Safety, Reliability and Risk of Structures, Infrastructures and Engineering Systems – Furuta, Frangopol & Shinozuka (eds)
© 2010 Taylor & Francis Group, London, ISBN 978-0-415-47557-0

A comparison of statistical models for visual inspection data

M.J. Kallen
HKV Consultants, Lelystad, the Netherlands

ABSTRACT: Large groups of structures like bridges, pavements and sewer systems, are often inspected visually and their condition is quantified based on a discrete scale. Markov chains have traditionally been used to model the uncertain rate at which these structures progress through such a condition scale. In order to determine optimal strategies for inspections and maintenance activities, these Markov chains must be fitted to the data obtained in the field. For this purpose, quite a few models and methods have been proposed in the past. These are reviewed here and references to applications in the field of civil engineering are given. A qualitative verdict of the performance and applicability of the individual models is given at the end of the paper.

1 INTRODUCTION

One of the primary uncertainties involved in asset management problems, is the uncertainty about the remaining lifetime of systems and structures. When the asset base is very large, visual inspections are commonly used to obtain information on the current state of the assets. These visual inspections generally classify the state of structures on an ordinal condition scale. They present an economically efficient way of gaining insight into the condition of a large number of systems or structures. Examples of assets which are often visually inspected are bridges, pavements, catenary systems and urban sewer systems. However, visual inspections may be very subjective and they may not be sufficient to determine the operational reliability and safety of a system.

In order to use the information gathered by visual inspections for estimating the remaining service life of a structure, one must apply a suitable probabilistic model to represent the uncertain deterioration of the state of the structure over time. Additionally, one must statistically estimate the parameters of such a model. Most commonly, a finite-state Markov process is used to model the uncertain deterioration. A number of statistical estimation methods for the estimation of transition probabilities have been proposed in the scientific literature. Some of these have been applied in asset management systems.

The following section gives a very short introduction to the theory of Markov chains and presents the notation that will be used throughout this paper. Section 3 subsequently presents the various models and their application to visual inspection data. Section 3.1 reviews regression models which do not use the principle of maximum likelihood. Those models which do, are presented in Section 3.2. Some references to models which do not fall within either one of these two categories are made in Section 3.3. Using a simple example, Section 4 illustrates the effect of ignoring dependence between condition states in a sequence of observations. Finally, Section 5 summarizes the findings and draws conclusions about the performance and applicability of each model.

2 THEORY OF MARKOV CHAINS

A few essential elements of the theory of Markov chains are outlined in this section. The primary purpose is to set up the necessary framework and to introduce the notation as it will be used in the remaining sections. For a more elaborate introduction, the reader is referred to any textbook on general probability theory or stochastic processes in particular.

As visual inspections generally classify the condition of a structure on a finite and discrete scale, the uncertain progression through these states is represented by a finite-state Markov process. This is commonly referred to as a Markov chain. We will only consider stationary Markov chains here, which have time-invariant transition probabilities or rates. Let $\{X(t), t \geq 0\}$ represent the state of the process at time t. If x_k is the observed state of the process at time

t_k, then for a set of observations at successive times $t_0 < t_1 < \cdots < t_k$, the Markov property states that:

$$\begin{aligned} &\Pr\{X_{k+1} = x_{k+1} \,|\, X_k = x_k, \ldots, X_0 = x_0\} \\ &= \Pr\{X_{k+1} = x_{k+1} \,|\, X_k = x_k\}, \end{aligned} \tag{1}$$

where X_k is a convenient shorthand notation for $X(t_k)$. The interval transition probability between a pair of states (i, j) during a time interval of length s is defined as

$$p_{ij}(s) = \Pr\{X(t+s) = j \,|\, X(t) = i\}, \tag{2}$$

with $s, t \geq 0$. All pairs of interval transition probabilities may be collected in a matrix denoted by $\mathbf{P}(s)$. The probability of being state i at time t is denoted by

$$p_j(t) = \Pr\{X(t) = i\}, \tag{3}$$

which may be collected in a vector representing the state distribution at time t: $\mathbf{p}(t)$. The expectation of the process at time t is simply determined by

$$\mathbb{E}[X(t)] = \sum_{\forall k} j \cdot p_k(t). \tag{4}$$

A differentiation is made between two common types of Markov chains: discrete- and continuous-time. The latter is a special type of semi-Markov process. The structure of a Markov chain and the uncertainty in the rate of transitions between the states, is defined by a transition probability matrix (denoted **P**) for discrete-time processes and by an intensity matrix (denoted **Q**) for continuous-time processes. This means that the latter is defined in terms of frequencies, whereas discrete-time transitions are defined by probabilities over a fixed period of time.

