Bayesian Networks to Quantify Transition Rates in Degradation Modeling: Application to a Set of Steel Bridges in The Netherlands

Alex Kosgodagan
Structural Reliability, TNO, The Netherlands - IRCCyN, École des Mines de Nantes, France

Oswaldo Morales-Nápoles
Faculty of Civil Engineering and Geosciences, TU Delft, The Netherlands - Structural Reliability, TNO, The Netherlands

Johan Maljaars
Structural Reliability, TNO, The Netherlands - Faculty of the Built Environment, TU Eindhoven, The Netherlands

Thomas Yeung
IRCCyN, École des Mines de Nantes, France

Bruno Castanier
LARIS, Université d'Angers, France

ABSTRACT: Bridge lifetime pose an important challenge in terms of maintenance for decision makers or asset managers. In this regard Markov chains have been used successfully in practice as models for bridge deterioration. However, one limitation of Markov chains can be the assessment of the transition probabilities. In this paper, we propose an approach based on Bayesian networks (BNs) to quantify the transition probabilities of the system state. One of the advantages of doing so is that the BN may be quantified through physical variables linked to the underlying degradation process in an intuitive way through expert judgment combined with field measurements. In addition, the possibility of using Bayesian inference allows updating the probabilities when observations become available that could provide different relevant views of the long-term degradation. An application to a hypothetical stock of steel bridges in the Netherlands is presented and illustrates the method.

1. INTRODUCTION

The ministry of Public Works and Water Management is responsible for approximately 3500 bridges in the Netherlands of which more than 300 are composed of steel. Steel structures are subject to fatigue degradation during their lifetime (Sangid 2013), and cracks occur frequently due to this degradation. As the whole set of bridges is tremendously large, the need of flexible models for degradation analysis for decision makers and asset managers has therefore risen over the past decades.

Existing literature offers a wide range of cross-disciplinary approaches when dealing with single and multiple structures that deteriorate. Physical models for deterioration have been commonly used in practice to date. They indeed can attain a great level of detail in the understanding of deterioration processes. Nevertheless, the uncertain behavior of quantities of interest naturally led to taking into account probabilistic and/or statistical features in such models (see for example Maljaars *et al.*(2012)). Kobayashi *et al.* (2014)) recently developed a pure-probabilistic-based method that principally describes the random evolution of three different types of cracks for civil infrastructures. The objective lies in describing the performance by selecting smartly the cracking condition through a competing stochastic process.

One of most challenging tasks in maintaining a network of sophisticated structures is to tackle it as a whole instead of seeing it as the multiplication of one individual element. Frangopol and Bocchini (2012) present a state-of-the-art review on transportation network performance focusing on a bridge network example. The main underlying difficulty is related to time consuming computations since existing models often require the quantification of a large number of parameters. Recent techniques to overcome this issue happen to be efficient though. Bocchini *et al.* (2013) suggest a numerical technique through a Gaussian random field for the enhancement of the computational efficiency of life-cycle analysis of transportation networks under uncertainty. They approached the problem of life-cycle analysis whereas this paper provides a survey on lifetime issues.

Reliability modelling also copes with this problem as Markov chains and Bayesian networks (BNs) have been frequently used. Examples of the latter are given in Langseth and Portinale (2007). In Straub (2009) a generic framework for stochastic modeling of deterioration processes is presented, based on dynamic Bayesian networks. Furthermore and specifically for bridges' degradation, statistic-based methods merged with Markov chains are put forward. They offer a logical way of evaluating transition probabilities since they rely on true collected data coming from inspections as pointed out in Madanat *et al.* (1995) and Morcous (2006). However, for dealing with a fleet of large and complex structures these methods can become inappropriate. The latter pointed out that they are subjected to the usual lack of data in particular concerning the assessment of small values for transition probabilities.

In this paper, we propose a model using BNs to derive transition rates for Markov chains in an intuitive manner. In order to do so, carefully selected uncertain physical quantities act as influence factors through the probabilistic scheme that Bayesian networks offer. These quantities are of particular importance as they are estimated by an expert elicitation procedure. This involves the use of results yielded by Markov chain theory that are both easy to derive and represent meaningful quantities to experts.

