Check for updates

Original Article

SHM Structural Health Monitoring

Structural Health Monitoring
1–23
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14759217251356401
journals.sagepub.com/home/shm
Sage

Test accuracy requirements for optimal decision-making in management of post-tensioned concrete bridges

Stefano Zorzi[1], Francesco Rossi[2,3], Enrico Tubaldi[2], John Quigley[2] and Daniele Zonta[1]

Abstract

The safety and reliability assessment of post-tensioned (PT) concrete bridges is critical to the management of the infrastructure. Assessing the structural health condition of PT concrete bridges is challenging due to the inaccessibility of prestressing systems. In combination with visual inspections, engineers rely on partially destructive and nondestructive testing to assess the health of the prestressing system of bridges. As a consequence, test outcomes often guide operator decisions on bridge management. However, the uncertainty of testing techniques can lead to suboptimal maintenance strategies and inefficient resource allocation. Thus, it should be taken into account in the decision-making process. This paper introduces a decision-making methodology based on Expected Utility Theory (EUT) to assess the accuracy requirements of tests to be convenient to operators. A practical application of the proposed methodology is shown in a real case study of a PT concrete viaduct in Italy. In this application, the assessment of the prestressing state of the viaduct is treated through strand-cutting tests.

Keywords

Structural health monitoring, post-tensioned reinforced concrete bridges, decision-making, non-destructive tests, expected utility theory, value of information

Introduction

The infrastructure system is central in the organization of nations and plays a critical role in their social and economic functionality.[1] After World War II, European nations expanded their highway networks, building a large number of bridges and viaducts.[2] Many of these are post-tensioned (PT) reinforced concrete bridges[3] and are approaching or exceeding their intended service life, raising concerns about structural health and safety.[4] In recent decades, about 50 structural bridge failures have occurred,[5] including catastrophic events such as the collapse of the Morandi bridge in Genoa, Italy.[6] Furthermore, the recent event in Dresden, Germany, confirmed this trend, making PT concrete bridges especially critical.[7] Aged infrastructure, increasing traffic demand, and frequent extreme events due to climate change[8,9] make the management of this infrastructure a central concern of civil society. Replacement of aging structures poses significant challenges in terms of financial, logistical, and environmental issues.[10] Therefore, infrastructure managers aim to preserve existing infrastructures through routine and extraordinary maintenance strategies.[11,12]

The assessment of PT concrete bridges with grouted duct is a significant issue for engineers. The main structural component of PT concrete bridges, the prestressing system, is hidden and cannot be monitored solely by visual inspections (VIs).[13,14] In this context, structural health monitoring plays a crucial role in ensuring the security of these key infrastructures. Engineers can employ structural monitoring as a way to identify eventual damage conditions in these structures.[15] More specifically, damage conditions can manifest

[1]University of Trento, Trento, Italy
[2]University of Strathclyde, Glasgow, UK
[3]SOCOTEC Italia Srl, Lainate, Italy

Corresponding author:
Enrico Tubaldi, University of Strathclyde, James Weir Building, Level 5, 75 Montrose Street, Glasgow G1 1XJ, UK.
Email: enrico.tubaldi@strath.ac.uk

themselves in trend variation of kinematic variables, which can be identified by long-term monitoring systems.[16] Alternatively, also in combination with structural monitoring,[17] engineers can make use of partial destructive tests (PDTs) and/or nondestructive tests (NDTs) to directly identify defects in the prestressing system (e.g., grout voids, reinforcing steel corrosion, strands/wires breaks).[18] This is confirmed by national guidelines on the management of bridges, such as the U.K. guidelines on the management of PT concrete bridges[19] and the Italian ministerial guidelines on the monitoring of existing bridges.[20]

In common practice, operators follow national guidelines, making use of PDT and NDT. In addition, operators tend to have high confidence in test outcomes. Therefore, test outcomes strongly influence operators' decisions on infrastructure management strategies. In other words, test outcomes drive infrastructure management strategies. However, recent studies have highlighted that even the most commonly used tests can be inaccurate and affected by significant uncertainties.[21–23] This constitutes a significant issue because the uncertainties involved could lead to suboptimal decisions, resulting in wasted time and resources.[12] Therefore, in the context of infrastructure management, the proper choice of the tests to be performed plays a key role. Accounting for test uncertainties is particularly important for the assessment of PT concrete bridges, as it is often not possible to complement test interpretation with information from VIs.

A rational approach to decision-making is the expected utility expected utility theory (EUT)[24,25] that provides a probabilistic-based method to guide decision-making from an economic perspective by accounting for uncertainty.[23] According to EUT, rational decision-making should account for (i) the accuracy (or the uncertainty) of inspection testing strategies and (ii) prior knowledge on the condition state of the infrastructure (i.e., knowledge before performing the test). Following a Bayesian perspective, prior knowledge plays a crucial role in decision-making.[26] Therefore, information on structure building technology, construction period, construction phase, mechanical characteristics, and condition state should be known for a rational decision-making process. In addition, test uncertainties must be considered in decision-making to avoid suboptimal management strategies (e.g., bridge closure based on a highly uncertain diagnostic test outcome when the bridge is actually safe).[27]

Based on this perspective and considering current practice in testing strategies, an important research question must be addressed: *under what conditions is a testing strategy cost-effective when the decision-making process is guided by the test outcomes?* Or, more simply, *under which conditions does it make sense to carry out a test?*

In this paper, the authors answer this relevant question which, to the best of their knowledge, represents a significant gap in the current literature. Specifically, the present paper presents an EUT-based framework for the marginal utility assessment of inspection tests. Following the proposed EUT-based framework, the present paper defines the accuracy requirements that tests must meet to be cost-effective. According to the EUT, these requirements are those that ensure that the economic gain of the test is positive. In other words, they are those that ensure that the value of information (VoI) is greater than the cost of the test.[16,28] VoI quantifies the utility of additional information to improve decision-making under uncertainty conditions and, in this context, represents the economic gain of the inspection testing strategy in decision-making.[29] The present paper focuses only on scenarios with a binary state condition of the structure (e.g., normal/abnormal condition).

The application of the proposed framework is illustrated by considering a real case study, the Alveo Vecchio viaduct. In the application, the focus is on the evaluation of prestress losses in PT cables through the strand-cutting test (SCT). In the absence of more reliable information, an expert knowledge elicitation (EKE) process[30] based on the Sheffield method[31] is developed to assess the expected accuracy of SCT and the prior knowledge on the state of prestress of the Alveo Vecchio viaduct.

The present paper is organized as follows: section "Problem statement and formulation" presents the problem statement and the formulation of the proposed EUT-based framework; section "Discussion" discusses the formulation by showing the reader the influence of each variable in the decision-making problem. At the end of this section, the essential points that should be taken into account when choosing a test are emphasized; to demonstrate how the proposed framework works in practice, section "Application" presents an illustrative application in the context of pretension losses in PT bridges; Finally, in section "Concluding remarks" the concluding remarks are presented.

Although the methodology is firmly grounded on theoretical concepts, it offers a practical approach for managers and operators to properly design the testing campaign in accordance with current national guidelines.

Problem statement and formulation

The present section illustrates the research problem and the procedure developed for solving it. In very simple

terms, the problem facing this paper can be defined by the following question:

Under what conditions does it make sense to carry out the test?

To answer this question, an EUT-based framework is proposed. Specifically, the test is assumed to be convenient if it is expected to bring a cost and/or loss reduction to the operator. From an EUT perspective, the test is marginally cost-effective if the *VoI* (i.e., the utility of additional information in terms of improved decision-making) is greater than the cost of the test. It is understood that the test is characterized by a specific level of uncertainty. The methodology allows identification of the minimum requirements for the accuracy of the test for it to be convenient. This methodology entails calculating first the prior expected cost (*prior stage*), corresponding to assuming that no test is performed. Subsequently, the pre-posterior expected cost resulting from all the possible test outcomes is calculated (*pre-posterior stage*). Finally, the two costs are compared to quantify the economic gain of the test and decide whether or not it is worth performing the test.

Before delving into the formal solution of the problem, the concepts and assumptions involved in this problem are presented.

Glossary

The present glossary formally defines the variables involved in the formulation.

State. The state, S, represents the health condition of the bridge. Each possible state is defined through a set of state parameters that can be continuous variables (e.g., the material strength or the load) and/or categorical variables (e.g., damaged or undamaged). In this paper, only binary categorical variables are considered, which are collected in the set Ω_S. In the present article, the state of the bridge can be: "normal" (N) or "abnormal" (A). N is the operational condition of the bridge, while A represents the unsafe condition of the structure. The states must be mutually exclusive and collectively exhaustive of each other. can Therefore, $\Omega_S = \{N, A\}$.

Action. The action, a, is a management strategy that the operator can undertake in response to the state of the bridge. Examples of actions are "do nothing," "repair," "replace," or "close the bridge." All possible actions are collected in the set Ω_a. In the present article, two possible actions are considered: "do nothing" (DN) and "repair" (R). Therefore, $\Omega_a = \{DN, R\}$.

Consequence. The consequence is the result or effect of an action a being the bridge in a specific state S. Examples of consequences in the context of civil infrastructure can be "nothing happens" or "the structure fails." In the present paper, only three consequences are considered: "nothing happens," "failure," and "the bridge is repaired." The consequence "nothing happens" occurs when the action is "do nothing" and the state of the structure is "normal." On the other hand, the consequence "failure" occurs when the action is "do nothing" despite the state of the structure being "abnormal." Specifically, the "failure" represents the exceeding of an operator-defined limit condition. Typically, this condition induces an increase in the probability of structural collapse of the bridge. Finally, the consequence "the bridge is repaired" occurs when the action is "repair." This last consequence is independent of the state. Furthermore, when the structure is "repaired," it is always returned to its "normal" state; therefore, no "failure" can occur.

Utility function. The utility function U is a mathematical representation that quantifies operator satisfaction with a specific consequence.[32] In economics, this satisfaction is usually called utility. Following the definition of the consequence, the utility is conditional on both state and action, that is, $u = U(a,\ S)$.

Cost. The cost, c, is the monetary quantification of the consequences. Costs can arise from various sources and take different forms. In the context of civil infrastructure, the cost can include financial losses or gains for the infrastructure owner, including the monetary quantification of casualties and social/environmental impacts arising from structural failure. Based on the definition of utility, the concept of cost has an opposite meaning to that of utility, u; therefore, cost and utility are often interchangeable, and the cost is a function of both state and action. In the present paper, the following costs are considered:

- $c_{DN|N}$ is the cost of the consequence "nothing happens," this cost is reasonably zero, that is, $c_{DN|N} = 0$
- $c_{DN|A} = c_F$ is the cost of the consequence "failure." This cost includes the potential losses that the operator incurs as a result of the failure. For example, these include the cost of reconstruction and the monetary quantification of potential casualties;

- c_R is the cost of the consequence "the bridge is repaired."c_R is independent of the state of the structure.

Test. In this paper, a test is defined as any procedure designed to gain knowledge on the state of the bridge. A test can consist of a single measurement of a combination of these within an experimental campaign. Examples of tests include NDTs, PDTs, and VIs. The cost of the test is c_{T}.

Test outcome. The outcome of the test, y, is the observation acquired from the test. It could be measurement (e.g., strain or load measurement) or a class (e.g., positive/negative). All possible test outcomes are collected in the set of test outcomes Ω_T. In the present paper, Ω_T is binary as Ω_S. Specifically, $\Omega_T = \{-, +\}$, where $-$ and $+$ stand for "negative outcome" and "positive outcome," respectively. Usually, tests are designed in such a way that the outcome of the test is a hint at the state. In the following, the "negative outcome" $(-)$ "suggests" that the state is "normal" state (N), while the "positive outcome" $(+)$ "suggests" that the state is "abnormal" state (A). The quantitative relationship between state and test outcome is properly discussed in the next paragraph.

Test accuracy. The test accuracy quantifies the capacity of tests to identify the state. In practice, the test accuracy can be defined as the proportion of states correctly classified by the test. In the case of binary test, outcome and binary state, the most commonly used accuracy parameters are the false positive rate (FPR) and the false negative rate (FNR). FPR measures the proportion of cases where the state is actually "normal" (N) and the test output is "positive" $(+)$. FNR measures the proportion of cases where the state is actually "abnormal" (A) and the test output is "negative" $(-)$. It is customary to represent the accuracy of the test through the test confusion matrix, defined as

$$\begin{bmatrix} 1-\mathrm{FPR} & \mathrm{FNR} \\ \mathrm{FPR} & 1-\mathrm{FNR} \end{bmatrix}. \tag{1}$$

The main diagonal defines the rates of having a correct classification of the state, while the cross-diagonal defines the rates of having a miss-classification of the state. Therefore, the test confusion matrix defines the likelihood of the test. This concept will be discussed throughout the following paragraphs. Finally, the test is ideal when the confusion matrix is the identity, meaning that the test always recognizes the actual state of the bridge.