Condition states are usually given numeric codes which represent the increasing (or decreasing) order of quality of the structure. For example, bridge inspections in the United States use a scale which ranges from 9 (excellent) to 0 (failed); see FHWA, p.38. With this kind of categorical data, it is somewhat unnatural to speak of an expectation of 1.3 at some time t. It could be interpreted as being somewhere between 1 and 2, with an inclination towards 1. If however, the states are given non-numeric codes, e.g. A, B, C, etc., then the expectation in Eq. (4) becomes meaningless.

3 REVIEW OF ESTIMATION METHODS

The use of Markov processes with a finite number of states has become quite common in civil engineering applications. In order to fit the deterioration process to the available data, several statistical models and corresponding estimation methods have been proposed to determine the optimal values of the model parameters. The parameters in a Markov process are the transition probabilities or intensities, depending on whether a discrete- or continuous-time process is used. This review is divided in three parts with the division being based on the method of estimation: regression-based estimation methods (Sec. 3.1), methods based on the maximum likelihood principle (Sec. 3.2), and less common methods like those using Bayesian statistics are also briefly mentioned (Sec. 3.3).

Classical statistical models include the linear and generalized linear models, which relate a response (or dependent) variable to one or more explanatory (or independent) variables using a linear function in the parameters. Generalized linear models form a broader class than the class of linear models. Besides the linear models as a special case, generalized linear models include the binary probit (logit), ordered probit (logit), and Poisson models amongst others. This area of statistical analysis is commonly known as regression analysis.

3.1 Regression-based methods

Fitting a Markov chain deterioration model by minimizing the distance between the observed states and the expectation of the model, is by far the most common approach found in the literature on infrastructure management. Assume that the condition of structures is modeled by the Markov chain $\{X(t), t = 0, 1, 2, \ldots\}$ and let $x_k(t)$ denote the k-th observation of a state at age t. In other words, the population of bridges is assumed to be homogeneous and for each t in a finite set of ages, there are one or more observations of the condition state. As the name suggests, the method of least squares minimizes the sum of squared differences between the observed state at age t and the expected state at the same age. This is formulated as follows:

$$\min_{p_{ij}} \sum_t \sum_k \left\{ x_k(t) - \mathbb{E}[X(t)] \right\}^2, \tag{5}$$

under the constraints $0 \leq p_{ij} \leq 1$ and $\sum_j p_{ij} = 1$. The expectation is defined in Eq. (4). The model in Eq. (5) is a nonlinear model as the expectation of $X(t)$ is a nonlinear function of the parameters, which are the transition probabilities.

The earliest references of the application of the least squares method in infrastructure management can be found in the area of pavement management. An overview of the early development is given by Carnahan et al. (1987) and Morcous (2006) also refers to Butt et al. (1987) as an example of the application in pavement management. Carnahan et al. (1987) and Morcous (2006) also discuss the use of the least absolute deviation regression, which minimizes the sum of the absolute value of the differences. A more re-

cent application to pavement management is given by Abaza et al. (2004) and the regression onto the state expectation is also applied to sewer system management by Wirahadikusumah et al. (2001).

In Cesare et al. (1994), least squares minimization is applied to a slightly different model compared to the one presented in (5). This approach consists of minimizing the weighted sum of squared differences between the observed proportion of states and the state distribution. The objective function is the following:

$$\min_{p_{ij}} \sum_t n(t) \sum_k \{y_k(t) - p_k(t)\}^2, \tag{6}$$

where $n(t)$ is the number of observed states at time t, y_k is the observed proportion of structures in state k, and $p_k(t)$ is the state distribution as in Eq. (3). The weights $n(t)$ are used to assign more weight to those proportions which have been determined with more observations.

Probably the most significant objection against using these approaches is the fact that successive observations are treated as being independent. This is in contradiction with the Markovian assumption of the underlying deterioration process. Another objection against the formulation of the model in (5), is the fact that the expectation of $X(t)$ depends on the definition of the condition scale. From this perspective, the model formulated in (6) is more appropriate.