Our objective is first to analyze the different possible configurations of BNs we construct with a view to pick the most adapted one with respect to the set of bridges we consider. Secondly, we are interested in thoroughly quantifying the BNs. The expert elicitation method cited above offers an innovative and intuitive mechanism. The whole process of predicting degradation in a probabilistic sense additionally benefits from it.

In order to achieve this, we divide the paper into the following sections. Section 2 provides a quick review of the basics on Bayesian networks and Markov chains as well as properties induced which are of importance to our problem. Section 3 presents the model formulation and poses the actual problem relying on the two above-mentioned methods. Section 4 introduces a numerical example on a set of hypothetical bridges and shows some of the results of interest. Finally Section 5 draws conclusions brought out from the numerical example and delivers perspectives and future work.

2. METHODS

In order to address the problem, this paper first includes two methods that rest upon probability theory, namely discrete Bayesian networks and discrete-time Markov chains.

Bayesian networks embed both graph and probability theory. They can be seen as directed acyclic graphs (DAG) where vertices (also called nodes) stand for univariate random variables and edges (also referred to as arcs) symbolize direct influences between nodes. The direct predecessors of a node are called "parents" and conversely "children" are immediate successors of a particular node. If the set of parents for a node is empty we call it a "source" node. A BN encodes the probability density or mass function on a set of variables by specifying a set of conditional independence statements in the DAG associated with a set of conditional probability functions. In addition, the possibility of using Bayesian inference in the model gives a dynamic feature when observations become available. A more exhaustive description on BNs can be found in Pearl (1988).

In discrete BNs, nodes represent discrete random variables. These models specify marginal distributions for source nodes, and conditional probability tables (CPTs) for child nodes. In practice, there exist two ways of filling out the CPTs, that is either make use of available data or exploit structured expert judgment methods. In our case the latter is employed as quantities of interest cannot be accessible through data nor be calculated. Next section discusses in more details the implemented procedure. In summary, BNs turn out to be a proficient approach when evaluating (un)conditional probability distributions. Associating them with Markov chains is therefore enticing to investigate.

Markov chains have demonstrated to be an adequate tool for the prediction of bridges conditions as reported notably in Mirzaei *et al.* (2012). Also, as outlined in Mašović and Hajdin (2014) for managing the Serbian bridge network, most worldwide bridge management systems have utilized Markov chains adopted them of addressing the problem. They indeed offer a simple broad frame to manipulate, especially when handling numerous and complex structures as bridges can be. In regards to deterioration issues, the validity of the Markov property has been at the same time praised (Scherer and Glagola (1994), Skuriat-Olechnowska (2005)) and criticized. Specifically concerning the transition from the best/initial condition state to the second best, the sojourn time is not exponentially distributed as the Markov property implicitly asserts. In order to overcome this issue Sobanjo (2011) used a semi-Markov model with a Weibull sojourn. Nonetheless, the choice of the semi-Markov model along with a Weibull sojourn time distribution with the parameters established using the maximum likelihood estimate were obtained using historical data for specific bridges. The model formulated in detail in the next section does not rely on this type of data. By consequence it is supposed throughout homogenous discrete-time Markov chains.

The degradation state space Ω varies generally from 3 to 9 states according to literature. For instance, Kallen (2007) operates with a 4- to 9-state degradation categories space while Morcous (2006) leans on a 6-state space. In addition, we assume that the deterioration process is only allowed at each time unit to either remain in the same state or move to the next state but cannot move backwards to better states. In other words, the worst state is an absorbing state for which renovation is highly encouraged but is not synonymous with a collapse and all the other states are transient. Assume n possible states for the Markov chain, the transition probability matrix is given by

$$\boldsymbol{P} = \begin{pmatrix} 1-p_{12} & p_{12} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1-p_{n-1n} & p_{n-1n} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

where p_{ij} is the well-known one-step transition probability.

