in this paper. The little discrepancy between the two costs is clearly due to the statistical uncertainty related to the limited sample size. A perfect matching would require theoretically an infinite sample size. These findings confirm the accuracy of the results obtained by applying the proposed methodology.

Table 3
Number of structures in each state at two consecutive inspections.

	Sample size	θ_1^{n+1}	θ_2^{n+1}	θ_3^{n+1}	θ_4^{n+1}	θ_5^{n+1}
$a\theta_1^n$	49	5	3	2	1	1
	151	17	9	5	1	1
	1375	161	81	41	1	1
$a\theta_2^n$	49	0	5	3	2	1
	151	0	17	9	5	1
	1375	0	161	81	41	1
$a\theta_3^n$	49	0	0	5	3	2
	151	0	0	17	9	5
	1375	0	0	161	81	41
$a\theta_4^n$	49	0	0	0	6	3
	151	0	0	0	21	9
	1375	0	0	0	201	81
$a\theta_5^n$	49	0	0	0	0	7
	151	0	0	0	0	25
	1375	0	0	0	0	241

6.1.2. Efficiency of the IM&R strategies prescribed by the proposed methodology

The IM&R strategy prescribed by standard POMDPs (i.e. obtained without considering epistemic uncertainties) has led to an average simulated cost of 7340 units (the calculated expected cost using standard POMDP is 5410) and a standard deviation of 1100 units. The significant discrepancy between the simulated cost and the POMDP cost will be explained later in this section.

These results show that while algorithmic calculation of classical POMDPs gives costs that are lower than those given by our approach (5410 vs 6380), the application of IM&R strategies that were obtained using classical POMDP results in actual costs that are much higher than those incurred by applying the results prescribed by our approach (7340 vs 6560). Hence, one can conclude that ignoring epistemic uncertainty in the present case will result in a 11.89 percent increase in the actual IM&R costs.

It should be remembered here the fact that both simulations (using IM&R strategies obtained respectively by classical POMDPs and our approach) were conducted under the assumption that the true deterioration matrix is not known precisely (i.e. there is epistemic

uncertainty). This explains the significant bias between the simulated cost and the POMDP calculated cost for the strategy prescribed by the POMDP. The high cost resulting from applying the IM&R strategy prescribed by classical POMDPs is due to the discrepancy between the assumptions with the simulated reality since dynamic programming in classical POMDP assumes a precisely known deterioration matrix. Notice also that the standard deviation obtained by using the proposed approach is significantly smaller than that obtained by using POMDPs without considering epistemic uncertainties (700 vs 1100).

6.2. Sensitivity analysis

Calculations were made in this section in the cases where the Dirichlet distributions for the deterioration transition matrix were generated from a small size sample, medium size sample, and large size sample (Table 3). Optimal strategies and minimum total expected costs for action-alone scheme and inspection-action scheme (Table 4) are computed for these different sample sizes. Since we are presuming that the epistemic uncertainties emanates from the limited available sample size, we assume in this example the stationary uncertainty model. We are thus assuming that there is one true transition matrix, which we do not know exactly.

It can be noted (see Table 4) that the more the epistemic uncertainty is large (smaller sample size) the higher the minimum expected costs will be. Obviously, not considering epistemic uncertainty is equivalent to assuming an infinite sample size. Note that the costs (7300 units for the action alone scheme and 5410 units for the inspection-action scheme) for that case are the smallest ones. Moreover, the optimal maintenance actions and prescribed inspection techniques, for a given initial belief state, are different for each sample size.

Similar calculations are performed for the cases where the matrices related to the maintenance actions and inspection techniques were estimated from samples of varying size for the case of an action-inspection scheme (Fig. 4). From this figure, it can be seen that the additional expected cost (due to the joint consideration of epistemic uncertainty for maintenance actions, inspections and degradation) with respect to the case where the epistemic uncertainties are not present (i.e. the case of an infinite sample size) is lower than the sum of additional expected costs due to each of these epistemic uncertainties considered separately. For example, for the case of a sample of

Table 4
Expected costs as a function of sample size used to estimate the deterioration model.