Assume that the condition of the component is modeled by a Markov chain $\{X(t), t = 0,1,2,\ldots\}$ and that at least two successive observations of the proportions, denoted by $\mathbf{y}(t-1)$ and $\mathbf{y}(t)$, are available. In the ideal situation, we would like to determine the transition probabilities p_{ij} such that $y_j(t) = \sum_i y_i(t-1)p_{ij}$ holds. In reality, we will have to accept that we will only be able to do so up to a certain error $e(t)$ at each time t. The objective function may be defined as follows:

$$\min_{p_{ij}} \sum_t \|\mathbf{y}(t) - \mathbf{P} \cdot \mathbf{y}(t-1)\|, \tag{7}$$

where $\mathbf{P}$ is the transition probability matrix over a single time unit and $\|\cdot\|$ is a vector norm.

For this model, Lee et al. (1970, Chapter 3) derive the classic least squares estimator $\hat{p} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ with appropriately defined matrices $\mathbf{X}$ and $\mathbf{Y}$. Unfortunately, this approach does not explicitly take into account the constraints for the transition probabilities. The row sum constraint holds, but $0 \leq p_{ij}$ may be violated. Alternative approaches are to use Lagrange multipliers (as it is an equality constrained optimization problem) or to use the method of maximum likelihood with an appropriate probability distribution for the errors $e(t)$. Intuitively, the model in Eq. (7) is quite appealing as it incorporates the progressive nature of the Markov process. It does so by directly relating an observed condition state to the condition state at the previous inspection, using the transition probability.

3.2 Maximum likelihood methods

In most situations, the method of estimating model parameters by maximizing the likelihood of the observations, is a possible approach. This is the case is if, for example, the error term in the model is assigned a probability distribution, or if the parameters are probabilities themselves.

3.2.1 Poisson regression

If an object has performed one or more transitions during the time between two periodic inspections, only the number of transitions and not the times of these transitions are known. In order to use count data to estimate transition probabilities, it is often assumed that the transitions are generated according to a Poisson process. A Poisson process is a stochastic process which models the random occurrence of events during a period of time. If the time between the occurrence of each event is exponentially distributed with parameter $\lambda > 0$, then the probability of n events occurring during a period with length $t \geq 0$ has a Poisson distribution. The probability density function of the Poisson distribution is given by $\Pr\{N(t) = n\} = (\lambda t)^n (n!)^{-1} \exp\{-\lambda t\}$, with mean λt such that the expected number of events per unit time is $\mathbb{E}[N(1)] = \lambda$. If there are $m = 1,2,\ldots$ independent observations (t_1, n_1), (t_2, n_2), $\ldots$, (t_k, n_m), the likelihood of these observations is given by

$$\begin{aligned} L(\mathbf{n} \mid \lambda) &= \Pr\{N(t_1) = n_1, \ldots, N(t_m) = n_m\} \\ &= \prod_{k=1}^m (\lambda t_k)^{n_k} (n_k!)^{-1} \exp\{-\lambda t_k\}. \end{aligned} \tag{8}$$

The maximum likelihood estimator for λ is simply $\hat{\lambda} = \sum_{k=1}^m n_k \big/ \sum_{k=1}^m t_k$.

The term 'Poisson regression' stems from the fact that the parameter λ is often assumed to depend on one or more covariates in a multiplicative model: $\lambda = \exp\{\beta'\mathbf{x}\}$, where $\mathbf{x}$ is a vector of covariates and β the vector of coefficients to be estimated. Poisson regression is therefore a generalized linear regression method with the logarithm as the link function; that is, $\log(\lambda) = \beta'\mathbf{x}$, which is also known as a log-linear regression model.

The Poisson process counts the number of events and does not account for different types of events. The assumption is therefore that each event is the same, namely a transition to the next state after an exponential waiting time. The model is therefore necessarily sequential (because it is not possible to distinguish between different target states) and the (random) waiting time in each state is the same. Another often men-

tioned limitation of the Poisson process is the fact that the variance of $N(t)$ is equal to its mean (and therefore increases when the mean increases), whereas the data may be more dispersed such that the variance should be greater than the mean.

Madanat and Wan Ibrahim (1995) used the likelihood in (8) while acknowledging the fact that $N(t)$ is actually finite for the model under consideration. They mention the possibility of truncating the Poisson distribution as a possible correction, but assert that observations of the last state are very rare such that they do not influence the resulting estimator significantly. To account for possible overdispersion, the authors suggest the use of the negative binomial distribution instead of the Poisson distribution for the count of transitions. Compared to the Poisson distribution, which it has as a special case, the negative binomial distribution includes an extra parameter which allows the variance to be adjusted independently of the mean. This is a common approach to account for overdispersion, see e.g. (Cameron and Travedi, 1998, Chapter 4). In a Bayesian framework, the negative binomial distribution is derived by assuming that the intensity λ is gamma distributed.

