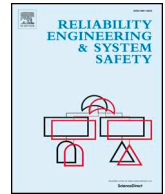




Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress



Multi-defect modelling of bridge deterioration using truncated inspection records

Gareth Calvert^{a,*}, Luis Neves^a, John Andrews^a, Matthew Hamer^b

^a Resilience Engineering Research Group, University of Nottingham, Nottingham, United Kingdom

^b Network Rail, Milton Keynes, United Kingdom

ARTICLE INFO

Keywords:

Bridge deterioration
Bridge management
Infrastructure
Multi-defect
Railway

ABSTRACT

Bridge Management Systems (BMS) are decision support tools that have gained widespread use across the transportation infrastructure management industry. The Whole Life Cycle Cost (WLCC) modelling in a BMS is typically composed of two main components: a deterioration model and a decision model. An accurate deterioration model is fundamental to any quality decision output.

There are examples of deterministic and stochastic models for predictive deterioration modelling in the literature, however the condition of a bridge in these models is considered as an ‘overall’ condition which is either the worst condition or some aggregation of all the defects present. This research proposes a predictive bridge deterioration model which computes deterioration profiles for several distinct deterioration mechanisms on a bridge.

The predictive deterioration model is composed of multiple Markov Chains, estimated using a method of maximum likelihood applied to panel data. The data available for all the defects types at each inspection is incomplete. As such, the proposed method considers that only the most significant defects are recorded, and inference is required for the less severe defects. A portfolio of 9726 masonry railway bridges, with an average of 2.47 inspections per bridge, in the United Kingdom is the case study considered.

1. Introduction

The functional operation of transportation networks is contingent on diverse asset portfolios including civil infrastructure. The railway in Great Britain includes over 26,000 bridges, which are constructed out of many different materials and are of various ages. Network Rail (NR) is the infrastructure asset manager for the railway network in England, Scotland, and Wales. Consequently, NR is responsible for the inspection, assessment, maintenance, and repair of this portfolio of bridges.

Ensuring the bridges are maintained to a suitable safety threshold is critical, as the consequence of structural failures would be enormous. The risk of structural failure is reduced by performing inspections and maintenance as per the industry guidelines [1]. However, it is also a significant challenge to achieve this within the budgetary constraints placed on railway infrastructure managers. To ensure that decisions made result in the optimal Whole Life Cycle Cost (WLCC), a modelling approach is employed. A decision support tool known as a Bridge Management System (BMS) is used to perform any required modelling and to prioritise and justify decisions to the regulator.

Any WLCC analysis can be roughly divided into two parts: a

deterioration model and a decision model [2]. The deterioration model is responsible for predicting the future condition of a bridge asset, whilst a decision model supports the decision making process in the development of maintenance strategies. The appropriateness of any decisions and the accuracy of cost outputs for maintenance strategies will be inherently flawed if the deterioration model does not have a sufficient prediction accuracy. This work presents a method to compute predictions of the deterioration profiles of multiple deterioration mechanisms opposed to the current practice of a single overall condition index. Moreover, existing condition records do not have observed conditions for all the deterioration modes at each inspection, and thus a score inference technique is developed to overcome this limitation of the data, so that transition rates between conditions can be estimated. A method of maximum likelihood applied to panel data, based on the works of Kalbfleisch and Lawless [3], and Kallen and van Noortwijk [4] is used to estimate parameters for the deterioration rates of each of the defects.

* Corresponding author.

E-mail address: gareth.calvert@nottingham.ac.uk (G. Calvert).

<https://doi.org/10.1016/j.ress.2020.106962>

Received 8 May 2019; Received in revised form 20 March 2020; Accepted 23 March 2020

Available online 26 March 2020

0951-8320/ © 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2. Bridge asset management

There are a variety of deterministic and stochastic modelling approaches for bridge asset management but Frangopol et al. [2] argued that stochastic modelling is more advantageous than all other modelling approaches for structural deterioration. Deterministic models can be used to predict values for measurable quantities [5], which is a difficult task to scale to large asset portfolios. Additionally, deterministic models provide limited scope to investigate the effects of uncertainty, and consequently these models are often used to predict the outcomes of 'worst case' scenarios [6]. Several factors that influence the safety of a structure have been parametrised using probabilistic frameworks opposed to deterministic alternatives [7,8]. Moreover, stochastic modelling has the capability to incorporate the inherent randomness of structural deterioration.

To model the degradation of bridge condition, there are two main sources of data: condition records from examinations and maintenance records, which outline the time between maintenance interventions. The use of maintenance records allows lifetime analysis and addresses the concerns of subjectivity of condition indices. However, maintenance data is often sparse and of poor quality [9], making its use unsuitable for many infrastructure managers.

The use of condition records from bridge inspections to model deterioration and estimate transition rates is more common [10]. As bridges degrade at a slower rate in comparison to other railway assets, for example mechanical assets, it is common for infrastructure managers to have considerably more condition records than maintenance records. Caution should be taken when using any condition index as it does not necessarily reflect the integrity of a load bearing structural element [11–13], although Frangopol and Liu [14] argue that maintenance interventions should typically be prioritized to civil infrastructure with unacceptable and poor condition rating levels. Note that the notion of a scale of 'overall' or 'worst' conditions are commonly defined for either data type for the purposes of modelling.

A performance study of 5700 bridges in Indiana, USA [15], is often cited as an early example of a Markov-based model being used for bridge deterioration modelling. Further studies made use of Markovian methods for predictive bridge deterioration modelling of a single bridge component or a whole bridge [16–18]. The Markov Chain approach is recognised as the most popular technique for stochastic deterioration models [2,19]. Many industry leading BMSs make use of the Markov chain approach for the predictive deterioration modelling in each system, including: AASHTOWare Bridge Management (formerly *Pontis*) developed in the United States; KUBA, used in Switzerland; Ontario Bridge Management System (OBMS) used in Ontario, Canada and Quebec Bridge Management System (QBMS) used in Quebec, Canada [20,21].

Other modelling frameworks have been used for the modelling of bridge deterioration: Semi-Markov [22–25], Petri Nets (PNs) [26–29] and Bayesian Belief Networks (BBN) [30]. Time-based approaches have been presented to obtain probability distribution parameters from historic condition data, including non-constant deterioration rates [31–35]. However, since bridges are inspected at intervals, rather than continuous monitoring, the exact transition event is not observed and hence assumptions are required to determine sojourn times for each condition state.

Markov chains do have limitations, including: the assumption of constant transition rates and the model size increasing exponentially with the increasing number of states [36]. Additionally, discrete condition states, as used in Markov chains, present cases which require expert judgment for the appropriate state classification, which can be subjective [2]. Moreover, when using condition records there is difficulty in ascertaining the effects of maintenance and the appropriate inclusion of records which exhibit an increase in condition [37]. Nevertheless, Morcous [18] observes that they are the most common stochastic technique used for bridge deterioration modelling, stating

that Markov chain models are implemented for their ability to predict future condition given the uncertainty of the deterioration process and their computational efficiency. In spite of the limitations of Markov chains, they are the most appropriate given the constraints of the industrial longitudinal data considered in this study, which does not record all the defect states and only covers a fraction of the structures lifetime.

In this study, Markov chain models are defined for bridge components, which can be used to predict lifetime indicators for overall bridges. These indicators for overall bridges can then be used to inform network/portfolio modelling. This approach enables the evaluation of a large variety of structural configurations, which may exist across a bridge portfolio. Nonetheless, it should be noted that there have been several studies performed which analyse maintenance scheduling at the network level using generalised bridge lifetime indicators [38,39].

Bridges can degrade under a variety of different deterioration mechanisms, consequently any scale that seeks to consolidate the different deterioration modes into one condition index will have a level of subjectivity and arbitrariness. Ceravolo et al. [40] proposed 'symptom-based' reliability models to overcome the limitations of ad hoc reliability indexes and to incorporate engineering knowledge gained from structural monitoring activities. However, such empirical measurements are often not available across large, diverse asset portfolios.

This research introduces a method to model the multiple different deterioration modes, such that more comprehensive predictions of bridge deterioration can be made. By modelling multiple deterioration modes simultaneously, it should lead to an improved accuracy in any predictive output. Additionally, by computing predictive deterioration profiles for each defect type, it would then be possible to produce a decision model that tests maintenance strategies based on particular defect types rather than the 'traditional' ambiguous repair actions (e.g. minor repair, major repair and replacement).

Defects can be grouped into different modes as they are different processes and impact the bridge component in different ways, e.g. mortar, brick surface, interactions between the bricks and the infill. The practical implication of multi-defect modelling is that it should not only facilitate more accurate and engineering based models but also provide models that can indicate specific future maintenance needs given the context of all the defects' extensiveness revealed at the most recent inspection.

The purpose of a model can be broadly categorised into three classes: Generator, Mediator and Predictor [41]. A generator model is used to generate hypotheses, mediator models are employed to make comparisons between competing strategies, and predictor models are used when a system is well understood and can provide accurate insights into future bridge condition states [42]. In bridge management, a mediator model can be used to investigate the benefits of different maintenance strategies. A predictor model can compare different maintenance strategies but could also affix accurate costings to any output, which is an objective of infrastructure asset managers. The prediction of multiple deterioration mechanisms enables the development of a decision model that could apply specific maintenance actions and consequently the ability to affix more accurate maintenance costs. Thus, a multi-defect deterioration model facilitates the development of a decision model which could be described as a predictor model.

It should be noted that the service life of civil infrastructure is characterised in part by the effects of progressive deterioration and sudden deterioration [43]. Progressive deterioration describes the development of various defect mechanisms and sudden deterioration is the result of hazards such as earthquakes, fires and floods, amongst others. The models described in this study are used to predict progressive deterioration behaviour and they do not model sudden deterioration outright. Nonetheless, the modelling of distinct defect mechanisms permits the evaluation of how vulnerable a structure may be to sudden deterioration.

Table 1
SevEx severity definitions for masonry bridge elements.

