Uncertainty in condition prediction of bridges based on assessment method – case study in Estonia.

Sander Sein, Juhan Idnurm

Tallinn University of Technology, Tallinn, Estonia

Jose C. Matos

University of Minho, Guimarães, Portugal

Contacting author: sander.sein@taltech.ee

Abstract

In this paper the uncertainty in condition assessment based on most common assessment methods, visual inspection and non-destructive testing, is investigated. For decision-making the averaged or estimated value is suitable, but if the basis of a decision is only a subjective visual inspection, then it could lead to a wrong decision. The second most traditional assessment method is non-destructive testing (NDT), which can give reliable results, but the interpretation of measurement is needed. To investigate the errors in both evaluations, benchmarking tests were carried out in Estonia within two groups, a group of experienced inspectors and a group of unexperienced students, to show how the importance of experience affects results. To present the influence of assessment uncertainty to condition prediction curves based on continuous-time Markov model are calculated and for updating, Bayesian inference procedure is used.

Keywords: reinforced concrete, non-destructive testing, visual inspections, assessment error, asset management, bridge assessment.

1. Introduction

The structural condition of bridges has a wide and direct impact for performance of the road network as a system. To keep the structural reliability high, bridges need to be regularly maintained. To optimize the maintenance strategies of existing bridge stock, their present condition needs to be assessed and determined. Based on COST Action TU1406 Working Group 1 Technical Report, most data are obtained by conducting visual inspection as an index form [1]. Visual inspection is a method that may yield subjective and unreliable results [2], but due to its simplicity and cost-effective data collection, this method will most likely remain the main aid for bridge assessment.

To improve the quality of acquired data, selected non-destructive tests (NDTs) are carried out additionally into regular bridge assessment practice. The NDTs are good for their repeatability without damaging the element under investigation, but to be compatible with visual inspection, they must be easy to use.

Unfortunately, all assessment systems are moreover database oriented and additional benefit can be added with Life Cycle Assessment by integrating deterioration models to predict the performance of these structures. Over the past twenty years, many models have been proposed including the ones-based on Markov chains [3], linear or non-linear probability functions [4], neural networks [5] and lifetime functions [6]. Although it is essential to obtain the assessment errors in the models, the amount of information makes the predictive models imprecise.

In this current work, Estonian bridge management system visual inspection methodology and most common standardized NDTs: sclerometer test carbonation depth and rebar depth, are tested by two groups with different expertise to clarify that visual inspections are unreliable method [7] for the bridge evaluation and more simple NDTs that are easy to carry out and not difficult to interpret. Results are then compared using Markov chain condition predictive models.

2. Visual inspections

Visual inspections in Estonia are carried out according to modified AASHTO methodology, where bridge evaluation is done on element basis. In this work, the bridge structure is divided into element groups that have been previously clustered by Sein et al. [8] as in Table 1.

Table 1. Classification of element groups [8]

Non-Structural Elements	Structural elements
Overlay	Deck plate
Barriers	Edge beam
Handrails	Piles and columns
Drainage	Supporting beam
Slopes	Wing wall, abutments
Deformation joints	Diaphragms
Other (river bed, signs etc.)	Main girder
Waterproofing	Bearings

Structural elements are subject of load carrying function to traffic and Non-Structural elements provide protection either to the structure or the users. The detailed list of bridge elements depends on the bridge structural type and is defined from pre-specified list of 145 elements during the bridge inspections.

Visual assessment of a bridge relies on inspecting every element unit of the bridge and evaluating each with a condition rating on scale from 1 to 4, based on the damage present and necessary rehabilitation method. Condition state 1 means, that element is in good condition and no rehabilitation is needed. Condition state 4 means, that element is in critical condition and major repair or replacement is needed. Condition states 2 and 3 describes the fair and bad condition respectively.

An overall element condition state is calculated based on the overall quantity of units and state factors [8] (1):

$$H_e = \frac{\sum_s k_s q_s}{\sum_s q_s} \cdot 100\%, \tag{1}$$

where H_e − condition state of element; s is condition state; k_s − coefficient of state and q_s is the number of units in current state.