3.2.2 *Multinomial model*

Assume that all structures are continuously monitored. For Markov chains this implies that each transition for every structure is observed. Let all observations be pooled by age $t \geq 0$ and let the set $\mathbf{N}_i(t) = \{N_{i1}(t), N_{i2}(t), \ldots, N_{in}(t)\}$ represent the random count of transitions to state $j = 1, \ldots, n$ from state i for all structures at age t. Because the deterioration process is continuously monitored, these counts are observed and are multinomially distributed for each state i. The probability of the observations $\mathbf{n}_i(t)$ at age t is given by a multinomial distribution.The maximum likelihood estimator for the transition probabilities is $\hat{p}_{ij} = n_{ij} / \sum_{j=1}^{n} n_{ij}$, where n_{ij} is the total number of observed transitions between states i and j over all ages of the structures.

This result was derived by Anderson and Goodman (1957) and Billingsley (1961). See also Lee et al. (1970) who refer to this type of data as 'micro data'. In the context of estimating bridge deterioration, Morcous (2006) referred to this method as the 'percentage prediction method'. The model has one fundamental disadvantage: it assumes that transitions are observed as they occur. In other words: the deterioration process must be continuously monitored. If this is not the case, the model should not be used.

3.2.3 *Panel data models*

When structures or systems are inspected periodically, we get a type of data which is commonly referred to as panel data. With this type of data, there is only information about the state of the structures at discrete points in time. There is no information about the exact timing of the transitions. Figure 1 shows a sample path of a Markov chain ("degrading" from state 0 to state 5) which is inspected at two points in time. Two transitions have taken place between the times of these inspections, but it is unknown at what time these transitions occurred.

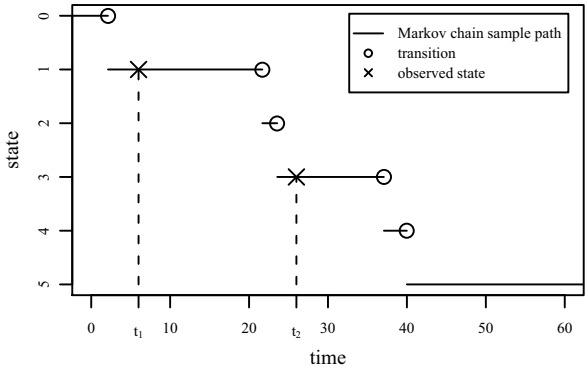

Figure 1. Sample path of a Markov chain with observed states at times t_1 and t_2.

If the random process of transitions is modelled by a Markov chain $X(t)$, the probability of each pair (i, j) of observed states, is given by the interval transition probability in Eq. (2). The likelihood of a sequence of observations on a single structure is given by the product over these interval transition probabilities:

$$L(\mathbf{P} \,|\, \mathbf{x}) = \prod_k p_{x_k, x_{k+1}}(t_k, t_{k+1}). \tag{9}$$

If structures within a larger group may be considered as mutually independent, the likelihood of all observations in the group of structures is obtained by taking the product of the likelihoods of the individual structures.

This method was used by Kallen and van Noortwijk (2006) in order to estimate the transition intensities in a continuous-time Markov chain for road bridges in the Netherlands. For example, Figure 2 shows a comparison between the expected value (Eq. 4) of two Markov chains fitted to a database containing 20 years of observed bridge conditions.

The weigted least-squares model used by Cesare et al. (1994) follows the sample average quite well, but results in an expected lifetime (if state 5 is considered to be a failure state) of over 200 years. The maximum likelihood model for panel data results in an expected lifetime of 45 years, which is closer to the lifetime at which bridges in the Netherlands generally require significant repairs. This difference is a direct consequence of the fact that the regression model ignores the observed transitions between states. Sec. 4 gives a simple example which illustrates the effect

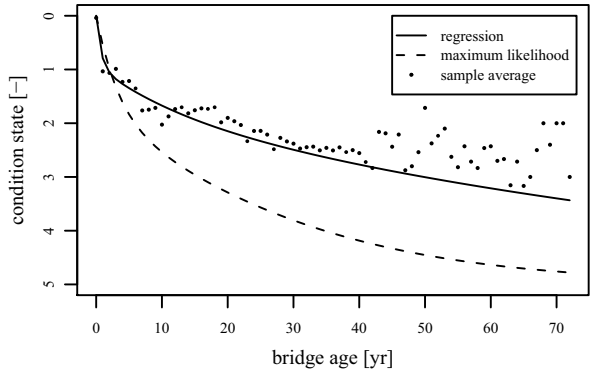

Figure 2. Expected value of a Markov chain for modelling the evolution of bridge conditions in the Netherlands using the weighted least-squares regression (Eq. 6) and the maximum likelihood model for panel data (Eq. 9)

of ignoring the dependences between successive state observations.