Score	Severity definitions
A	No visible defects to masonry. (Cracks are scored separately.)
B	Brickwork - depth of spalled and weakened/softened material < 10 mm. Stonework - depth of spalled and weakened/softened material < 20 mm. Or any evidence of the presence or effect of water (defined as percolation, run-off, etc).
C	Deterioration of pointing. (Record the maximum and typical depth lost.)
D	Brickwork - depth of spalled and weakened/softened material ≥ 10 mm but less than the depth of a header. Stonework - depth of spalled and weakened/softened material ≥ 20 mm but less than the depth of a block.
E	Hollowness/Drumming. (Not associated with B or D.)
EX	Includes all incidences of: loose/wedged bricks/blocks - not displaced, loose/wedged bricks/blocks - displaced but not to the full depth of the structural element or missing brick/blocks - one or more, but not to the full depth of the structural element.
F	Choose most extensive from: bulging, distortion tilting (vertical alignment), displacement: loose and/or wedged displaced bricks/blocks to the full depth of the element or missing brick/block to the full depth of the element

3. Inspection records

NR manage a diverse range of bridge assets in their portfolio, with each bridge denoted as having a primary material which includes: concrete, metal, masonry, timber and composites. Masonry bridges are the case study considered as they are the most populous type in the NR portfolio. NR use an alpha-numeric scoring system called Severity Extent (SevEx) to record the observed condition of a bridge during a detailed examination. The definition for each severity score for a masonry bridge element aligns with a different defect mode. The possible defects that can be observed are: shallow spalling, deterioration of pointing, deep spalling, hollowness/drumming, loose or missing block work from the surface of bridge element and displaced or missing blockwork to the full depth of the element. The extent score details the coverage of the observed defect on the bridge element. Full definitions for the severity and extent scores can be found in Tables 1 and 2 respectively.

At inspection, each element is assessed and any defects can be assigned an alpha-numeric score on the SevEx scale. All the possible masonry SevEx scores range from A1 (no visible defect) to F6 (over 50% of the element surface having displaced or missing blocks), forming

$$C = \{A1, B2, \dots, B6, C2, \dots, C6, \dots, D2, \dots, D6, E2, \dots, E6, EX2, \dots, EX6, F2, \dots, F6\}. \quad (1)$$

At each inspection, for each bridge element, the two ‘worst’ scores are recorded alongside the date of the inspection. When several inspections have taken place on the same bridge, the repeated observations form a longitudinal study. The longitudinal study can be split into ‘exam pairs’ and stated in the following format for n exam pairs,

Record	Inspection 1		Inspection 2		Time interval
	SevEx 1	SevEx 2	SevEx 1	SevEx 2	(Years)
r_i	$\alpha_{i,1}$	$\beta_{i,1}$	$\alpha_{i,2}$	$\beta_{i,2}$	t_i

where $i = 1, \dots, n$ and $\alpha_{i,j}$ is the worst score at the j^{th} inspection, $\beta_{i,j}$ is the second worst score at the j^{th} inspection, $\alpha_{i,j}, \beta_{i,j} \in C$ and $t_i \in \mathbb{R}$, $0 < t < 18$. This process is repeated for all bridge elements that have had multiple inspections and the resulting tables are pooled into one data set.

The historic condition records are filtered to only include records which exhibit stationary behaviour or deterioration, i.e. $\{\alpha_{i,1},$

Table 2
SevEx extent definitions for masonry bridge elements.

Score	Extent definitions
1	No visible defects to masonry (cracks are scored separately).
2	Localised defect due to local circumstances (such as mechanical damage).
3	Defect occupies less than 5% of surface of the structural element.
4	Defect occupies 5% to 10% of the surface of the structural element.
5	Defect occupies 10% to 50% of the surface of the structural element.
6	Defect occupies more than 50% of surface of the structural element.

$\beta_{i,1}\} \leq \{\alpha_{i,2}, \beta_{i,2}\}$. Pairs of inspections that exhibit extreme cases of deterioration are also omitted under the assumption that they are the result of non-standard behavior, for example, bridge strikes, fires etc.

3.1. Multi-defect condition states

To leverage the SevEx condition scale for any potential multi-defect model, some adaptations to the scale are required. The SevEx scale has the overall state of no defect defined as A1, whereas for multi-defect modelling a no defect state is required for each defect i.e. $A1 \rightarrow \{B1, C1, D1, E1, EX1, F1\}$. Thus, the SevEx conditions states for each defect type are as follows: $\mathcal{B} = \{B1, \dots, B6\}$, $\mathcal{C} = \{C1, \dots, C6\}$, $\mathcal{D} = \{D1, \dots, D6\}$, $\mathcal{E} = \{E1, \dots, E6\}$, $\mathcal{EX} = \{EX1, \dots, EX6\}$ and $\mathcal{F} = \{F1, \dots, F6\}$. The inspection records for a multi-defect panel should be in the format as shown in Table 3, for all n records ($i = 1, \dots, n$), where B_{ij} is the condition of Defect B in the j^{th} inspection of the i^{th} record with $B_{ij} \in \mathcal{B}$. C_{ij} , D_{ij} , E_{ij} , EX_{ij} and F_{ij} are described similarly. However, at each inspection only the two worst severity scores are recorded.

3.2. Score ranking

The current format of inspection records contains the two ‘worst’ scores; for any score inference, a definition of what ranking is used to determine ‘worst’ scores is required. However, how the worst scores are defined is unclear, and consequently two candidate rules were considered:

- **Rule One:** The SevEx scores are ranked according to severity score, followed by the extent score, and thus the rule has the following order of precedence: $A1 < B2 < \dots < B6 < \dots < F2 < \dots < F6$.
- **Rule Two:** The SevEx scores have a numerical weight which is used in a Bridge Condition Marking Index (BCMI) calculation. Using the BCMI weight a 1D integer condition scale could be created to rank the different SevEx scores. The integer scale value for each SevEx score is shown in Table 4. Under this ranking, there are still possible cases where a tie-break rule would need to be developed.

Documentation compiled by NR describing the condition scores of bridges used at examinations states that the two most severe defects should be recorded at each inspection, and that the same severity rating can not be used more than once [44]. Moreover, the guidance for bridge inspectors states that with an ageing masonry bridge stock it would be rarely appropriate to categorise a bridge element as A1, and thus a minor defect, even if it is with little structural significance, should be reported. After consultation with NR bridge engineers it was determined that the bridge inspectors adhere to rule one when determining the two worst scores to record for a bridge element at inspection. The use of the rule one avoids the arbitrary conversion weightings which rule two relies on. Moreover, the proposed methodology is applicable to any score ranking method.

Table 3Multi-Defect score panel format, where $B_{i,j}$ is the condition of Defect B in the j th inspection.

Record	Inspection 1						Inspection 2						Interval (Years)
	B	C	D	E	EX	F	B	C	D	E	EX	F	
r_i	$B_{i,1}$	$C_{i,1}$	$D_{i,1}$	$E_{i,1}$	$EX_{i,1}$	$F_{i,1}$	$B_{i,2}$	$C_{i,2}$	$D_{i,2}$	$E_{i,2}$	$EX_{i,2}$	$F_{i,2}$	t_i

Table 4

Network Rail assigned integer condition scale score for each SevEx score.

Severity	Extent					
	1	2	3	4	5	6
B	1	2	2	3	4	5
C	1	4	4	5	6	7
D	1	7	7	8	11	12
E	1	7	9	11	13	14
EX	1	7	9	11	13	14
F	1	7	10	12	14	15

3.2.1. Multi-defect score inference

Score inference can be performed on the historic inspection records, to express the recorded score panel as a multi-defect inspection panel as fully as possible. Recall that for the i th record, the worst score recorded at inspection j is given by $\alpha_{i,j}$, and the second worst score is denoted by $\beta_{i,j}$. Then, $Sev(\alpha_{i,j})$ denotes the severity score of the SevEx score of $\alpha_{i,j}$, similarly $Ex(\alpha_{i,j})$ denotes the extent score of the SevEx score of $\alpha_{i,j}$.

If the score panel at inspection is reported as, $\alpha_{i,j} = A1$ and $\beta_{i,j} = A1$ then the multi-defect panel will be $\{B1, C1, D1, E1, EX1, F1\}$. In the situation that a score panel of $\beta_{i,j} = A1$ and $\alpha_{i,j} \neq A1$ is reported at inspection, the multi defect panel would have five defects that have an extent of one, the severity score of these five would be all γ that satisfy, $(\gamma \in \{B, C, D, E, Ex, F\}) \wedge (Sev(\alpha_{i,j}) \neq Sev(\gamma))$. (2)

For the one defect that has an extent score greater than one, which is $Sev(\alpha_{i,j})$, the extent score would be $Ex(\alpha_{i,j})$.

Finally, if an inspection is recorded such that $\alpha_{i,j} \neq A1$ and $\beta_{i,j} \neq A1$, it is still possible to make some assertions on the unobserved defects. The score inference relies on the assumption that, if an inspection score panel does not contain a high severity score, it must be due to the high severity defect being absent, otherwise the bridge examiner should have recorded the presence of that defect instead of the lower severity defect. The ranking rule selected states that: $B < C < D < E < EX < F$. Consider the i th record at the j th inspection, where $\alpha_{i,j}$ and $\beta_{i,j}$ are known SevEx scores and they have two different severity scores, i.e. $Sev(\alpha_{i,j}) \neq Sev(\beta_{i,j})$. If a candidate score value is denoted as $\gamma_{i,j}$, then there are four possible severity scores that were not recorded at inspection but could possibly be inferred, i.e.

$$(Sev(\gamma_{i,j}) \notin \{Sev(\alpha_{i,j}), Sev(\beta_{i,j})\}) \wedge (Sev(\gamma_{i,j}) \in \{B, C, D, E, EX, F\}). \quad (3)$$

For each of the four severity values $\gamma_{i,j}$ can assume, an attempt of inferring the extent score can be sought by using the following inference rules:

$$\text{If } Sev(\gamma_{i,j}) > Sev(\alpha_{i,j}), \text{ then } Ex(\gamma_{i,j}) = 1. \quad (4)$$

$$\text{If } Sev(\gamma_{i,j}) < Sev(\alpha_{i,j}) \text{ and } Sev(\gamma_{i,j}) > Sev(\beta_{i,j}), \text{ then } Ex(\gamma_{i,j}) = 1. \quad (5)$$

Alternatively, if $Sev(\gamma_{i,j}) < Sev(\alpha_{i,j})$ and $Sev(\gamma_{i,j}) < Sev(\beta_{i,j})$ then $Ex(\gamma_{i,j})$ is unknown, otherwise $Ex(\gamma_{i,j}) = 1$. In other words, a more severe defect is not recorded if it is absent, while a less severe defect could be absent independent of its extent score.