	Scheme	Small sample size (49)	Medium sample size (151)	Large sample size (1375)	Non epistemic uncertainty
Time period 1 expected direct cost	Action alone	0	800	800	800
	Inspection-action	564	500	547	564
Time period 2 expected direct cost	Action alone	799	218	469	419
	Inspection-action	503	440	474	493
Time period 3 expected direct cost	Action alone	507	556	441	588
	Inspection-action	517	414	464	464
Time period 4 expected direct cost	Action alone	574	447	564	545
	Inspection-action	520	416	407	425
Time period 5 expected direct cost	Action alone	567	423	332	324
	Inspection-action	521	416	390	380
Time period 6 expected direct cost	Action alone	668	603	676	666
	Inspection-action	513	410	379	365
Time period 7 expected direct cost	Action alone	307	49	4	3
	Inspection-action	304	174	125	100
Total expected cost	Action alone	7996	7440	7309	7300
	Inspection-action	6380	5657	5498	5410
Optimal action during the first stage	Action alone	a_1	a_3	a_3	a_3
	Inspection-action	i_4	i_3	i_4	i_4

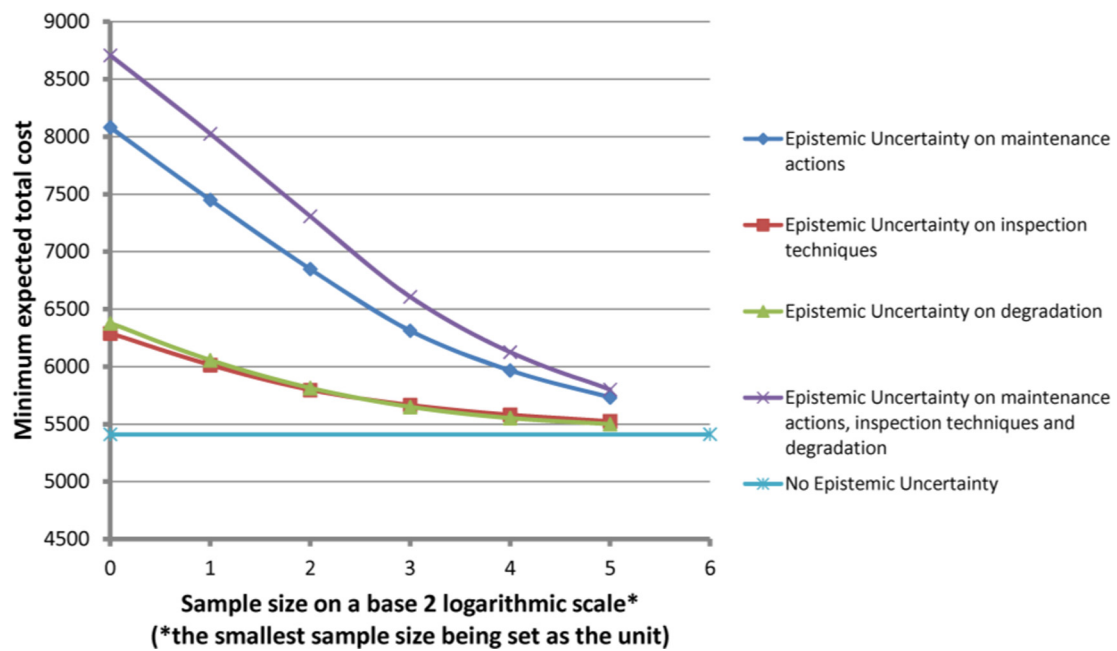


Fig. 4. Effect of sample size on the minimum expected cost.