3. MODEL FORMULATION

We first consider a set of m hypothetical steel bridges. To each bridge k it is coupled a Markov chain $\{X_t^k\}_{t>0}$ having n number of states ($n \geq 2$). Define the stochastic process $\{\gamma_t\}_{t>0}$ that serves as a global indicator in regards to the health of the entire stock. The outcome of the latter is obtained by an aggregation of each bridge's condition. In addition we introduce the uncertain physical parameters (*covariates*) influencing the transition probabilities of the Markov chains such as load solicitation, bridge's inner geometries, traffic density, etc. While we could end up with 20 of them for a single structure (Maljaars *et al.*(2012)), the problem of studying a fleet of bridges forces us to sharply decrease this number and conserve the most relevant ones. These covariates act subsequently as nodes in the BN which in turn implies a dependence relationship between them.

3.1. Construction of the BNs

Deterministic factors inherent to each bridge are analyzed in view to construct the dependence structure between the bridges. Among these we can cite the geographical location, the highway they are part of, their surface area and so forth. Once a set of bridges have been picked based on these criteria, we build a BN according to each selected set. Hence we end up with dissimilar BNs each corresponding to a particular selection. These differences are expressed in terms of the arcs connecting the covariates. Relative to the value of these known factors there is also the possibility to link some of the covariates belonging to different bridges. The advantage of doing so is twofold. The first added value is directly related to the inability to completely monitor the entire network. Measured data would say something about the remainder through the mechanism of inference.

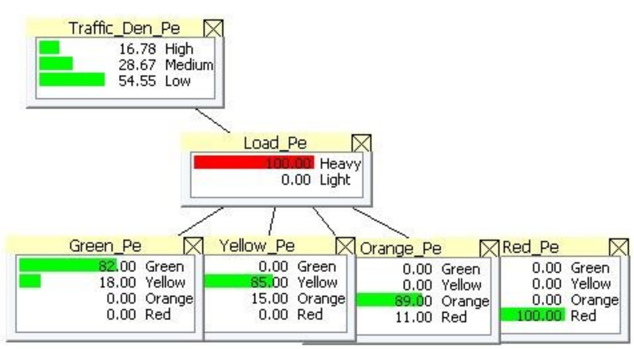

Figure 1: Exact inference in a discrete BN

In Bayesian network theory, inference is a central concept (Pearl 1988). Available information is inserted and propagates through the BN in both top-down and bottom-up manner. The distribution is therefore updated which eventually gives a dynamic feature to the model. For discrete BNs, discrete Bayes' rule is invoked in order to perform and propagate inference. This is illustrated in Figure 1 where evidence has been observed in state "Heavy" for node *Load_Jo*. This translated in a probabilistic way amounts to saying that $P(Load\_Pe = Heavy) = 1$.

3.2. Quantification

In order to quantify properly the CPTs of each physical random quantity, an expert elicitation method is summoned based on an innovative and indirect bottom-up approach. The choice of soliciting experts about quantities which are familiar to them seems rather obvious. However it turns out that transition probabilities prove difficult to elicit directly whereas expected first passage time for sequential condition states does (Cooke 1991). The link between these quantities and transition probabilities allows us then to perform the elicitation. Let us first provide with the distribution of the first time of passage from state i to state j which is defined as

$$f_{ij}(t) = P(X_t = j, X_{t-1} \neq j, \dots, X_1 \neq j | X_0 = i)$$

and can be expressed as the following recursive equation

$$f_{ij}(t) = \begin{cases} \sum_{k \neq j} p_{ik} f_{kj}(t-1), t > 1 \\ p_{ij}, t = 1 \end{cases} \quad (1)$$

The expected first arrival time is thereafter assessed by the following formula

$$E[T_{ij}] = 1 + \sum_{k \neq j} E[T_{kj}] p_{ik} \quad (2)$$

where $T_{ij} = \inf\{s : X_s = j | X_0 = i\}$.