3.2.2. Inference examples

Consider the example panel data in Table 5, which can be explicitly defined as a multi-defect panel using the score inference rules, shown in Table 6. Both Records 1 and 2 can be explicitly defined as multi-defect panels, with Record 1 using the first score inference rule, (4) and Record 2 using both score inference rules, (4), (5). From Records 3 and 4 it

Table 5

Example bridge inspection panel data using NR's SevEx condition scale.

Record	Exam Inspection 1		Inspection 2		Time interval (Years)
	$\alpha_{i,1}$	$\beta_{i,1}$	$\alpha_{i,2}$	$\beta_{i,2}$	
1	C3	B2	C5	B3	6.34
2	E2	B3	E3	B6	6.12
3	EX2	D3	EX3	E4	5.79
4	F2	EX4	F2	EX5	7.26

Table 6

Inferred multi-defect panel, using the score inference rules on the example bridge inspection panel data from Table 4. Dashes denote unknown scores.

Record	Inspection 1						Inspection 2						Time interval (Years)
	B2	C3	D1	E1	EX1	F1	B3	C5	D1	E1	EX1	F1	
1	B3	C1	D1	E2	EX1	F1	B6	C1	D1	E3	EX1	F1	6.34
2	—	—	D3	E1	EX2	F1	—	—	—	E4	EX3	F1	6.15
3	—	—	—	—	EX4	F2	—	—	—	—	EX5	F2	5.79
4	—	—	—	—	—	—	—	—	—	—	—	—	7.26

can be observed that the multi-defect panel may not always be explicitly defined. For example, for inspection 1 of record 3, defects with severity scores D and EX were found. Consequently, any defect with severity score B or C would be excluded for being less severe, independently of extent. In the cases where a multi-defect panel is not explicitly defined, the inspection pair for a severity score can only be used in any data analysis if an extent score exists for the defect type at both inspections.

Generally, the 'lower' severity scores will become unobserved when the bridge element exhibits 'higher' severity scores. Thus, any model will make the assumption that the rate of deterioration estimated for the lower severity scores which were observed continues to hold true when they become unobserved. As current industry practice is to base maintenance scheduling models off of the worst score, this is seen as a reasonable assumption given the data available. However, in the future, NR intend to record inspections by tracking particular defects by a unique identifier. This updated recording regime will make the whole multi-defect score panel observable all of the time and result in the score inference rules and deterioration behaviour assumption no longer being required.

3.3. Merging of extent scores

Whilst analysing the inspection records, it became clear that an extent score of 2, i.e. $\{B2, C2, D2, E2, EX2, F2\}$, was underutilised by bridge examiners. It was also apparent that the low number of observed records with extent score equal to 2 was common across all severity scores. The under reporting of this score could be due to the sojourn time of this condition state being considerably shorter than any inspection interval.

However, it was determined that a more likely explanation was the fact that an extent score of 2 and 3 are very similarly defined; extent score 2 is defined as, 'Localised defect due to local circumstances' and extent score 3 defined as 'Defect occupies less than 5% of surface of the structural element'. Thus, if a defect is not present it would be assigned $Ex(\alpha_{i,j}) = 1$, whereas if there is some defect but its coverage is less than 5% of the surface, bridge inspectors are being cautious and assigning an extent score of 3.

Table 7
The extent scores used in this study in reference to the NR extent scores.

	NR Extent Score					
	1	2	3	4	5	6
Study Extent Score	1	2	2	3	4	5

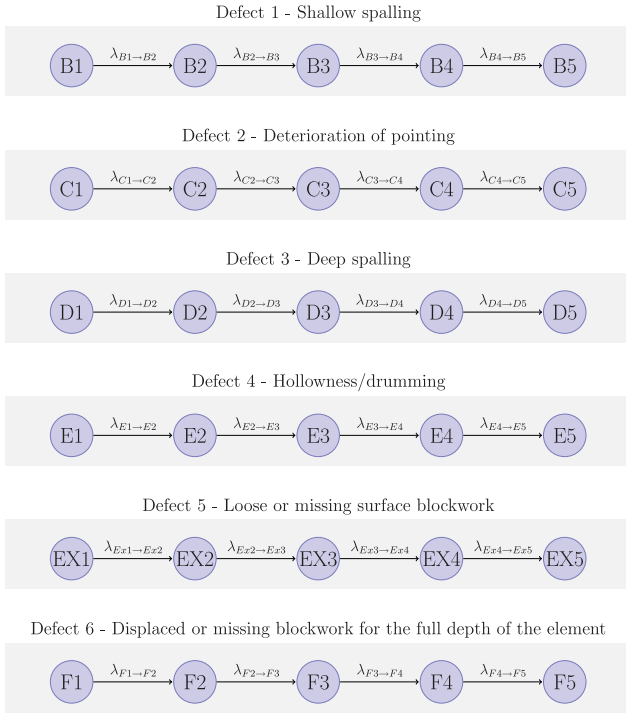


Fig. 1. Multi-defect deterioration Markov model .

To address the potential for any erroneous errors due to this, the extent scores of 2 and 3 were merged, with the extent scores used in this study defined in reference to the NR extent scores, as shown in Table 7.

4. Proposed model

Due to the constraints of the data, discussed further in Section 4.1, it was determined that a continuous-time Markov chain would be the most appropriate modelling technique. The proposed multi-defect deterioration model is shown in Fig. 1. The predictive model reports the probability of an extent score for each of the six severity scores, which for masonry bridges aligns with the extent of each of the six different defect types.

The transition rate matrix for severity B is described by,

$$Q_B = \begin{matrix} & \begin{matrix} B_{1,1} & B_{1,2} & B_{1,3} & B_{1,4} & B_{1,5} \end{matrix} \\ \begin{matrix} B_{1,j} \\ B_{2,j} \\ B_{3,j} \\ B_{4,j} \\ B_{5,j} \end{matrix} & \begin{pmatrix} -\lambda_{B1,B2} & \lambda_{B1,B2} & 0 & 0 & 0 \\ 0 & -\lambda_{B2,B3} & \lambda_{B2,B3} & 0 & 0 \\ 0 & 0 & -\lambda_{B3,B4} & \lambda_{B3,B4} & 0 \\ 0 & 0 & 0 & -\lambda_{B4,B5} & \lambda_{B4,B5} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad (6)$$

The model described by (6) makes the assumption that a bridge element can only degrade instantaneously to an extent score of one more than the current extent score. The inability to make an instantaneous transition of more than one extent score is considered to be a more realistic representation of the physical process of bridge deterioration, as the defects will exhibit continuous growth. Moreover, the inability of the model to make ‘state jumps’ is deemed to be a helpful

attribute to avoid the model being over-fitted to the data.

The transition rate matrices, Q_C , Q_D , Q_E , Q_{EX} and Q_F for severity scores C , D , E , EX and F respectively, are similarly described as Q_B . Thus, the entire model is described by the following transition rate matrix,

$$Q_{MD} = \begin{bmatrix} [Q_B] & & & & & \\ & [Q_C] & & & & \\ & & [Q_D] & & & \\ & & & [Q_E] & & \\ & & & & [Q_{EX}] & \\ & & & & & [Q_F] \end{bmatrix} \quad (7)$$

The continuous time Markov chain approach as used in this study, assumes that there is no additional information of the bridge condition, or timing of condition transitions between the discrete observation times. However, the model does not implicitly assume that the bridge element will remain in its most recently observed condition state until an inspection reveals it to be otherwise. This model assumption is deemed to be reflective of the physical reality of continuous bridge deterioration. Notwithstanding, in a model that applied maintenance strategy, an inspection regime must be considered to reveal condition rather than assume a continuously reviewed state. However, the purpose of this model is to parameterise the deterioration mechanisms under a do-nothing maintenance strategy.

Methods such as partially-observable Markov processes can be used to incorporate the variability of inspections [2]. However, the quantification of the inspection variability was deemed to be beyond the scope of this study. Additionally, there are several other organisations and agencies globally that use bridge condition scales akin to SevEx, however they may have different inspection regimes. The purpose of this model is to be as general as possible for maximum applicability as well as provide insight into the novel idea of modelling bridge degradation by defect group opposed to the traditional single condition index approach.

4.1. Parameter estimation

NR has a vast portfolio of bridges and the time and expense required to inspect is significant. The earliest record inspection record that NR have for a bridge asset is from 1999. Between 1999 and 2017, of the bridge elements that have had multiple inspections, 57.25% have had two inspections, 34.28% have had three and 8.47% have had four or more inspections.

Considering a structural element that has been inspected multiple times over a period of time; a record can be produced detailing the element's condition over time. An example of is shown in Fig. 2a. The time-based approach considers the time it required to move from one condition to another, so the specific element records are used to determine the number of condition transitions for each observed time interval. An example of amalgamated records is shown in Fig. 2b.

Many of the masonry bridges in the NR portfolio were constructed during the Victorian era in the 19th century [45], and thus have had an active service life of over 100 years in most cases. As there is an extensive gap between the construction of the masonry bridges in the NR portfolio and the first recorded inspections as well as the maintenance interventions records, it was deemed that the use of bridge age to compute time-dependent transition rate matrices as shown in [4], would be inappropriate for the available data.

