

Performance Assessment Model of Non-destructive Technologies in Inspecting Concrete Bridge Decks

Sherif Abdelkhalek*, Tarek Zayed

Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Abstract

Non-destructive testing (NDT) methods are typically utilized to thoroughly inspect existing defects in the concrete bridge deck and tackle the limitations of common inspection practices (e.g., visual inspection). Nevertheless, the reliability of inspection outcomes crucially depends on choosing the most appropriate NDT technologies. In this regard, a comprehensive Performance Assessment Model (PAM) was developed. The developed model incorporated 40 parameters to precisely assess the performance of different NDT technologies from diverse perspectives (e.g., defect detection capability, ease of use, speed, and cost). The required data were collected through a survey questionnaire. The model utilized the Analytic Network Process (ANP) technique to calculate the importance weight of each parameter, whereas the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) was exploited to compute the performance index (PI) of NDT technologies. The outcomes of this study clearly illustrate the disparity in the performance of different NDT technologies. Furthermore, it was shown that none of these technologies could either exhibit the best performance in all the proposed parameters or efficiently identify all types of defects. Based on the PAM results, a selection model was proposed to assist bridge authorities and consultants in choosing the most efficient NDT technologies for inspection purposes.

Keywords: Concrete bridge deck; Inspection; Non-destructive technologies; Analytic Network Process (ANP); TOPSIS; Performance assessment.

* Corresponding author at the Department of Building and Real Estate (BRE), Block Z, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. ORCID: <https://orcid.org/0000-0002-4408-5814>. Email: sherif.abdelkhalek@connect.polyu.hk (S. Abdelkhalek).

1. Introduction

Bridge deck provides the primary function of bridges (i.e., driving surface). Maintaining this valuable part to meet the standard service level costs between 50% and 80% of the overall bridge's expenses (Gucunski *et al.* 2013). To effectively reduce its operational costs, it is crucial to conduct all necessary maintenance works on time (i.e., performing timely maintenance is a critical step to mitigating its deterioration) and avoid unnecessary maintenance activities. Accordingly, regular inspections are typically scheduled during the service life of a bridge to identify minor and major deficiencies in this valuable part (Popescu *et al.* 2019). These deficiencies are exemplified by different types of defects, such as cracks, scaling, spalling, delamination, and corrosion. The inspection process aims to identify the types and locations of defects and quantify their extent and severity. Therefore, a periodical and timely diagnosis of these defects can help decision-makers to identify the most appropriate intervention on time.

In this regard, visual inspection is often conducted by experienced inspectors within a 24-month interval or less, depending on the current condition of the bridge. Inspectors usually use simple tools, such as measuring tape, marker, chalk, and flashlight (Huston *et al.* 2011), to identify surface flaws' locations and quantify their extent and severity. Additionally, they may incorporate simple non-destructive techniques in the inspection, such as hammer sounding and chain drag, to detect delaminated areas (Tinkey and Olson 2010, Abdelkhalek and Zayed 2019).

However, with the suspicion of severe defects, more advanced NDT technologies may be recommended. These include ground-penetrating radar (GPR), impact echo (IE), ultrasonic surface wave (USW), ultrasonic pulse-echo (UPE), half-cell potential (HCP), electrical resistivity (ER), polarization resistance (PR), and infrared thermography (IRT). Furthermore, images captured by a camera are considered an efficient alternative to the visual inspection

technique. Recently, the image-based technique has been implemented in inspecting various types of structures, such as bridges and buildings (Duque *et al.* 2018). Unmanned aerial vehicle (UAV) equipped with a camera increases the benefits of this technique by increasing inspection speed, reducing safety risks, and accessing hard-reached elements (Seo *et al.* 2018a, Seo *et al.* 2018b).

Incorporating NDT technologies in the inspection process assuages the drawbacks of common inspection practices (e.g., visual inspection), which include subjectivity, uncertainty, and limited capability in detecting subsurface defects. Therefore, NDT techniques provide a more accurate condition assessment and higher confidence in the inspection outcomes (Alsharqawi *et al.* 2020). However, to ensure the reliability and efficiency of the inspection process, it is crucial to choose the most appropriate techniques that can effectively identify the existing defects and their severity. To achieve this, decision-makers should have adequate knowledge about the performance of these techniques with respect to some critical inspection criteria, such as defect detection capability, accuracy, and speed.