Notice that for fixed j in Eq. pik (2) we have to solve a system of equations that has $card(\Omega)$ equations and unknowns. Complexity in solving this equation is therefore directly related to the chosen number of bridge condition states.

The process in quantifying the BN is according to following 3-step procedure:

1. Ask experts their uncertainty distribution over the expected first passage time between sequential condition states, namely $E[T_{12}], \dots, E[T_{n-1n}]$. This can be done for example using Cooke's (Cooke 1991).
2. Retrieve the transition probability matrix by solving the system of equation in Eq. p_{ik} (2).

3. Compute the CPT's of each node in the BN using Bayes' rule along with the marginal distribution of nodes yielding the transition probabilities.

According to the complexity of the BN (number of arcs and states) the above-mentioned procedure has to include the assessment of marginal distribution as well as sets of conditional probabilities for nodes.

4. NUMERICAL EXAMPLE

We consider here a BN comprised of $m = 5$ steel bridges labeled as *Ry, Pe, Ma, Jo* and *Co*. All bridges are assumed to be located in the same highway in the Netherlands and each possessing 2 covariates which stand for the uncertainty on traffic density and load solicitation. Both refer to an annual period and are displayed in Figure 2 labeled as *Traffic_Den_"label_of_the_bridge"* and *Load_"label_of_the_bridge"* respectively. Traffic density represents the number of vehicles per kilometer per lane averaged over the total number of lanes. In this example we use a distribution having 3 states, namely *High, Medium* and *Low* which correspond to standstill, queued and free flow traffic respectively. Node *Load* has in turn 2 states, *Heavy* and *Light* which are defined with respect to the maximum load solicitation capacity each bridge can bear.

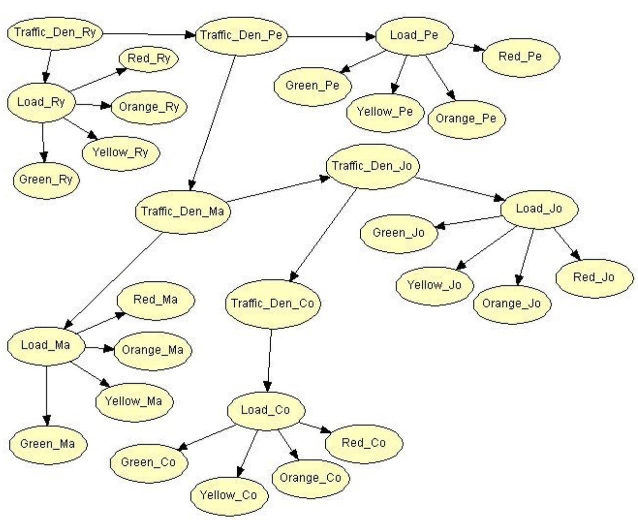

Figure 2: BN with 5 bridges sequentially linked through nodes Traffic_Den

Furthermore we operate here with $n = 4$ states for space $\Omega = \{G, Y, O, R\}$, where G stands for *Green* and represents an "as good as new" condition, Y for *Yellow*, O for *Orange* and R for *Red* in successive deterioration states. State *Red* does not indicate total collapse of the structure but rather heavy maintenance required. In an attempt to give experts the best possible description in regards to the degradation mechanisms that are involved, these conditions are better qualitatively and quantitatively defined.

The stochastic process γ_t defining the *global indicator* takes value in a 3-state space:

- State *1:* at least 60% of the stock is in state *Green* and the rest in any other state but *Red*
- State *2* : all other possible combinations not contained in state *1* and *3*
- State *3:* at least 60% of the stock is in state *Red* and the rest in any other state but *Green*

In this case, bridges are linked consecutively through nodes *Traffic_Den*. This choice comes from the fact that they belong to the same highway and are closely located between one to another. It can be thus plausibly inferred that there exists a correlation between their respective traffic density. Notice there is here only one source node, namely *Traffic_Den_Ry,* which means its marginal distribution is not calculated with conditional probabilities.