3.2.4 Probit and logit models

The binary probit and ordered probit models are linear regression models in which a continuous latent (unobservable) variable is observed to be in two (binary) or more (ordered) discrete categories. These models are appealing for the application in maintenance modeling, as the condition states are often assumed to be related to some underlying deterioration process which can not be measured directly.

Let the unobservable amount of deterioration be given by the random variable Y, then the probit model regresses this variable onto a linear model with standard normal errors ε:

$$Y = \beta'\mathbf{x} + \varepsilon, \quad \text{with } \varepsilon \sim N(0,1), \tag{10}$$

where $\mathbf{x}$ is the vector of the explanatory variables (also referred to independent or exogeneous variables) with error ε. The row vector β contains the coefficients to be estimated and the first parameter, which is β_0, is usually taken as the intercept by setting $x_0 = 1$. Now let Z be a discrete random variable which represents the actual observed states, then $Z = 0$ if $Y \leq \tau$ and $Z = 1$ otherwise. The threshold τ is a model parameter and must be determined by the estimation procedure. For the ordered probit model with $n+1$ states it follows from

$$Z = \begin{cases} 0 & \text{if } Y \leq \tau_1, \\ 1 & \text{if } \tau_1 < Y \leq \tau_2, \\ \vdots & \\ n-1 & \text{if } \tau_{n-1} < Y \leq \tau_n, \\ n & \text{if } \tau_n < Y. \end{cases} \tag{11}$$

Because $Y \in (-\infty, \infty)$ for all $k = 1, 2, \ldots$, the thresholds τ and $\tau_i, i = 1, 2, \ldots$ for the state conditions must be located between $-\infty$ and ∞. Note that the thresholds do not have to be equidistant. From these relationships, the probability of each observation can be determined. For the binary probit model this is simply $\Pr\{Z = 1 \,|\, \mathbf{x}\} = \Pr\{Y > \tau \,|\, \mathbf{x}\}$, where $\Pr\{Y > \tau\} = \Pr\{\beta'\mathbf{x} + \varepsilon > \tau\} = \Pr\{\varepsilon > \tau - \beta'\mathbf{x}\} = 1 - \Phi(\tau - \beta'\mathbf{x})$. Here $\Phi(x)$ is the cumulative standard normal distribution function. The notation $\Pr\{Z = 1 \,|\, \mathbf{x}\} = \Phi(\beta'\mathbf{x} - \tau)$ is also often used, which is equivalent as the normal distribution is symmetric with $\Phi(-x) = 1 - \Phi(x)$. Obviously $\Pr\{Z = 0 \,|\, \mathbf{x}\} = 1 - \Pr\{Z = 1 \,|\, \mathbf{x}\}$. Similarly, for the ordered probit model, the probabilities of observing each condition state is given by

$$\begin{aligned} \Pr\{Z = 0 \,|\, \mathbf{x}\} &= \Phi(\tau_1 - \beta'\mathbf{x}) \\ \Pr\{Z = 1 \,|\, \mathbf{x}\} &= \Phi(\tau_2 - \beta'\mathbf{x}) - \Phi(\tau_1 - \beta'\mathbf{x}) \\ &\vdots \\ \Pr\{Z = n-1 \,|\, \mathbf{x}\} &= \Phi(\tau_n - \beta'\mathbf{x}) - \Phi(\tau_{n-1} - \beta'\mathbf{x}) \\ \Pr\{Z = n \,|\, \mathbf{x}\} &= 1 - \Phi(\tau_n - \beta'\mathbf{x}) \end{aligned}$$

Under the assumption that the observations are independent, the likelihood function for the coefficients β and the thresholds τ, given the observations $\mathbf{z}$ and the explanatory variables $\mathbf{x}_k$ is simply

$$L(\beta, \tau \,|\, \mathbf{z}) = \prod_k \Pr\{Z = z_k \,|\, \mathbf{x}_k\}, \tag{12}$$

which can be maximized to estimate the unknown coefficients and thresholds. Before doing so, the model must be 'identified' by setting either the intercept β_0 or one of the thresholds τ_i equal to zero or some other constant. Fixing either the intercept or one of the thresholds will influence the other, but not the probability of the outcome z_k; see (Long, 1997, pp.122–123).