Assumptions

The assumptions based on the present paper are:

- The bridge can have only two different states: "normal" (N) and "abnormal" (A). The set of states is $\Omega_S = \{N, A\}$. The states in Ω_S are mutually exclusive and collectively exhaustive of each other;
- The operator can select only two different actions: DN or "repair" (R). The set $\Omega_a = \{DN, R\}$;
- All consequences are monetizable in terms of economic costs, c;
- The utility function is linear and equal to the negative of economic costs of the consequences, that is, $u = -c$. In practice, based on the present assumption, cost and utility are interchangeable, and the optimal action is the one with the minimum cost (that is equivalent to the maximum utility);
- It is assumed that $c_R < c_F$;
- The decision-making process in the *prior stage* is represented graphically by the decision tree depicted in Figure 1. A decision tree is a logic framework to concatenate states, actions, and costs.[25] It employs chance nodes (represented by circles) and decision nodes (represented by squares). The chance nodes represent the marginalization of the costs of each action given each possible state with respect to the probability of the states. In other words, in the chance nodes, the expected costs are calculated. The decision nodes represent the process of selecting the optimal action, which is the action related to the minimum expected cost (or, equivalently, to the maximum expected utility);
- In the *posterior stage* (i.e., when it is known the outcome of the test), it is assumed that the operator will automatically follow the recommendation of the test outcome. Specifically, when the test outcome is "negative" $(-)$, the operator always chooses the action DN, and when the test outcome is "positive" $(+)$, the operator always chooses to R the bridge. Although the rule above is introduced here as an assumption, in Appendix C it is formally demonstrated that it is a logical necessity that follows the application of EUT. The decision-making process in the *posterior stage* is represented graphically in Figure 2;
- The test is characterized by only two test outcomes: negative $(-)$ or positive $(+)$. Recall that the negative outcome $(-)$ represents the expectation of having a normal state (N), while the positive outcome $(+)$ represents the expectation of having an abnormal state (A). The set of test outcomes is $\Omega_T = \{-, +\}$;

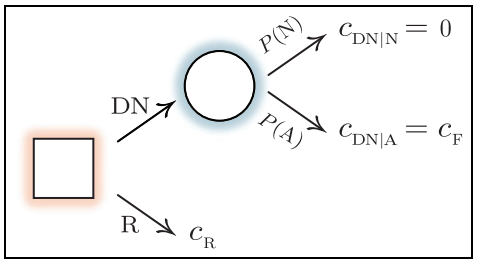


Figure 1. Decision tree of the EUT-based framework. DN: "do nothing" R: "repair" the bridge.

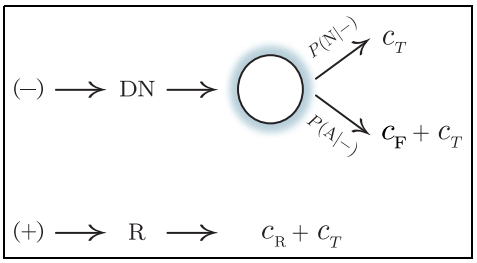


Figure 2. Decision-making process in the *pre-posterior stage*. DN: "do nothing" R: "repair" the bridge.

- The determinant of the test confusion matrix (see Equation (1)) is greater than or at most equal to zero. In other words, the confusion matrix is a positive semi-definite matrix. The reasons behind this assumption will be clarified throughout the discussion.

Prior expected cost c_p

In the *prior stage*, the decision-making process is the one represented in Figure 1. Since in the *prior stage*, the test information is not considered, the decision is made based on the prior probabilities of the states. Specifically, $P_N = P(N)$ is the prior probability of having the "normal" condition of the structural system, while $P_A = P(A)$ is the prior probability of having its "abnormal" condition. Since N and A are mutually exclusive and collectively exhaustive, $P_N = 1 - P_A$. In a decision-making process, the operator chooses the optimal action, a^*, in the set of actions, Ω_a. The optimal action is the action with the minimum expected cost given P_A (and P_N). This minimum expected cost is defined as the prior expected cost c_p. Since the cost of the action "repair" is independent of the state, the *prior* expected cost of the action R, $c_R^{(I)}$, is obviously equal to c_R. On the other hand, the *prior* expected cost of the action "do nothing," $c_{DN}^{(I)}$, is defined as

$$c_{DN}^{(I)} = (1 - P_A)\, c_{DN|N} + P_A\, c_{DN|A} = P_A\, c_F. \qquad (2)$$

Based on $c_{DN}^{(I)}$ and $c_R^{(I)}$, the optimal action, a^*, is the one that minimizes the cost. The *prior* expected cost, c_p, is the expected cost of the optimal action, a^*, that is,

$$c_p = \min\left[c_{DN}^{(I)}, c_R^{(I)}\right] = \min[P_A\, c_F, c_R]. \qquad (3)$$

Figure 3 shows the result of the decision-making process a priori. Specifically, this figure shows c_{DN} and c_R for all the possible values of P_A. Highlighted in green is the optimal cost, which is related to the optimal action, a^*. It can be observed that there is a value of P_A, denoted as $\overline{P}_A = c_R/c_F$, for which the operator changes its decision. More specifically, for $P_A < P_A$, the optimal action a priori is DN, otherwise, for $P_A > \bar{P}_A$, the optimal action a priori is R. The condition $P_A = \overline{P}_A$ is a condition of indifference.

Pre-posterior expected cost c_{pp}

At the *posterior stage*, the decision-making process is the one represented in Figure 2. Unlike the *prior stage*, the chance nodes are now governed by the posterior probabilities of the state, which is the probability of having each state given a specific test outcome. In this case, the posterior probabilities are $P(N\,|\,y)$ and $P(A\,|\,y)$, where y represents the test outcome, which can be negative $(-)$ or positive $(+)$. In particular, the decision-making process in Figure 2 explicitly reports $P(N\,|-)$ and $P(A\,|-)$, which are, respectively, the probability of having a "normal" and an "abnormal" condition given a negative outcome of the test. The posterior probabilities are calculated by means of Bayes' formula as

$$P(\mathrm{N}\,|\,y) = \frac{P(y\,|\,\mathrm{N})\,P_N}{P(y)}; \qquad P(\mathrm{A}\,|\,y) = \frac{P(y\,|\,\mathrm{A})\,P_{\mathrm{A}}}{P(y)}\ . \quad (4)$$

In practice, with Bayes' formula, the prior information on the state is integrated with the test information (the likelihood) in order to infer the state a posteriori. In Equation (4), P_N and P_A are the prior probabilities of the state ("normal" and "abnormal," respectively), and $P(y\,|\,N)$ and $P(y\,|\,A)$ are the likelihood terms. Recall that the likelihood terms represent the test accuracy and are contained in the test confusion matrix (see Equation (1)). In particular, $P(+\,|\,\mathrm{N}) = \mathrm{FPR}$ and $P(-\,|\,\mathrm{A}) = \mathrm{FNR}$. Finally, the denominator of the Bayes formula, named evidence or marginal likelihood, is the marginal probability of the test outcome y. This term assumes the role of normalization factor.

Similarly to the *prior stage*, the *posterior* expected cost of the action "repair," $c_R^{(II)}$, does not depend on the state. However, the cost of the test has now to be

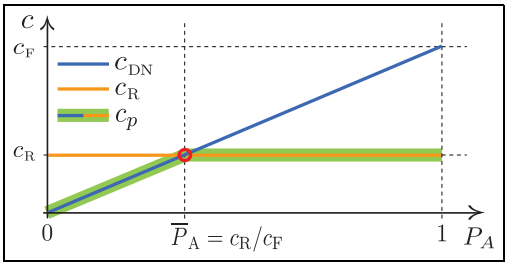


Figure 3. Prior expected cost c_p, corresponding to the optimal action given P_A.

considered, that is, $c_R^{(II)} = c_R + c_T$. On the other hand, the *posterior* expected cost of the action "do nothing," $c_{DN}^{(II)}$, is now defined as

$$c_{DN}^{(II)} = P(\mathrm{N} \mid -)\, c_T + P(\mathrm{A} \mid -)\, (c_F + c_T). \tag{5}$$

$c_{DN}^{(II)}$ and $c_R^{(II)}$ are the costs after knowing the test outcome (i.e., a posteriori). Now, to estimate the *posterior* expected cost before actually performing the experiment, a *pre-posterior* analysis is proposed. Specifically, the *pre-posterior* expected cost, c_{pp}, is obtained by marginalizing $c_{DN}^{(II)}$ and $c_R^{(II)}$ with respect to the marginal distribution of observations $P(-)$ and $P(+)$. In formula,

$$c_{pp} = P(-)\, c_{DN}^{(II)} + P(+)\, c_R^{(II)}. \tag{6}$$

Based on Equations (4) and (6), the *pre-posterior* expected cost can be rewritten as

$$c_{pp} = (1 - P_A)\, c_{N,pp} + P_A\, c_{A,pp}, \tag{7}$$

where $c_{N,pp}$ and $c_{A,pp}$ are the expected cost a *pre-posteriori* when the state is "normal" and "abnormal," respectively. $c_{N,pp}$ and $c_{A,pp}$ are defined as

$$c_{N,pp} = \mathrm{FPR}\, c_R + c_T; \quad c_{A,pp} = \mathrm{FNR}\, c_F + (1 - \mathrm{FNR})\, c_R + c_T. \tag{8}$$

Please refer to Appendix A for further details on its derivation. Figure 4 illustrates the result of the decision-making process *pre-posteriori*.

Figure 4(a) and (b) clarifies the practical meaning of costs $c_{N,pp}$ *and* $c_{A,pp}$. Specifically, Figure 4(a) shows the expected cost a *pre-posteriori* when the state is "normal," $c_{N,pp}$. In this condition, the minimum expected cost is obtained when $\mathrm{FPR} = 0$. In fact, the operator always chooses the optimal action: "do nothing"; therefore $c_{N,pp} = c_T$. Oppositely, the maximum expected cost is obtained when $FPR = 1$. Now, the operator always chooses the suboptimal action "repair" the bridge; therefore $c_{N,pp} = c_R + c_T$. In general, $c_{N,pp}$ is a linear interpolation between c_T and $c_R + c_T$ with respect to FPR. Similar reasoning can be followed to explain Figure 4(b), showing the expected cost a *pre-posteriori* when the state is "abnormal," $c_{A,pp}$. In this condition, the minimum expected cost is obtained when $FNR = 0$. In fact, the operator always chooses the optimal action: "repair" the bridge; therefore $c_{A,pp} = c_R + c_T$. Oppositely, the maximum expected cost is obtained when $FNR = 1$. In this case, the operator always chooses the suboptimal action "do nothing"; therefore $c_{A,pp} = c_F + c_T$. In general, $c_{A,pp}$ is now a linear interpolation between $c_R + c_T$ and $c_F + c_T$ with respect to *FNR*. Finally, Figure 4(c) represents the variation of the *pre-posterior* expected cost, c_{pp}, with respect to P_A, given $c_{N,pp}$ and $c_{A,pp}$. In this figure, the black line represents the variation of c_{pp} with respect to P_A. Observe that c_{pp} is the linear interpolation between $c_{N,pp}$ and $c_{A,pp}$ defined in Equation (6). The *prior* expected cost, c_p, is now represented in the figure by a green curve.

Required accuracy domain

Having defined the *prior* expected cost and the *pre-posterior* expected cost, it is now possible to tackle the question posed at the beginning of the present section. As already discussed, the test is assumed to be convenient if it is marginally cost-effective to the operator; therefore, the test is convenient if the economic gain of the test is greater than zero. The economic gain is denoted by the symbol G_T and is defined as

$$G_T = c_p - c_{pp}. \tag{9}$$

For readers familiar with the concept of VoI,[28,29] it can be observed that, under the linearity conditions of the utility function, $\mathrm{VoI} = G_T + c_T$. In addition, referring to Figure 4(c), it is worth observing that the economic gain G_T is the vertical distance between the black curve, c_{pp}, and the green curve, c_p.

Now, the shaded area in Figure 4(c) represents the range of P_A where the test is convenient for the operator (i.e., $G_T > 0$). Developed based on Equation (9), this range is defined as $\overline{P}_{A,1} < P_A < \overline{P}_{A,2}$, where

$$\begin{aligned} P_{A,1} &= \frac{\mathrm{FPR} + \gamma_T}{\gamma_F\,(1 - \mathrm{FNR}) - (1 - \mathrm{FPR} - \mathrm{FNR})}; \\ \overline{P}_{A,2} &= \frac{(1 - \mathrm{FPR}) - \gamma_T}{\gamma_F\,\mathrm{FNR} + (1 - \mathrm{FPR} - \mathrm{FNR})}. \end{aligned} \tag{10}$$

When P_A is out of this range, the operator would not follow the test result but would always choose the optimal action in the prior stage; therefore, the test is not convenient. Specifically, when $P_A < \overline{P}_{A,1}$ the operator chooses to "do nothing" (DN), while when $P_A > \overline{P}_{A,2}$ the operator chooses to "repair" the bridge (*R*).