The quantification of the CPTs is then executed by means of the expert elicitation technique. Per bridge, it is asked uncertainty distributions over three expected transitions between successive states. We assume that when newly constructed, bridges are in condition *Green*. More precisely, the following type of question is solicited to the expert:

"*We are looking at the 5 highway steel bridges at the time of their construction. Could you provide with the 5th, 50th and 95th quantiles of your uncertainty distribution about the expected years that it takes for the bridge considered to transit between each consecutive state?"*. Expert subsequently fills out his/her uncertain estimates with the help of relevant information on each bridge, such as the span of the bridge, its age, the type of deck, the main joining method applied (welded, riveted or bolted), and so on. For bridge *Ry* estimates are detailed in Table 1.

Table 1: Expert elicitation estimates (years) for bridge Ry (hypothetical estimates).

Transitions/Quantiles	*5th*	*50th*	*95th*
Green → Yellow	*10.2*	*13.9*	*17.6*
Yellow→ Orange	*9.4*	*13.7*	*16.2*
Orange → Red	*9.2*	*12.6*	*15.9*

Next, the transition probability matrices at time zero which corresponds to the year of construction are computed. Below it is shown the matrix for the 50th quantile for bridge *Ry*.

$$\boldsymbol{P_{Ry}} = \begin{pmatrix} 0.9281 & 0.0719 & 0 & 0 \\ 0 & 0.927 & 0.0730 & 0 \\ 0 & 0 & 0.9206 & 0.0794 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Finally, step 3 of the procedure is carried out. An example of the CPTs used is exhibited in Table 2. For instance, first cell in Table 2 reads

$$P\left(X_t^{Ry} = G \middle| X_{t-1}^{Ry} = G, Load\_Ry_{t-1} = Heavy\right) = 0.8714$$

Other conditional probabilities in the table may be read in the same way. Notice that, in addition to the questions asked regarding expected transition times, we add questions to quantify the conditional probabilities of certain nodes in order to completely parameterize the BN. Other CPTs are interpreted similarly during the quantification.

Table 2: CPT of node Green_Ry

	Heavy	*Light*
Green	*0.8714*	*0.9437*
Yellow	*0.1286*	*0.0563*
Orange	*0*	*0*
Red	*0*	*0*

Conditionally, on available evidence coming from monitoring for instance, we are able to distinguish between conditional and unconditional distributions. Here, for illustration purposes, inference on high traffic density has been inserted for bridge *Ry* taking the value of the 50th quantile. In Figure 3 both distributions are plotted and we can observe that a continuous high traffic density accelerates the time to reach the failure state and thus reduces its lifetime. In terms of the plotted distribution this translates into a more abrupt slope.

Confidence interval stemming from the 5th and 95th quantiles shows a relatively large uncertainty especially expressed around year 30. Figure 3 results from a single expert estimates. Typically, combining experts individual estimates according to performance in the Cooke's method would result in uncertainty bands in the order of individual experts. Notice also that the effect of conditionalizing on high traffic is smaller than expert's uncertainty.

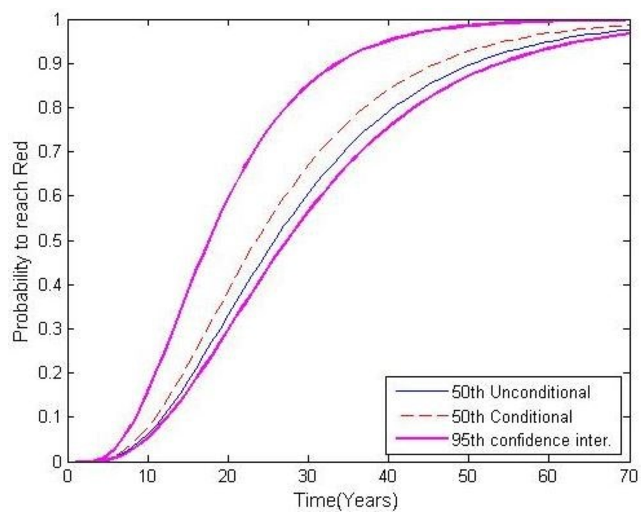

Figure 3: Expert uncertainty on lifetime distribution for bridge Ry

Beyond the fact of having forecasts on each individual elements of the network, we are also able to establish how the random aggregated global indicator behaves. Set $\Omega_R = \Omega \setminus \{R\}$, for notation purposes, we use *1* to denote bridge *Ry*, *2* for bridge *Pe*, and so on. Concerning the distribution of the process γ_n being in state *1*, we are seeking the following:

$$P(\gamma_t = 1) = P\left[\bigcup_{i=1}^{m-1}\left\{\bigcup_{j=i+1}^{m}\left\{\{X_t^i \in \Omega_R\} \cap \{X_t^j \in \Omega_R\} \cap \left\{\bigcap_{k\neq i,j}\{X_t^k = G\}\right\}\right\}\right\}\right]$$

Recall that if the process γ_t is in state *1*, it means that the network has at least 60% of bridges being in condition G*reen* and the rest of the network is in any other state but *Red*. To compute this probability, we first apply the principle of inclusion and exclusion. It is the well-known formula giving the distribution of the union of events which applies to our case. Along with it we make use of the directional separation property (see Pearl 1988). This allows us to separate variables (or nodes) being conditionally independent giving another variable. We end up computing products of univariate conditional distributions instead of dealing with multidimensional joint distributions. For instance we typically use that the set of variables {*Green_Ry, Yellow_Ry, Orange_Ry Red_Ry*} is conditionally independent of {*Green_Pe, Yellow_Pe, Orange_Pe, Red_Pe*} giving {*Traff_Den_Ry*}. Using the same methodology, we calculate the distributions of state *2* and *3* which are all plotted in Figure 4.

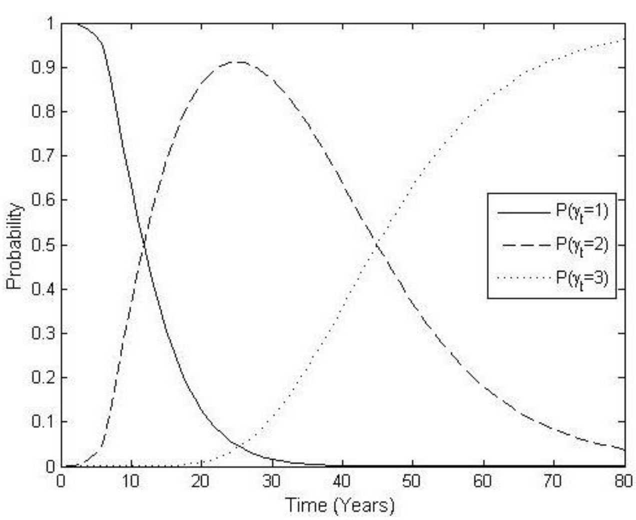

Figure 4: Distribution of states 1, 2 and 3 of γ_t

From the latter we can distinguish three different phases. First we can assert that until year 12 the entire network should remain in a good condition. Prior to this date it is likely that most of the bridges should be in a *Green* or *Yellow* condition state as probability is equally distributed between states *1* and *2* at year 12. Then from approximately that year to year 45, the distribution mass lies predominantly in state *2* with a peak reached around year 25. This phase is of interest as state *2* regroups most of the total possible number of combinations (see previous definition). Thus the uncertainty concerning the state in which each bridge will be at between year 12 and 45 is high. This information will be relevant for maintenance purposes for example, hence particular attention should be drawn to the time span corresponding to this phase. Finally in year 45 a crossover occurs between state *2* and state *3* distributions. From that year on this leads the indicator to have a steadily increasing probability to reach state *3* than to any other state.