Accordingly, the present research aims to: 1) review previous studies to identify the key parameters affecting the performance of NDT technologies; 2) develop a comprehensive Performance Assessment Model (PAM) that can precisely assess the performance of different NDT technologies; 3) implement the developed model to prioritize NDT technologies according to the assessment parameters proposed in the model; 4) develop a tool that can support decision-maker in selecting the most efficient technologies for inspection. The outcomes of this study provide a useful resource that can assist bridge authorities and consultants in selecting the most efficient technologies for inspection purposes. Consequently, the inspection process will afford a more accurate assessment of the condition of the bridge deck, which is a paramount factor in determining a cost-effective maintenance strategy.

Background and Research Gaps

Several studies have investigated the performance of NDT technologies from different perspectives (Scott *et al.* 2003, Yehia *et al.* 2007, Vaghefi *et al.* 2012, Gucunski *et al.* 2013, Oh *et al.* 2013, Pailes and Gucunski 2015, Hesse *et al.* 2017, Omar *et al.* 2017, Vemuri and Atadero 2017, Lin *et al.* 2018, Sultan and Washer 2018, Azari and Lin 2019, Janků *et al.* 2019). Accordingly, various criteria were proposed to fulfil this goal, such as the capability of an NDT technology to detect different types of defects, accuracy, inspection speed, inspection cost, and simplicity of data collection and analysis. Two approaches were adopted to measure these criteria: qualitative and quantitative. In the qualitative approach, the data were collected from the literature and NDT experts (Vaghefi *et al.* 2012, Hesse *et al.* 2017, Omar *et al.* 2017). On the other hand, in the quantitative approach, the data were obtained from laboratory and field tests. In the laboratory test, specimens with prefabricated defects were used to assess the performance of different NDT techniques (Yehia *et al.* 2007, Azari *et al.* 2014, Lin *et al.* 2018). As for the field test, real bridges were inspected using these technologies and the results were verified using a core testing method (Scott *et al.* 2003, Rens Kevin *et al.* 2005, Gucunski *et al.* 2013, Oh *et al.* 2013, Agdas *et al.* 2016, Sultan and Washer 2018).

For example, Yehia *et al.* (2007) investigated the capability and accuracy of IE, GPR, and IRT to detect cracks, delamination, and voids. In their study, defects of different sizes were built at various depths in laboratory specimens to test the defect detection capability and accuracy of the aforesaid devices. Oh *et al.* (2013) evaluated the performance of IRT, chain drag, and two types of air-coupled IE on a real bridge and verified the findings using a core testing method. The researchers proposed eleven criteria to measure the performance of the four NDT technologies. These criteria included capital cost, operational cost, operation time, analysis time, sensitivity to noise and environmental conditions, and operator skill. Others were

portability in terms of device size and weight, potential use without lane closure, need for surface preparation, delamination detection accuracy, and repeatability.

Another research group utilized a two-stage laboratory test to assess the capability of nine NDT methods (e.g. USW, IE, GPR, ER, HCP, and IRT) to detect six types of prefabricated defects, i.e., delamination, honeycombing, vertical crack, void, corrosion, and overlay debonding (Lin *et al.* 2018). In the first stage, the specimens were tested without any type of overlay layers. However, in the second stage, the specimens were covered with seven types of overlay layers (i.e., epoxy, latex-modified concrete, silica fume–modified concrete, polyester polymer, asphalt with a liquid membrane, asphalt with a sheet membrane, and asphalt without a membrane). The overlay layers in all the specimens were constructed so that half of the overlay layer was bonded, while the other half was unbonded. Each defect type was created in both bonded and unbonded halves so as to assess the defect detection capability of the selected technologies.

Abdelkhalek and Zayed (2020b) developed a hybrid simulation model to estimate the duration and cost of the inspection process using individual or multi-NDT technology. Agent-based and discrete event simulation approaches were integrated to build this model. Six cost items were considered in the model. These items included technician salary, transportation cost, traffic control cost, bridge user delay cost, and other sources.