The logit model takes the same approach as the probit model, but assumes that the errors have a standard logistic distribution. Like the standard normal distribution, the standard logistic distribution is symmetric, but has a slightly larger variance.

The ordered probit model is generally not used to estimate transition probabilities in a Markov process, but Madanat et al. (1995) made some assumptions in order to apply this method to a Markov chain for modeling bridge deterioration. The first assumption is that the Markov chain is progressive (sometimes also referred to as monotonic, which means that it only proceeds in one direction). Then, the observations z_k are assumed to be the number of transitions between two consecutive inspections: $z_k = j - i$. A different Z is defined for each row except for the last in the transition probability matrix, thus allowing for different deterioration mechanisms in each (transient) state. Therefore, $\Pr\{Z_i = z\} = \Pr\{X(1) = i + z \,|\, X(0) = i\}$ for $i =$

$0, 1, \ldots, n-1$ and the authors introduce additional notation to allow the transition probabilities to be estimated for each individual bridge. Also, Madanat et al. (1995) use a log-linear model instead of the linear model in (10) to ensure that the unobserved condition is non-negative: $\log(y_k) = \beta' \mathbf{x}_k + \varepsilon_k$. Then, the latent variable Y has a lognormal distribution with support $[0, \infty)$ and the thresholds for the condition states are also within this range. The software that the authors have used for the estimation, identified the model by setting the first threshold equal to 0, which corresponds to setting $\log(\tau_1) = 0 \Rightarrow \tau_1 = 1$. This model was later extended by Madanat et al. (1997) to a random effects model by the inclusion of another error term to reflect the differences (heterogeneity) between structures.

The approach suggested by Madanat et al. (1995), which was later applied by Baik et al. (2006) to the problem of modeling deterioration of wastewater systems, has a number of shortcomings. In what they see as an advantage, the option to estimate transition probability matrices for individual bridges requires a significant amount of inspection data and the suggested averaging of transition probabilities to obtain transition matrices for groups of bridges is faulty. Transition probabilities for groups of bridges should be directly estimated using the inspection data from all bridges within the group and not by averaging the transition probabilities of the individual bridges. A more fundamental shortcoming is related to the dependence of transition probabilities on bridge ages. The authors state that 'the transition probabilities are explicitly ... nonstationary', because they are a function of time or the age of the bridge. The truth is that the aspect of time is included as an explanatory variable in the linear model and it is used to estimate a transition probability matrix of a stationary Markov chain. For example, take

$$Y(t) = \beta_0 + \beta_1 t + R, \text{ with } R \sim N(0,1) \tag{13}$$

as a simple model to describe the uncertainty in deterioration over time t. The probability of no transition between time $t_0 = 0$ and the first inspection at time t_1 is given by the probability that the amount of deterioration at time t_1 has not exceeded the first threshold τ_1: $\Pr\{Z = 0 \,|\, t_1\} = \Pr\{Y(t_1) \leq \tau_1\}$. However, this probability is taken as the probability p_{00} of no transition out of the initial state 0 during a unit time. Subsequently, p_{01} is the probability that the amount of deterioration is somewhere between τ_1 and τ_2, p_{02} that it is somewhere between τ_2 and τ_3, etc. Therefore, each transition probability in a row of the transition probability matrix is related to a different age, but they are used in a transition matrix for a single unit time which is used to model transitions at all ages.

Under the assumption that the Markov chain is sequential (monotonic and without skipping a state), Bulusu and Sinha (1997) proposed to use a binary probit model for fitting the Markov chain to inspection data. A restrictive requirement for the application of this model is that only one transition occurs during the time between two successive inspections. The problems previously described for the probit model suggested by Madanat et al. (1995) are further aggrevated by the inclusion of a binary random variable in the linear model from (10), which equals one if a transition took place in the previous inspection interval and zero otherwise. This attempt at incorporating time dependence into the model, directly violates the Markov property, which must hold if a Markov chain is used.

Similar to Bulusu and Sinha (1997), Yang et al. (2005) apply the same model, but without the extra binary variable and the errors R are assumed to have a logistic distribution. This is therefore a binary logit model as described earlier. The authors refer to this approach as 'logistic regression', although this terminology is also used by some to refer to regression with a log-linear model.