The approach used to produce Fig. 2b is unsuitable for NR records due to the size of the inspection intervals. When a second inspection shows deterioration from the first inspection, the inspection interval can not be assumed to be the exact time it took for that degradation to occur. As the inspection interval can be several years, it is impossible to ascertain how long the bridge element has been in the worse condition state before the inspection took place. An example of this is shown in Fig. 3a.

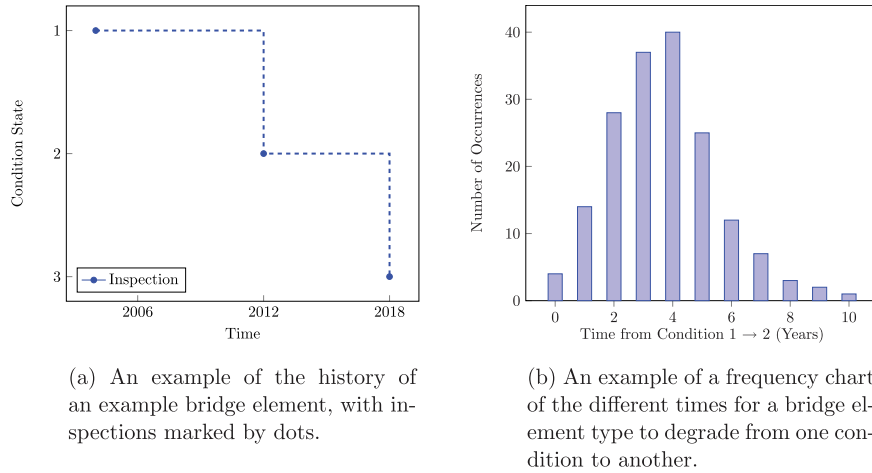


Fig. 2. An example between using historic inspection records to analyse a specific elements and a group of elements.

Additionally, due to the large inspection intervals and lack of continuous monitoring, one can not deduce the route between the initial inspection and the second, if there is a score difference greater than one. For example, if the first inspection recorded a 1, and the second a 3, the route of deterioration is unrevealed. Moreover, as the deterioration route is unrevealed, one does not know whether the bridge condition degraded from condition 1 to condition 2 to condition 3 or from condition 1 directly to condition 3, see Fig. 3b.

The censoring of time intervals and unknown deterioration paths are due to the data being a form of panel data or a longitudinal study. To address these issues a memory-less distribution is employed which does not require information on the previous histories of condition. A common implementation of this is to use discrete time Markov chains.

The estimation of parameters of a distribution describing bridge deterioration need to consider the defined frequency of inspections. In some organisations and jurisdictions the inspection intervals are a fixed size, which allows for the Transition Probability Matrices (TPM) to be computed. The number of records that show a transition from condition i to condition j is denoted as $n_{i,j}$. The probability $p_{i,j}$ of a transition from condition i to j can be computed by the following,

$$p_{i,j} = \frac{n_{i,j}}{n_i}, \quad (8)$$

where n_i is the sum of all inspections pairs which have an initial condition of i . When time is known to be both fixed and constant for all observations, $p_{i,j}$ has been shown to be a Maximum Likelihood Estimator (MLE) [46]. The changes in the probability distribution of a portfolio, C , with N condition states, from time 0, to t , can be derived from the Chapman-Kolmogorov equation,

$$[c_1^t \ c_2^t \ \dots \ c_N^t] = [c_1^0 \ c_2^0 \ \dots \ c_N^0] \begin{bmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,N} \\ 0 & \ddots & \ddots & p_{2,N} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}^t, \quad (9)$$

where c_i^t is the probability of being in state i at year t .

For the NR bridge portfolio the size of interval between detailed inspections is determined by the condition of the overall bridge at its previous inspection or if observations made at an annual visual inspection result in a detailed inspection being required. For example, stone bridges are categorised into lower, medium and high risk categories with medium and high risk bridges inspected every 6 years and low risk bridges every 12 years [1]. Additionally, curved or straight masonry bridges of four or more spans with RA10¹ loading have maximum inspection intervals of 3 and 6

years respectively. An example distribution of inspection intervals is shown in Fig. 4. Thus, any estimation technique will be required to analyse pairs of inspections with varying interval times.

4.2. Maximum likelihood estimation approach

An alternative parameter estimation method, more suitable for when the time intervals between inspections vary significantly was proposed by Kalbfleisch and Lawless [3]. The technique has been used to estimate parameters for continuous-time Markov chains in deterioration modelling of structural assets [4,47]. In this approach, the parameters of the model, see (6) and (7) are computed by maximising the likelihood of the observed inspection results.

Consider the observed discrete variable data, $\{x_1, x_2, \dots, x_n\}$, the Likelihood function $L(\theta)$ is defined as the joint probability mass function of the observed data given θ . If $\{x_1, x_2, \dots, x_n\}$ is a random sample from a distribution with probability function $f(x|\theta)$ then the Likelihood function is given by

$$L(\theta) = \prod_{i=1}^n f(x_i|\theta). \quad (10)$$

The Maximum Likelihood Estimator (MLE) is the value of θ which maximises $L(\theta)$. The log-likelihood function, $F(\theta)$, is defined to be the natural logarithm of the likelihood function $L(\theta)$, as,

$$F(\theta) = \sum_{i=1}^n \ln(f(x_i|\theta)). \quad (11)$$

Note, that the natural logarithm is a monotonically increasing function, and thus maximising the likelihood function is equivalent to maximising the log-likelihood function.

4.2.1. Estimation of the optimal transition rate matrix

The likelihood is the predicted probability of the occurrence of the observed condition transitions:

$$L(Q_{MD}) = \prod_{r=1}^N p_r, \quad (12)$$

where N denotes the number of observed condition transition records for all severity scores and

$$p_r = p_{i,j,t}, \quad (13)$$

where i is the condition score at the first inspection of record r , j is the condition score at the second inspection of record r and t is the size of the inspection interval between the first and second inspection of record r . The $p_{i,j,t}$ value is found from the $(i, j)^{th}$ element from the

¹ RA10 refers to the route availability, which defines the axle weight that can be transported over a bridge.

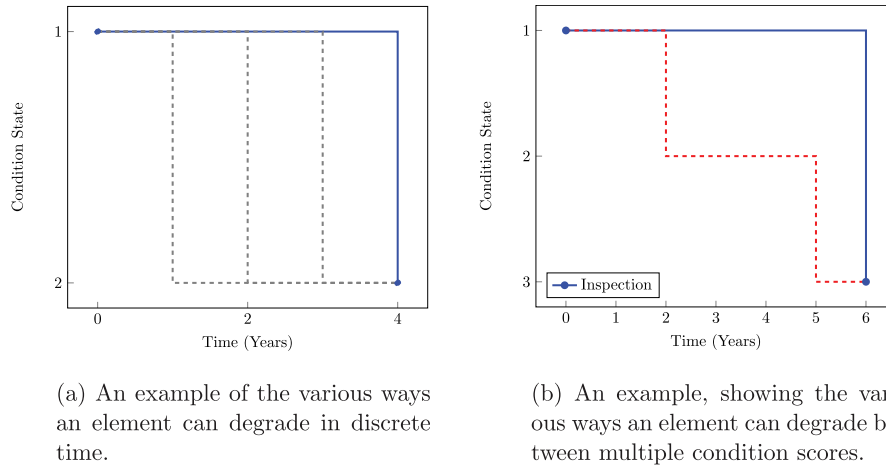


Fig. 3. Example plots showing the uncertainty in deterioration paths during the inspection interval and between the revealed condition states.

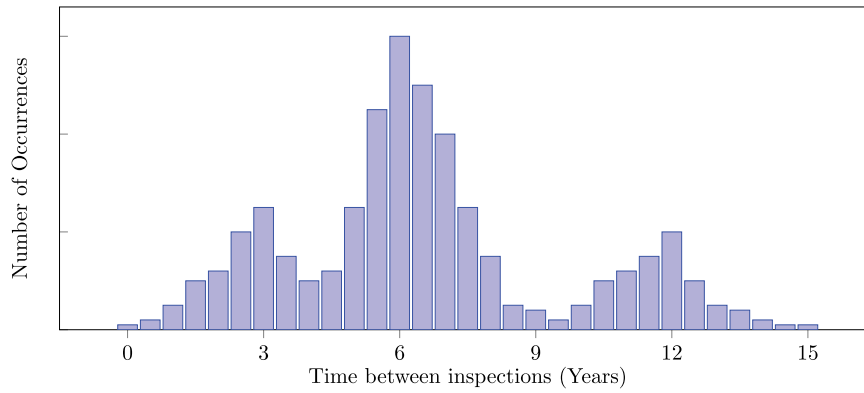


Fig. 4. A typical distribution of time intervals between two inspections in the NR inspection records. The distribution of inspection intervals has local maxima at 3 years, 6 years and 12 years.

appropriate transition probability matrix, $P(t)$, which is calculated as,

$$P(t) = e^{Q_{MD} \cdot t}, \quad (14)$$

where t is the time interval between inspections. Moler and Van Loan describe methods to compute the matrix exponential [48].

4.3. Optimising the maximum likelihood estimator approach

The MLE approach seeks to determine the set of parameters, θ , such that the transition rate matrix, Q_{MD} , maximises the following objective function,

$$F(\theta) = \log(L(Q_{MD})) = \sum_{i=1}^N \log(p_i). \quad (15)$$

Additionally, there is a constraint on all the parameters in the upper diagonal in the transition rate matrices, Q_B , Q_C , Q_D , Q_E , Q_{EX} and Q_F in Q_{MD} , that they must be positive.

Typically the MLE parameter values can be determined by taking derivatives of the log-likelihood function; in this problem, a derivative-free approach was undertaken as future work will also require these techniques. Rios and Sahinidis authored a review paper of the algorithms and software implementations of derivative-free optimisation [49]. Simulation optimisation methods can be used with a ‘black-box’ objective function which does not require derivative information: Amaran et al. reviewed the algorithms and applications of simulation optimisation [50].