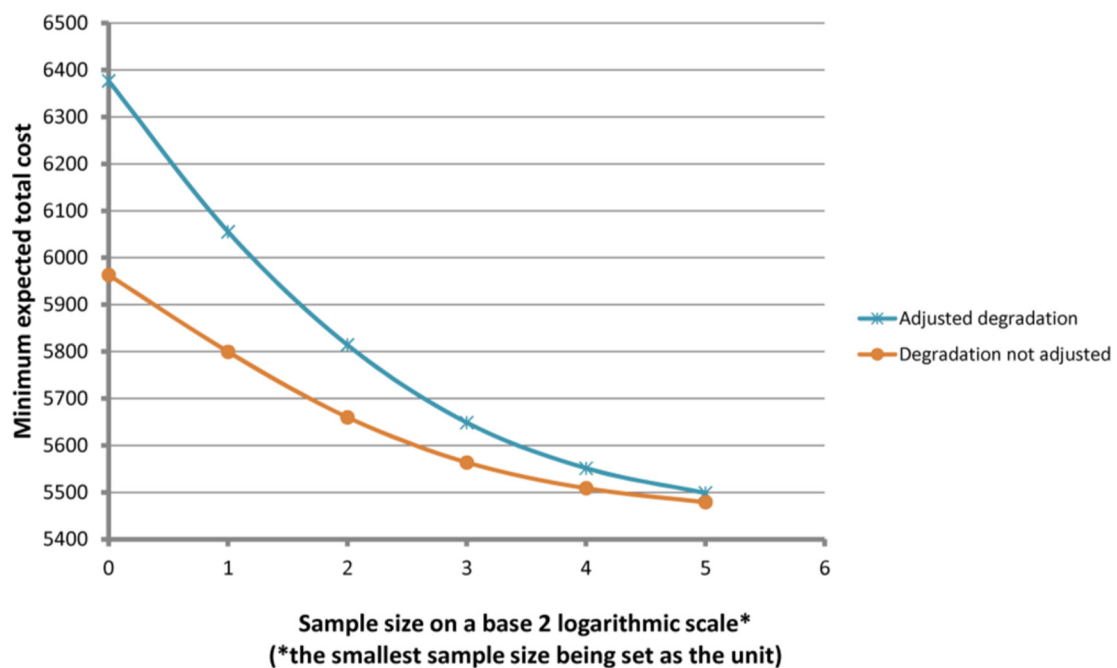


Fig. 5. Effect of sample size for degradation on the minimum expected cost.

unit size ($n = 49$ i.e. 1 on the \log_2 scale), the additional expected cost due to the joint consideration of epistemic uncertainty for maintenance actions, inspections and degradation is equal to $(8024 - 5410 = 2614$ units). This cost is lower than the sum of additional expected costs due to each of these epistemic uncertainties considered separately, i.e. $(7448 - 5410) + (6055 - 5410) + (6013 - 5410) = 3286$ units.

In Fig. 5, the results corresponding to the probabilistic epistemic uncertainty are presented for the cases where the degradation transition matrices are generated with and without taking into account the physical impossibility to have non-zero elements below the diagonal (cf. Section 4). As one would expect, using an adjusted deterioration transition matrix results in a higher total expected cost. Hence, the

probability distributions resulting from Dirichlet matrices (pseudo-counts) can and must be modified when additional information relative to the evolution process is available.

The preceding results (Table 4, Figs. 4 and 5) show the large influence of the epistemic uncertainties on the minimum expected total cost. This fact underlines the importance of considering such kind of information in the optimization process. Special care should therefore be given to identify the epistemic uncertainty model.

7. Conclusion

The use of Markov Decision Processes requires the estimation of transition matrices to describe the stochastic deterioration process,

the imperfect maintenance actions and the imperfect inspection techniques. Such estimations are affected by high level of uncertainties, due to the lack of data and to uncertain and fuzzy expert judgments. We proposed in this paper a methodology for including all available information related to the epistemic uncertainty shrouding the true values of the transition matrices in the decision process of a finite horizon MDP. The use of Dirichlet distributions to model the uncertainty of the estimated transition matrices was shown to be a natural choice which answers much of the drawbacks exhibited by the point estimation procedures found in the literature. We showed how the dynamic programming framework reduces drastically the added computational complexity that Monte Carlo simulations would bring to the problem. A simulation analysis was conducted in order to validate the expected costs calculated by our methodology against those obtained by simulation under the assumption that the transition matrices are uncertain. The simulated costs resulting from the use of the IM&R strategies prescribed by our model and the ones resulting from the IM&R strategies prescribed by the classical POMDPs (i.e. without considering the epistemic uncertainties) were presented and discussed. In Appendices A.1 and A.2, we treated the problems of data gathering using: (i) imperfect inspections; (ii) inspections at unequal time intervals. A numerical application is presented to highlight the impact of including epistemic uncertainties in POMDPs on the estimated expected costs, and naturally on the optimal strategies. It is shown that the use of all the information available to the manager (from historical data for example) will lead to more realistic expected life cycle costs, and will lead consequently to more appropriate decisions concerning the maintenance and/or the inspection of then asset.

Appendix A

The estimation of the transition matrices used by a POMDP for IM&R optimization of civil engineering structures usually relies on historical data gathered by the managers of these structures during periodic inspections or non-periodic inspections, such as those prescribed by a GPOMDP. Hence, these inspections are performed at time intervals having lengths which might differ from that of a single time period of the POMDP (as assumed in Section 4), and are in general imperfect, in the sense that given the true state of the structure, their results are characterized by probability distributions. In the following two subsections, these concerns are addressed.

