

Incorporating Bayesian Networks in Markov Decision Processes

R. Faddoul, Ph.D.¹; W. Raphael²; A.-H. Soubra³; and A. Chateauneuf⁴

Abstract: This paper presents an extension to a partially observable Markov decision process so that its solution can take into account, at the beginning of the planning, the possible availability of free information in future time periods. It is assumed that such information has a Bayesian network structure. The proposed approach requires a smaller computational effort than the classical approaches used to solve dynamic Bayesian networks. Furthermore, it allows the user to (1) take advantage of prior probability distributions of relevant random variables that do not necessarily have a direct causal relationship with the state of the system; and (2) rationally take into account the effects of accidental or rare events (such as seismic activities) that may occur during future time periods of the planning horizon. The methodology is illustrated through an example problem that concerns the optimization of inspection, maintenance, and rehabilitation strategies of road pavement over a 14-year planning horizon. DOI: [10.1061/\(ASCE\)IS.1943-555X.0000134](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000134). © 2013 American Society of Civil Engineers.

CE Database subject headings: Optimization; Markov process; Maintenance; Decision making; Bayesian analysis; Infrastructure.

Author keywords: Optimization; Markov process; Maintenance; Decision making; Bayesian analysis; Infrastructure.

Introduction

The life-cycle cost of civil engineering assets incorporates their design and construction costs, their inspection maintenance and rehabilitation (IMR) costs during their projected life cycle, and eventually, the costs necessary for their decommissioning. However, for optimal decisions, one must include the users' costs in the IMR costs. The users' costs usually involve those incurred by the users as a result of an imperfect performance of the asset and those related to the failure risk of the asset, which are expressed in monetary units. Hence, an accurate evaluation of the optimal IMR costs for civil engineering assets is essential for the correct estimation of the entire life-cycle costs of these assets. As such, the accurate evaluation of IMR costs is vital not only for the management optimization of existing assets, but also for a rational choice of the type and characteristics of new assets.

Designing a mathematical model for a real life problem usually entails the need to pursue two conflicting objectives: (1) fidelity to reality in the sense that the mathematical model must include, as much as possible, all of the variables that are relevant to the real problem and that the mathematical structure of the model must relate these variables to each other in a way that emulate the real life problem; (2) efficiency and tractability of the proposed

mathematical model in the sense that the data needed as input for the model must be easily accessible and the solution method of the transcribed mathematical problem must be known and its computation time must not be too expensive. This paper proposes an extension of the partially observable Markov decision process (POMDP) models used for the IMR optimization of civil engineering structures, so that they will be able to take into account the possibility of free information that might be available during each of the future time periods. It is supposed that such information has a Bayesian network (BN) structure. The proposed model can be viewed as a dynamic Bayesian network (DBN) model, in which the decision variables can be optimized by dynamic programming and decision trees. As such, the proposed approach is able to tackle the problem of computational complexity that existing DBN methodologies stumble upon.

Literature Review

After it was proposed by Bellman (1961), dynamic programming (DP) was readily adopted as an efficient and intuitive algorithmic framework to solve for optimal strategies in sequential decision problems. In these problems, the state of the system can be chosen so that it possesses the Markov property; i.e., knowing the present state of the system is all that is needed to optimally decide for future strategy, regardless of the history of the system. During the last decades, several extensions and generalizations of the basic DP algorithms (deterministic and probabilistic DPs, as defined by Bellman) were proposed to model real life problems. In the early 1960s, POMDPs were introduced (Eckels 1968; Monahan 1982). In a POMDP, the system has an uncertain state (belief state), which can be, for example, a probability distribution of the state variable θ . Methodologies have been suggested for including in a Markov decision process (MDP) epistemic uncertainty (Faddoul et al. 2009a), non-Markovian effects of actions and/or deterioration processes (Robelin and Madanat 2007), and the effects of resource constraints in the case of several simultaneous POMDPs (Robelin and Madanat 2007; Faddoul et al. 2010, 2013). Moreover, several

¹Saint-Joseph Univ., Mar Roukos, PoB. 11-154, Riad El Solh, Beyrouth, Lebanon. E-mail: faddoul_rific@yahoo.com

²Head of the Civil and Environmental Engineering Dept., ESIB, Saint-Joseph Univ., Mar Roukos, PoB. 11-154, Riad El Solh, Beyrouth, Lebanon (corresponding author). E-mail: wassim.rafael@fi.usj.edu.lb

³Professor, Univ. of Nantes, Bd. de l'université-BP 152, 44603 St-Nazaire, Cedex, France. E-mail: Abed.Soubra@univ-nantes.fr

⁴Professor, LaMI, Polytech Clermont-Ferrand, BP 206, 63174 Aubière, Cedex, France. E-mail: alaa.chateauneuf@cust.univ-bpclermont.fr

Note. This manuscript was submitted on February 1, 2012; approved on October 12, 2012; published online on October 13, 2012. Discussion period open until May 1, 2014; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Infrastructure Systems*, Vol. 19, No. 4, December 1, 2013. © ASCE, ISSN 1076-0342/2013/4-415-424/\$25.00.

investigations that extend POMDPs to include inspection planning were proposed in the literature during the last decade (Madanat and Ben-Akiva 1994; Corotis et al. 2005; Faddoul et al. 2011, 2009b; Frangopol et al. 2012). In these models, the inspection costs motivated the inclusion of inspection decisions in the optimization pertaining to sequential decision making of the IMR of civil engineering structures. If the state of a system is described by a variable θ that cannot be observed freely and unerringly at the beginning of each time period, the previously mentioned methodologies will allow an optimum choice of one or several inspection techniques during each time period of the planning horizon. The common purpose to all of the previously mentioned extensions to the original DP formulation was to include as much of the relevant information to the problem as possible to decrease the bias and the variability of the obtained optimal solutions with respect to the true unknown solution. It is in this line of reasoning that this paper proposes a methodology (which can be viewed as a DBN) that allows a more accurate (although computationally tractable) stochastic degradation modeling in POMDPs.