Based on the definition of c_p and c_{pp}, G_T is a function of the test accuracy parameters (FPR and FNR),

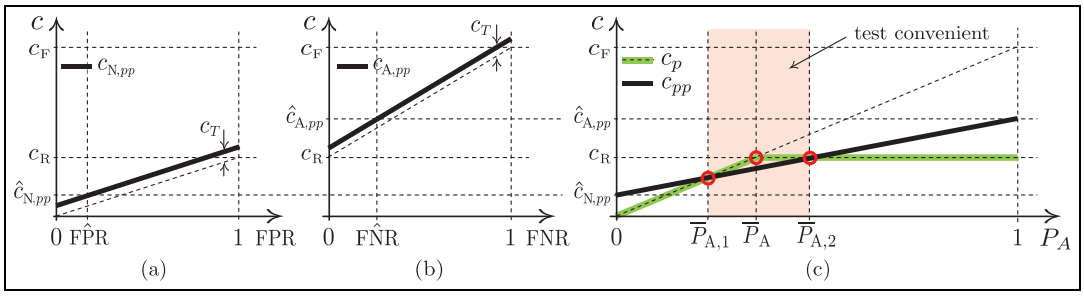


Figure 4. Pre-posterior expected cost, c_{pp}. (a) $c_{N,pp}$ with respect to FPR. (b) $c_{A,pp}$ with respect to FNR. (c) c_{pp} with respect to P_A. The black line represents the pre-posterior expected cost, c_{pp}. The green line represents the prior expected cost, c_p. The shaded area represents the range of P_A where the test is convenient.
FPR: false positive rate; FNR: false negative rate.

as well as the prior probability of the states (P_A and P_N) and of the costs (c_F, c_R, and c_T). Now, the *required accuracy domain*, D_{RA}, is defined as the subspace of the accuracy parameters (FPR and FNR) such that the economic gain of the test is positive, that is, $G_T>0$. In practice, if FPR and FNR are in D_{RA}, the test is convenient for the operator. Furthermore, the boundary of D_{RA} represents the minimum requirements of accuracy for which the test is convenient for the operator.

Based on Equation (9), FPR and FNR are in the *required accuracy domain*, D_{RA}, if

$$\text{FPR} < q_0 + q_1 \,\text{FNR}, \tag{11}$$

where q_0 (the intercept) and q_1 (the slope) are defined as

$$\begin{aligned} q_0 &= min\left[1, \frac{P_A}{1-P_A}(\gamma_F - 1)\right] - \frac{\gamma_T}{1-P_A}; \\ q_1 &= -\frac{P_A}{1-P_A}(\gamma_F - 1). \end{aligned} \tag{12}$$

In this equation, γ_F and γ_T are the normalized costs defined as

$$\gamma_F = \frac{c_F}{c_R}; \qquad \gamma_T = \frac{c_T}{c_R}. \tag{13}$$

Appendix B illustrates the details of this derivation. In addition, the reader should remember that $0 \leqslant \text{FPR} \leqslant 1$ and $0 \leqslant \text{FNR} \leqslant 1$. The *required accuracy domain*, D_{RA}, is represented graphically in Figure 5. In particular, the shaded area is the *required accuracy domain*, that is, the domain of FNR and FPR for which the test is convenient. The shape of this domain is governed by the relative costs (γ_F and γ_T) and the prior probability of the states (P_A) as defined in Equations (11) and (12).

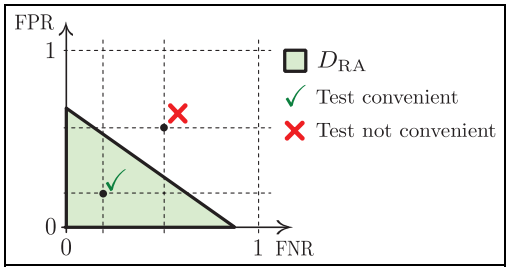


Figure 5. Graphical representation of the *required accuracy domain*, D_{RA}.

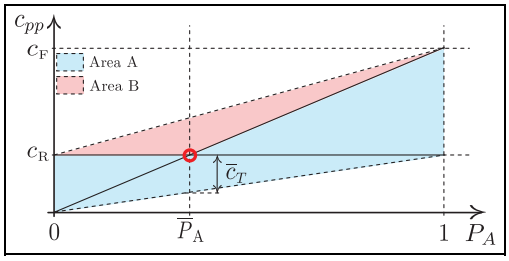


Figure 6. Possible configurations of c_{pp} when $c_T = 0$. $\bar{c}_T$ is the maximum cost defined in Equation (15).
FPR: false positive rate.

Discussion

The present section illustrates the implications of the proposed EUT-based framework.

Figure 6 shows the possible configurations of c_{pp} when $c_T = 0$. Specifically, depending on P_A, c_{pp} can be located in the union of Areas A and B.

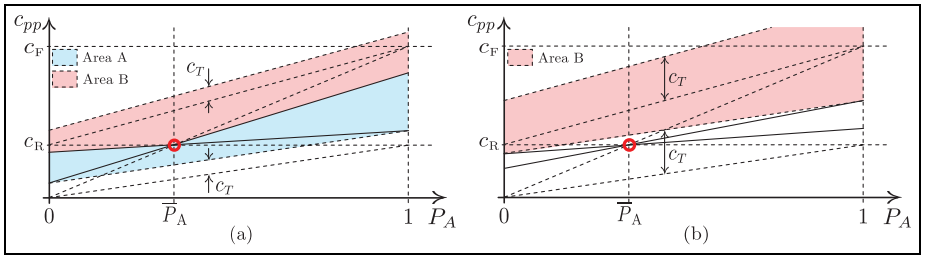


Figure 7. Possible configurations of c_{pp} when $c_T>0$. (a) Possible configurations c_{pp} when $c_T<\bar{c}_T$. (b) Possible configurations of c_{pp} when $c_T>\bar{c}_T$.
FPR: false positive rate.

In practice, the union of Areas A and B is obtained on the basis of Equation (7), considering all possible combinations of FPR and FNR. In particular, the lower edge of Area A represents the lower bound of c_{pp}, which is reached when the test information is perfect (FPR = 0 and FNR = 0). Now, based on Equation (7), it is straightforward to prove that c_{pp} can be located in Area B only if FPR + FNR>1 (the determinant of the test confusion matrix is negative). Note that the determinant of the test confusion matrix is negative when the φ-coefficient of the test is negative. In statistics, the φ-coefficient (also called Matthews' correlation coefficient) is a measure of correlation for two binary variables.[33] Specifically, the φ-coefficient measures the correlation between the test outcomes and the actual states. Since the test accuracy is known *a-priori*, if the φ-coefficient is negative, the operator can always reverse the decision (or the test outcomes), obtaining a positive φ-coefficient and, therefore, a positive determinant of the test confusion matrix. This motivates the assumption of semi-positive definite test confusion matrix (see the subsection "Assumptions"). As a consequence, c_{pp} cannot be located in Area B, but only in Areas A of Figure 6.

Now, Figure 7 illustrates the general case where the test has a cost greater than zero. The observations and considerations regarding the trends of Figure 6 also apply to the case where $c_T> 0$. However, in this case, the cost of the test c_T also plays a crucial role. Specifically, c_{pp} can now be located in Area B only if

$$\mathrm{FPR}+\mathrm{FNR}>1-\gamma_T\frac{\gamma_F}{\gamma_F-1}. \tag{14}$$

In practice, as shown in Figure 7(a), the cost of the test reduces Areas A and increases Area B. Area A completely disappears when $c_T>\bar{c}_T$, where the upper limit of the test cost, $\bar{c}_T$, is defined as

$$\bar{c}_T=\left(1-\frac{c_R}{c_F}\right)c_R. \tag{15}$$

This upper limit cost corresponds to the upper bound of VoI obtained when, a priori, a condition of indifference occurs (see Figure 3) and the test information is perfect.[34] In this condition, as shown in Figure 7(b), the test is never convenient for the operator for any combination of P_A, FNR, and FPR.

It has been shown how the parameters that govern the problem affect c_{pp}. Now, a discussion of how these parameters affect the *required accuracy domain*, D_{RA}, is proposed. Recall that the required accuracy domain is the subspace of the accuracy parameters where the test is convenient. Developed based on Equations (11) and (12), Figure 8 illustrates how the prior probability of the state (P_A) and the relative costs (γ_F and γ_T) affect the required accuracy domain.

First, note that the domain is represented by a shaded area and is bounded by a line. This is due to the assumption of a linear utility function. The influence of the cost of the test, c_T, is represented in Figure 8(a). The cost c_T reduces the required accuracy domain by shifting the boundary downward. This shift is equal to $\gamma_T/(1-P_A)$ (see Equations (11) and (12)). Figure 7(b) and (c) shows how P_A and c_F influence the slope of the boundary of D_{RA}. In practice, when P_A tends to be very small, the operator is highly confident a priori that the bridge condition is "normal" (N); therefore, the operator is strongly inclined to "do nothing" (a = DN). In this condition, the operator decides to change the action (i.e., to "repair" the bridge) only if the test is characterized by a very low FPR. Otherwise, there would be a high probability of unnecessary repairs, leading to wasted economic resources. For this reason, when $P_A<\overline{P}_A$, tests should be characterized by a low FPR to be cost-effective for the operator. In other terms, the convenience of the test is mainly controlled

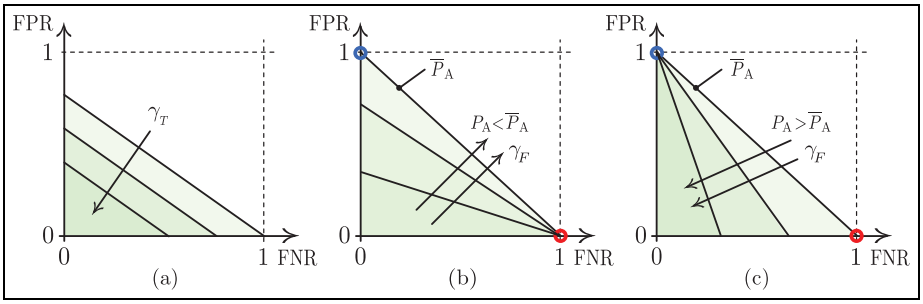


Figure 8. Possible configurations of the required accuracy domain, D_{RA}. (a) Variation of D_{RA} when γ_T variates. (b) Variation of D_{RA} when γ_F or P_A variates. Since $P_A<\overline{P}_A$, the boundary of D_{RA} rotates around the red pole ($FNR = 1$). (c) Variation of D_{RA} when γ_F or P_A variates. Since $P_A>\overline{P}_A$, the boundary of D_{RA} rotates around the blue pole ($FPR = 1$).
FPR: false positive rate.

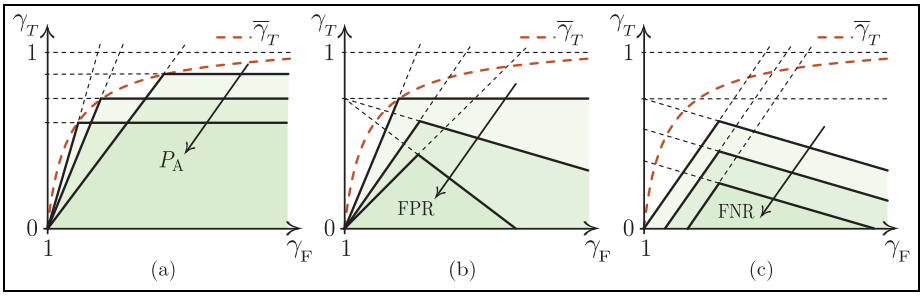


Figure 9. Representation of the required accuracy domain with respect to the costs given P_A, FPR, and FNR. The shaded area represents the domain where the test is convenient. The dashed red line represents the trend of the normalized maximum test cost as defined in Equation (15). (a) Variation of the domain when P_A variates. (b) Variation of the domain when FPR variates. (c) Variation of the domain when FNR variates.
FPR: false positive rate; FNR: false negative rate.

by FPR; therefore, when FPR is low, FNR may also be high. The reader can observe graphically this evidence in Figure 8(b), where $P_A<\overline{P}_A$ and $\gamma_T = 0$; specifically, observe that the domain boundary rotates around the red pole. On the other hand, when P_A tends to be very large, the operator is highly confident a priori that the bridge condition is "abnormal" (A); therefore, the operator is strongly inclined to proceed with "repair" the bridge ($a = \mathrm{R}$). Now, the operator decides to change the action (i.e., to "do nothing") only if the test is characterized by a very low FNR. Otherwise, there would be a high probability of "doing nothing" when it would be necessary to "repair" the bridge. For this reason, when $P_A>\overline{P}_A$, tests should be characterized by a low FNR to be cost-effective for the operator. In other terms, the convenience of the test is mainly controlled by FNR; therefore, when FNR is low, FPR may also be high. The reader can now observe graphically this evidence in Figure 8(c), where $P_A>\overline{P}_A$ and $\gamma_T = 0$; specifically, observe that the domain boundary rotates around the blue pole

Developed based on Equation (11), Figure 9 provides an alternative representation of the *required accuracy domain*; specifically, the figure shows that domain in terms of the relative costs γ_F and γ_T, given P_A, FNR, and FPR.

This alternative domain can be considered to assess under what combinations of cost a specific test is convenient. This alternative domain is defined through the following inequalities:

$$\gamma_{F,min}<\gamma_F<\gamma_{F,max}\ , \tag{16}$$

where

$$\gamma_{F,min} = 1 + \frac{(1-P_A)\ \mathrm{FPR} + \gamma_T}{P_A(1-\mathrm{FNR})};$$
$$\gamma_{F,max} = 1 + \frac{(1-P_A)\ (1-\mathrm{FPR}) - \gamma_T}{P_A\ \mathrm{FNR}}\ . \quad (17)$$

Figures 9(a) to (c) show the effect of P_A, FPR, and FNR on this domain, respectively. In these figures, the dashed red line represents $\bar{\gamma}_T = \bar{c}_T/c_R$ (see Equation (15)) as γ_F varies. Developed based on Equation (16), the maximum relative cost considering the test accuracy is derived as

$$\gamma_{T,max} = (1-P_A)(1-\mathrm{FPR}-\mathrm{FNR}). \quad (18)$$

As expected, $\gamma_{T,max}$ occurs in correspondence with $\gamma_F = 1/P_A$ (the state indifference a priori). The use of this alternative domain is clarified at the end of the case study of this paper.