In this research Particle Swarm Optimisation (PSO) methods and constrained non-linear optimisation active sets algorithms were

implemented for comparison. PSO is a population based method and was first introduced by Kennedy and Eberhart [51–53]. Active set algorithms are useful when the fitness function is evaluated using an analytical expression rather than a numerical estimation.

A MATLAB script was developed to determine the MLE for the historic inspection records. The script made use of the *pswarm* and *fmincon* functions in MATLAB, [54]. The functions are variations of PSO [55], and active set algorithms [56,57], respectively. Each implementation seeks to minimise the objective function, thus the objective function of maximising $F(\theta)$ was found by minimising $-F(\theta)$.

5. Validation of the multi-defect model

To ensure that the score inference and multi-defect model are both accurate and robust, a series of validation checks were identified. The verification and validation checks for the multi-defect model are:

1. Verify the multi-defect model using synthetic records. This requires the use of data, produced using known distributions, as well as the ordering rule to infer six scores from the observed two at each inspection.
2. Validate the multi-defect model using historic inspection records. The historic inspection records are split into training and test sets: to estimate transition rates and analyse the goodness of fit of the model, respectively.

The subsequent sections will explain the methods used to address the points above.

5.1. Verifying the multi-defect model, using synthetic data

The multi-defect model is verified using synthetic records which was produced using known distributions. The process is as follows:

1. Assign values for the parameters that populate each of the severity score transition rate matrices.
2. Generate a number of samples of bridge element inspections by using Monte Carlo simulation of the defect model with the known transition rates. Moreover, the inspection time interval for each simulated synthetic record will be sampled from $N(6, 1)$, and the initial condition for each severity score was sampled from a distribution reflecting observed initial conditions in the NR data. Each of the six severity scores are recorded at both the first and second inspections.
3. Using the score ranking rules, determine the worst two scores and record in the same format as NR historic inspection records.
4. Using the score inference rules, infer a multi-defect inspection record for each inspection.
5. Using the MLE parametric statistical inference method, estimate values for each transition rate from the inferred multi-defect inspection records.
6. Compare the estimated parameters from the synthetic data to the known parameters used to produce the synthetic data. As the number of synthetic records increases, the estimate parameter values should converge to the known values.

5.1.1. Example

The values of the parameters used to synthesise records are shown in Table 8. There were 25,000 records synthesised as described in Section 5.1, and then transition rates were estimated from the synthetic records using the MLE approach. The values of the estimated transition rates are shown in Table 8.

It can be observed that the estimated rates are a good approximation

Table 8
Transition rates estimated from the synthetic data using MLE approach.

	B1 → B2	B2 → B3	B3 → B4	B4 → B5
Known	0.1500	0.0800	0.0850	0.0450
Estimated	0.1480	0.0790	0.0835	0.0443
% Error	1.4	1.2	1.7	1.5
	C1 → C2	C2 → C3	C3 → C4	C4 → C5
Observed	0.0810	0.0500	0.0600	0.0240
Estimated	0.0862	0.0490	0.0600	0.0240
% Error	6.4	2.0	0.05	0.1
	D1 → D2	D2 → D3	D3 → D4	D4 → D5
Known	0.0240	0.0600	0.0525	0.0140
Estimated	0.0248	0.0596	0.0512	0.0146
% Error	3.2	0.6	2.4	4.2
	E1 → E2	E2 → E3	E3 → E4	E4 → E5
Known	0.0090	0.0820	0.0400	0.0225
Estimated	0.0095	0.0828	0.0365	0.0209
% Error	5.4	0.9	8.8	7.1
	EX1 → EX2	EX2 → EX3	EX3 → EX4	EX4 → EX5
Known	0.0070	0.0435	0.0410	0.0555
Estimated	0.0070	0.0437	0.0381	0.0579
% Error	0.6	0.4	7.0	4.3
	F1 → F2	F2 → F3	F3 → F4	F4 → F5
Known	0.0055	0.1160	0.0860	0.0415
Estimated	0.0055	0.1138	0.0842	0.0439
% Error	0.3	1.9	2.1	5.8

of the known transition rates, shown in Fig. 5. The lower severity scores are more prone to estimation errors due to the lack of coverage of those scores in the inference rules when higher severity scores are present. This can be observed in the example estimation of severity C, however given the complexity of the problem being considered, this estimation was still deemed to be sufficiently accurate. Moreover, with the supply of more records, one would find further convergence to the known values. However, for the example, 25,000 records were synthesised, as that number represents a typical sample size of records for a bridge element in the NR data.

5.2. Validating the multi-defect model, using historic inspection records

To validate the multi-defect model using historic inspection records, the data set is split into training and testing sets. The proposed random split between the two sets is a 3:1 ratio between training and test sets. The training set is used to estimate the values for transition rates from the observations and the test set is used to evaluate the goodness-of-fit of the model and the estimated transition rate values.

Masonry is commonly modelled as a homogeneous class of material; however, it can be easily sub-divided into two materials types: brick and stone. Analysis of the NR data suggested that there were differences in the deterioration rate between brick and stone bridge elements. Moreover, there are subtle differences in the definition of the extent score for severity B and D, see Table 1. For the multi-defect model it was deemed that these two materials should have their records split into two cohorts for the purposes of parameter estimation. Additional factors have been shown to alter the deterioration profile of a bridge [58], including local, structural and material characteristics. However, such cohort based studies reduce the amount of data available to calibrate each model. In this study, no further cohort analysis beyond material type was performed, to maximise the amount of data available for the severe defects, which are of rare occurrence.

Bridges are extremely heterogeneous and the structural hierarchy of bridges varies greatly. At inspection, NR bridges have a score panel recorded for each structural element of the bridge e.g. abutment, spandrel wall, arch barrel etc. An example deterioration profile output of the multi-defect model for a brick, underbridge, spandrel wall is shown in Fig. 6, with its transition rates shown in Table 9. A spandrel wall is a masonry wall that is positioned on the arch barrel and retains the back-fill [45]. A railway underbridge is a bridge which carries the railway over a road, river etc.

5.3. Assessing goodness of fit

Pearson's Chi-squared goodness of fit test is a type of hypothesis testing which is commonly used to assess the fit of models estimated using categorical panel data. To be able to use Pearson's chi-squared goodness of fit test, the events must be mutually exclusive and be from a random sample.

Consider n observations from sample data that are arranged in a frequency histogram having k class intervals. Let O_i be the observed frequency in the i th interval and E_i as the corresponding expected frequency as predicted by the fitted distribution from the observed data. The test statistic is expressed as,

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}. \quad (16)$$

It is common that the goodness of fit test is conducted at the 5% significance level, although this should be taken as merely convention and not definitive [59].

As shown in [3], the size of the time interval between inspections must be considered in the assignment of the intervals for the calculation of the test statistic. However, with the varying inspection intervals in the NR data, there is an imbalance of interval values. This causes low

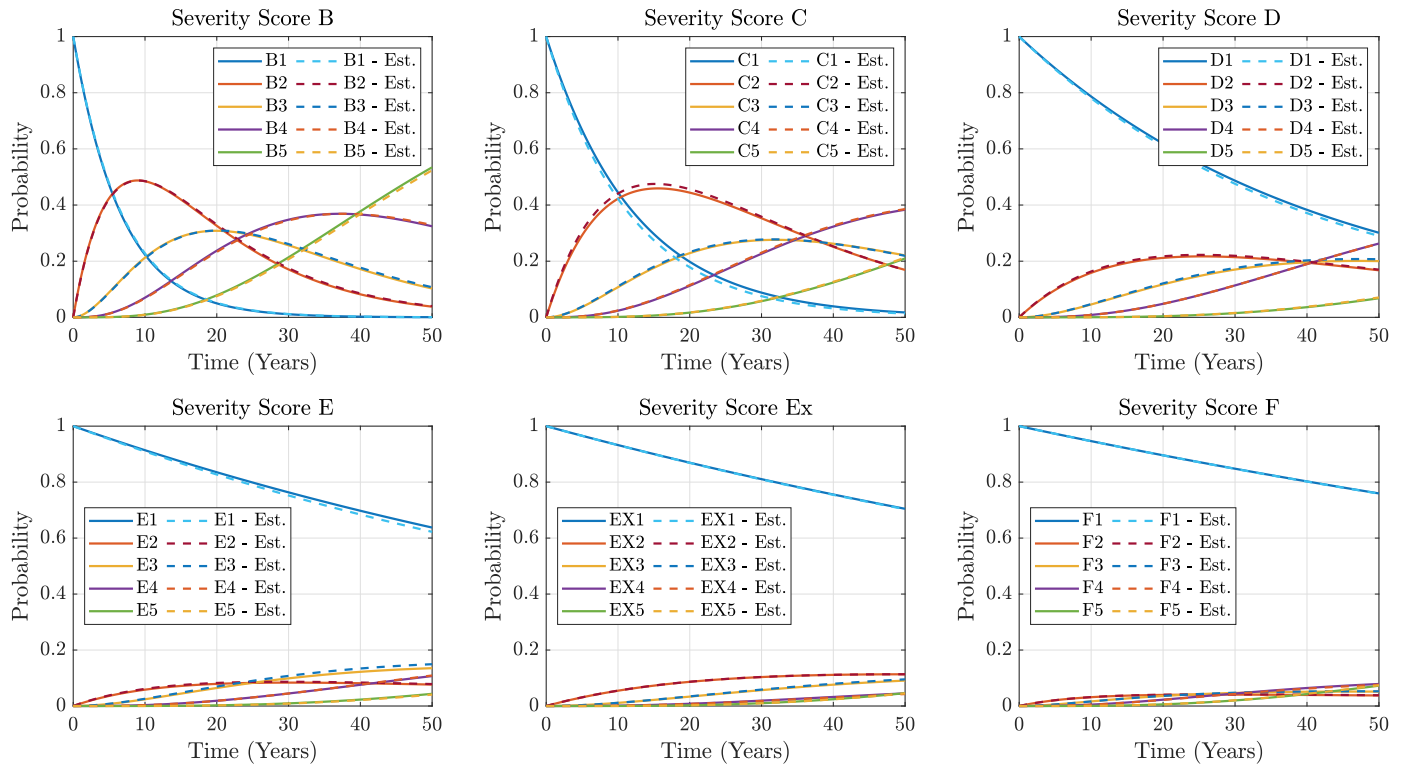


Fig. 5. Deterioration profiles for defect types with the known distribution for the synthetic data and the estimated distributions.