DBNs are a special class of Bayesian networks that can be used for modeling time series data and represent stochastic processes. They consist of a sequence of time slices (which are often repetitive). Each of these slices consists of one or more BN nodes connected by directed edges, and thus, can be considered to be a primal BN. Several authors have investigated the use of BNs and DBNs for maintenance planning and contrasted them to the use of other deterioration models (Celeux et al. 2006; Francois et al. 2008; Jones et al. 2010; Straub 2009). Francois et al. (2008) presented a DBN modeling for the IMR of a railway in which the BN structure was used to perform Monte Carlo simulations to choose the optimal IMR parameters. Straub (2009) proposed a DBN for modeling the stochastic deterioration process. This study provided an efficient (polynomial in the number of nodes) algorithm to update the model parameters based on available evidence from inspections; however, the problem of efficiently searching for optimal values for the decision nodes was not addressed. Also, this model did not allow the user to optimally choose one (or more) optimal inspection techniques between several available techniques. Attoh-Okine and Bowers (2006) used a BN to model bridge deterioration. Their model allowed for the conditional computation of the deterioration of bridge members based on the state of other elements. Manoj (2009) suggested the use of BNs to predict the probability of terrorist attacks on critical transportation infrastructure facilities. Langseth and Portinale (2007) discussed the very important concerns related to the actual process of building the necessary BNs (such from expert elicitation and/or the use of empirical data) in the context of reliability and deterioration modeling. These issues were tackled, in more general settings, in the studies by Cooper and Herskovits (1992) and Heckerman et al. (1995). These subjects will not be addressed in this paper.

Cooper (1988) presented a general method for using BNs as decision networks; the complexity of the algorithm is exponential in the number of decision nodes.

On the other hand, Cooper (1990) showed that the probabilistic inference by using general BN is NP-hard and concluded that research should be directed toward the design of efficient special case algorithms that build on the special characteristics of each particular BN. Numerous and more efficient special case algorithms were proposed for particular settings (Zhang et al. 1992, 1994; Jensen et al. 1994; Zhang 1998).

To summarize, except for a few efficient algorithms for BN/DBN decision problems with special structures, the computational complexity of the classical algorithms used to solve decision problems modeled by BN/DBN arises from: (1) the inference step (for a

specific set of values taken by the decision variables), which was shown by Cooper (1990) to be NP-hard; and (2) the optimization step, which is generally exponential in the number of decision nodes (Cooper 1988).

This paper presents a DBN for the IMR of civil engineering structures. In this model, the variables related to the decision nodes can be optimized by DP and decision trees. The decision variables concern the maintenance actions and types of inspections during each time period. The proposed model extends the classical POMDPs and the generalized partially observable Markov decision process (GPOMDP) (Faddoul et al. 2011, 2009b) so that it can take into account free available information during future time periods. It improves the existing IMR models using DBN (Cooper 1988, 1990; Zhang et al. 1992, 1994; Jensen et al. 1994; Zhang 1998) in that: (1) it builds on the efficiency of DP to tackle the problem of exponential complexity connected with the number of decision variables; and (2) it divides the original DBN into smaller BNs to tackle the NP-hardness of the problem.

Although the methodology proposed in this paper is motivated by the IMR optimization of civil engineering infrastructures and is illustrated for a GPOMDP, its applicability is quite general and its use in any classical POMDP is straightforward.

The next sections first formulate the problem at hand as a DBN and present the DP and the BN methodologies. This is followed by a combined approach that introduces the BNs in POMDPs. Next, a numerical example that involves the IMR of a road pavement is presented and discussed. Finally, the paper concludes with a discussion on the computational aspect of the method.

Problem Formulation as a DBN

One may consider that the IMR manager of a civil engineering asset had run a POMDP algorithm at the beginning of the planning horizon and that they are complying with the decisions prescribed by that algorithm. At the beginning of each time period, the belief state of the asset is known to the manager. This belief state is the outcome of the deterioration model, which is possibly updated via Bayesian techniques by using the results of planned inspections. One may suppose further that at the beginning of a time period n , free information relevant to the state of the asset is available to that manager. This information will be generally uncertain. If one assumes that this uncertainty can be expressed by a probabilistic distribution, an expected attitude of the structure manager will be to enhance the optimality of their decisions by updating the belief state of the structure during the time period n by using Bayesian techniques. However, this paper will show that the optimality of the manager's decisions can be enhanced if the possibility of using such free information was planned in advance at the beginning of the planning horizon by using prior probability distributions of the observable nodes of the BN.

Modeling the joint probability distribution of the freely observable random variables as a BN is particularly useful because (1) it allows one to easily infer the state of the structure by collecting information concerning the realization of certain relevant random variables, even if these variables do not have a causal relationship with the state of the structure (technique of information back-propagation); and (2) it allows one to rationally take into account, at the beginning of the planning horizon, the effects of accidental events such as seismic activities or terrorist attacks (Manoj 2009). The present methodology does not merely use a BN to construct the deterioration transition matrix. Instead, the actual BN is inserted into a POMDP to form a DBN that can be solved by DP and decision tree analysis.

Dynamic Programming

In a POMDP (Eckels 1968; Monahan 1982), the state of the system at the beginning of each time period cannot be fully observed. The manager of the system must rely on the characterization of a partially observed state; i.e., a belief state, which is usually described by probability 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 v_i^n ($i = 1, \dots, k$) are the probabilities associated with the different states θ_i , 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 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_j^n to the state θ_i^n if the maintenance action a is implemented 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_j^n to the state θ_i^{n+1} as a result of the degradation process.

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

As it is well known, the solution of a finite time horizon POMDP by DP consists of recursively calculating 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 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 the action a was applied while in the belief state v^n at the beginning of stage n :

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

Hence, in classical POMDPs used for IMR optimization of civil engineering structures, the recursive relation consists of minimizing the expected cost by choosing the appropriate maintenance action. Recently, a more general POMDP (called GPOMDP) was presented by Faddoul et al. (2011, 2009a, b). In a GPOMDP, the structure manager has the opportunity, at the beginning of each time period, to decide for an optimal sequence of decisions to be implemented during that stage. This sequence of decisions usually consists of one or several inspections and/or actions, applied sequentially. For example, a sequence of decisions consisting of two inspection decisions followed by one action decision is suitable for maintenance problems when a more precise and costly inspection is implemented on the basis of the results of a relatively cheap inspection. It is also suitable when specialized inspection technologies in detecting some of the states of the structure are implemented on the basis of the results given by an inspection technology, which is efficient over the entire state space of the structure. In the GPOMDP by Faddoul et al. (2011, 2009a), the optimal planning of one or a sequence of imperfect inspections and/or maintenance actions is possible. The inspections are imperfect in the sense that, given the true state θ^n of the system and the inspection technology i , their results will be characterized by conditional probability