Summary of the discussion

In this section, the authors have been provided with a discussion of the role of all variables within the proposed EUT-based framework. The main considerations are listed:

- In general, the condition for the test to be convenient is given by Equation (11), which depends on the prior probabilities and the costs. If this condition is met, the economic gain from the test is positive (i.e., $G_T>0$);
- When the test cost is zero, there is always at least a combination of P_A, FNR, and FPR such that $G_T>0$;
- There exists an upper limit of the test cost, $\bar{c}_T$, defined in Equation (15), such that the test is not always convenient, regardless of all other parameters;
- The test is convenient if relative cost γ_F is included between $\gamma_{F,min}$ and $\gamma_{F,max}$, which depends on the costs and the accuracy of the test (see Equations (16) and (17));
- If P_A is lower than the indifference point in the *prior stage* (i.e., $P_A < \overline{P}_A$), the convenience of the test is mainly controlled by the FPR. This means that when the FPR is low, the FNR may also be high;
- If P_A is greater than the indifference point in the *prior stage* (i.e., $P_A > \overline{P}_A$), the convenience of the test is mainly controlled by the FNR. This means that when the FNR is low, the FPR may also be high.

In the following section, an illustrative application of the proposed EUT-based framework is presented.

Application

The present section illustrates an illustrative application of the proposed EUT-based framework to a real case study. The context of this application is the monitoring of the level of prestressing of the cables of the Alveo Vecchio viaduct girders. Among the testing technologies available for assessing the prestress state, this application considers the SCT. To the best of the authors' knowledge, specific studies on the accuracy of SCT are lacking in the literature. To overcome this limitation, the authors employed EKE to estimate the accuracy of SCT. Before delving into the results of the application, an overview of the Alveo Vecchio viaduct, the SCT, and the EKE procedure for estimating the test accuracy is provided.

Case study: Alveo Vecchio viaduct

The Alveo Vecchio viaduct is a PT reinforced concrete bridge designed and built between 1966 and 1968 along the *A16 Napoli–Canosa* highway near the municipality of Candela, Apulia, south of Italy. The viaduct was decommissioned in 2005 after a landslide that caused a pier to settle and the deck across a span to collapse. The viaduct construction technique is very common and can be considered representative of approximately 50% of the viaducts built in Europe.

The Alveo Vecchio viaduct consists of two structurally independent carriageways (one for each traffic direction). The viaduct piers are 3.30 m high and have deep foundations consisting of eight 23-m-long piles with a diameter of 1.2 m. The viaduct abutments are built on six piles with a diameter of 1.2 m. Each carriageway is composed of three decks, each 35.5 m long. The decks are independent of each other and are simply supported on the piers and abutments. They are placed on a slight slope (1.45%). Each deck consists of four PT reinforced concrete girders connected by five cross-girders and a 20-cm-thick reinforced concrete slab. Figure 10 shows the structural components of the Alveo Vecchio viaduct.

The girder prestressing system is composed of 14 PT parabolic cables. The initial prestress amounts to 1250 MPa (average tension) and is applied on site using hydraulic jacks. Each cable consists of 12 parallel high-strength steel wires with a nominal ultimate strength of 1700 MPa and a nominal yield strength of 1450 MPa. The diameter of each wire is equal to

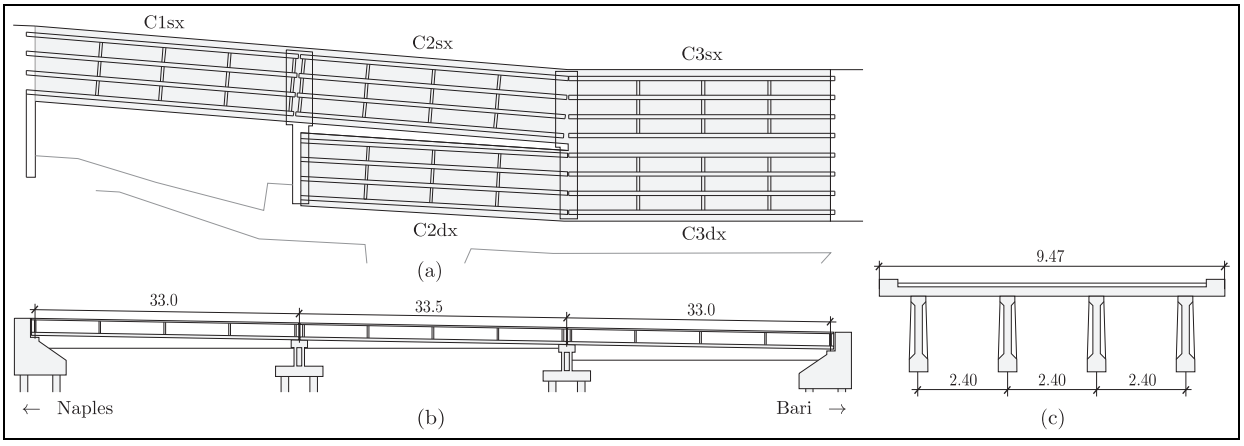


Figure 10. Top view (a), lateral view (b), and cross section (c) of the Alveo Vecchio viaduct.

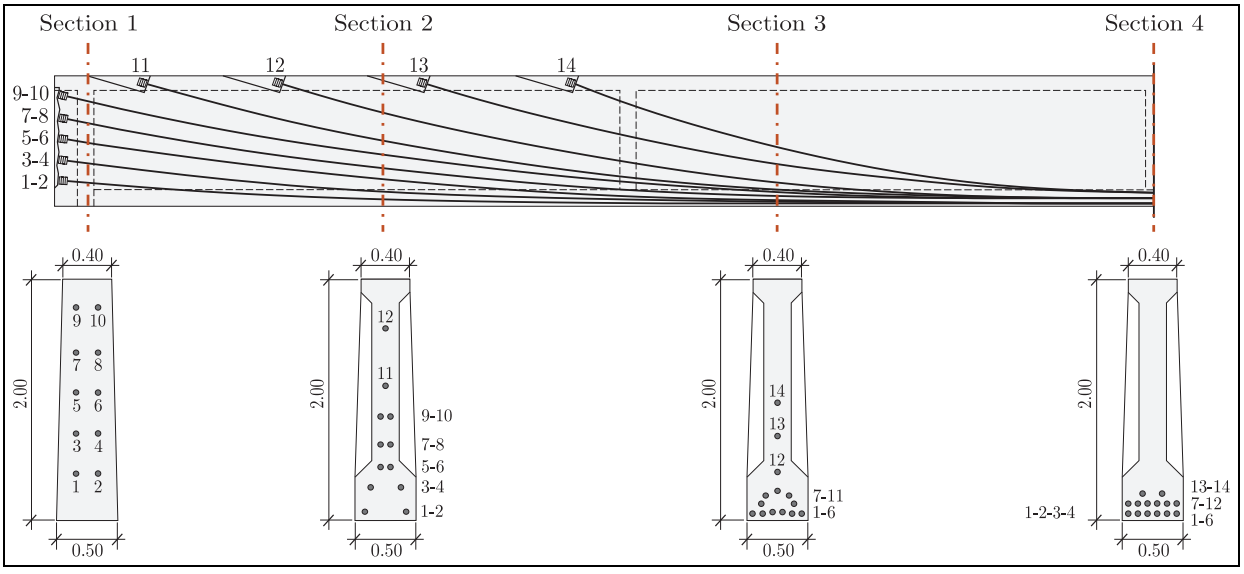


Figure 11. Longitudinal section and cross sections of the girders of the Alveo Vecchio viaduct.

7 mm. The 12 wires were placed in a corrugated metal duct and bonded to the girder by high-pressure grout injection. Figure 11 presents the prestressing system.

Specifically, the figure shows the longitudinal layout of the cables within the girder and the location of the cables in four cross sections. Further details about the Alveo Vecchio viaduct are provided in the study by Tonelli et al.[35]

The context of the present application is the estimation of the state of prestressing of the viaduct girders. The loss of prestress is a common and pivotal issue in PT reinforced structures, especially bridges and viaducts.[36] In fact, loss of prestress can be related to corrosion phenomena in prestressing steel[37] and can lead to excessive deformation and cracking of structural components. Cracks reduce the flexural stiffness of the deck, leading to a further decrease in service performance.[38] In addition, cracks can accelerate the degradation process of materials, particularly reinforcement steel, further compromising the long-term safety of the structure.

Actions and consequences

In this application, the authors imagine that the bridge is in an operational condition. The possible states of the bridge are two: "normal" condition (N) and "abnormal" condition (A); therefore, the set of states is $\Omega_S = \{N, A\}$. In this application, the authors assume that the state of the viaduct is considered "abnormal"

(A) when the mean residual pretension in the cables is below 10% of the design tension, $\sigma_{p,d}$. This tension value constitutes the threshold of the "abnormal" condition σ_p, where the subscript $_p$ stands for "prestress" and the overline indicates a threshold. It should be noted that the prestress level is, in principle, a continuous quantity. However, in the present work, the prestressing state is discretized in a binary form by means of the threshold σ_p. This simplification is introduced due to the difficulty in defining the consequences of misclassification for continuous states. Obviously, the state of the viaduct is considered "normal" (N) when the mean residual pretension in the tendons is above σ_p. The design prestress, $\sigma_{p,d}$, is defined as the difference between the initial prestress, $\sigma_{p,0}$, and the expected tension losses $\Delta\sigma_p$ at the time of the assessment. Tension losses are derived from the design documentation and include short-term losses and losses due to rheological phenomena (concrete creep and shrinkage and steel relaxation). Specifically, the initial tension $\sigma_{p,0}$ is equal to 1250 MPa, while the expected tension loss $\Delta\sigma_p$ is equal to 350 MPa. As a result, the design prestress $\sigma_{p,d}$ is equal to 900 MPa, and the threshold of "abnormal" condition σ_p is equal to 810 MPa.

The actions that the operator can perform are DN and "repair" the viaduct (R); therefore, the set of actions is $\Omega_a = \{DN, R\}$. In this context, the action "repair" involves conducting additional investigations, preventive interventions such as limiting traffic load, and possibly performing external pretension interventions.

The consequences of the action defined above are the following: "nothing happens" when the action is DN and the state is "normal" (N); "failure" when the action is DN and the state is "abnormal" (A); and "the viaduct is repaired" when the action is "repair" (R). Note that the consequence "failure" does not correspond to the "collapse" of the viaduct (of part of this). In fact, excessive prestress loss is a degradation phenomenon that does not directly result in collapse. By "collapse," the authors refer to any mechanism that may lead to structural failure of the bridge due to excessive prestress loss, resulting in the traffic interruption. These failure mechanisms include shear and bending failures. In particular, shear failure mechanisms, which are the most relevant in the case of prestress losses, are caused by the loss of the confinement effect provided by axial prestressing forces.[39] On the other hand, bending failure mechanisms are mainly associated with local corrosion phenomena (pitting), which can be correlated with excessive prestress loss.[40] In this application, the authors assume that the probability of collapse given the excessive prestress loss, $P_{C|A}$, is equal to 2.5% (capital C stands for "collapse"). On the other hand, the probability of collapse in case of a normal state $p_{C|N}$ is assumed to be negligible. It is important to note that these assumptions made by the authors have a direct impact on the definition of the consequence costs involved in the decision-making problem.

The monetary quantification of the action "repair," c_R is the direct cost of restoring the safety of a viaduct deck. This cost is assumed to be 0.12 k€/m^2· 320 m^2 = 38.4 k€. In the present application, the cost of "failure," c_F is obtained as the product between the cost of collapse c_C and the probability of collapse given the "abnormal" state, $P_{C|A}$. The cost of collapse, c_C, can be assumed as the sum of two contributions: the direct cost of reconstructing a deck of the viaduct, $c_{C,1}$, and the indirect cost of the expected casualties, $c_{C,2}$. The cost of reconstruction, $c_{C,1}$, is assumed to be 4.0 k€/m^2· 320 m^2 = 1280.0 k€. Based on Brighenti,[41] the cost of casualties, $c_{C,2}$, can be calculated as

$$c_{C,2} = w_{C,2} \cdot n_{C,2}, \tag{19}$$

where $w_{C,2}$ is the unit cost of casualty, while $n_{C,2}$ is the expected number of casualties. The first is assumed to be 3840.0 k€.[28] The expected number of casualties is calculated using the following formula[41]:

$$n_{C,2} = \frac{ADT}{\bar{v}} L_{max}\, n_{pv}\, \lambda_v, \tag{20}$$

where ADT is the average daily traffic, $\bar{v}$ is the average speed of the traffic, L_{max} is the maximum length of the viaduct spans, n_{pv} is the expected number of people per vehicle, and λ_v is the lethality of the collapse. In the present application, assuming the viaduct is still in operation, ADT = 65,570 vehicles/day,[42] $\bar{v}$ is assumed to be 90 km/h, L_{max} = 32.0 m, n_{pv} is assumed to be 1.33 $person/vehicle$, and λ_v is assumed to be 0.5. As a result, $n_{C,2}$ = 0.647 and, therefore, $c_{C,2}$ = 2482.5 k€. The monetary quantification of the collapse is $c_C = c_{C,1} + c_{C,2} = 3762.5$ k€. The cost of "failure" is $c_F = P_{C|A} \cdot c_C$ = 0.025·3762.5 k€ = 94.1 k€. Finally, the monetary quantification of the consequences of "nothing happens" is obviously equal to zero. The costs of the consequences involved in the present application are summarized in Table 1.