frequencies to occur for some test statistic intervals. Generally, the values for each expected state to state movement should have occurrences that are greater than 5; if the occurrences are less than 5, then

neighbouring intervals are merged. For the test set of NR data, this required the merging of a considerable amount of intervals and as such, it was deemed that the Pearson Chi-squared test statistic was an

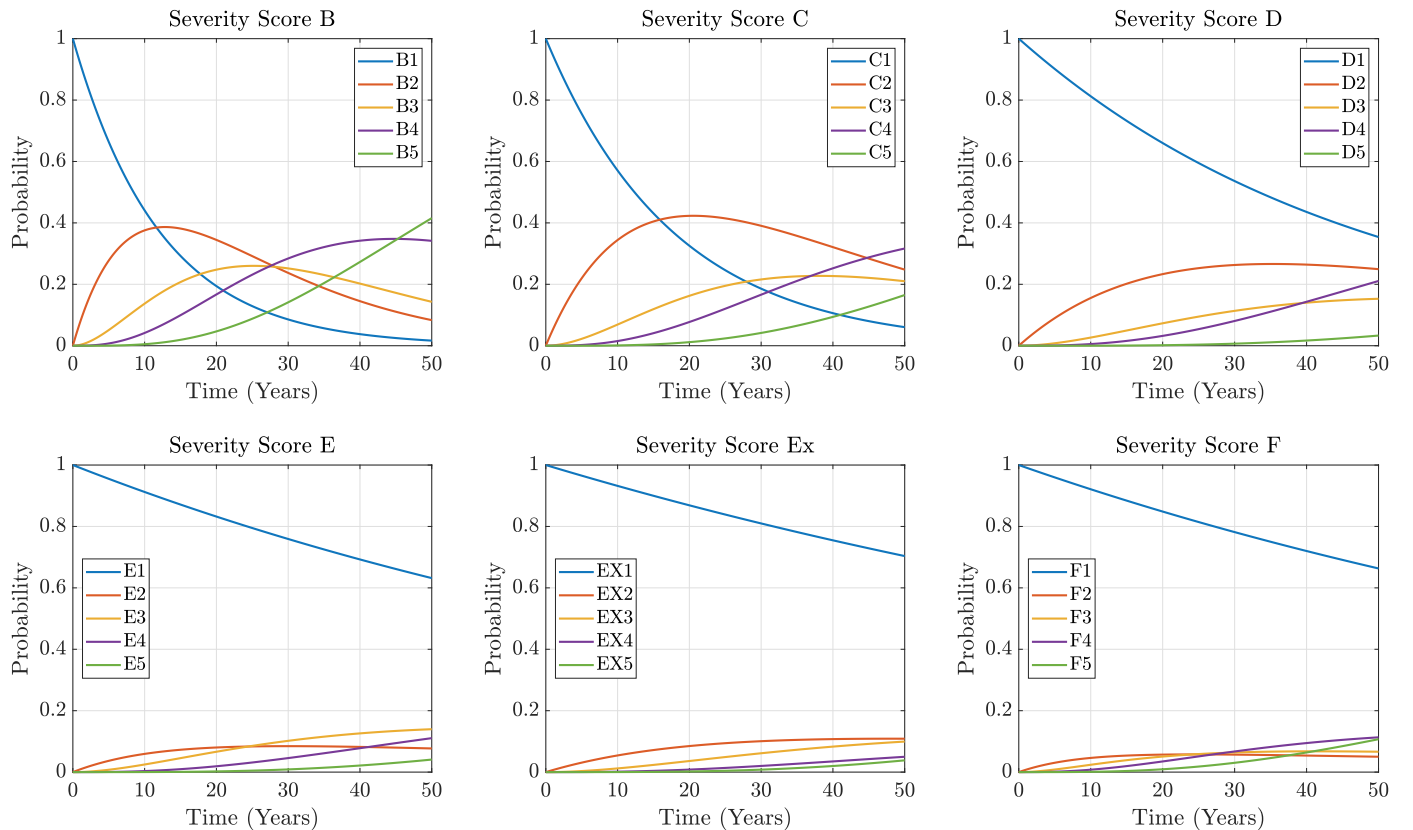


Fig. 6. Deterioration Profiles for all defect modes for a brick, underbridge, spandrel wall.

Table 9

Transition rates estimated from NR inspection records for a brick, underbridge, spandrel wall.

	B1 → B2	B2 → B3	B3 → B4	B4 → B5
Estimated	0.0820	0.0741	0.0826	0.0413
	C1 → C2	C2 → C3	C3 → C4	C4 → C5
Estimated	0.0561	0.0418	0.0616	0.0248
	D1 → D2	D2 → D3	D3 → D4	D4 → D5
Estimated	0.0208	0.0375	0.0568	0.0092
	E1 → E2	E2 → E3	E3 → E4	E4 → E5
Estimated	0.0092	0.0826	0.0390	0.0206
	EX1 → EX2	EX2 → EX3	EX3 → EX4	EX4 → EX5
Estimated	0.0070	0.0461	0.0366	0.0434
	F1 → F2	F2 → F3	F3 → F4	F4 → F5
Estimated	0.0082	0.1166	0.0919	0.0413

inappropriate measure of fit for the data.

5.4. Comparing the observed and predicted final inspection

To assess the goodness-of-fit, a comparison between the observed final condition and predicted final condition was performed. The records in the training set are used to estimate the values of each of the transition rates and the records in the test set are used in the comparison. The process for the comparison requires:

- A total number for each condition state in the final inspection for all the observed records.
- The probabilities for each condition state at the final inspection given the condition at initial inspection for each observed record.
- The sum of all the probabilities for each condition state for all predicted final conditions.

For a brick underbridge, spandrel wall, the error rate for each condition state can be found in Table 10, with the mean percentage error and weighted mean percentage error for each severity score shown in Table 11. The weighted mean percentage error, W , for a particular severity score is calculated as follows:

$$W = \frac{\sum_{i=1}^5 n_i \cdot e_i}{N}, \quad (17)$$

where n_i is the number of a observations in extent score i , e_i is the percentage error for extent score i from the model and N is the total number of observations for the severity score being considered.

5.5. Results discussion

Variability in recorded bridge inspection conditions is well documented in literature [60–65], and thus it would be inappropriate to not consider the impact of this in any model analysis. Consequently, the values for the errors and the mean errors are considered to be sufficiently low to be content that the model represents an appropriate goodness of fit.

For all the severe defects, i.e. D , E , EX and F , the absorbing state of extent score 5, is deemed to be a poor condition state, which would require immediate maintenance intervention. Moreover, for E , EX and F , this is true for some of the preceding states. Thus, maintenance activity would typically be scheduled in advance of defects progressing to such states and few observations would be made for these absorbing

Table 10

Errors between Observed and Predicted Final Conditions, for a brick underbridge, spandrel wall.

	B1	B2	B3	B4	B5
Observed	488	615	360	240	127
Predicted	555.9	594.8	332.3	240.0	107.0
% Error	13.9	3.3	7.7	9.0×10^{-5}	15.7
	C1	C2	C3	C4	C5
Observed	1127	1073	350	209	59
Predicted	1039.4	1153.7	364.2	194.7	66.0
% Error	7.8	7.5	4.1	6.8	11.9
	D1	D2	D3	D4	D5
Observed	2304	848	209	88	19
Predicted	2262.0	853.9	231.5	107.2	13.5
% Error	1.8	0.7	10.7	21.8	29.2
	E1	E2	E3	E4	E5
Observed	3584	255	148	103	39
Predicted	3565.1	259.6	171.8	96.8	35.8
% Error	0.5	1.8	16.1	6.1	8.1
	EX1	EX2	EX3	EX4	EX5
Observed	4231	207	43	15	5
Predicted	4214.5	219.6	47.8	15.3	3.8
% Error	0.4	6.1	11.3	2.3	24.9
	F1	F2	F3	F4	F5
Observed	3651	175	168	180	147
Predicted	3653.2	188.1	140.4	183.8	155.6
% Error	0.1	7.5	16.4	2.1	5.8

Table 11

Mean percentage errors and weighted mean percentage errors for each severity score for a brick, underbridge, spandrel wall.

	B	C	D	E	EX	F
Mean % Error	8.12	7.61	12.86	6.51	8.99	6.37
Weighted Mean % Error	7.42	7.23	2.74	1.37	0.79	1.28

states. The low observations for these states can result in high prediction errors (e.g. $D5$, $EX5$), however these are rare events and represent less than 1% of the final observations for those particular defects.

From Table 10, it can be seen that there is a low prediction error shown for $D1$, $E1$, $EX1$ and $F1$, and thus the model is accurate at predicting defect absence and presence. Furthermore, for these defects that pose the greatest risk to the structural integrity of the bridge, a Pearson's chi squared test of the predicted final observations shows that the model is accurate to a statistical significance of 5%. Due to the data censoring, the errors for B and C are more significant. Nonetheless, the model is a useful tool for asset managers to predict and schedule specific maintenance interventions.

The weighted mean is used to analyse the fit of the model without the metric being adversely impacted by low frequencies. From Table 10 it can be observed that the trend for the weighted mean percentage error is that the higher severity scores have smaller values. This would be expected as the score inference rule favours the higher severity scores over the lower scores, hence there are more revealed extent scores for the higher severity scores which results in an improved parameter estimation.