A.1. Inspections at different time intervals

As stated in Section 4, if data is available on the true state of several structures at the beginning of two consecutive time periods, then this data is readily used as pseudo-counts which form the elements of the Dirichlet matrices (matrices having the parameters of Dirichlet distributions as rows). In Section 5, we stated that, in order to compute the expected value of $*c(v^{n+1})$, Monte Carlo sampling is convenient. In other terms, the Dirichlet matrices are first used to generate several transition matrices. Then, these transition matrices are used to compute several updated belief states v^{n+1} in order to evaluate the mean value of the associated cost. If the data is available on the true state of several structures at the beginning of two non-consecutive time periods which are p time periods apart; then, this data can be used to determine a Dirichlet matrix which can be used to generate several transition matrices for use in Monte Carlo sampling. However, these matrices describe the evolution of the system during a time length of p periods, and thus, cannot be used directly in our model. A direct consequence of Chapman–Kolmogorov equations is that the transition matrix of a Markov process for a time length of p periods is simply the matrix product of p transition matrices for one time period. Hence, after being generated from the Dirichlet matrix,

the p_{th} roots of each matrix are calculated in order to use them in our model.

In practice, the available data originates from inspections performed at varying time interval. In such case, the matrices intended to be used for Monte-Carlo sampling have to be generated from a mixture of Dirichlet matrices. The weight assigned to each of these Dirichlet matrices in such a mixture can be chosen to be proportional to the relative size of the collection of the associated inspection results.

A.2. Imperfect inspections

The data used by the manager of a civil engineering structure, usually originates from imperfect inspection techniques. The uncertainty pertaining to the results of these inspections are usually characterized by conditional probability distributions. Using Bayes equation, we can derive the conditional probability distribution $\Pr[\theta_k^n | r, i]$ of the state of the structure given the inspection technique i and the result r . Hence, if we ignore completely the state of the structure at the time of the first inspection, getting a result p from an inspection technique i_1 will allow us to calculate an exactly defined belief state $v_{i_1}^p$. As for the belief state vector of the structure at the next inspection, two cases must be considered:

- 1 A prior model of the degradation process is available to the manager (this model can either take the form of a transition matrix or be a continuous time deterioration model which can be translated into a transition matrix (Mishalani & Madanat, 2002)). In this case, the belief state $v_{i_2}^q$ of the structure after the second inspection i_2 will be a Bayesian update of the belief state vector $(v_{i_1}^p \times M)$ based on the result q of the inspection i_2 , where M is the deterioration transition matrix.
- 2 Non prior model of degradation is available and the manager is building his first Dirichlet matrix, or he is simply adding pseudo-counts to an already existing Dirichlet matrix; i.e. he is updating his belief about the probability distributions of the multinomial parameters present in the transition matrix in a Bayesian fashion since the rows of the Dirichlet matrix are the natural conjugates of the multinomial distributions in the transition matrix. In this case, $v_{i_2}^q$ will be calculated as $v_{i_1}^p$; i.e. calculated solely based on the result q of the inspection i_2 .

Consequently, the probability that a certain structure has evolved from a true state θ_h to a true state θ_l between two consecutive inspections i_1 and i_2 which yielded the results p and q respectively, will be $P = v_{i_1}^p[h] \times v_{i_2}^q[l]$.

Therefore, if the numbers of structures o_{ij} (inspected by inspection i_1 , then by inspection i_2 at a later stage), which were observed in state r_i and evolved to the observed state r_j , are organized in a square matrix O , the matrix S of Section 4 can be estimated using the following relation for each of its elements:

$$s_{hl} = \sum_p \sum_q o_{pq} \times v_{i_1}^p[h] \times v_{i_2}^q[l].$$

References

- Avrachenkov, K. E., Filar, J. A., & Haviv, M. (2001). A survey on singular perturbations of Markov chains and decision processes. In E. Feinberg, & A. Shwartz (Eds.), *Handbook of Markov decision processes: Methods and applications* (pp. 113–150). Netherlands: Kluwer.
- Bagnell, J., Ng, A., & Schneider, J. (2009). Solving uncertain Markov decision problems. Technical report CMU-RI-TR-01-25. Robotics Institute, Carnegie Mellon University.
- Bellman, R. (1961). *Adaptive control processes: A guided tour*. Princeton (N.J.): Princeton University Press.
- Congressional Budget Office (CBO)(1999). *Trends in public infrastructure spending*. Washington, DC: CBO.
- Corotis, R. B. (2009). Risk communication with generalized uncertainty and linguistics. *Structural Safety*, 31(2), 113–117.