3.3 Other methods

Other methods which will not be further treated here are non-parametric models and estimation methods using a Bayesian approach. An example of the application of the non-parametric Kaplan-Meier estimator to bridge condition states is presented in DeStefano and Grivas (1998). The problem of estimating the transition probabilities in the multinomial model (see Sec. 3.2.2), using a Bayesian approach, is described in Lee et al. (1970). Bulusu and Sinha (1997) compare the application of this approach to the application of a binary probit model to bridge deterioration modeling. Finally, Micevski et al. (2002) use a different formulation of the Bayesian likelihood function for the estimation of transition probabilities in a Markov model for the condition of storm waterpipes.

4 IGNORING THE DEPENDENCE BETWEEN TRANSITIONS

We have seen that a number of estimation models ignore the dependence between transitions. The following simple example shows how this may give very different results. Take a discrete-time Markov chain $\{X(t), t \geq 0\}$ with transition probability matrix

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \end{bmatrix},$$

where we assume that the process starts in state 2 and ends in state 0. Figure 3 shows how this process is observed at times t=0, 1 and 4 to be in states 2, 1 and 1 respectively. A model which ignores the dependence between successive states is a model which ignores the fact that the probability of being in state 1 at time $t = 4$ depends on the process previously being in state 1 at time $t = 1$.

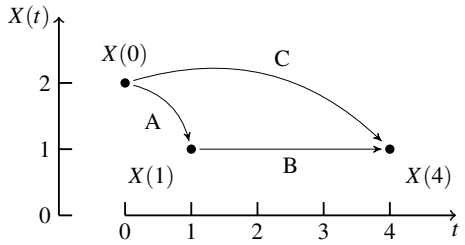

Figure 3. Example of a discrete-time Markov chain $X(t)$ observed at times $t = 1, 2$ and 5.

The probability of the individual transitions A, B and C is 0.5, 0.125, and 0.8125 respectively. A model which treats the transitions as dependent, would assign a likelihood of $\Pr\{A\} \cdot \Pr\{B\} \approx 0.1$, whereas a model which ignores this dependence would result in a likelihood of $\Pr\{A\} \cdot \Pr\{C\} \approx 0.4$. In other words, the latter model would assign a likelihood approximately four times greater to this particular sequence of state observations. What we see is that there are many possible sequences which result in transition C and the sequence formed by transitions A and B is just one of these.

5 SUMMARY AND CONCLUSIONS

Since the 1980's, there have been many different models and methods which have been applied for the estimation of Markov chain models in civil engineering problems. Most of these have been reviewed here and references to publications describing their application are given. Because visual inspections almost always rate the quality of structures on a discrete scale, Markov chains have been the traditional choice for representing the uncertain progression of these structures through these condition scales.

A direct comparison of the performance of these models and methods is not possible as they operate under different assumptions. The primary difference between each of these methods is the type of information that is available to the modeller. Three types of data may be distinguished:

Type I: observations of the state itself and represented by realizations $x(t)$ of the process $X(t)$,

Type II: aggregated data in the form of relative fractions of proportions represented by $y(t)$, and

Type III: count data in the form of the number of transitions represented by realizations $n(t)$ of some counting process $N(t)$.

Although they have been used for the purpose of estimating transition probabilities in Markov chains, both the multinomial model described in Sec. 3.2.2 and the (ordered) probit or logit models described in Sec. 3.2.4 are not suited for this purpose. The multinomial model requires continuous monitoring of the structure and even then it is difficult to determine when a transition actually takes place, because the interpretation of the different condition catergories is subjective. The probit and logit models are suited for categorial data, but their application to data generated by a Markov chain is unnatural.

Ideally, an estimation model will account for the fact that successive observations are dependent. In Sec. 4 we have given an example which shows the effect of ignoring this dependence. These are the regression model represented by the objective function in Eq. (7) and the maximum likelihood methods described in Sections 3.2.1 and 3.2.3. All three of these approaches represent the three different types of data that may be available to the modeller, namely types II, III and I respectively. The two other regression models represented by the objective functions in Eqs. (5) and (6) do not explicitely account for this dependence. They are therefore less suited for the purpose of estimating transition probabilities or intensities.

With the storage capability and the flexibility of today's databases, it is possible to store the actual states of individual structures after each inspection. Aggregating this kind of detailed information would be a shame and therefore there is no reason not to use a statistical estimation method that can cope with this kind of information. In this respect, the maximum likelihood-based methods for panel and count data described in Sections 3.2.1 and 3.2.3 are most suited for this purpose. In general, maximum likelhood methods allow for great flexibility in model building. Transitions probabilities or intensities may depend on covariates which may include time- or age-dependency.