Strand-cutting test

SCT is a widely used technique to determine the residual prestress level in prestressed concrete structures.[43] The measurement of residual prestress is performed by recording the strain release of one or more strands when they are cut. The setup of SCT is represented in Figure 12.[44,45]

Specifically, the strain release is measured by means of one or more strain gauges glued to the surface of the wire, close to the location of the cut. This strain

Table 1. Costs of consequences.

Consequences	Costs	Relative costs	Description
Nothing happens	0	—	Expected cost of "nothing happens"
Repair the deck	c_R = 38.4 k€	γ_R = 1	Expected cost of repair the deck
Failure	$c_{C,1}$ = 1280.0 k€	—	Expected cost of reconstruction after collapse
	$c_{C,2}$ = 2482.5 k€	—	Expected cost of casualties due to collapse
	$c_C = c_{C,1} + c_{C,2}$ = 3762.5 k€	—	Total expected cost of collapse
	$c_F = P_{C\|A} \cdot c_C$ = 94.1 k€	γ_F = 2.45	Expected cost of failure
Perform the test	c_T = 1.5 k€	γ_T = 0.04	Expected cost for SCT performance

SCT: strand-cutting test.

measurement is converted into residual prestress value by Hooke's law as

$$\Delta\sigma_p = E_p \, \Delta\varepsilon_{\mathrm{p}}, \tag{21}$$

where $\Delta\sigma_p$ is the residual prestress, $\Delta\varepsilon_p$ is the strain variation, and E_p is the Young's modulus of steel wires.[46] Although this technique may appear straightforward in both execution and interpretation, several critical aspects must be considered. The selection of measurement points must be carefully made by the operator to ensure a sufficient transfer length.[47] The prestress is not uniformly distributed along the tendon layout due to differential stress losses caused by anchorage effects. Moreover, test results are highly dependent on the quality of execution: strands must be properly isolated from the grout to allow full elastic deformation, and shock due to the release of strand tension can introduce systematic errors in measurement instruments.[48] Finally, since the accuracy of prestress force estimation depends on the knowledge of the Young's modulus of steel wires. For this reason, SCT should be accompanied by one or more tensile tests on steel wires.[45,47]

SCT is classified as an indirect test, as prestress is inferred by strain measurements rather than direct assessment. Furthermore, SCT is a PDT, as its execution requires local removal of the concrete cover, exposure of the strands within the conduit, and cutting of one or more wires.[45] Consequently, SCT must be carefully designed to mitigate its impact on structural integrity.

Although qualitative concerns about the accuracy of SCT have been raised, to the best of the authors' knowledge, no studies in the literature provide a quantitative characterization of its accuracy. Since assessing the economic value of test information is infeasible without a quantitative understanding of test accuracy, in the absence of more reliable data, the authors propose employing EKE techniques to estimate the expert expectation on SCT accuracy.

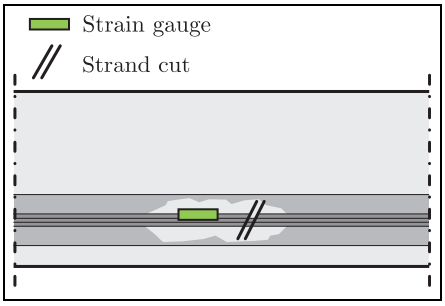


Figure 12. Setup of the strand-cutting test.[44,45]

Expert knowledge elicitation

EKE is a structured process used to collect, quantify, and formalize expert judgments, particularly in contexts where empirical data are scarce, incomplete, or uncertain. It employs systematic methods to extract and represent expert opinions in a form that supports decision-making.[30]

In the present paper, the authors adopt a process inspired by the Sheffield elicitation method,[31] which is based on the principles of behavioral aggregation. In this method, one or more facilitators guide a group of experts through a structured workshop to exchange information and reach a consensus on the probability distributions of the quantities of interest. The conventional Sheffield elicitation approach consists of three main phases. First, a preliminary training phase introduces the experts to the quantities of interest, outlines key assumptions, and clarifies the statistical concepts involved. Experts are also encouraged to share insights based on their domain knowledge in this phase. Next, an elicitation phase is conducted, typically using a structured questionnaire to capture expert judgments on the probability distributions of the quantities of interest. Finally, in the aggregation phase, the elicited probability distributions are shared with all experts, fostering discussion until a consensus-based pooled distribution is reached.

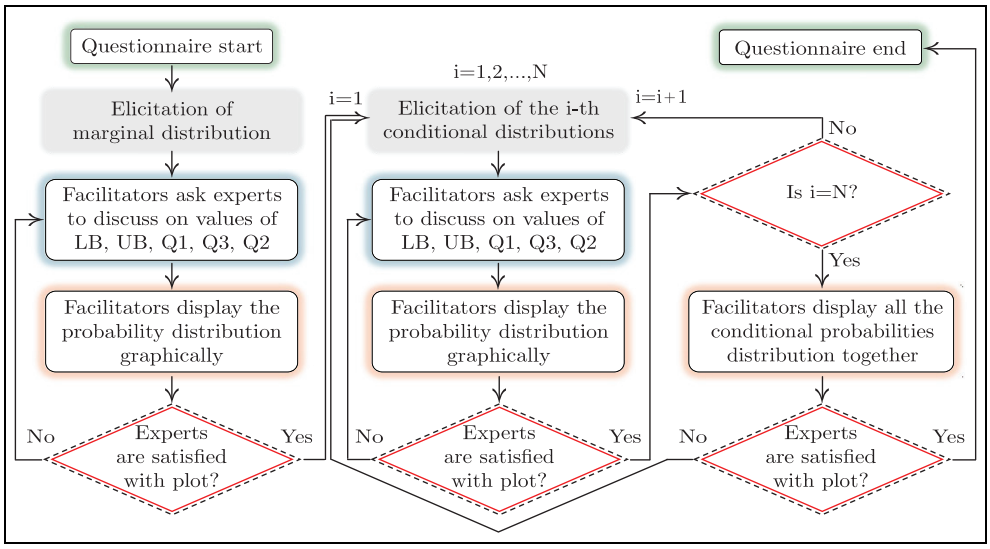


Figure 13. Flow chart of expert elicitation.

In the present application, the goal of the EKE process is to quantify the expert expectation on the accuracy of the SCT and the prior state of prestress in the Alveo Vecchio viaduct girders. In practice, the goal of the EKE process is to assess FPR, FNR, and P_A. These quantities are calculated from the joint distribution $P(\sigma_p, \tilde{\sigma}_p)$, where σ_p is the actual level of prestress on the steel wires, and $\tilde{\sigma}_p$ is the level of prestress measured by SCT. According to the total probability law,

$$p(\sigma_p, \tilde{\sigma}_p) = p(\tilde{\sigma}_p \mid \sigma_p)\, p(\sigma_p), \tag{22}$$

where $p(\tilde{\sigma}_p \mid \sigma_p)$ is the conditional distribution of $\tilde{\sigma}_p$ given σ_p, and $p(\sigma_p)$ is the marginal distribution of σ_p. Therefore, the quantities of interest of the EKE process are σ_p (the actual residual prestress) and $\tilde{\sigma}_p \mid \sigma_p$ (the prestress measured by SCT conditional on σ_p).

Unlike the conventional Sheffield elicitation method, in the present study, the pooled distributions are derived from a single questionnaire administered to experts during a single workshop. In this workshop, experts interact with each other until they converge on a pooled probability distribution. Therefore, knowledge aggregation occurs directly during the questionnaire phase through the interaction guided by two facilitators. This represents the main divergence from the conventional Sheffield elicitation method, where the elicitation phase and the aggregation phase occur at different times. The EKE process of the present application consists of three main phases:

a. Briefing. In this phase, experts are instructed on the context of the Alveo Vecchio viaduct (geometry, materials, structural scheme, construction period, etc.), the details of the prestressing system (technology, design tension, expected losses), and the details of the SCT (test characteristics, location, and number of tests);
b. Training. In order for all experts to share a common understanding of the problem, experts are informed of the quantities of interest. In addition, experts are informed of the ways in which EKE is conducted. Finally, experts are trained in the use of probabilistic quantities for defining quantities of interest (percentile concept, distribution concept, skewness concept, etc.);
c. Expert elicitation. In this phase, two facilitators guide experts to interact with each other in order to converge on a shared distribution with respect to the quantities of interest. First, experts are invited to introduce themselves and share with everyone their background with respect to the elicitation context. In addition, they are asked to declare any conflicts of interest. The facilitators then propose the questionnaire to the experts. The questionnaire phase is an iterative process, described graphically in the flow chart in Figure 13. First, the marginal distribution of the actual residual prestress $p(\sigma_p)$ is elicited. Next, the conditional distribution of the predisposition measured by SCT $p(\tilde{\sigma}_p \mid \sigma_p)$ is elicited conditional on the upper and lower bounds and on the 25th, 50th, and 75th percentiles of the marginal distribution. The elicitation process is based on the quartile method.[49] To mitigate central bias,[50] quartiles are elicited from the experts following this sequence: lower bound (LB), upper bound (UB), 25th quartile (Q1), 75th quartile (Q3), and 50th quartile (Q2, the median). The probability distributions are obtained as a fit of the quartiles using a Beta distribution. Among distributions, Beta is upper and

Table 2. Result of the EKE process: quartiles.

Quartile	σ_p	$\tilde{\sigma}_p\|0$	$\tilde{\sigma}_p\|250$	$\tilde{\sigma}_p\|550$	$\tilde{\sigma}_p\|850$	$\tilde{\sigma}_p\|1100$	$\tilde{\sigma}_p\|1350$	$\tilde{\sigma}_p\|1500$
LB	0	0	150	400	600	800	1100	1200
25th quartile (Q1)	600	1	200	450	700	950	1200	1300
50th quartile (Q2)	800	2	230	510	800	1050	1250	1400
75th quartile (Q3)	1100	3	260	550	850	1100	1350	1500
UB	1500	10	300	650	950	1250	1500	1700

EKE: expert knowledge elicitation; LB: lower bound; UB: upper bound.

Table 3. Optimal beta distribution parameters for marginal and conditional distributions.

Quartile	σ_p	$\tilde{\sigma}_p\|0$	$\tilde{\sigma}_p\|250$	$\tilde{\sigma}_p\|550$	$\tilde{\sigma}_p\|850$	$\tilde{\sigma}_p\|1100$	$\tilde{\sigma}_p\|1350$	$\tilde{\sigma}_p\|1500$
Parameter α	2.247	1.519	1.607	1.193	2.653	2.119	1.376	1.196
Parameter β	1.858	5.396	1.432	1.593	2.109	1.956	1.870	1.692
LB	0	0	150	400	600	800	1100	1200
UB	1500	10	300	650	950	1250	1500	1700
Mean	821.0	2.296	229.3	507.0	795.0	1034	1270	1407
Standard deviation	330.4	1.472	37.26	63.57	72.43	99.80	95.93	124.9
Coefficient of variation (%)	40.2	64.1	16.2	12.5	9.1	9.7	7.6	8.9

LB: lower bound; UB: upper bound.

lower bounded and is particularly flexible, offering many possibilities for skewness. Each distribution obtained on the basis of the above quartiles is shown graphically to the experts to foster debate until a consensus is reached. As pointed out in the study by Gosling,[31] experts may not be familiar with interpreting the plots of density functions. To mitigate this issue, facilitators provide experts with information on the mean and coefficient of variation of density functions. These quantities are in common use in the engineering field and are easily interpreted by the experts. In addition, facilitators ensure that experts agree on the symmetry/asymmetry of the distributions to make sure that the effect of skewness of distributions is properly taken into account.

To ensure complete coverage of the relevant areas of expertise,[51] in the present application the following three experts were involved: (i) a Professor from the University of Trento specializing in PT structure design; (ii) a Professor from the University of Chieti-Pescara with academic expertise in the use of SCT; and (iii) an operator from a multinational company operating in engineering diagnostics, representing the practical execution of SCT in the field. Note that each expert covers a specific area of SCT accuracy knowledge.

Results of EKE and assessment of the accuracy of SCT

The results of EKE are reported in terms of quartile values in Table 2. The quartiles of each column in Table 2 are modeled using a Beta distribution, $\text{Beta}(\alpha, \beta, \text{LB}, \text{UB})$, to analytically define the marginal distribution of σ_p and the conditional distributions of $\tilde{\sigma}_p$ given σ_p.

The optimal Beta distribution parameters (i.e., α, β, LB, and UB), the mean, the standard deviation, and the coefficient of variation for each distribution are provided in Table 3.

Based on the values in Table 3, the joint distribution, $p(\sigma_p, \tilde{\sigma}_p)$, is computed using the formula in Equation (22). Figure 14(a) shows graphically the joint distribution, $p(\sigma_p, \tilde{\sigma}_p)$ in the shape of a contour plot.