- Corotis, R. B., Ellis, J. H., & Jiang, M. (2005). Modeling of risk-based inspection, maintenance and life-cycle cost with partially observable Markov decision process. *Structure and Infrastructure Engineering*, 1, 75–84.
- Cranahan, J. V. (1988). Analytical framework for optimizing pavement maintenance. *Journal of Transportation Engineering*, 114(3), 307–322.
- Delage, E., & Mannor, S. (2009). Percentile optimization for Markov decision processes with parameter uncertainty. *Operations Research*, 58, 203–213.
- Der Kiureghian, A., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter?. *Structural Safety*, 31(2), 105–112.
- DeStefano, P. D., & Grivas, D. A. (1998). A method for estimating transition probability in bridge deterioration models. *ASCE Journal of Infrastructure Systems*, 4(2), 56–62.
- Droesbeke, J. J., Fine, J., & Saporta, G. (2002). Méthodes bayésiennes en statistique 331–348. Paris: Société Française De Statistique, Editions Technip.
- Durango, P., & Madanat, S. (2008). Optimization of inspection and maintenance decisions for infrastructure facilities under performance model uncertainty: A quasi-Bayes approach. *Transportation Research, Part A*, 42, 1074–1085.
- Eckels, J. E. (1968). Optimum maintenance with incomplete information. *Operations Research*, 16, 1058–1067.
- Faddoul, R., Raphael, W., & Chateaufort, A. (2009). A Generalized Partially Observable Markov Decision Process updated by decision trees for maintenance optimization. *Structure and Infrastructure Engineering*, 5.
- Givan, R., Leach, S. M., & Dean, T. (2000). Bounded-parameter Markov decision processes. *Artificial Intelligence*, 122(1–2), 71–109.
- Hofer, E. (1996). When to separate uncertainties and when not to separate. *Reliability Engineering System Safety*, 54, 113–118.
- Iyengar, G. (2005). Robust dynamic programming. *Mathematics of Operations Research*, 30(2), 1–21.
- Kuhn, K., & Madanat, S. (2006). Robust maintenance policies for Markovian systems under model uncertainty. *Computer-Aided Civil and Infrastructure Engineering*, 21(3), 171–178.
- Kallen M. J. (2007). *Markov processes for maintenance optimization of civil infrastructure in the Netherlands* (Ph.D. thesis). Delft: Delft University of Technology.
- Madanat, S. M., Mishalani, R. C., & Wan Ibrahim, W. H. (1995). Estimation of infrastructure transition probabilities from condition rating data. *ASCE Journal of Infrastructure Systems*, 1(2), 120–125.
- Minka T. (2003). Estimating a Dirichlet distribution <http://research.microsoft.com/en-us/um/people/minka/papers/dirichlet/minka-dirichlet.pdf>.
- Mishalani, R. G., & Madanat, S. M. (2002). Computation of infrastructure transition probabilities using stochastic duration models. *ASCE Journal of Infrastructure Systems*, 8(4), 139–148.
- Monahan, G. E. (1982). A survey of partially observable Markov decision processes; theory, models, and algorithms. *Management Science*, 28, 1–16.
- Neves, L. A. C., Frangopol, D. M., & Cruz, P. J. S. (2006). Probabilistic life-time oriented multiobjective optimization of bridge maintenance: Single maintenance type. *Journal of Structural Engineering*, ASCE.
- Nilim, A., & El Ghaoui, L. (2005). Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, 53(5), 780–798.
- Orcesi, A., & Cremona, C. (2010). A bridge network maintenance framework for Pareto optimization of stakeholders users costs. *Reliability Engineering and System Safety*, 95, 1230–1243.
- Pham, H., & Wang, H. (1996). Imperfect maintenance. *European Journal of Operational Research*, 94(3), 425–438.
- Revelle, C., Whitlatch, E., & Wright, J. (1997). Civil and environmental systems engineering. Edgewood Cliffs, NJ: Prentice Hall.
- Robelin, C. A., & Madanat, S. (2007). History-dependent optimization of bridge maintenance and replacement decisions using Markov decision processes. *Journal of Infrastructure Systems*, 13, 195–201.
- Satia, J. K., & Lave, R. L. (1973). Markov decision processes with uncertain transition probabilities. *Operations Research*, 21(3), 728–740.
- Wang, H. (2002). A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3), 469–489.
- White, C. C., & Eldeib, H. K. (1994). Markov decision processes with imprecise transition probabilities. *Operations Research*, 42(4), 739–749.
- Wicker, N., Muller, J., Kalathur, R. K. R., & Poch, O. (2008). A maximum likelihood approximation method for Dirichlet's parameter estimation. *Computational Statistics & Data Analysis*, 52(3), 1315–1322.