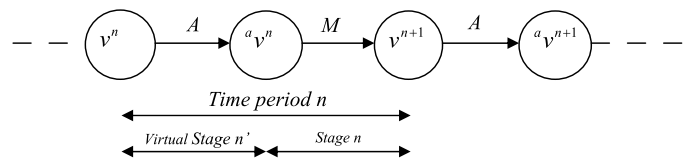


Fig. 1. Classical POMDP

distributions ($\Pr[r_1|\theta^n, i], \Pr[r_2|\theta^n, i], \dots, \Pr[r_m|\theta^n, i]$). Finally, in the GPOMDP, a decision tree is used as the recursive relation required by the POMDP (Fig. S2). The decision tree applied to each belief state at the beginning of each time period n can be considered to be a function of two variables: (1) the belief state v^n ; (2) the optimal discounted costs for all belief states at the beginning of time period $n+1$, i.e., $\{{}^*c(v^{n+1})v^{n+1}\}$. Because the decision tree, given v^n and $\{{}^*c(v^{n+1})v^{n+1}\}$, will give an optimal expected cost and an optimal sequence of decisions concerning inspection and maintenance types to be applied, it is used in GPOMDP as the recursive relation required by DP. For a classical POMDP, a GPOMDP makes use of the deterioration transition matrix M to calculate the belief state v^{n+1} given the belief state $^a v^n$. It is assumed that the effect of a maintenance action is immediate and takes place at the beginning of the time period in which the decision was made. Concerning the effect of the deterioration process, it is assumed that its effects take place at the end of each time period. However, the proposed methodology can be applied straightforwardly with minor modifications to other schemes; for example, in which the effects of maintenance actions and deteriorations take place at the end of each time period. In view of this, given that the maintenance action a was applied at the beginning of time period n , the belief state vector $^a v^n$ can be thought of as the available belief state at the end of a virtual stage n' of time length 0 (i.e., at the beginning of stage n') (Fig. 1).

During the stage n , one can view the effect of the deterioration process as that of an action for which there is no alternative and that brings the structure from the belief state $^a v^n$ to the belief state v^{n+1} at the beginning of time period $n+1$ [i.e., at the beginning of the virtual stage $(n+1)'$]. Hence, a POMDP (or GPOMDP) having a planning horizon with N time periods will be viewed as one with a planning horizon of $2N$ time periods with a virtual stage at the beginning of each time period. Although such a point of view can be considered superfluous in describing a classical POMDP or a GPOMDP, it will be shown in a later section that it is convenient for integrating BNs in a POMDP framework. In the following sections, virtual stages (n') will be referred to as "stage n' " and actual stages (n) as "time period n ".

Bayesian Networks

A BN, also called a belief net, is a directed acyclic graph in which with the nodes are random variables and the directed edges indicate conditional probabilistic dependence between children nodes and their parents. A conditional probability distribution giving the probabilities of the random variable, in each particular state for each possible combination of the parent nodes, is associated with each of the child nodes of the graph. Unconditional probability distributions are associated with each of the root nodes (i.e., nodes without parents). The conditional and the unconditional probability distributions can be chosen to be deterministic; i.e., they assign all of the probability mass to a single value. For the current purpose, BNs that have decision nodes will not be considered; i.e., nodes for which the associated variables can be fixed by a decision from

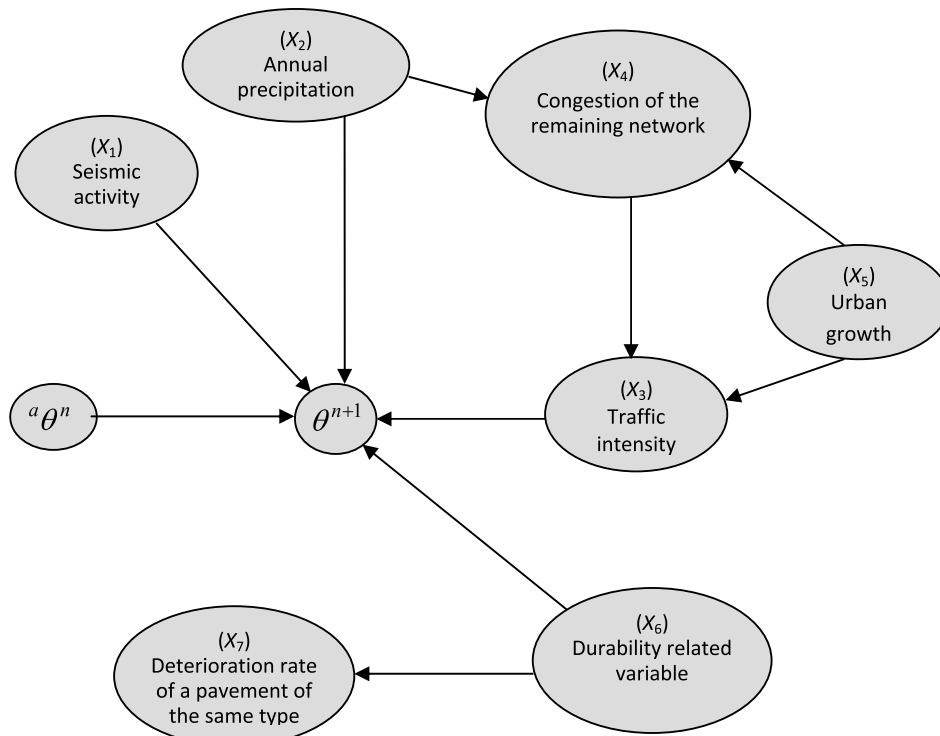


Fig. 2. Example of a BN for pavement deterioration

the manager of the structure. Fig. 2 illustrates a simple Bayesian network with nine nodes describing the causal influence of various relevant variables on the state θ^{n+1} of a road pavement at the beginning of time period $n + 1$.

The state variable $a\theta^n$ is the state of the pavement after the application of the action a during the time period n . The structure of a BN represents a kind of knowledge that one may have about the joint probability distribution of the random variables represented by the nodes. The lack of arcs represents conditional independence assumptions among the random variables represented by the nodes. Stated in other terms, a node is independent of its nondescendants given its parents; more generally, two disjoint sets of Nodes A and B are conditionally independent given C if C separates A and B. That is, if along every undirected path between a node in A and a node in B there is a Node d such that: (1) d has converging arrows and neither d nor its descendants are in C; or (2) d does not have a converging arrow and d is in C (Pearl 1986). For example, the joint probability distribution of discrete random variables associated with the BN of Fig. 2 will be written according to the chain rule of factorization (if the structure of the BN is disregarded), as follows:

$$\begin{aligned}
 P(\theta^{n+1}, a\theta^n, X_1, X_2, X_3, X_4, X_5, X_6, X_7) \\
 = P(\theta^{n+1} | a\theta^n, X_1, X_2, X_3, X_4, X_5, X_6, X_7) \\
 \times P(a\theta^n | X_1, X_2, X_3, X_4, X_5, X_6, X_7) \\
 \times P(X_1 | X_2, X_3, X_4, X_5, X_6, X_7) \times P(X_3 | X_2, X_4, X_5, X_6, X_7) \\
 \times P(X_4 | X_2, X_5, X_6, X_7) \times P(X_2 | X_5, X_6, X_7) \\
 \times P(X_5 | X_6, X_7) \times P(X_7 | X_6) \times P(X_6)
 \end{aligned} \quad (2)$$