Condition Index (CI) is calculated like Health Index in Pontis [9]. The result is expressed with only one number between 0-100 and it is calculated based on element condition and weight factor. The CI shows the need for intervention [3] and it is agreed in Estonia, that an optimal level is reached, when bridge will be repaired before the CI is under 70. With CI less than 33, closing of the structure should be considered. In this current work, both element- and system level information is compared.

2.1 Inspections

The inspections were carried out on 3 common road bridge with similar structural typology, but conditions varied from good to almost critical. The overall information is presented in Table 2.

Table 2. Inventory information of visually inspected bridges

Bridge name	Lagedi	Assaku	Saku
Typology	Girder		
Length (m)	95.20	36.20	67.20
Spans	5	3	4
Material	Reinforced concrete		
Condition index	44	65	57
Year of Construction	1970	1989	1974
Year of repair	-	-	2005

The test inspection involved two groups of inspectors, where first group of 7 inspectors consisted of bridge experts with different background in bridge engineering and second group of 5 inspectors without any expertise.

In first group only 3 inspectors had previous expertise in the bridge assessment and familiar with the methodology, other inspectors had only read the inspection manual, but done similar bridge assessments previously. First group inspected two bridges, Lagedi and Assaku (Figure 1) viaduct.

Second group consisted of engineering Master students, who had no experience in bridge assessment. The methodology was introduced in one 1.5-hour lecture and they had read the inspection manual. Second group inspected Saku viaduct.

Both inspections were carried out in October and during all inspections it was rainy, which may influence the overall results and inspector motivation, but since all inspectors were in similar conditions, then it is considered that the weather condition didn't affect the differences of final evaluations.

2.2 Inspection results

The results are presented in Table 3 with every rating to bridge condition, overall mean and standard deviation (SD) of all assessments.

Table 3. Overall results of all inspections

Inspector	Lagedi CI	Assaku CI	Saku CI
1	48.9	66.9	49.3
2	36.7	66.7	74.0
3	38.2	66.8	19.3
4	36.7	87.0	76.1
5	46.0	75.1	70.1
6	39.8	87.0	-
7	55.3	88.3	-
Mean	43.1	76.8	57.8
SD	6.6	9,6	21,5

The overall mean of Lagedi viaduct is close to previous inspection result and SD shows, that 6 inspectors out of 7 assessed the bridge in a condition, where reconstruction is needed. Based on SD of results, the evaluations were in the same range. SD of assessment results of inspectors 2 and 3, who had experience more than 2 years, was 0.8.

Assaku bridge had a higher score for condiiton and overall inspection results were more scattered. The overall mean is more than 10 points higher than previous evaluation, but results are in similar range. Filtering out only experienced inspectors' results, then mean is almost the same as previous inspection result and SD of most experienced inspectors, 2 and 3, results was 0.1.

Saku viaduct was inspected by inexperienced inspectors and although the mean value is close to previous inspection result, only one inspector was close to the value. The SD of results shows, that results are scattered and using only one result can end with a wrong decision.

Further investigation is based on the element level data of Assaku viaduct. Three most experienced inspectors of first group, inspectors 1 to 3, have evaluated the CI between 66,7 to 66,9. The results are presented as maximum and minimum values, to show most extreme differences (Table 4).

Table 4. Extreme values of element level assessment

Element group	MAX (Inspector)	MIN (Inspector)
Overlay	66.7 (3)	60.5 (1)
Barriers	75.0 (3)	54.2 (1)
Handrails	66.7 (-)	66.7 (-)
Drainage	33.3 (3)	0.0 (2)
Slopes	64.1 (1)	40.3 (2)
Deformation joints	33.3 (3)	23.3 (2)
Other	100.0 (-)	100.0 (-)
Waterproofing	90.0 (3)	66.7 (2)
Deck plate	100.0 (2)	91.3 (1)
Edge beam	16.7 (1;2)	0.0 (3)
Piles and columns	66.7 (2;3)	62.5 (1)
Supporting beam	95.4 (2)	62.1 (1)
Abutments	66.7 (3)	59.3 (1)
Main girder	56.9 (1)	56.2 (2.3)
Bearings	91.1 (1)	55.2 (3)

Non-structural elements represent more elements with bigger area or quantity and due to that, the evaluations differ notably.