It can be observed that, according to experts, the SCT is highly accurate. In fact, since the distribution seems to be very tight, there is a high correlation between σ_p and $\tilde{\sigma}_p$. Furthermore, the distribution is slightly skewed toward the bottom left. This means that the SCT tends to slightly underestimate the actual tension of the steel wires. During the elicitation phase, the three experts discussed with each other until they agreed on the probability distribution of each quantity of interest. Facilitators found that the discussion was often governed by the most charismatic expert. However, each expert contributed significantly to the construction of each pooled distribution. Moreover, focusing on the distributions conditional on high values of σ_p, it can be observed that experts converge on a coefficient of variation of about 10%. At the end of the elicitation phase, the experts expressed a strong consensus on the configuration of the pooled joint distribution, the final expected result of the EKE process. However, this result may have been influenced by the

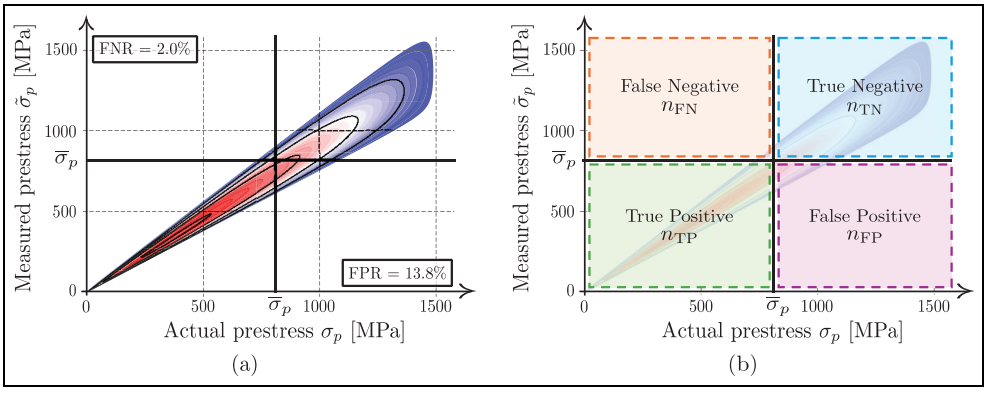


Figure 14. Result of the EKE process. (a) Contour plot of the joint probability distribution $P(\sigma_p, \tilde{\sigma}_p)$. The dashed bold black line represents the threshold of the "abnormal" condition, $\bar{\sigma}_p$. The results in terms of FPR and FNR are represented. (b) Graphical representation of true/false negatives and true/false positives.
EKE: expert knowledge elicitation; FPR: false positive rate.

limited number of experts involved in the elicitation. In addition, the pooled joint distribution might be affected by the correlation arising from the exposure of experts to the same context. In fact, all experts are Italian, and two of them come from academia. However, the Sheffield method offers participants the opportunity to use their own judgment about the correlations between experts' evaluations, although it is difficult to evaluate how effectively they can take advantage of this opportunity.[52] In principle, involving a larger number of experts from different backgrounds may lead to more reliable pooled distributions. Nevertheless, these limitations affect only the application of the proposed methodology, not the methodology itself. These aspects could be addressed in future developments of this work.

Now, based on the joint distribution $p(\sigma_p, \tilde{\sigma}_p)$, FPR and FNR can be calculated as

$$\begin{aligned} \mathrm{FPR} &= \frac{\int_0^{\bar{\sigma}_p} \int_{\bar{\sigma}_p}^{\infty} p(\sigma_p, \tilde{\sigma}_p)\, \mathrm{d}\sigma_p\, \mathrm{d}\tilde{\sigma}_p}{\int_0^{\infty} \int_{\bar{\sigma}_p}^{\infty} p(\sigma_p, \tilde{\sigma}_p)\, \mathrm{d}\sigma_p\, \mathrm{d}\tilde{\sigma}_p}; \\ \mathrm{FNR} &= \frac{\int_{\bar{\sigma}_p}^{\infty} \int_0^{\bar{\sigma}_p} p(\sigma_p, \tilde{\sigma}_p)\, \mathrm{d}\sigma_p\, \mathrm{d}\tilde{\sigma}_p}{\int_0^{\infty} \int_0^{\bar{\sigma}_p} p(\sigma_p, \tilde{\sigma}_p)\, \mathrm{d}\sigma_p\, \mathrm{d}\tilde{\sigma}_p}, \end{aligned} \tag{23}$$

where $\bar{\sigma}_p$ is the threshold of the "abnormal" condition (810 MPa). Equation (23) is solved numerically by means of the Metropolis–Hastings algorithm,[53] which is used to generate a set of representative samples from the joint distribution $P(\sigma_p, \tilde{\sigma}_p)$ given in Equation (22). For this purpose, a proposal Gaussian distribution with zero mean and an uncorrelated covariance matrix with standard deviation equal to 150 MPa for both σ_p and $\tilde{\sigma}_p$ is considered. An increasing number of samples is generated until convergence, which is achieved after a total of 10^5 samples. The first 10% of these (50,000 samples) is neglected to burn the initial path. Based on the generated set of samples, FPR and FNR are estimated as

$$\mathrm{FPR} = \frac{n_{FP}}{n_{TN} + n_{FP}}; \qquad \mathrm{FNR} = \frac{n_{FN}}{n_{TP} + n_{FN}}, \tag{24}$$

where, based on Figure 14(b), $n_{FP} = 32{,}262$ is the number of samples in the false positive area (bottom right), $n_{TP} = 201{,}616$ is the number of samples in the true positive area (bottom left), $n_{FN} = 4254$ is the number of samples in the false negative area (top left), and $n_{TN} = 211{,}868$ is the number of samples in the true negative area (top right). Finally, the prior probability of having an "abnormal" state is calculated as

$$P_A = \frac{n_{TP} + n_{FN}}{n_{TN} + n_{FP} + n_{TP} + n_{FN}}. \tag{25}$$

As a result, the false positive ratio is equal to FPR = 13.8%, the false negative ratio is equal to FNR = 2.0% and the prior probability of having an "abnormal" state is equal to $P_A = 48.0\%$.

Decision-making

Based on the results of the EKE process, the optimal action at the *prior stage* (without test information) is "repair" the viaduct (R). In fact, the expected cost of the action DN is equal to $c_{DN} = P_A\, c_F = 45.2$ k€, which is greater than the cost of "repair," $c_R = 38.4$ k€. Therefore, the *prior* expected cost, c_p, corresponds to the cost of "repair," that is, 38.4 k€.

At this stage, the operator may wonder whether performing an SCT campaign may be convenient. The

cost of this test campaign is equal to $c_T = 1.5$ k€ and, according to experts, the SCT is characterized by FPR = 13.8% and FNR = 2.0%. Based on Equation (7), the *pre-posterior* expected cost, c_{pp}, is equal to 23.2 k€. Since $c_{pp} < c_p$, the test is convenient for the operator. Specifically, the expected economic gain of SCT is positive, that is, $G_T = c_p - c_{pp} = 15.2$ k€. The above results are represented graphically in Figure 15.

Finally, based on Equation (11), the required accuracy domain, D_{RA}, is defined. Recall that the required accuracy domain, D_{RA}, fully defines the accuracy requirements that the test should meet to be cost-effective according to EUT. As expected, the accuracy parameters of the SCT are abundantly included in D_{RA}, as shown in Figure 16(a). Figure 16(b) shows the cost domain within which the SCT is cost-effective for the operator, given fixed values of FPR, FNR, and P_A. In practice, this domain expresses the sensitivity of the cost parameters with respect to the economic gain of the SCT. In particular, the SCT is convenient if the relative cost of γ_F is between 1.24 and 43.5. Specifically, when γ_F is less than 1.24, it is always cost-effective to DN. On the other hand, when γ_F is greater than 43.5, it is always cost-effective to "repair" the viaduct (R). Finally, the maximum relative cost of the test γ_T is 0.434, corresponding to 16.7 k€. As expected, this maximum cost is equal to the VoI.

In this section, the authors present the application of the proposed EUT-based framework to the real case study of the Alveo Vecchio viaduct. In particular, the SCT is considered to assess the prestress losses in the viaduct tendons. In the absence of reliable data regarding the test accuracy, an EKE is carried out, through which the test's FPR and FNR are estimated. The prestressing state is binarized using a threshold value σ_p. This simplification allows the definition of the consequences of misclassification and the identification of combinations of FNR and FPR for which the test is economically convenient for the operator, given the expected consequences (see Figure 16(a)). Future studies may allow for the consideration of continuous or multi-class states through a more general definition of consequences, with particular reference to failure-related outcomes.

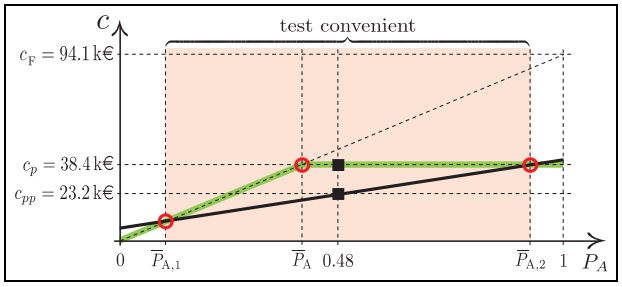


Figure 15. Representation of c_p and c_{pp} given $c_F = 94.1$ *k*, $c_T = 1.5$ *k*, and $P_A = 48.0\%$.
EKE: expert knowledge elicitation; FPR: false positive rate; FNR: false negative rate.

Concluding remarks

Partially destructive or non-destructive testing on PT concrete bridges is essential to understand the structural health of these structures, whose reliability is strongly controlled by the state of the prestressing system. The test outcomes guide the operator's decisions on bridge management. Uncertainties in testing can lead to suboptimal maintenance strategies; therefore, these must be considered in the decision-making process. In the present paper, the authors have proposed a decision-making methodology based on EUT to assess

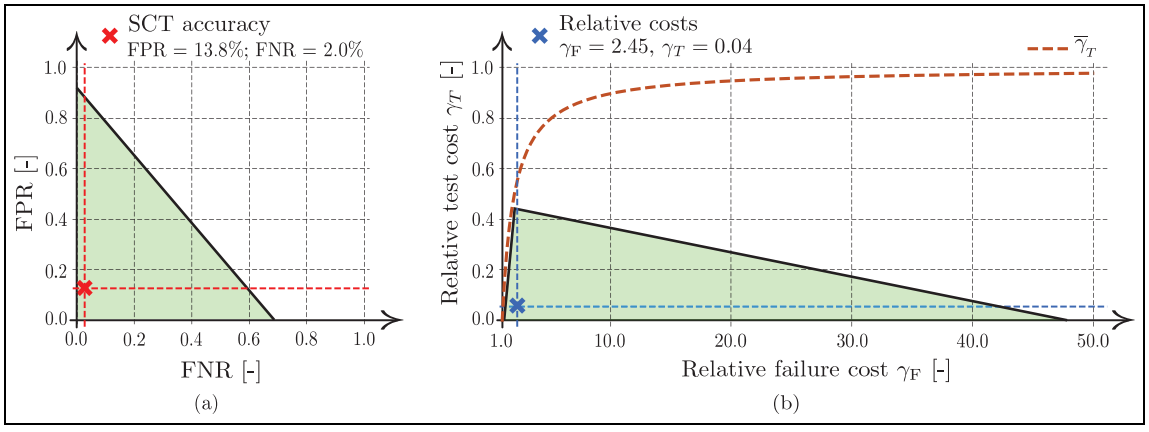


Figure 16. Results of the EKE process. (a) Required accuracy domain given $\gamma_F = 2.45$, $\gamma_T = 0.04$, and $P_A = 48.0\%$; (b) Combinations of relative costs given FPR = 13.8%, FNR = 2.0%, and $P_A = 48.0\%$.
EKE: expert knowledge elicitation; FPR: false positive rate; FNR: false negative rate.

the accuracy requirements for tests to be convenient to operators. To overcome the absence of specific studies on the efficacy of this test, an EKE process is considered for the evaluation of FPR and FNR, in addition to prior knowledge on the state of the Alveo Vecchio viaduct.

The main findings and observations of the paper are briefly summarized below:

- A straightforward formula has been derived to define the test accuracy requirements for binary state scenarios. These requirements have been proposed in the form of a subspace of FPR and FNR, named the required accuracy domain. If FPR and FNR belong to the required accuracy domain, the test is cost-effective for the operator;
- A discussion of the role of all variables within the proposed EUT-based methodology has been proposed. Specifically, limitations have been identified in terms of the cost of consequences arising from the test uncertainty. In addition, the conditions for which a large FPR is acceptable and those for which a large FNR is acceptable have been discussed;
- A demonstrative case study application has been provided to show the reader how the proposed methodology could work in practice. In this application, the SCT has been considered for the indirect evaluation of the prestressing state of the Alveo Vecchio viaduct. Nevertheless, this framework can also be applied with other tests commonly used for assessing the condition of PT prestressed concrete bridges;

The main limitations and possible future developments of the proposed methodology are as follows:

- Due to a lack of quantitative information about the prior knowledge and the accuracy of the test, recourse was made to EKE. However, this approach is difficult to apply in practice; further research is essential to obtain reliable and ready-to-use information on the test accuracy, not only for the SCT but also for other tests commonly used to evaluate the condition assessment of prestressed concrete bridges.
- Limitation of the current paper is the binary representation of state and a linear utility function. Although this represents a limitation, it is a necessity in this first work. In fact, greater emphasis has been placed on the role of each variable in the decision-making process instead of the generality of the method. Future developments should consider multi-class/continuous representations of the state. Furthermore, the potential integration with degradation models should be explored in order to optimize inspections over time based on EUT. Further research could also investigate the role of aleatory uncertainties in defining test accuracy requirements.