By taking into account the structure of the BN in Fig. 2, the joint probability distribution can be simplified to the following form:

$$\begin{aligned}
 P(\theta^{n+1}, a\theta^n, X_1, X_2, X_3, X_4, X_5, X_6, X_7) \\
 = P(\theta^{n+1} | a\theta^n, X_1, X_2, X_3, X_6) \times P(a\theta^n) \times P(X_1) \\
 \times P(X_3 | X_4, X_5) \times P(X_4 | X_2, X_5) \times P(X_2) \times P(X_5) \\
 \times P(X_7 | X_6) \times P(X_6)
 \end{aligned} \quad (3)$$

If one or more of the variables of the BN are observed; i.e., if the probability distributions of those variables are set to be the deterministic distributions assigning all of the probability to single observed values, the probability distributions of the remaining variables are updated accordingly. More specifically, for this particular example, if one is given the probability distribution of the node $a\theta^n$ (i.e., the belief state ν^{n+1}), then for each combination of the values x_1, x_2, x_5 , and x_6 taken by the nodes X_1, X_2, X_5 , and X_6 , respectively, there will be a different belief state ν^{n+1} , which will be calculated by using one of the various available inference techniques used for BNs. Eq. (4) gives the standard inference technique by simply using the Bayes formula and marginalization (Pearl 1986) for the case in which the variables X_1, \dots, X_i are observed and the variables X_{i+1}, \dots, X_n are not observed:

$$P(\theta^{n+1} | a\theta^n, X_1, \dots, X_i) = \frac{P(\theta^{n+1}, a\theta^n, X_1, \dots, X_i)}{P(a\theta^n, X_1, \dots, X_i)} \quad (4)$$

where

$$P(\theta^{n+1}, a\theta^n, X_1, \dots, X_i) = \sum_{X_{i+1}} \dots \sum_{X_n} P(\theta^{n+1}, a\theta^n, X_1, \dots, X_n)$$

and

$$P(a\theta^n, X_1, \dots, X_i) = \sum_{\theta^{n+1}} \sum_{X_{i+1}} \dots \sum_{X_n} P(\theta^{n+1}, a\theta^n, X_1, \dots, X_n)$$

Because the random variables X_1 , X_2 , X_5 , and X_6 are independent, the probability P of the events x_1 , x_2 , x_5 , and x_6 occurring simultaneously will be

$$P(x_1, x_2, x_5, x_6) = P_1(x_1) \times P_2(x_2) \times P_5(x_5) \times P_6(x_6) \quad (5)$$

where $P_1(x)$, $P_2(x)$, and $P_5(x)$ = unconditional probability distributions associated with the root nodes; and $P_6(x)$ = prior probability distribution associated with the variable X_6 , which quantifies the capacity of the type of structures to withstand the effects of degradation. The variable X_6 , the durability related variable is a variable that is not directly observable and has an impact on the deterioration rate of the pavement. Such a variable can be, for example, an uncertain constant in the physical deterioration model from which the Bayesian network and/or the deterioration transition matrix are derived. The usefulness of X_6 and X_7 stems from the fact that they allow one to benefit from costly inspections conducted for other pavement sections and result in an observation of the deterioration rate, and consequently, in an updated probability distribution of X_6 .

Therefore, if one is given the belief state ν^n , then for each belief state ν^{n+1} produced by each particular combination of the values x_1 , x_2 , x_5 , and x_6 (used by the nodes X_1 , X_2 , X_5 , and X_6 , respectively), there will be an associated probability P . The same reasoning is valid for the case in which some or all of the variables included in the BN are continuous. In this case, one simply has to replace probabilities with densities. Hence, if a probability distribution is available for each of the evidence nodes and for θ^n , then the variability of θ^{n+1} will be described by a set of probability distributions (which is finite and discrete if all of the evidence variables are finite and discrete). In other words, based on prior probability distributions for the evidence nodes, the BN will generate a set of belief states ν^{n+1} according to a probability distribution, which can be calculated by using Eq. (5).

The variable X_4 measures the traffic congestion of the remaining network linking the same locations as the road containing the pavement under study. The future prior probability distributions of this variable for each time period are assumed to be made available by elicitation of the network expert. This elicitation is supposed to take into consideration, among others, the expected IMR activity on the network. Such an approach constitutes an approximation of the real problem of interdependencies (economics, logistics, and traffic disturbances) that arise in the IMR optimization of a network.

The BN that was analyzed in this section may be extended and fine-tuned. For example, to account for the random variability of the seismic demand on different components of the transportation system (e.g., the remaining network) that is not in the same location, one must (1) insert a node (representing the seismic demand on the pavement subject to IMR) in the path linking X_1 to θ^{n+1} ; and (2) add a path to the graph that links X_1 to X_4 . This path would have a node between X_1 and X_4 that represents the seismic demand on the parallel road for example (Fig. S1). The two added demand nodes are considered to be evidence nodes in the sense that their prior probability distribution is replaced by observed values when time period n is reached.

There may exist causality dependences between the nodes that may violate the assumptions of the BN. Many of the acyclicity violation cases may be addressed by adding hidden variables (nodes) to the network. For example, a learning algorithm may find a causal relationship, direct or indirect, from Node A to Node B and from Node B to Node A. This is often because the learning algorithm has overlooked one or more important variables that are influencing both Nodes A and B. However, there may exist cases in which the causal relationship is truly cyclic (Sprites et al. 1993; Sprites 1995;

Pearl and Dechter 1996). For example, urban growth can affect traffic intensity; inversely, traffic intensity can affect urban growth, hence violating the acyclicity assumption. Although less popular and more complex than BNs, learning and inference algorithms for probabilistic models represented as directed cyclic graphs (DCG) are becoming more available.

Finally, the BN parameters can be usually determined (1) from data by using available learning algorithms or (2) by using a direct knowledge elicitation method from experts.

Inclusion of Bayesian Networks in POMDPs

A classical POMDP (or GPOMDP) that has a stochastic degradation process behaves during each time period n like a classical deterministic MDP, although the first works over belief states whereas the second works over exact true states. In such a model, the result of degradation (which is probabilistic in the original state space) is unique and well defined in the belief state space (i.e., space of probability distributions). This is because the outcomes are belief states that are exactly defined and expressed by probability vectors. At the beginning of each time period n , a classical POMDP recursive relationship (i.e., the decision tree procedures in a GPOMDP) will call for optimal costs $^*c(\nu^{n+1})$ calculated for the updated belief states at the beginning of time period $n + 1$.