It is also interesting, that inspector number 3 tend have more maximum values and no minimal values, inspector number 2 have the opposite pattern.

Figure 1. Side view of Assaku viaduct (bms.teed.ee)

For structural elements, the most notable difference is the evaluation of bearings, which were easy to inspect on abutments, but without proper ladder it was not visible on piers (Figure 1). Bearings were the element group with maximum SD 16.7 between the results. The mean of element assessment ratings SD is 8.5, which is lower than SD of overall CI in first group.

3. Non-destructive testing

To increase the reliability of collected data, the visual assessment information should be updated with non-destructive testing information. During the comparison of different assessments, only few most common tests are investigated. Although two of the tests are standardized, the results of rebound test hammer have still questionable reliability [10]. All of tests are commonly used to detect material properties: compressive strength of concrete cover according to EVS-EN 12504-2, thickness of rebar cover and carbonation depth according to EVS-EN 14630.

All the tests are suitable only for reinforced or pre-stressed concrete structures, their test duration is short, and it is easy to obtain the results.

3.1 Bridges

Two reinforced concrete bridge were investigated and the main criteria for the selection was the casting technique of concrete: one should have vibrated and second not. Both selected bridges are common in Estonia. First tested bridge was Alliku bridge, constructed in 1975 and reconstructed in 2010. It is simply supported reinforced concrete slab bridge, it was renovated with strengthening of sub- and superstructure and concrete casting was done using nowadays methods.

The benchmarking test places were selected on two different elements: abutment and slab. Areas were marked and numbered.

The second bridge was Karutiigi, constructed in 1980, was simply supported reinforced concrete slab. It was reconstructed due to widening of the road in the beginning of 2018. Due to height of superstructure and flow rate of river, only one place was selected.

3.2 Inspectors and equipment

The benchmarking of NDT involved also two groups of inspectors like in visual assessment. First group, who carried out tests on both bridges, consisted of bridge experts with previous experience and knowledge in testing. Unfortunately, only 3 inspectors did the tests and there are only 2 different results to compare in each test. First group tests were done in period of May to June. Second group consisted of 8 engineering Master students, who had no previous experience in NDTs. Second group tested only Alliku bridge and tests were done in September. During both tests the temperature were over 15°C and weather was dry.

Sclerometer tests were carried out with SilverSchmidt Type-N concrete test hammer by Proceq. Values are presented as “Q”-value of impact instead of using conversion curves, which should consider the carbonation depth and will give rather conservative results for concrete strength. The thickness of rebar cover was measured with Proceq Profoscope+ cover meter and carbonation depth was measured using phenolphthalein solution and caliper. Values are presented as an average result.

3.3 Test results

Test results are presented in Table 6. Experts, marked as E, are numbered separately from students, who are marked as S.

In overall the rebound test hammer results are in one range and show small deviation. The second test area had smaller “Q”-value, but SD is 1.8. The third test area of Karutiigi bridge have only two different results with mean “Q”-value 49.7 and SD 0.8.

Table 6. Test results of Alliku bridge

Tester	Rebound value Q	Carbonation depth [mm]	Average cover depth [mm]
E 1	66.5	9	45
E 2	67.7	8	45
S 1	66.1	6	43
S 2	63.7	7	40
S 3	64.5	7	41
S 4	62.8	7	45
S 5	63.8	6	42
S 6	67.3	7	46
S 7	66.6	7	43
S 8	65.2	7	41
Mean	65.4	7.1	43.1
SD	1.6	0.83	2.0

The carbonation depth test results of area 2 show that experts obtained higher results compared to students although the results were in one range. In test area 1, the results were opposite – experts obtained smaller results than students. In both cases the SD is relatively high considering the relation to overall result.

The results are scattered because investigated test method requires decent cleaning of the drill hole and measurement should be taken within 30 seconds after the application of solution. Nevertheless, it is doubtful that with bigger carbonation depth the SD ratio to Mean value is similar.