Acknowledgments

The authors wish to acknowledge S Perno (Sapienza University of Rome), G Brando (University "G.d'Annunzio" of Chieti-Pescara), A Bonelli (University of Trento), FA Faccia (SOCOTEC Italia Srl) for their collaboration in the EKE process. The authors would also like to acknowledge the people who contributed to the success of this project, including P Migliorino (Italian Ministry of Infrastructure and Transport); A Selleri, A Marchiondelli, M Cicolani, M Conte, L Tripoli, P Anfosso, M Di Napoli, M Perna, and D Sena (Autostrade per l'Italia SpA).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study presented in this paper was funded by: the ReLUIS Interuniversity Consortium under the agreement DPC-ReLUIS and the Superior Council of Public Works stipulated pursuant to art. 3 of the Decree of the Minister of Infrastructure no. 578 of December 17, 2020 (ReLUIS 2020-2022) and DPC-ReLUIS 2024-2026; the European Union under Next Generation EU, Mission 4 Component 2 CUP E53D23003560006 (HORUS). The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union, the European Research Executive Agency, the Italian Superior Council of Public Works, the Italian Department of Civil Protection (DPC) or the ReLUIS Council's position and assessments; neither the European Union nor any of the granting authority can be held responsible for them.

ORCID iDs

Stefano Zorzi https://orcid.org/0000-0002-3537-4649
Enrico Tubaldi https://orcid.org/0000-0001-8565-8917

Data availability statement

All data generated and analyzed during the EKE process are fully available within this article. No additional datasets were created or analyzed beyond those included in the published text.

References

1. Bhide S, Culmo M, Ma J, et al. The Interstate Highway System and the development of prestressed concrete bridges. *Interstate Struct* 2006; 49.

2. Billington DP. Historical perspective on prestressed concrete. *PCI J* 2004; 49(1): 14–30.
3. Rodriguez S. Design of long span concrete box girder bridges: challenges and solutions, in: structures. *Structures* 2004; 2004: 1–11.
4. Paulìk P, Makita T, Bathen L, et al. *Management of post-tensioned bridges, fib Bulletins, fib*. The International Federation for Structural Concrete, 2023.
5. Bado MF and Casas JR. A review of recent distributed optical fiber sensors applications for civil engineering structural health monitoring. *Sensors* 2021; 21(4): 1818.
6. Calvi GM, Moratti M, O'Reilly GJ, et al. Once upon a Time in Italy: the Tale of the Morandi Bridge. *Struct Eng Int* 2019; 29(2): 198–217.
7. Reynodls J. *Shocking moment Dresden bridge collapses into the river below just moments after tram had gone over the crossing*. Daily Mail, September 2024.
8. Henning TFP; World Bank. *Integrating climate change into road asset management*. World Bank, 2017.
9. Ng W-S. Impact of climate change on infrastructure. In: Leal Filho W, Azul AM, Brandli L, et al. (eds.) *Industry, innovation and infrastructure*. Cham: Springer International Publishing, 2021, pp.489–497.
10. Regier R and Hoult NA. Distributed strain behavior of a reinforced concrete bridge: case study. *J Bridge Eng* 2014; 19(12): 05014007.
11. Poli F, Bado MF, Verzobio A, et al. Bridge structural safety assessment: a novel solution to uncertainty in the inspection practice. *Struct Infrastruct Eng* 2023; 21: 421–435.
12. Zorzi S, Bado MF and Zonta D. Optimizing on-site testing for post-tensioned concrete bridges. *Struct Infrastruct Eng* 26 October 2024. DOI: 10.1080/15732479.2024.2407951.
13. Zatar W, Ruiz FM and Nghiem H. Revolutionizing concrete bridge assessment: implementing nondestructive scanning for transformative evaluation. *Appl Sci* 2024; 14(24): 11590.
14. Hurlebaus S, Hueste MBD, Karthik MM, et al. *Inspection guidelines for bridge post-tensioning and stay cable systems using NDE methods*. Washington, DC: Transportation Research Board, 2017.
15. Farrar CR and Worden K. An introduction to structural health monitoring. *Philos Trans R Soc A Math Phys Eng Sci* 2007; 365(1851): 303–315.
16. Pozzi M and Der Kiureghian A. Assessing the value of information for long-term structural health monitoring. *Proc SPIE* 2011; 7984: 79842W.
17. Zhang W-H, Qin J, Lu D-G, et al. VoI analysis of temporally continuous SHM information in the context of adaptive risk-based inspection planning. *Struct Saf* 2022; 99: 102258.
18. Latte Bovio F, Chichi F, Celati S, et al. Assessment of inspection procedures for pre-stressed concrete bridges with post-tensioned cables. In: *Proceedings of the 1st Conference of the European Association on Quality Control of Bridges and Structures*, vol. 200. Cham: Springer International Publishing, 2022, pp. 1090–1099.
19. Highways England. *CS 465–Management of post-tensioned concrete bridges*. Version 0. Design Manual for Roads and Bridges, Technical Report, National Highways Birmingham, UK, 2020.
20. Ministero delle Infrastrutture e dei Trasporti, Consiglio Superiore dei Lavori Pubblici. *Linee guida per la classificazione e gestione del rischio, la valutazione della sicurezza ed il monitoraggio dei ponti esistenti*. Roma, Italy: Ministero delle Infrastrutture e dei Trasporti, Consiglio Superiore dei Lavori Pubblici, 2020.
21. Mundell C, Hendy C, Argyle T, et al. Advancements in bridges NDT: National Highways' "Structures Moonshot"—Paper No. 1. In: *Bridge maintenance, safety, management, digitalization and sustainability*, 1st ed. London: CRC Press, 2024, pp.3653–3662.
22. Mundell C, Hendy C, Argyle T, et al. Advancements in Bridges NDT: National Highways' "Structures Moonshot"—Paper No. 2. In: *Bridge maintenance, safety, management, digitalization and sustainability*, 1st ed. London: CRC Press, 2024, pp.3663–3672.
23. Zorzi S, Rossi F, Tonelli D, et al. Metrological validation of Non-Destructive Testing techniques for assessing the condition state of pre-stressing systems. In: *Proceedings of the 14th International Workshop on Structural Health Monitoring 2023: Designing SHM for Sustainability, Maintainability, and Reliability*, 1st ed. London: DES tech Publications, Inc., 2023, pp.2417–2424.
24. Von Neumann J and Morgenstern O. *Theory of games and economic behavior: 60th anniversary commemorative edition*. Princeton University Press, 2007.
25. Raiffa H and Schlaifer R. *Applied statistical decision theory*. Boston, MA: Harvard University, 1961.
26. Cappello C, Zonta D, Pozzi M, et al. Impact of prior perception on bridge health diagnosis. *J Civil Struct Health Monit* 2015; 5(4): 509–525.
27. Cappello C, Zonta D and Glisic B. Expected utility theory for monitoring-based decision making. *Proc IEEE* 2016; 104(8): 1647–1661.
28. Zonta D, Glisic B and Adriaenssens S. Value of information: impact of monitoring on decision-making. *Struct Control Health Monit* 2014; 21(7): 1043–1056.
29. Degroot MH. Changes in utility as information. *Theory Decis* 1994; 17(3): 287–303.
30. Dias LC, Morton A and Quigley J (eds.). *Elicitation: The Science and Art of Structuring Judgement, Vol. 261 of International Series in Operations Research & Management Science*. Cham: Springer International Publishing, 2017.
31. Gosling JP.SHELF: The Sheffield Elicitation Framework. In: Dias LC, Morton A and Quigley J (eds.) *Elicitation*, vol. 261. Cham: Springer International Publishing, 2017, pp 61–93.
32. Arrow KJ. Utilities, attitudes, choices: a review note. *Econometrica* 1958; 26(1): 1.
33. Matthews B. Comparison of the predicted and observed secondary structure of T4 phagelysozyme. *Biochim Biophys Acta* 1975; 405(2): 442–451.
34. Bertola NJ, Proverbio M and Smith IFC. Framework to approximate the value of information of bridge load

testing for reserve capacity assessment. *Front Built Environ* 2020; 6: 65.
35. Tonelli D, Rossi F, Brighenti F, et al. Prestressed concrete bridge tested to failure: the Alveo Vecchio viaduct case study. *J Civil Struct Health Monit* 2023; 13(4–5): 873–899.
36. Bonopera M, Chang K-C and Lee Z-K. State-of-the-art review on determining prestress losses in prestressed concrete girders. *Appl Sci* 2020; 10(20): 7257.
37. Podolny W. Corrosion of prestressing steels and its mitigation. *PCI J* 1992; 37(5): 34–55.
38. Wang L.Flexural behaviors of corroded post-tensioned concrete beams. In: *Strand corrosion in prestressed concrete structures*. Springer Nature Singapore, 2023, pp.193–223.
39. Granata MF, Recupero A, Culotta G, et al. Influence of axial force and corrosion on failure of prestressed concrete structures considering M-V interaction. *Eng Struct* 2021; 242: 112552.
40. Yu Q-Q, Gu X-L, Zeng Y-H, et al. Flexural behavior of corrosion-damaged prestressed concrete beams. *Eng Struct* 2022; 272: 114985.
41. Brighenti F. *Risk-based bridge asset management*. PhD Thesis, University of Trento, 2025.
42. Associazione Italiana Società Concessionarie Autostrade e Trafori. AISCAT Information – Semiannual Edition 1-2/2023, Technical Report. 1-2/2023, Rome, Italy, 2023.
43. Baran E, Shield C and French C.A comparison of methods for experimentally determining prestress losses in pretensioned prestressed concrete girders. In: Russell B and Gross S (eds.) *Special Publication on Historic Innovations in Prestressed Concrete: SP-231*. American Concrete Institute, 2005, pp.161–180.
44. Bagge N, Nilimaa J and Elfgren L. In-situ methods to determine residual prestress forces in concrete bridges. *Eng Struct* 2017; 135: 41–52.
45. Kralovanec J and Prokop J. Indirect methods for determining the state of prestressing. *Transport Res Proc* 2021; 55: 1236–1243.
46. Abdel-Jaber H and Glisic B. Monitoring of prestressing forces in prestressed concrete structures—an overview. *Struct Control Health Monit* 2019; 26(8): 2374.
47. Agredo Chàvez J, Gonzalez-Libreros C, Wang L, et al. Assessment of residual prestress in existing concrete bridges: the Kalix bridge. *Eng Struct* 2024; 311: 118194.
48. Halsey JT and Miller R. Destructive testing of two forty-year-old prestressed concrete bridge beams. *PCI J* 1996; 41(5): 84–93.
49. Quigley J and Walls L. Reconciling experts opinion concerning the value of testing using Bayesian networks: a bridge too far? In: *5th International ASRANet Conference*, Edinburgh, UK, 14 June -16 June 2010
50. Seaver DA, Winterfeldt DV and Edwards W. Eliciting subjective probability distributions on continuous variables. *Organ Behav Hum Perf* 1978; 21: 379–391.
51. Bolger F. The selection of experts for (probabilistic) expert knowledge elicitation. In: Dias LC, Morton A and Quigley J (eds.) *Elicitation*, vol. 261. Cham: Springer International Publishing, 2017, pp.393–443.
52. Wilson KJ and Farrow M. Combining judgements from correlated experts. In: Dias LC, Morton A and Quigley J (eds.) *Elicitation*, vol. 261. Cham: Springer International Publishing, 2017, pp.211–240.
53. Bolstad WM. *Understanding computational Bayesian statistics*. Wiley series in computational statistics. Hoboken, NJ: Wiley, 2010.

Appendix

Appendix A

Pre-posterior expected cost, c_{pp}.