Fig. 6 shows the deterioration profiles for a brick, underbridge, spandrel wall, with severity scores B , C and D exhibiting more rapid deterioration than severity scores E , EX and F . Severity scores B , C and D represent shallow spalling, deterioration pointing and deep spalling,

respectively, which are faster acting defects than hollowness, loose block work and fully displaced block work, which are denoted by severity scores *E*, *EX* and *F* respectively. For faster acting defects, the accuracy of the rate of deterioration aids in the appropriate scheduling and budgeting of minor interventions. It is critical that the processes of hollowness, loose block work and displaced block work are also well understood, as whilst they are slower acting, they are the defects that are deemed to represent the most risk to the structural integrity of a bridge.

The comparison of different maintenance strategies requires an evaluation on the WLCC. The contribution of this research is that the deterioration model enables specific defect predictions, which occur at different stages in an asset life cycle and require different resources and expense to resolve. This defect specific approach which provides additional condition indicators is novel for a bridge engineering application.

A limitation of the presented model is that it does not account for any interactions between the different defects modes. Future work will evaluate how this can be modelled taking into account the limitations of the available data. However, for bridge managers, an understanding of the extensiveness of defects is more critical and hence they have been modelled independently for this case study.

6. Conclusions

This paper presented a multi-defect bridge deterioration model, which outputs multiple predictive condition profiles. The case study of masonry bridge elements, deteriorate through a variety of different defect modes, and thus the proposed model offers more versatility than current predictive bridge deterioration models which provide a single condition index.

An ideal inspection regime would record a complete multi-defect inspection panel, however in many organisations this is not the case and so a score inference technique using logic rules was also introduced to utilise existing NR data. A proven MLE parameter estimation technique can then be applied to estimate the transition rates between condition states.

The main objective of this paper was to develop a multi defect model for bridge deterioration and evaluate whether this model could be reliably calibrated using incomplete inspection data, where only the two most severe defects are recorded. The results indicate that, although the model can be calibrated using the available data, the quality of fit is influenced by the censored data, particularly for less severe condition states.

The current approach in asset management of bridges involves an inspection regime which is based on the overall condition of the bridge. As such, the time interval between inspections is reduced as the condition worsens, but the same inspection procedure is followed on all bridges of the same class (e.g. masonry arch bridges). However, the cost of detecting different defects varies very significantly. For example, cracks can be observed at a distance, using for example drones, while hollowness can only be detected by touch, and thus requires expensive lifting equipment and longer possession times. The multi-defect deterioration model allows the definition of targeted inspections defined in terms of the risk of occurrence of each defect, thus enabling an optimised inspection policy.

Moreover, different deterioration mechanisms require different repair actions, with different durations and costs. By predicting the deterioration of bridges in terms of different defects, it will be possible to predict, in further detail, what actions will be required and what their costs will be.

In future work, the interactions between different deterioration and failure mechanisms should be evaluated. It is expected that the condition of a component in terms of one defect may alter the rate of evolution of other defect mechanisms. Such trends would impact on predictions of service life of the bridge element but additionally on the

condition dynamics upon maintenance intervention. An advantage of the multi-defect approach to deterioration modelling is the additional condition indicators the model outputs. The additional outputs enable the development of specific and targeted maintenance strategies, which in conjunction with any identified interactions could facilitate the development of targeted maintenance strategies and the quantification of the benefits of early interventions.

Moreover, the current implementation of the model allows infrastructure agencies to monitor the development of 'severe' defect mechanisms that pose a large risk to the structural integrity of the bridge. In future work an analysis should be performed to identify possible precursor condition events, which could improve the inspection regime, condition recording and strategy development.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

John Andrews is the Royal Academy of Engineering and Network Rail Professor of Infrastructure Asset Management. He is also Director of the Lloyd's Register Foundation (LRF) Resilience Engineering Research Group at the University of Nottingham. Luis Neves is an Assistant Professor in Structural Engineering and Infrastructure Asset Management at the University of Nottingham. Gareth Calvert is a PhD student at the University of Nottingham. This work was supported by the Engineering and Physical Sciences Research Council [EP/N50970X/1]. Additional support was received from Network Rail. Matthew Hamer is a Whole Life Cycle Costing Manager at Network Rail. The authors gratefully acknowledge the support of these organizations.

References

- [1] Network Rail. Handbook for the examination of structures - part 1c: Determining the examination regime. Tech. Rep., Network Rail; 2014.
- [2] Frangopol DM, Kallen M-J, van Noortwijk JM. Probabilistic models for life-cycle performance of deteriorating structures: review and future directions. *Prog Struct Mater Eng* 2004;6(4):197–212. <https://doi.org/10.1002/pse.180>.
- [3] Kalbfleisch JD, Lawless JF. The analysis of panel data under a Markov assumption. *J Am Stat Assoc* 1985;80(392):863–71.
- [4] Kallen MJ, Noortwijk JMV. Statistical inference for Markov deterioration models of bridge conditions in the Netherlands. *Third International Conference on Bridge Maintenance, Safety and Management (IABMAS)*. 2006. p. 535–6. 0415403154 (ISBN); 9780415403153 (ISBN)
- [5] Nicolai RP, Dekker R, van Noortwijk JM. A comparison of models for measurable deterioration: an application to coatings on steel structures. *Reliab Eng Syst Saf* 2007;92(12):1635–50. <https://doi.org/10.1016/j.res.2006.09.021>.
- [6] Stewart MG. Reliability-based assessment of ageing bridges using risk ranking and life cycle cost decision analyses. *Reliab Eng Syst Saf* 2001;74(3):263–73. [https://doi.org/10.1016/S0951-8320\(01\)00079-5](https://doi.org/10.1016/S0951-8320(01)00079-5).
- [7] Yang SI, Frangopol DM, Neves LC. Service life prediction of structural systems using lifetime functions with emphasis on bridges. *Reliab Eng Syst Saf* 2004;86(1):39–51. <https://doi.org/10.1016/j.res.2003.12.009>.
- [8] Yang SI, Frangopol DM, Kawakami Y, Neves LC. The use of lifetime functions in the optimization of interventions on existing bridges considering maintenance and failure costs. *Reliab Eng Syst Saf* 2006;91(6):698–705. <https://doi.org/10.1016/j.res.2005.06.001>.
- [9] Le B, Andrews J. Modelling railway bridge degradation based on historical maintenance data. *The Journal of the Safety and Reliability Society* 2015;35(2):32–55.
- [10] Frangopol DM, Dong Y, Sabatino S. Bridge life-cycle performance and cost: analysis, prediction, optimisation and decision-making. *Struct Infrastruct Eng* 2017;13(10):1239–57. <https://doi.org/10.1080/15732479.2016.1267772>.
- [11] Neves LC, Frangopol DM. Condition, safety and cost profiles for deteriorating structures with emphasis on bridges. *Reliab Eng Syst Saf* 2005;89(2):185–98. <https://doi.org/10.1016/j.res.2004.08.018>.
- [12] Liu M, Frangopol DM. Probability-Based bridge network performance evaluation. *J Bridge Eng* 2006;11(5):633–41. [https://doi.org/10.1061/\(asce\)1084-0702\(2006\)11:5\(633\)](https://doi.org/10.1061/(asce)1084-0702(2006)11:5(633)).
- [13] Saydam D, Frangopol DM, Dong Y. Assessment of risk using bridge element condition ratings. *J Infrastruct Syst* 2013;19(3):252–65. [https://doi.org/10.1061/\(ASCE\)IS.](https://doi.org/10.1061/(ASCE)IS.)
- [14] Frangopol DM, Liu M. Maintenance and management of civil infrastructure based