In practical situations, the data will include a lot of noise. This may be due to typing or other kinds of errors made during the registration of inspection results. This may also be due to differences of opinion between inspectors. Together with the omission of maintenance activities in a database, the latter will result in quite a few condition improvements which can not be accounted for. We are primarily interested in estimating the rate at which a structure progresses through condition states due to aging and usage. It is therefore advised to include a flag in the database, which

the inspector may use to identify the (suspected) reason for the condition improvement. He may suspect that maintenance was performed since the previous inspection or he may disagree with the condition evaluated at the previous inspection. Such a flag makes it easier for a modeller to filter out data which is not of interest to him.

REFERENCES

Abaza, K.A., Ashur, S.A., & Al-Khatib, I.A., 2004. Integrated pavement management system with a Markovian prediction model. *Journal of Transportation Engineering*, 130(1):24–33.

Anderson, T.W. & Goodman, L.A., 1957. Statistical inference about Markov chains. *Annals of Mathematical Statistics*, 28:89–110.

Baik, H.S., Jeong, H.S., & Abraham, D.M., 2006. Estimating transition probabilities in Markov chain-based deterioration models for management of wastewater systems. *Journal of Water Resources Planning and Management*, 132(1):15–24.

Billingsley, P., 1961. Statistical models in Markov chains. *Annals of Mathematical Statistics*, 32:12–40.

Bulusu, S. & Sinha, K.C., 1997. Comparison of methodologies to predict bridge deterioration. *Transportation Research Record*, 1597:34–42.

Butt, A.A., Shahin, M.Y., Feighan, K.J., & Carpenter, S.H., 1987. Pavement performance prediction model using the Markov process. *Transportation Research Record*, 1123:12–19.

Cameron, A.C. & Travedi, P.K., 1998. *Regression analysis of count data*. Cambridge University Press, Cambridge, United Kingdom.

Carnahan, J.V., Davis, W.J., Shahin, M.Y., Keane, P.L., & Wu, M.I., 1987. Optimal maintenance decisions for pavement management. *Journal of Transportation Engineering*, 113(5):554–572.

Cesare, M., Santamarina, J.C., Turkstra, C.J., & Vanmarcke, E., 1994. Risk-based bridge management: optimization and inspection scheduling. *Canadian Journal of Civil Engineering*, 21(6):897–902.

DeStefano, P.D. & Grivas, D.A., 1998. Method for estimating transition probability in bridge deterioration models. *Journal of Infrastructure Systems*, 4(2):56–62.

Kallen, M.J. & van Noortwijk, J.M., 2006. Statistical inference for Markov deterioration models of bridge conditions in the Netherlands. In Cruz, P.J.S., Frangopol, D.M., & Neves, L.C., (eds.), *Bridge Maintenance, Safety, Management, Life-Cycle Performance and Cost: Proceedings of the Third International Conference on Bridge Maintenance, Safety and Management, 16-19 July, Porto, Portugal, 2006.*, pages 535–536, London. Taylor & Francis.

Lee, T.C., Judge, G.G., & Zellner, A., 1970. *Estimating the parameters of the Markov probability model from aggregate time series data*. North-Holland, Amsterdam.

Long, S.J., 1997. *Regression models for categorical and limited dependent variables*. Sage Publications, Thousand Oaks, CA.

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

Madanat, S. & Wan Ibrahim, W.H., 1995. Poisson regression models of infrastructure transition probabilities. *Journal of Transportation Engineering*, 121(3): 267–272.

Madanat, S.M., Karlaftis, M.G., & McCarthy, P.S., 1997. Probabilistic infrastructure deterioration models with panel data. *Journal of Infrastructure Systems*, 3(1):4–9.

Micevski, T., Kuczera, G., & Coombes, P., 2002. Markov model for storm water pipe deterioration. *Journal of Infrastructure Systems*, 8(2):49–56.

Morcous, G., 2006. Performance prediction of bridge deck systems using Markov chains. *Journal of the Performance of Constructed Facilities*, 20(2):146–155.

FHWA, 1995. *Recording and coding guide for the structure inventory and appraisal of the nation's bridges. Report No. FHWA-PD-96-001*. U.S. Department of Transportation, Federal Highway Administration, Washington, D.C.

Wirahadikusumah, R., Abraham, D., & Iseley, T., 2001. Challenging issues in modeling deterioration of combined sewers. *Journal of Infrastructure Systems*, 7(2): 77–84.

Yang, J., Gunaratne, M., Lu, J.J., & Dietrich, B., 2005. Use of recurrent Markov chains for modeling the crack performance of flexible pavements. *Journal of Transportation Engineering*, 131(11):861–872.