The extension of the classical POMDP model to the case of a probabilistic POMDP (so that it behaves like a probabilistic MDP over the belief states during each time period n) leads to the fact that the result of degradation will be a probabilistic distribution of belief states (Fig. 3). This probability distribution will be provided by a particular BN defined by the manager of the structure, as described in the previous section. In other words, to find the optimal cost $^*c(\nu^n)$, a probabilistic POMDP recursive relation (the decision tree procedure in a GPOMDP) will call for the expected value of $^*c(\nu^{n+1})$, knowing that the vector ν^{n+1} has a probability distribution over the entire belief state space at the beginning of time period $n + 1$ (as defined by a BN), rather than evaluating the cost of exactly updated belief states (as defined by a simple transition matrix).

Fig. 3 illustrates the proposed probabilistic POMDP in contrast to a classical POMDP for the particular case in which the belief state space is two-dimensional. In the probabilistic approach, the expected cost associated with the probability distribution of ν^{n+1} will be

$$E[^*c(\nu^{n+1})] = \int_{\text{state space}} ^*c(\nu^{n+1}) \times f(\nu^{n+1}) \quad (6)$$

where $f(\cdot)$ = probability distribution of the belief state ν^{n+1} resulting from the BN; and $E[\cdot]$ = mathematical expectation operator; $f(\cdot)$ will be calculated by using Eq. (5). Fig. 4 illustrates the GPOMDP by using a BN instead of a transition matrix.

The current approach evaluates $E[^*c(\nu^{n+1})]$ rather than $^*c[E(\nu^{n+1})]$. These two terms are not generally equal unless the function $^*c(\cdot)$ is a linear function of the state ν^{n+1} , which is not generally the case. Hence, the BN cannot simply be used to construct a mean transition matrix, as would be the case if one has only to evaluate $^*c[E(\nu^{n+1})]$.

The model presented in this section can be considered to be a DBN model, which makes use of the DP algorithm to solve the optimization problem related to the actions and inspections to be performed during each time period.

However, for the current approach to work, only one edge, linking the nodes representing the states of the system, must connect the consecutive time slices of the original DBN (Fig. 4). Hence, the

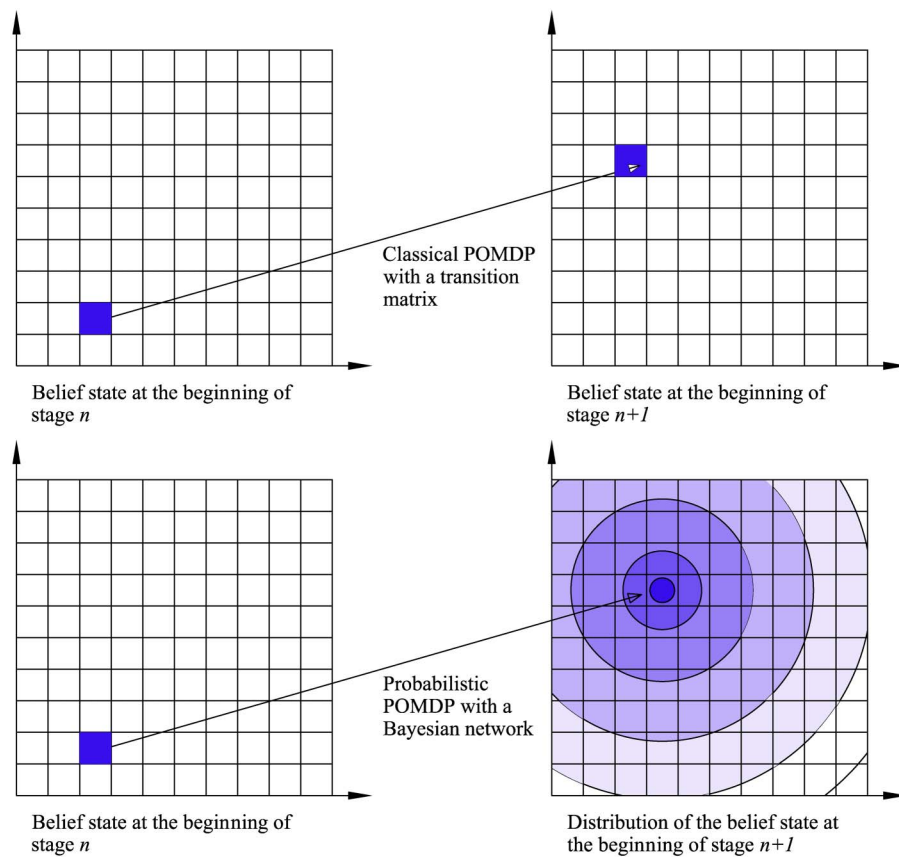


Fig. 3. Probabilistic versus classical POMDPs in the case of a two-dimensional state space

nodes of the BN during each time period, except those representing the state of the structure, must be: (1) independent from nodes in other time slices; or (2) independent from the decisions made by the manager of the system (i.e., their future prior probability distributions for each time period must be available and can be computed separately from the IMR optimization problem). In this latter case, the edges linking the nodes among the different time slices can be removed during the backward calculation of DP without losing information.

For example, the variable X_5 (urban growth) in Fig. 2 depends on a slow process; thus, the node X_5 is not independent from nodes in previous time periods. However, because it is independent from the decisions of the manager, and because its prior probability distributions for each time period can be computed separately, the edges linking the nodes X_5 among the different time

slices can be removed during backward calculation without losing information.

Example Application

This example considers a section of an interstate highway pavement subject to deterioration. The goal is to establish an optimal IMR planning for this pavement. The planning horizon is set to 14 years. The length of each time period is two years long and the discount rate is: $\alpha = 0.049$. It is assumed that the performance of the pavement may be described by five states (Table 1). Because user cost $cs^a\theta_i^n$ models for infrastructure facilities are sometimes not readily available, this study follows Madanat (1993) and replace them, in this example, by constraining the performance of the pavement to be above a specified minimum allowable threshold: $\theta < \theta_5$. The proposed mathematical model can take into account such a constraint by heavily penalizing the state θ_5 ; i.e., associating quasi-infinite user cost to this state and zero user

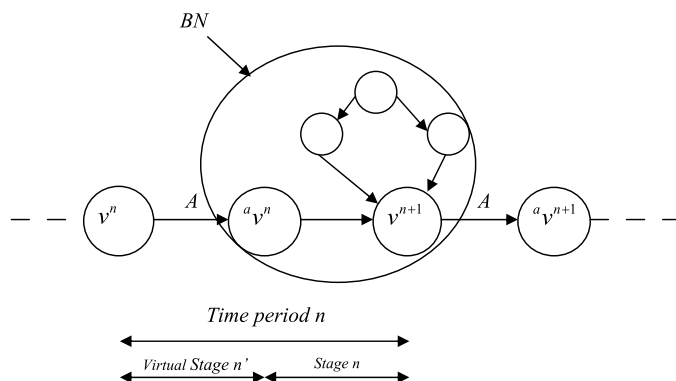


Fig. 4. POMDP using BN instead of a deterioration matrix

Table 1. Relationship between International Roughness Index and Pavement Serviceability Rating

Condition	Interstate PSR ^a	Interstate IRI ^b (m/km)
θ_1 (very good)	≥ 4	< 0.94
θ_2 (good)	3.5–3.9	0.94–1.48
θ_3 (fair)	3.1–3.4	1.49–1.87
θ_4 (mediocre)	2.6–3.0	1.89–2.68
θ_5 (poor)	≤ 2.5	> 2.68

Note: Data from Federal Highway Administration (2008).