Overall average results of concrete cover measurements are in one range (Table 6), but the average result is based on 5 to 10 measurements, which were in the range between 38 to 50 mm. In test area 2, the measurements were in smaller range and in test area 3 it was not possible to measure the cover depth, due to missing reinforcement.

In conclusion of NDT benchmarking tests, it is clear, that errors are smaller in comparison to visual inspections and non-experts can achieve higher accuracy without any previous experience in simple tests. It is still important to translate the results into quantitative scale like, for example, Mateus and Bragança [10] to start using these tests for regular assessment or as additional evaluation to visual assessment. In this research, only statistical errors and predictive model updating are investigated, but it for the future developments it is suggested to combine the mean values of NDTs with visual inspection outcome.

4. Predictive models and updating

For modelling, a probabilistic condition degradation model of an abutment of specific bridge is developed using Monte Carlo method. The model is based on historical information of Estonian national road bridges.

The Markov chains are stochastic processes, that are widely used for modelling information of existing bridges. Most extensively these models are based on a discrete scale, where transition between states is defined as (2) [11]:

$$\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_i \end{bmatrix}^T_{t+\Delta t} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_i \end{bmatrix}^T_{t} \times \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1j} \\ 0 & p_{22} & \cdots & p_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & p_{ij} \end{bmatrix}_{\Delta t} \quad (2)$$

Where $C_{t+\Delta t}$and C_t are condition vectors at time $t+\Delta t$ and t, respectively. Vectors are defined as the probability of an element being in each performance state, C_j. Probability of transition between state i and j from instant t and $t+\Delta t$ is defined by p_{ij}, which is an element of a matrix P. If the intervals between inspections are not regular, as in Estonia, then continuous-time Markov process (CTMP) using transition intensity matrix have been proposed and defined as in (3) and (4) [12]:

$$\frac{\partial}{\partial t} P = P \times Q \quad (3)$$

$$P = e^{Q \times \Delta t} = \sum_{n=0}^{\infty} \frac{(Q \times \Delta t)^n}{n!} \quad (4)$$

Where, P is the transition matrix and Q is the intensity matrix, which represents the instantaneous probability of transition between the state i and j, where $j \neq i$. The intensity matrix for the deterioration process is present in (5) is calculated using state-dependent and time-independent model. Broader investigation of different CTMP formulations were done by Kallen and Noortwijk [12] using bridge condition data for statistical estimation.

$$Q = \begin{bmatrix} -\theta_1 & \theta_1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -\theta_i & \theta_i \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (5)$$

Where θ_i is the instantaneous transition probability between adjacent state i and j. The initial estimate of matrix Q is calculated through (6) [13]:

$$\theta_i = q_{ij} = \frac{n_{ij}}{\sum \Delta t_i} \tag{6}$$

Where n_{ij} is the number of elements that moved from state i to state j, and $\sum \Delta t_i$ is the sum of intervals between observations.

The intensity matrix of abutment element group based on historical information Estonian bridges have following transition probabilities [0.0375;0.0230;0.0100].

Information is updated using Bayesian updating with informative prior. This approach has been introduced by Neves and Frangopol [3], by combining Bayesian updating with simulation for improving expert judgment. The obtained results showed significant impact on the prediction, when including the information obtained from inspections.

Based on the Bayes theorem, the probability density function of condition including the results of inspection can be defined as (7) [14]:

$$f^{"}(C_T) = K \times L(C_T) \times f^{'}(C_T) \tag{7}$$

Where $f^{"}(C_T)$ is the probability density function of the condition at the time T considering both inputs, that are present in posterior distribution, $f^{'}(C_T)$ is the probability density function of the condition at the time T considering only assessment, $L(C_T)$ is the likelihood function. K is a normalizing constant defined with (8) [14]:

$$K = \frac{1}{\int_{-\infty}^{\infty} L(C_T) \times f^{'}(C_T) dC_T} \tag{8}$$

For the Monte-Carlo simulation, the mean and standard deviation of assessments were put into the scale of 1-4. The CI at time τ, can be calculated as (9), (10) [15], [16]:

$$\mu_C^\tau = \frac{\sum_{i=1}^n C_\tau^i \times L(C_T^i)}{\sum_{i=1}^n L(C_T^i)} \tag{9}$$

$$\sigma_C^\tau = \sqrt{\frac{\sum_{i=1}^n C_\tau^i \times L(C_T^i)}{\sum_{i=1}^n L(C_T^i)} - \left(\frac{\sum_{i=1}^n C_\tau^i \times L(C_T^i)}{\sum_{i=1}^n L(C_T^i)}\right)^2} \tag{10}$$