In the present appendix, the authors derive the final form of Equation (6). The starting form of Equation (6) is

$$c_{pp} = P(-)\, c_{DN}^{(II)} + P(+)\, c_R^{(II)}. \tag{A.1}$$

The steps to obtain the final form are:

- Replace $c_{DN}^{(II)}$ and $c_R^{(II)}$ with their explicit definition. The result is

$$c_{pp} = P(-)\, P(A \mid -)\, c_F + P(+)\, c_R + c_{\mathrm{T}}. \tag{A.2}$$

- Replace the posterior probabilities . and . with the explicit form of Bayes' theorem. The result is

$$c_{pp} = P(-)\, \frac{P(- \mid A)\, P(A)}{P(-)}\, c_F + P(+)\, c_R + c_{\mathrm{T}}. \tag{A.3}$$

- Simplify $P(-)$ The result is

$$c_{pp} = P(- \mid A)\, P(A)\, c_F + P(+)\, c_R + c_{\mathrm{T}}. \tag{A.4}$$

- Replace P(+) with P(+ | N) P(N) + P(+ | A) P(A) using

$$\begin{aligned} c_{pp} = {} & P(- \mid A)\, P(A)\, c_F + P(+ \mid N)\, P(N)\, c_R \\ & + P(+ \mid A)\, P(A)\, c_R + P(A)\; c_{\mathrm{T}} + P(N)\; c_{\mathrm{T}}. \end{aligned} \tag{A.5}$$

- Replace P_N with $(1 - P_A)$, . with false positive rate (FPR), . with $(1 - \mathrm{FNR})$, . with $(1 - \mathrm{FPR})$, . with false negative rate (FNR), and collect the terms with respect to $(1 - P_A)$ and P_A. The result is

$$\begin{aligned} c_{pp} = {} & (1 - P_A)\, FPR\, c_R + P_A\, (FNR\, c_F + (1 - FNR)\, c_R) \\ & + P(A)\; c_{\mathrm{T}} + (1 - P_A)\; c_{\mathrm{T}}. \end{aligned} \tag{A.6}$$

- Finally, collect the terms $c_{N,pp}$ and $c_{A,pp}$ (see their definition in Equation (7)). The result is the researched final form of Equation (6).

$$c_{pp} = (1 - P_A)\, c_{N,pp} + P_A\, c_{A,pp}\ . \tag{A.7}$$

Appendix B

Required accuracy domain derivation

In the present appendix, the required accuracy domain is derived, based on Equation (11). The required accuracy domain, D_{RA}, is the subspace of the test accuracy parameters (FPR and FNR) where VoI$>c_T$, or, equivalently, $G_T>0$. As detailed in the present paper, G_T is defined as

$$G_T = c_p - c_{pp} > 0, \tag{B.1}$$

where c_p and c_{pp} are the prior expected cost and the pre-posterior expected cost, respectively. Based on Equations (3) and (6), Equation (B.1) can be expanded as

$$G_T = \min[P_A\, c_F, c_R] - P_N\, c_{N,pp} - P_A\, c_{A,pp} > 0. \tag{B.2}$$

Next, based on Equation (7), Equation (B.2) can be further expanded as

$$\begin{aligned} G_T = \min[P_A\, c_F, c_R] - P_N\, FPR\, c_R - P_A\, FNR\, c_F \\ - P_A\,(1 - FNR)\, c_R - c_{\mathrm{T}} > 0. \end{aligned} \tag{B.3}$$

Note that Equation (B.3) includes a minimization operator. Therefore, Equation (B.3) can be divided into two different cases:

- *Case* $P_A < \overline{P}_A$. Considering only the first term of the minimization operator, Equation (B.3) becomes

$$\begin{aligned} G_T = P_A\, c_F - P_N\, FPR\, c_R - P_A\, FNR\, c_F \\ - P_A\,(1 - FNR)\, c_R - c_{\mathrm{T}} > 0. \end{aligned} \tag{B.4}$$

- Equation (B.4) can be rewritten as

$$G_T = P_A\,(1 - FNR)\,(c_F - c_R) - P_N\, FPR\, c_R - c_{\mathrm{T}} > 0. \tag{B.5}$$

- Next, the inequality in Equation (B.5) can be rewritten in terms of *FPR* and as

$$FPR < \frac{P_A}{1 - P_A}\,\frac{c_F - c_R}{c_R} - \frac{1}{P_N}\,\frac{c_{\mathrm{T}}}{c_R} - \frac{P_A}{P_N}\,\frac{c_F - c_R}{c_R}\,\mathrm{FNR}. \tag{B.6}$$

- *Case* $P_A > P_A$. Considering only the second term of the minimization, Equation (B.3) becomes

$$\begin{aligned} G_T = c_R - P_N\, FPR\, c_R - P_A\, FNR\, c_F - P_A\,(1 - FNR) \\ c_R - c_{\mathrm{T}} > 0. \end{aligned} \tag{B.7}$$

- Equation (B.7) can be rewritten as

$$G_T = P_N\, c_R - P_N\, FPR\, c_R - P_A\, FNR\,(c_F - c_R) - c_{\mathrm{T}} > 0. \tag{B.8}$$

- Next, similarly to the previous case, the inequality in Equation (B.8) can be rewritten in terms of *FPR* and as

$$FPR < 1 - \frac{1}{1 - P_A}\,\frac{c_T}{c_R} - \frac{P_A}{1 - P_A}\,\frac{c_F - c_R}{c_R}\,FNR. \tag{B.9}$$

Now, one can note that Equations (B.6) and (B.9) differ only for one term. Therefore, these equations can be collected in only one inequality as

$$\begin{aligned} FPR < min\left[1\,,\ \frac{P_A}{1 - P_A}\,\frac{c_F - c_R}{c_R}\right] \\ - \frac{1}{1 - P_A}\,\frac{c_{\mathrm{T}}}{c_R} - \frac{P_A \mathrm{p}}{1 - P_A}\,\frac{c_F - c_R}{c_R}\,FNR. \end{aligned} \tag{B.10}$$

Finally, by defining the relative costs γ_F and γ_T as

$$\delta_F = \frac{c_F - c_R}{c_R}, \quad \text{and} \quad \delta_T = \frac{c_{\mathrm{T}}}{c_R}, \tag{B.11}$$

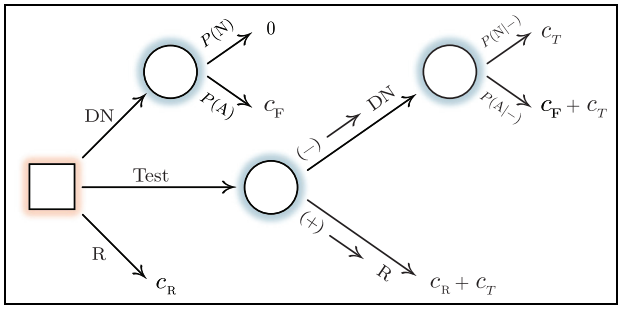
Equation (B.10) can be rewritten as reported in Equation (10), that is,

$$\begin{aligned} \mathrm{FPR} < \min\left[1, \frac{P_A}{1 - P_A}\,(\gamma_F - 1)\right] - \frac{\gamma_{\mathrm{T}}}{1 - P_A} \\ - \frac{P_A}{1 - P_A}\,(\gamma_F - 1)\,\mathrm{FNR}. \end{aligned} \tag{B.12}$$

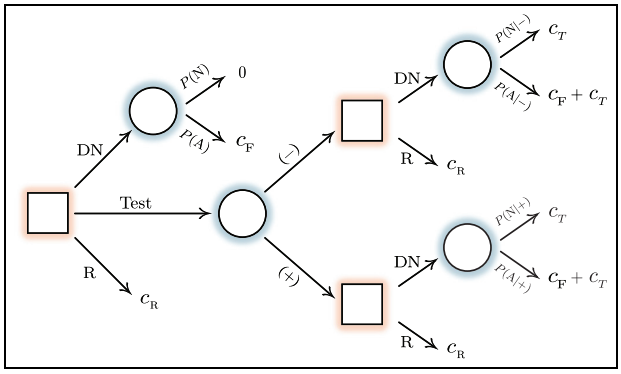
Appendix C

Decision-making process

In the present paper, the authors represent the decision-making process using two separate approaches: first, a prior process, which does not take into account test information, and, second, a *pre-posterior* process, which incorporates test information. By comparing the expected total costs of these two approaches (c_p and c_{pp}, respectively), the authors identify the test accuracy conditions under which testing is



Appendix Figure C.1. Decision tree in case (i).



Appendix Figure C.2. Decision tree in case (ii).

marginally cost-effective. The present appendix clarifies that:

(1) The chosen decision-making process aligns with the decision tree in Appendix Figure C.1
(2) Since test accuracy is known beforehand, the chosen process corresponds to the more general decision tree in Appendix Figure C.2, where actions are not deterministically dependent on test outcomes.

Case (i)

The tree shown in Appendix Figure C.1 formally corresponds to the procedure described in section "Problem statement and formulation." In fact, the test is considered cost-effective when the expected total cost with the test information is less than the expected total cost without the test information. Therefore, the test is convenient when its expected total cost is less than both the expected costs of the actions "do nothing" (DN) and "repair" the bridge (R).

As expected, the solution of this decision tree corresponds exactly to the one presented in section "Problem statement and formulation." In particular, given FPR and FNR, the test is convenient if $\overline{P}_{A,1} < P_A < \overline{P}_{A,2}$.

Case (ii)

The tree in the Appendix Figure C.2, corresponds to the general case where the action to be taken as a result of the test is not directly related to the test outcome. Observe that in this case, the decision tree is more complex and includes three decision nodes.

Under the assumption of a positive determinant of the test confusion matrix, the solution of this decision tree is the same as that obtained in section "Problem statement and formulation," that is, the test is convenient if $\overline{P}_{A,1} < P_A < \overline{P}_{A,2}$.

The expected cost of the action "do the test" (Test) is equal to

$$c_{Test} = \min \begin{cases} c_{Test,a} = P_A c_F + c_T \\ c_{Test,b} = c_R + c_T \\ c_{Test,c} = (1 - P_A)\, c_{N,pp} + P_A\, c_{A,pp} \\ c_{Test,d} = (1 - P_A)\, \hat{c}_{N,pp} + P_A\, \hat{c}_{A,pp} \end{cases}. \quad \text{(C.1)}$$

where

$$c_{N,pp} = FPR\, c_R + c_T; \quad c_{A,pp} = FNR\, c_F + (1 - FNR)\, c_R + c_T, \quad \text{(C.2)}$$

$$\tilde{c}_{N,pp} = (1 - FPR)\, c_R + c_T; \quad \tilde{c}_{A,pp} = (1 - FNR)$$
$$c_F + FNR\, c_R + c_T . c_{N,pp} = FPR\, c_R + c_T; \quad \text{(C.3)}$$
$$c_{A,pp} = FNR\, c_F + (1 - FNR)\, c_R + c_T.$$

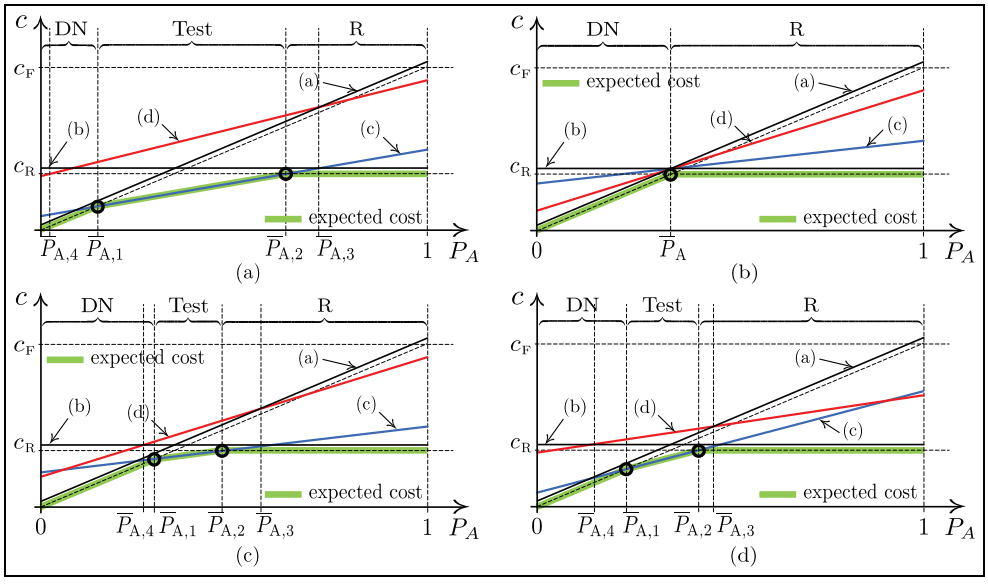
However, under the assumption of a positive determinant of the test confusion matrix, the solution of this decision tree is the same as that obtained in section "Problem statement and formulation," that is, the test is convenient if $\overline{P}_{A,1} < P_A < \overline{P}_{A,2}$.

Note that the marginal cost $c_{Test,c}$ is equal to c_{pp} (see sections "Problem statement and formulation" and "Concluding remarks"). On the other hand, the marginal cost $c_{Test,d}$ has the same structure as c_{pp} and can be obtained by replacing FPR with TNR (i.e., $1 - FPR$) and FNR with TPR (i.e., $1 - FNR$). This corresponds to the permutation of the rows of the confusion matrix of the test (i.e., the permutation of the test outcomes).

The marginal cost $c_{Test,d}$ is lower than $c_{Test,a}$ if $P_A > P_{A,3}$, where

$$P_{A,3} = \frac{1 - FPR}{\gamma_F\, FNR + (1 - FPR - FNR)} .. \quad \text{(C.4)}$$

Furthermore, the marginal cost $c_{Test,d}$ is lower than $c_{Test,b}$ if $P_A > P_{A,4}$, where



Appendix Figure C.3. Configurations of $c_{Test,a}$ (a), $c_{Test,b}$ (b), $c_{Test,c}$ (c), and $c_{Test,d}$ (d) with respect to P_A.

$$P_{A,4} = \frac{\text{FPR}}{\gamma_F\,(1-FNR)-(1-FPR-FNR)}.. \qquad \text{(C.5)}$$

Since (i) $c_R < c_F$ and (ii) the test confusion matrix has a positive determinant, the marginal cost $c_{Test,d}$ is simultaneously lower than both $c_{Test,a}$ and $c_{Test,b}$ if $P_{A,3} < P_{A,4}$. This condition occurs if and only if the determinant of the test confusion matrix is negative (FPR + FNR > 1). Since this contradicts the assumptions above, it mathematically demonstrates that the marginal cost $c_{Test,d}$ is never simultaneously lower than the other costs ($c_{Test,a}$, $c_{Test,b}$, and). Moreover, even if one was to admit test confusion matrices with negative determinants, the marginal cost $c_{Test,c}$ would never be simultaneously lower than the others. Therefore, for any combination of FPR and FNR, one between $c_{Test,c}$ and $c_{Test,d}$ is always excluded in $0 \leqslant P_A \leqslant 1$. The above can be seen graphically in Appendix Figure C.3. Specifically, note that $c_{Test,d}$ (the red curve labeled with the letter "d") is never the lowest cost.