^aPavement serviceability rating.

^bInternational roughness index.

Table 2. Transition Probabilities and Costs of the Inspection Techniques

Inspection technique	SD (m/km)	$P(r = j - 2 \theta^n = j)$	$P(r = j - 1 \theta^n = j)$	$P(r = j \theta^n = j)$	$P(r = j + 1 \theta^n = j)$	$P(r = j + 2 \theta^n = j)$	Cost (US\$/m ²)
i_4	0.08	0	0.05	0.9	0.05	0	0.108
i_3	0.16	0	0.1	0.8	0.1	0	0.054
i_2	0.24	0	0.15	0.7	0.15	0	0.036
i_1	0.32	0	0.2	0.6	0.2	0	0.027
i_0	∞	0.2	0.2	0.2	0.2	0.2	0

cost to the other states. The study also follows Madanat (1993) and considers four hypothetical inspection techniques. The adapted measurement errors related to these inspection technologies are assumed to be normally distributed with zero mean and SDs of 0.08, 0.16, 0.24, and 0.32 m/km. These distributions were transformed into discrete measurement probabilities by using basic theorems of probability. Table 2 indicates, for each inspection technique, the SD of the measurement error and the probabilities of obtaining the different results, $r = i$ ($i = 1, \dots, 5$), knowing that the true state of the pavement is $\theta^n = j$. The average cost associated with these inspection techniques (Table 2) is based on studies conducted for the Federal Highway Administration (FHWA) (Hudson et al. 1987). In addition to the four inspection techniques, a fifth inspection technique is included that has an infinite SD, which stands in this model for the case in which no inspection will be done (i.e., the study entirely relies on the prediction of the degradation model), and its cost is nil. Also, it is assumed that only two imperfect maintenance actions can be employed. The transition matrices related to these maintenance actions are adapted from the work of Madanat (1993). The cost of action a_0 (no maintenance) is nil, the cost of action a_1 (two-inch overlay) is US\$6.56 per m³ and the cost of action a_2 (reconstruction) is US\$21.71 per m³. The uncertainties associated with the different maintenance actions are expressed by the matrices of Tables 3, 4, and 5.

The stochastic deterioration of the pavement is assumed to be described by a Markov chain with state space $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$.

Table 3. Transition Matrix for Action a_0

State of the structure	$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
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

Table 4. Transition Matrix for Action a_1

State of the structure	$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
θ_1^n	0.916	0.084	0	0	0
θ_2^n	0.852	0.148	0	0	0
θ_3^n	0.788	0.212	0	0	0
θ_4^n	0.736	0.264	0	0	0
θ_5^n	0.632	0.368	0	0	0

Table 5. Transition Matrix for Action a_2

State of the structure	$a\theta_1^n$	$a\theta_2^n$	$a\theta_3^n$	$a\theta_4^n$	$a\theta_5^n$
θ_1^n	0.94	0.06	0	0	0
θ_2^n	0.94	0.06	0	0	0
θ_3^n	0.94	0.06	0	0	0
θ_4^n	0.94	0.06	0	0	0
θ_5^n	0.94	0.06	0	0	0

It is supposed that the initial belief state of the pavement section is $v = [0.1 \ 0.2 \ 0.55 \ 0.15 \ 0]$. The deterioration during each time period is modeled by a BN having the structure illustrated in Fig. 2. For brevity and without any loss of generality, this example neglected the effect of seismic activity. The effects of traffic intensity [in thousands of equivalent standard axle loads (ESALs)], annual amount of precipitation (millimeters of rain per year), and the durability related variable [thickness of the hot mix asphalt concrete (HMAC) in centimeters] were assessed by using data extracted from the long-term pavement program (LTTP) (FHWA 1997).

It is assumed that the variables X_2 , X_5 , and X_7 are the only observable nodes in this example. The results were computed for two cases; namely, the transition of the pavement state during each time period as a result of deterioration is modeled by (1) a BN having the structure illustrated in Fig. 2; and (2) a transition matrix M (Table 6), which was calculated to be the mean transition of the BN. In other words, M was calculated by using the same data used to construct the BN, but by neglecting its structure.

Table 7 presents the different values that are given by the various variables of the BN, their meanings, and the associated probabilities. In Table S1, the transition probabilities are presented for different combinations of the values given by the parent nodes at the beginning of time period $n + 1$. Also, Tables S2 and S3 provide the conditional probability distributions for the variable X_4 , representing the state of a parallel road, and for the variable X_3 , representing the traffic intensity. The variable X_4 has an indirect causal effect on the state of the pavement under study. This is because the traffic (X_3) on the pavement whose maintenance is to be optimized depends on the performance of parallel roads (linking the same

Table 6. Transition Probabilities of the Degradation Process

State of the structure	θ_1^{n+1}	θ_2^{n+1}	θ_3^{n+1}	θ_4^{n+1}	θ_5^{n+1}
$a\theta_1^n$	0.839	0.121	0.039	0	0
$a\theta_2^n$	0	0.787	0.142	0.070	0
$a\theta_3^n$	0	0	0.708	0.192	0.099
$a\theta_4^n$	0	0	0	0.578	0.421
$a\theta_5^n$	0	0	0	0	1

Table 7. Mathematical Modeling of the Influence of the Parent Nodes on the Belief State of the Structure

Variables	Possible values	Description	Probability
X_2 (annual precipitations)	−1	<500 mm	0.3
	0	500 < 1,000	0.4
	1	>1,000	0.3
X_3 (traffic intensity)	−1	<110 KESAL/year	$P(-1 X_5, X_4)$
	0	110 < 180 KESAL/year	$P(0 X_5, X_4)$
	1	>180 KESAL/year	$P(1 X_5, X_4)$
X_6 (durability related variable: thickness of the HMAC base)	0	>50 mm	0.6
	1	<50 mm	0.4

Note: KESAL = kilo equivalent single axle load.