Where μ_C^τ and σ_C^τ are the mean and standard deviation of the CI at the time τ including model and assessment, C_τ^i is the CI at time τ connected to sample i, C_T^i C_τ^i is the CI at time T connected to sample i and n is the number of samples.

4.1 Results

For correct comparison, the standard deviation of different assessments is expressed as a ratio from the mean value. In addition, it is assumed that standard deviation of degradation model is 0.1. Highest standard deviation is from visual inspections carried out by master students, the standard deviation in the condition index scale is 1.49. Lowest standard deviation of visual inspection is 0.5, which is understandable, because if one evaluates element in condition 3, then the state can be between 2.5 to 3.5.

In comparison, the most scattered NDT results have scaled standard deviation 0.47. Most precise assessment results were obtained using rebound hammer, the scaled standard deviation was 0.10.

There are three aspects of differences that were compared: deviation in lowest assessment precision (Figure 2), best visual assessment (Figure 3) and highest assessment precision (Figure 4).

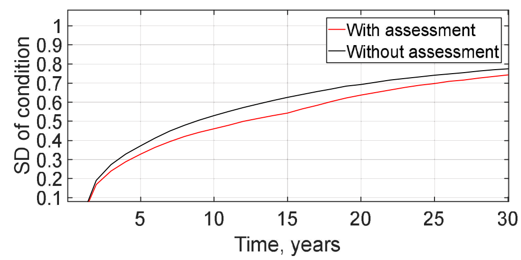

Figure 2. Standard deviation of condition with visual assessment carried out by inspector without any previous experience.

To present the influence, only one situation is visualized. The situation visualizes the change of standard deviation during 30 years of new element, which was put into operation in year 1, the assessment is made in year 15.

Even with lowest precision there is visible difference in two scenarios, which means that even without experienced inspectors, it is better for the owner to visually assess the bridges instead of using just degradation models.

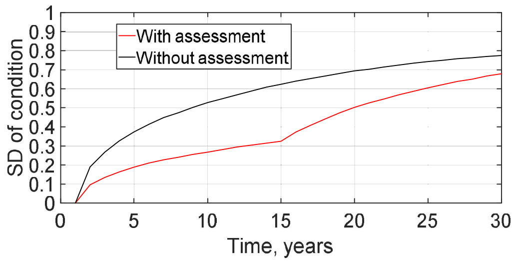

Figure 3. Standard deviation of condition with visual assessment carried out by inspector with previous experience.

In comparison of experienced and inexperienced inspectors, it is clear, that first ones can assess the elements more precisely. The difference of SD in year 15 is 0.22.

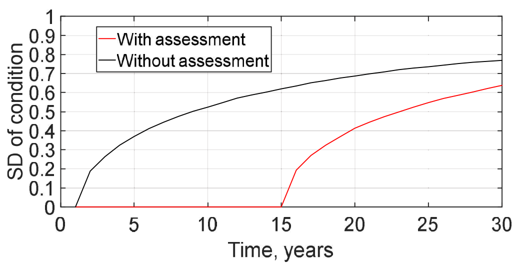

Figure 4. Standard deviation of condition with non-destructive assessment carried out by novice or expert.

Finally, in comparison of NDT and visual assessment it is also clear that in proper interpretation, the NDT is better option for precise assessment. In addition, for novice inspectors, it is more reliable to carry out tests than do visual inspection. Although, in case of testing, the places must be clarified previously.