- on condition, safety, optimization, and life-cycle cost. *Struct Infrastruct Eng* 2007;3(1):29–41. <https://doi.org/10.1080/15732470500253164>.
- [15] Jiang Y, Saito M, Sinha KC. Bridge performance prediction model using the Markov chain. *Transp Res Rec* 1988;1180(1):25–32.
 - [16] Cesare MA, Santamarina C, Turkstra C, Vanmarcke EH. Modeling bridge deterioration with Markov chains. *J Transp Eng* 1992;118(6):820–33. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1992\)118:6\(820\)](https://doi.org/10.1061/(ASCE)0733-947X(1992)118:6(820)).
 - [17] Scherer WT, Glagola DM. Markovian models for bridge maintenance management. *Journal of Transport Engineering* 1994;120(1):37–51.
 - [18] Morcos G. Performance prediction of bridge deck systems using Markov chains. *J Perform Constr Facil* 2006;20(2):146–55.
 - [19] Agrawal A, Kawaguchi A, Chen Z. Bridge element deterioration rates. Tech. Rep. Albany, NY, USA: New York State DoT; 2009.
 - [20] Thompson PD, Small EP, Johnson M, Marshall AR. The Pontis bridge management system. *Struct Eng Int* 1998;8(4):303–8.
 - [21] Mirzaei Z, Adey BT, Klatter L, Thompson PD. The IABMAS Bridge Management Committee - Overview of Existing Bridge Management Systems 2014. Tech. Rep. IABMAS; 2014.
 - [22] Ng S-K, Moses F. Bridge deterioration modeling using semi-Markov theory. *Struct Saf Reliab* 1998;113–20.
 - [23] Kleiner Y. Scheduling Inspection and Renewal of Large Infrastructure Assets. *Journal of Infrastructure Systems* 2001;7(4):136–43. 10.1061/(ASCE)1076-0342(2001)7:4(136)
 - [24] Mishalani RG, Madanat SM. Computation of infrastructure transition probabilities using stochastic duration models. *J Infrastruct Syst* 2002;8(4):139–48. [https://doi.org/10.1061/\(ASCE\)1076-0342\(2002\)8:4\(139\)](https://doi.org/10.1061/(ASCE)1076-0342(2002)8:4(139)).
 - [25] Thomas O, Sobanjo J. Semi-Markov models for the deterioration of bridge elements. *Journal of Infrastructure Systems* 2016;22(3):1–12. 10.1061/(ASCE)IS.1943-555X.0000298.
 - [26] Rama D, Andrews JD. A reliability analysis of railway switches. *Proc Instit Mech Eng Part F* 2013;227(4):344–63. <https://doi.org/10.1177/0954409713481725>.
 - [27] Le B, Andrews J. Petri net modelling of bridge asset management using maintenance-related state conditions. *Struct Infrastruct Eng* 2016;12(6):730–51. <https://doi.org/10.1080/15732479.2015.1043639>.
 - [28] Le B, Andrews J, Fecarotti C. A Petri net model for railway bridge maintenance. *J Risk Reliab* 2017;231(3):306–23. <https://doi.org/10.1177/1748006X17701667>.
 - [29] Yianni PC, Rama D, Neves LC, Andrews JD, Castlo D. A Petri-net-based modelling approach to railway bridge asset management. *Struct Infrastruct Eng* 2017;13(2):287–97. <https://doi.org/10.1080/15732479.2016.1157826>.
 - [30] Attoh-Okine NO, Bowers S. A Bayesian belief network model of bridge deterioration. *Proceedings of the Institution of Civil Engineers (ICE): Bridge Engineering* 2006;159(2):69–72. <https://doi.org/10.4319/lo.2013.58.2.0489>.
 - [31] DeStefano PD, Grivas DA. Method for estimating transition probability in bridge deterioration models. *J Infrastruct Syst* 1998;4(2):56–62. [https://doi.org/10.1061/\(ASCE\)1076-0342\(1998\)4:2\(56\)](https://doi.org/10.1061/(ASCE)1076-0342(1998)4:2(56)).
 - [32] van Noortwijk JM, Klatter HE. The use of lifetime distributions in bridge maintenance and replacement modelling. *Comput Struct* 2004;82(13–14):1091–9. <https://doi.org/10.1016/j.compstruc.2004.03.013>.
 - [33] Sobanjo J, Mtenga P, Rambo-Roddenberry M. Reliability-Based modeling of bridge deterioration hazards. *J Bridge Eng* 2010;15(6):671–83. [https://doi.org/10.1061/\(asce\)be.1943-5592.0000074](https://doi.org/10.1061/(asce)be.1943-5592.0000074).
 - [34] Agrawal AK, Kawaguchi A, Chen Z. Deterioration rates of typical bridge elements in New York. *J Bridge Eng* 2010;15(4):419–29. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0000123](https://doi.org/10.1061/(ASCE)BE.1943-5592.0000123).
 - [35] Wellalage NKW, Zhang T, Dwight R. Modified Weibull approach to deterioration modelling: an application to bridge elements. *Proc. of the Second Intl. Conf. on Advances In Civil, Structural and Mechanical Engineering- CSM* 2014. 2014. p. 247–51.
 - [36] Norris JR. *Markov chains*. 1st Cambridge University Press; 1997.
 - [37] Robelin C-A, Madanat SM. History-Dependent bridge deck maintenance and replacement optimization with Markov decision processes. *J Infrastruct Syst* 2007;13(3):195–201. [https://doi.org/10.1061/\(ASCE\)1076-0342\(2007\)13:3\(195\)](https://doi.org/10.1061/(ASCE)1076-0342(2007)13:3(195)).
 - [38] Orcesi AD, Cremona CF. A bridge network maintenance framework for pareto optimization of stakeholders/users costs. *Reliab Eng Syst Saf* 2010;95(11):1230–43. <https://doi.org/10.1016/j.res.2010.06.013>.
 - [39] Bocchini P, Frangopol DM. A probabilistic computational framework for bridge network optimal maintenance scheduling. *Reliab Eng Syst Saf* 2011;96(2):332–49. <https://doi.org/10.1016/j.res.2010.09.001>.
 - [40] Ceravolo R, Pescatore M, De Stefano A. Symptom-based reliability and generalized repairing cost in monitored bridges. *Reliab Eng Syst Saf* 2009;94(8):1331–9. <https://doi.org/10.1016/j.res.2009.02.010>.
 - [41] Heath B, Hill R, Ciarallo F. A survey of agent-based modeling practices (January 1998 to July 2008). *Journal of Artificial Societies and Social Simulation* 2009;12(4).
 - [42] Bush SJW, Henning TFP, Raith A, Ingham JM. Development of a Bridge Deterioration Model in a Data-Constrained Environment. *Journal of Performance of Constructed Facilities* 2017;31(5):1–8. 10.1061/(ASCE)CF.1943-5509.0001074.
 - [43] Yang DY, Frangopol DM. Life-cycle management of deteriorating civil infrastructure considering resilience to lifetime hazards: a general approach based on renewal-reward processes. *Reliab Eng Syst Saf* 2019;183:197–212. <https://doi.org/10.1016/j.res.2018.11.016>.
 - [44] Network Rail. Handbook for the examination of structures part 2c: Condition marking of bridges. Tech. Rep. Network Rail; 2010.
 - [45] McKibbens LD, Sawar CMN, Gaillard CS. Masonry arch bridges: condition, appraisal and remedial treatment. Tech. Rep. London, UK: CIRIA; 2006.
 - [46] Frankel E. *Systems reliability and risk analysis*. 2nd Kluwer Academic Publishers; 1988.
 - [47] Ferreira C, Neves L, Silva A, de Brito J. Stochastic Petri net-based modelling of the durability of renderings. *Autom Constr* 2018;87:96–105. <https://doi.org/10.1016/j.autcon.2017.12.007>.
 - [48] Moler C, Van Loan C. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev* 2003;45(1):3–49. <https://doi.org/10.1137/S00361445024180>.
 - [49] Rios LM, Sahinidis NV. Derivative-free optimization: a review of algorithms and comparison of software implementations. *J Global Optim* 2013;56(3):1247–93. <https://doi.org/10.1007/s10898-012-9951-y>.
 - [50] Amaran S, Sahinidis NV, Sharda B, Bury SJ. Simulation optimization: a review of algorithms and applications. 4OR - A Quarterly Journal of Operations Research 2014;12(4):351–80. <https://doi.org/10.1007/s10479-015-2019-x>.
 - [51] Eberhart R, Kennedy J. A new optimizer using particle swarm theory. *Proceedings of the 6th International Symposium on Micro Machine and Human Science* 1995:39–43. <https://doi.org/10.1245/s10434-016-5403-0>.
 - [52] Kennedy J, Eberhart R. Particle Swarm Optimization. *Proceedings of the IEEE international conference on neural networks*. 1995. p. 1942–8. <https://doi.org/10.1109/ICNN.1995.488968>.
 - [53] Shi Y, Eberhart R. A Modified Particle Swarm Optimizer. *Proceedings of IEEE international conference on evolutionary computation*. 1998. p. 69–73. <https://doi.org/10.1109/ICEM.2007.4350772>.
 - [54] MATLAB and Optimization Toolbox Release 2018a. The MathWorks, Inc.; Natick, Massachusetts, United States; 2018.
 - [55] Mezura-Montes E, Coello Coello CA. Constraint-handling in nature-inspired numerical optimization: past, present and future. *Swarm Evol Comput* 2011;1(4):173–94. <https://doi.org/10.1016/j.swevo.2011.10.001>.
 - [56] Han SP. A globally convergent method for nonlinear programming. *J Optim Theory Appl* 1977;22(3):297–309. <https://doi.org/10.1007/BF00932858>.
 - [57] Schittkowski K. NLPQL: A fortran subroutine solving constrained nonlinear programming problems. *Ann Oper Res* 1986;5(2):485–500. <https://doi.org/10.1007/BF02022087>.
 - [58] Yianni PC, Neves LC, Rama D, Andrews JD, Dean R. Incorporating local environmental factors into railway bridge asset management. *Eng Struct* 2016;128:362–73. <https://doi.org/10.1016/j.engstruct.2016.09.038>.
 - [59] Sterne JA, Smith GD, Cox DR. Sifting the evidence what's wrong with significance tests? *Br Med J* 2001;322(7280):226–31. <https://doi.org/10.1136/bmj.322.7280.226>.
 - [60] Madanat S. Optimal infrastructure management decisions under uncertainty. *Transportation Research Part C: Emerging Technologies* 1993;1(1):77–88. [https://doi.org/10.1016/0968-090X\(93\)90021-7](https://doi.org/10.1016/0968-090X(93)90021-7).
 - [61] Ben-Akiva M, Humplick F, Madanat S, Ramaswamy R. Infrastructure management under uncertainty: latent performance approach. *J Transp Eng* 1993;119(1):43–58. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1993\)119:1\(43\)](https://doi.org/10.1061/(ASCE)0733-947X(1993)119:1(43)).
 - [62] Phares BM, Washer GA, Rolander DD, Graybeal BA, Moore M. Routine highway bridge inspection condition documentation accuracy and reliability. *J Bridge Eng* 2004;9(4):403–13. [https://doi.org/10.1061/\(ASCE\)1084-0702\(2004\)9:4\(403\)](https://doi.org/10.1061/(ASCE)1084-0702(2004)9:4(403)).
 - [63] Moore M, Phares B, Graybeal B, Rolander D, Washer G. Reliability of visual inspection for highway bridges, Volume i: Final report. Tech. Rep. Federal Highway Administration; 2001.
 - [64] Neves L, Frangopol D. Optimization of bridge maintenance actions considering combination of sources of information: inspections and expert judgment. *Bridge Maintenance, Safety, Management and Life-Cycle Optimization - Proceedings of the 5th International Conference on Bridge Maintenance, Safety and Management* 2010:1913–9.
 - [65] Berrade MD, Scarf PA, Cavalcante CA, Dwight RA. Imperfect inspection and replacement of a system with a defective state: a cost and reliability analysis. *Reliab Eng Syst Saf* 2013;120:80–7. <https://doi.org/10.1016/j.res.2013.02.024>.