Table 8. Use of a BN versus the Use of a Transition Matrix for Modeling the Degradation in the DP Recursive Relation

Year	Expected IMR costs (US\$): transition matrix (one inspection followed by a maintenance action)	Expected IMR costs (US\$): BN (one inspection followed by a maintenance action)
0	5.95	5.93
2	1.85	1.79
4	1.70	1.52
6	1.50	1.41
8	1.49	1.37
10	1.45	1.34
12	0	0
Recommended decision for the first period	i_3	i_2
Total cost	12.11	11.66

locations as the road to which the pavement under study belongs). If it is not possible to directly observe the value of X_4 , it can be inferred by using conditional probability distributions on the observed values of the variables X_2 (weather conditions) and X_5 (urban growth). The latter also has an effect on X_3 . Knowing the state of a pavement with the same type (related to durability) as that of the pavement under study allows inferences to be made about the uncertain state of the latter. The results were computed by using specialized graphical user interface (GUI) software for GPOMDP that was developed by the authors. Table 8 presents the expected costs obtained for the cases in which (1) a BN is used for modeling the degradation; and (2) a transition matrix is used (which is calculated as the expected mean of the BN). The solution (i.e., optimal expected costs and prescribed decisions) differs by 3.8%, depending on whether a BN or simply a transition matrix is used to model the degradation process in the DP recursive relation.

The prescribed inspection type for the first time period, in the case of using the BN, is i_2 , whereas it is i_3 that has a smaller SD (hence, it is costlier) for the case of using a transition matrix. This result is because the BN will contribute to the decision process during each time period by introducing relevant available information. Thus, there is less of a need for costly inspection techniques. The expected direct costs decrease sharply after the first period. This is because the initial belief state of the pavement is very poor. Starting from the second time period, a continuous gradual decrease is observed, ending with a zero costs for the final period. This is because the manager of the pavement is less concerned with potential future costs. This result may change if a specified state at the end of the planning horizon was imposed on the manager. A total of 100 simulations were performed to test the obtained prescribed IMR strategy for the two cases: (1) using a simple transition matrix; and (2) using a BN. The evolution of the state of the pavement was generated randomly by using the two degradation models. For the obtained belief state at the beginning of each time period, the IMR strategy was implemented that was prescribed by the solution of the problem obtained by using the BN methodology and the mean transition matrix methodology respectively. The results obtained by using the IMR strategy of the BN had an average expected cost of US\$11.78 per m^3 and an SD of US\$1.36 per m^3 . The results obtained by using the mean transition matrix IMR strategy had an average expected cost of US\$12.32 per m^3 and an SD of US\$2.12 per m^3 . These findings confirm the results obtained by applying the proposed methodology. Also, the SD obtained by using the DBN methodology is smaller than that obtained by using the classical transition matrix.

Discussion

The model presented in this paper can be considered to be a DBN model that makes use of the efficient DP algorithm to solve the optimization problem related to the actions and inspections to be performed during each time period. As such, the solution procedure has a computational complexity that is polynomial in the number of decision nodes. However, it has been shown that POMDPs are polynomial space (PSPACE) complete (Papadimitriou and Tsitsiklis 1987). Hence, the computation time grows exponentially in the current case with the number of possible states of the system.

Any available calculation methodology can be used for POMDPs. However, in the numerical example of this paper, the standard methodology was implemented of discretizing the continuous belief state space (for this particular numerical application, there are 759,375 discrete states during each time period) recursive DP was applied. The current approach implies calculating a belief state for the next time period for each combination of the values used by the nodes of the BN.

For the calculation of the BN, any of the available methodologies can be applied. In the numerical application, the standard marginalization methodology was applied (Pearl 1986). Thus, the extra computational cost required by using a BN instead of a degradation matrix is exponential for the number of nodes of the BN.

Nevertheless, the usual number of states adopted in practice for civil engineering assets is below 10. For example, in the AASHTO bridge inspection manual (AASHTO 2011), the number of states of the bridge elements is limited to four. Additionally, the current approach divides the original DBN into smaller BNs (one for each time period). Therefore, the advantages in computation time obtainable by the proposed approach far outbalance the classical DBN algorithms. The computation time for the numerical application was approximately 50 s when using a degradation matrix and 2 min when using a BN.

The purpose of the method is to allow the manager to make use of a more detailed deterioration model. The additional information obtained by the manager is included via the parameters of the model (i.e., the observable variables of the BN). In fact, for each possible combination of the observed variables, there is a different transition matrix. The proposed model allows the manager to know which transition matrix is relevant during each time period; hence, to benefit from the added value of information which contributes in decreasing the expected cost.

Conclusion

This paper presented a methodology that takes advantage of the structure of the input data used in modeling the stochastic degradation process of a POMDP. It is assumed that this structure has a BN form. This methodology allows one to account, at the beginning of the planning, for the free information relevant to the state of the system that might be available to the manager during future time periods. The primary advantages of using this modeling are described in the following. First, the variables related to the decision nodes (inspections and maintenance actions) of the DBN are optimized by DP and decision trees. As such, the proposed approach has a smaller computational complexity (computation time) than the classical approaches used to solve DBNs, because in these approaches, the complexity is usually exponential in the number of decision nodes. Second, it allows the user to rationally take into account the effects of future possible accidental events such as seismic activities or terrorist attacks. The usefulness is illustrated of exploiting information regarding the realization of events that do not have a causal relationship with the state of the system.

An example application showed that modeling the degradation process as a BN results in an optimal solution that differs from a solution obtained by using a mean transition matrix, which was estimated by using the same data used for the BN.

Notation

The following symbols are used in this paper:

- A_{an} = transition matrix from θ^n to ${}^a\theta^n$ related to the action a^n ;
- a^n = action to be applied at the beginning of stage n ;
- ${}^*a^n$ = optimal action to be applied at the beginning of stage n ;
- $c(\nu^n)$ = expected cost at the beginning of stage n , associated with the belief state ν^n ;
- ${}^*c(\nu^n)$ = optimal expected cost at the beginning of stage n , associated with the belief state ν^n ;
- M = Markov chain transition matrix;
- α = discount factor;
- θ = state of the system;
- θ^n = state of the system at the beginning of stage n ;
- ${}^a\theta^n$ = state of the system during stage n , after the application of a maintenance action a ;
- ν^n = belief state of the system at the beginning of stage n ; and
- ${}^a\nu^n$ = belief state of the system during stage n , after the application of a maintenance action.