5. Conclusions for discussion

The paper investigates the uncertainty of assessment carried out by people with different level of expertise, to present how the standard deviation of results influence condition prediction. The survey of visual assessments shows person without previous knowledge about bridge condition assessment can have almost three times higher deviation in their results than experienced ones. In visual assessment, inspectors with previous expertise can obtain results with small deviation in assessing the overall bridge condition, but element level assessment results have higher errors. In comparison of non-destructive testing methods three common and simple methods were compared. In case of test results, there are no clear difference between experienced and novice testers. In comparison of assessment methods, the most imprecise non-destructive test results are more accurate than visual assessment results.

For further investigation, it is suggested to put the non-destructive test results in similar scale to visual assessment results, add costs of assessment and analyze one method in simultaneous inspections to aggregate the number of assessments and costs during life-cycle of a bridge. In getting the most optimal condition control plan it is important to find the best combination of different methods to keep the standard deviation under desired level and minimize costs.

6. Acknowledgement

The authors acknowledge the supporting of COST Action TU1406 supported by EU Framework Program Horizon 2020, Archimedes Foundation and Professor Karl Õiger Scholarship Foundation.

7. References

[1] Strauss A, Ivankovic AM, Matos JC, Casas JR. WG1 technical report: Performance indicators for roadway bridges of COST Action 1406.

[2] Kušar M, Šelih J. Analysis of bridge condition on state network in Slovenia. Građevinar. 2014 Oct 10;66(09.):811-22.

[3] Neves, L.C. and Frangopol, D.M., 2005. Condition, safety and cost profiles for deteriorating structures with emphasis on bridges. Reliability engineering & system safety, 89(2), pp.185-198.

[4] Miyamoto A, Kawamura K, Nakamura H. Bridge management system and maintenance optimization for existing bridges. Computer-Aided Civil and Infrastructure Engineering. 2000 Jan;15(1):45-55.

[5] Yang SI, Frangopol DM, Neves LC. Optimum maintenance strategy for deteriorating

bridge structures based on lifetime functions. Engineering structures. 2006 Jan 1;28(2):196-206.

[6] Phares BM, Washer GA, Rolander DD, Graybeal BA, Moore M. Routine highway bridge inspection condition documentation accuracy and reliability. Journal of Bridge Engineering. 2004 Jul;9(4):403-13.

[7] Roberts J, Shepard R. Bridge management for the 21st century. Transportation Research Record: Journal of the Transportation Research Board. 2000 Jan 1(1696):197-203.

[8] Sein S, Matos JC, Idnurm J. Statistical analysis of reinforced concrete bridges in Estonia. Baltic Journal of Road & Bridge Engineering. 2017 Dec 1;12(4).

[9] Alwash M, Breysse D, Sbartaï ZM, Szilágyi K, Borosnyói A. Factors affecting the reliability of assessing the concrete strength by rebound hammer and cores. Construction and Building Materials. 2017 Jun 1; 140:354-63.

[10] Mateus R, Bragança L. Sustainability assessment and rating of buildings: Developing the methodology SBToolPT–H. Building and environment. 2011 Oct 1;46(10):1962-71.

[11] Scherer WT, Glagola DM. Markovian models for bridge maintenance management. Journal of Transportation Engineering. 1994 Jan;120(1):37-51.

[12] Kallen MJ, Van Noortwijk JM. Statistical inference for Markov deterioration models of bridge conditions in the Netherlands. InProceedings of the Third International Conference on Bridge Maintenance, Safety and Management (IABMAS) 2006 Jul (pp. 16-19).

[13] Jackson C. Multi-state modelling with R: the msm package. Cambridge, UK. 2007 Oct 1.

[14] Ang AH, Tang WH. Probability concepts in engineering: emphasis on applications in civil & environmental engineering. New York: Wiley; 2007.

[15] Chen MH, Shao QM, Ibrahim JG. Monte Carlo methods in Bayesian computation. Springer Science & Business Media; 2012 Dec 6.

[16] Frangopol DM, Neves LC. Structural performance updating and optimization with conflicting objectives under uncertainty. InStructures Congress 2008: 18th Analysis and Computation Specialty Conference 2008 (pp. 1-10).