5. CONCLUSIONS

This paper has proposed a method to describe the lifetime distribution of bridges. Incentives fostering the use of this model first lie in its tractability and its applicability as it appears one could export it to different civil assets. Indeed the most challenging task remains the adequate use of covariates, this model requires relatively small amounts of data which is in general easily available. The last two features also allow the model to be operational in practice. As for extending it to the whole stock of highway steel bridges in the Netherlands, it is much likely to be feasible since dependencies handled through the BN structure corresponds to some extent to the geographical location of the bridges. Still a general challenge in using discrete BNs is that their quantification grows rapidly with the number of nodes and states. This model may be used to update uncertainty on the bases of monitoring or inspections. By examining the distribution of the global health state indicator, asset managers may be able to describe the general "health" of the stock of bridges. Increasing or defining states according to the requirements of particular applications, this would reduce the uncertainty around state *2* for instance and would eventually aim at optimizing maintenance costs.

Future research extending the results presented up to now include the elicitation to with a panel of experts with the objective to combine expert opinions. Extending the model to account for maintenance actions shall be taken into account in view to come up with a method that would cover partial or complete renovations as well. The effects of maintenance actions with respect to expert's uncertainty (as shown in Figure 3 for example) should also be investigated in the future.

6. REFERENCES

Bocchini, P., Frangopol, D.M. and Deodatis, G. (2011). "A random field based technique for the efficiency enhancement of bridge network life-cycle analysis under uncertainty." *Engineering Structures,* 33, 3208-3217

Cooke, R.M. (1991). "Experts in Uncertainty: Opinion and Subjective Probability in Science."*Oxford University Press.*

Frangopol, D.M. and Bocchini, P. (2012). "Bridge network performance, maintenance and optimization under uncertainty: accomplishments and challenges." *Structure and Infrastructure Engineering,* 8(4), 341-356.

Kallen, M. J. (2007). "Markov processes for maintenance optimization of civil infrastructure in the Netherlands." *Ph.D. Thesis*, TU Delft.

Kobayashi, K., Kaito, K. and Lethanh, N. (2014). "A competing Markov model for cracking prediction on civil structures", *Transportation Research Part B,* 68, 345-362.

Langseth H. and Portinale, L. (2007). "Bayesian networks in reliability". *Reliab Engineering and System Safety,* 92(1), 92-108.

Madanat, S., Mishalani, R. and Wan Ibrahim, W. H. (1995). "Estimation of infrastructure transition probabilities from condition rating data". *Journal of Infrastructure Systems*, 1(2):120–125.

Maljaars, J., Steenbergen, H.M.G.M. and Vrouwenvelder, A.C.W.M. (2012). "Probabilistic model for fatigue crack growth and fracture of welded joints in civil engineering structures." *International journal of fatigue*, 38, 108–117.

Mašović, S. and Hajdin, R. (2014). "Modelling of bridge elements deterioration for Serbian bridge inventory." *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, 10:8, 976-987.

Mirzaei, Z., Adey, B.T., Klatter, L., and Kong, J.S. (2012). "Overview of existing bridge management systems.", *IABMAS report*.

Morcous, G. (2006). "Performance Prediction of Bridge Deck Systems Using Markov Chains." *Journal of performance of constructed facilities © ASCE*, 38, 146–155.

Pearl, J. (1988) "Probabilistic reasoning in intelligent systems: networks of plausible inference". San Francisco California: Morgan Kaufmann.

Sangid, M.D. (2013). "The physics of fatigue crack initiation." *International journal of fatigue,* 57, 58-72.

Scherer, W.T. and Glagola, D.M. (1994). "Markovian models for bridge maintenance management." *Journal of Transportation Engineering*, 120(1), 37-51.

Skuriat-Olechnowska, M. (2005). "Statistical inference and hypothesis testing for Markov chains with interval censoring – Application to bridge condition data in the Netherlands." *A thesis submitted to the Delft University of Technology; Delft, the Netherlands.*

Sobanjo, J. (2011). "State transition probabilities in bridge deterioration based on Weibull sojourn times." *Structure and Infrastructure Engineering,* 7(10), 747–764.

Straub, D. (2009). "Stochastic Modeling of Deterioration Processes through Dynamic Bayesian Networks." *Journal of Mechanical Engineering,* 135,1089-1099.