Supplemental Data

Tables S1–S4 and Figs. S1–S2 are available online in the ASCE Library (www.ascelibrary.org)

References

- AASHTO. (2011). *Bridge element inspection manual*, 1st Ed., Washington, DC, 172.
- Attoh-Okine, N. O., and Bowers, S. (2006). "A Bayesian belief network model of bridge deterioration." *Proc. ICE—Bridge Eng.*, 159(2), 69–76.
- Bellman, R. (1961). *Adaptive control processes: A guided tour*, Princeton University Press, Princeton, NJ.
- Celeux, F., Corset, A., Lannoy, B., and Ricard, G. (2006). "Designing a Bayesian network for preventive maintenance from expert opinions in a rapid and reliable way." *Reliab. Eng. Syst. Saf.*, 91(7), 849–856.
- Cooper, G. F. (1988). "A method for using belief networks as influence diagrams." *4th Workshop on Uncertainty in Artificial Intelligence*, Corvallis Press, Albany, OR, 55–63.
- Cooper, G. F. (1990). "Probabilistic inference using Bayesian belief networks." *Artif. Intell.*, 42(2–3), 393–405.
- Cooper, G. F., and Herskovits, E. (1992). "A Bayesian method for the induction of probabilistic networks from data." *Mach. Learn.*, 9(4), 309–347.
- Corotis, R. B., Ellis, J. H., and Jiang, M. (2005). "Modeling of risk-based inspection, maintenance and life-cycle cost with partially observable Markov decision process." *Struct. Infrastruct. Eng.*, 1(1), 75–84.
- Eckels, J. E. (1968). "Optimum maintenance with incomplete information." *Oper. Res.*, 16(5), 1058–1067.
- Faddoul, R., Raphael, W., and Chateaneuf, A. (2011). "A generalized partially observable Markov decision process updated by decision trees for maintenance optimization." *Struct. Infrastruct. Eng.*, 7(10), 783–798.
- Faddoul, R., Raphael, W., Chateaneuf, A., and Soubra, A. (2009a). "Optimal planning of inspection and maintenance by including epistemic uncertainties." *Proc., 1st Int. Conf. on Soft Computing Technology in Civil, Structural and Environmental Engineering*, Civil-Comp Press, Stirlingshire, UK.
- Faddoul, R., Raphael, W., Chateaneuf, A., and Soubra, A.-H. (2009b). "Processus généralisé de décision markovien partiellement observable utilisant les arbres de décisions." *Proc., 41st Journées de Statistique, Société Française de statistiques (SfS)*, Paris (in French).
- Faddoul, R., Raphael, W., Chateaneuf, A., and Soubra, A. (2010). "Processus de décisions markoviens parallèles soumis à des contraintes budgétaires." *JFMS'10*, Hermes Science Publishing Ltd, London, UK (in French).
- Faddoul, R., Soubra, A.-H., Raphael, W., and Chateaneuf, A. (2013). "Extension of dynamic programming models for management optimization from single structure to multi-structures level." *Struct. Infrastruct. Eng.*, 9(5), 432–447.
- Federal Highway Administration. (1997). "Investigation of development of pavement roughness." *Rep. No. FHWA-RD-97-147*, Washington, DC.
- Federal Highway Administration. (2008). "2008 Status of the nation's highways, bridges, and transit: Conditions and performance." *Rep. to Congress*, Washington, DC.
- Francois, O., Bouillaut, L., Aknin, P., Leray, P., and Dubois, S. (2008). "Dynamic Bayesian network modeling maintenance strategies: Prevention of broken rails." *World Congress on Railway Research (WCRR'08)*, World Congress on Railway Research, Sydney.
- Frangopol, D. M., Saydam, D., and Kim, S. (2012). "Maintenance, management, life-cycle design and performance of structures and infrastructures: A brief review." *Struct. Infrastruct. Eng.*, 8(1), 1–25.
- Heckerman, D., Geiger, D., and Chickering, D. M. (1995). "Learning Bayesian networks: The combination of knowledge and statistical data." *Mach. Learn.*, 20(3), 197–243.
- Hudson, W., Elkins, Cl., Uddin, W., and Reilly, K. (1987). "Improved methods and equipment to conduct pavement distress surveys." *Final Rep., FHWA-TS-87-213*, FHWA, Washington, DC.
- Jensen, F., Jensen, F. V., and Dittmer, S. L. (1994). "From influence diagrams to junction trees." *Proc., 10th Conf. on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, San Francisco, CA, 367–373.
- Jones, B., Jenkinson, I., Yang, Z., and Wang, J. (2010). "The use of Bayesian network modeling for maintenance planning in a manufacturing industry." *Reliab. Eng. Syst. Saf.*, 95(3), 267–277.
- Langseth, H., and Portinale, L. (2007). "Bayesian networks in reliability." *Reliab. Eng. Syst. Saf.*, 92(1), 92–108.
- Madanat, S. (1993). "Optimal infrastructure management decisions under uncertainty." *Transp. Res. Part C*, 1(1), 71–80.
- Madanat, S., and Ben-Akiva, M. (1994). "Optimal inspections and repair policies for infrastructure facilities." *Transp. Sci.*, 28(1), 55–62.
- Manoj, K. J. (2009). "Dynamic Bayesian network for predicting the likelihood of a terrorist attack at critical transportation infrastructure facilities." *J. Infrastruct. Syst.*, 10.1061/(ASCE)1076-0342(2009)15:1(31), 31–39.
- Monahan, G. E. (1982). "A survey of partially observable Markov decision processes; Theory, models, and algorithms." *Manage. Sci.*, 28(1), 1–16.
- Papadimitriou, C. H., and Tsitsiklis, J. N. (1987). "The complexity of Markov decision processes." *Math. Oper. Res.*, 12(3), 441–450.
- Pearl, J. (1986). "Fusion, propagation and structuring in belief networks." *Artif. Intell.*, 29(3), 241–288.
- Pearl, J., and Dechter, R. (1996). "Identifying independence in causal graphs with feedback." *Proc., 12th Annual Conf. on Uncertainty in Artificial Intelligence*, Morgan and Kaufmann publishers, Burlington, MA, 420–426.
- Robelin, C. A., and Madanat, S. (2007). "History-dependent optimization of bridge maintenance and replacement decisions using Markov decision processes." *J. Infrastruct. Syst.*, 10.1061/(ASCE)1076-0342(2007)13:3(195), 195–201.
- Sprites, P. (1995). "Directed cyclic graphical representations of feedback models." *Proc., 11th Annual Conf. on Uncertainty in Artificial Intelligence*, Morgan and Kaufmann publishers, Burlington, MA, 491–498.
- Sprites, P., Glymour, C., and Scheines, R. (1993). "Causation, prediction and search." *Lecture Notes in Statistics*, Vol. 81, Springer-Verlag, Berlin.

- Straub, D. (2009). "Stochastic modeling of deterioration processes through DBN." *J. Eng. Mech.*, 10.1061/(ASCE)EM.1943-7889.0000024, 1089–1099.
- Zhang, N. L. (1998). "Probabilistic inference in influence diagrams." *Comput. Intell.*, 14(4), 475–497.
- Zhang, N. L., Qi, R., and Poole, D. (1992). "Stepwise-decomposable influence diagrams." *Proc., 3rd Conf. on Knowledge Representation*, Cambridge, MA.
- Zhang, N. L., Qi, R., and Poole, D. (1994). "A computational theory of decision networks." *Int. J. Approximate Reasoning*, 11(2), 83–158.