Some studies also proposed two or more sub-criteria to accurately measure the performance of NDT technologies under the main criteria. For instance, Gucunski *et al.* (2013) evaluated the performance of nine technologies (IE, USW, GPR, HCP, PR, ER, IRT, hammer sounding, and chain drag). Six main criteria and twenty-five sub-criteria were considered in evaluating these technologies. The main criteria included defect detection capabilities, accuracy, precision, ease of use, speed, and cost. Defect boundaries detectability, defect depth detectability, time needed for data collection, analysis and interpretation, and potential for

automation were some examples for the proposed sub-criteria. The study utilized both field and laboratory tests to assess the performance of the technologies. Findings of the field test were verified using a core testing method. On the other hand, two kinds of specimens were used in the laboratory test, i.e., a concrete slab with prefabricated defects and a part of a bridge deck. A similar study by Omar *et al.* (2017) qualitatively assessed the performance of five NDT technologies (IE, UPE, GPR, IRT, and HCP). Five criteria and fifteen sub-criteria were used in the assessment. Data for the study were obtained from a survey questionnaire distributed among experts on bridge inspection and NDT. The Fuzzy Analytic Hierarchy Process (FAHP) was used to determine the weights of the proposed criteria and sub-criteria.

Abdelkhalek and Zayed (2020a) reviewed the literature on the criteria and sub-criteria that were proposed to assess the performance of NDT technologies. The authors identified six main criteria and twenty-eight sub-criteria. Their findings showed that there is a need to extend the proposed performance criteria to cover other critical aspects that have a significant impact on the selection of NDT technologies. For instance, the study proposed the inclusion of the ability to compare previous test results with the current ones and the availability of a standard scale to identify the severity of defects. Accordingly, there is a critical need to comprehensively identify and incorporate these parameters into an encompassing framework that can thoroughly assess the performance of different NDT technologies. Furthermore, the performance of the regular camera (i.e., hereinafter referred to as camera) and dependency between performance factors (e.g., the relationship between data collection speed and cost) have not been studied in the literature. By considering the aforesaid aspects, the reliability of assessing NDT technologies will be greatly improved.

2. Research Methodology and Model Design

Performance assessment of NDT technologies is considered a Multi-Criteria Decision-Making (MCDM) problem. Accordingly, this study adopted two MCDM techniques to develop

the proposed model. The techniques included the Analytic Network Process (ANP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). The ANP and TOPSIS techniques have been widely used in several applications in civil engineering (El-Abbasy *et al.* 2015, Yadollahi *et al.* 2015, Salehi *et al.* 2018, El-Abbasy *et al.* 2019, Rong *et al.* 2020). In the current study, the ANP technique was employed to calculate the importance weights of the performance parameters.

The Analytic Hierarchy Process (AHP) and the ANP techniques were developed by Professor Thomas L. Saaty. The AHP is a powerful technique that can handle complex and multi-criteria decision-making problems. It breaks down a complex problem into several hierarchical levels. The first level of the hierarchy includes the decision goal, while its decision criteria are represented in the subsequent levels. Furthermore, by assuming that the decision elements (e.g., criteria and parameters) are independent of each other, the AHP applies pairwise comparisons between the decision criteria to determine the relative importance of each.

On the contrary, the ANP technique adopts a more generalized concept by eliminating the need to assume that (i) higher-level elements are independent of lower-level elements and (ii) elements in the same level are independent of each other (Saaty 2004b). With this background, obvious interdependency among different performance parameters makes the ANP suitable for our model.

On the other hand, TOPSIS has been used to determine the Performance Index (PI) of different NDT technologies. TOPSIS was developed by Hwang and Yoon (1981). The main concept in TOPSIS is to rank different alternatives based on their distances from the positive ideal solution (hypothetical best alternative) and the negative ideal solution (hypothetical worst alternative). The best alternative is the one that has the shortest distance to the positive ideal solution and the farthest distance to the negative ideal solution. Detailed procedure to calculate

the performance index (PI) according to this concept are discussed in the data analysis section. The required data for the model were collected using a survey questionnaire. Figure 1 presents the main stages and detailed steps adopted to build this model. Further details about these stages and steps to achieve each are elucidated in the next subsections.

[Figure 1 near here]

2.1. Parameters in the Literature

Parameters proposed to assess the performance of NDT technologies were collected from the literature. Five main criteria were used to categorize these parameters, i.e., capability, performance under different environments, ease of use, cost, and speed (Table 1). Further thirteen sub-criteria were proposed under the main criteria as shown in Table 1. This categorization represented the backbone structure for the model developed in the current study. Further details about this categorization are provided in the next section.

[Table 1 near here]

2.2. Performance Assessment Model

The performance assessment model (PAM) was developed following three steps: introducing new parameters that were needed to improve the reliability of the performance assessment model, building a PAM network, and selecting suitable software for the analysis. The details of each step are provided next.

2.2.1. New Parameters

Six groups of new parameters (to make a total of 13 parameters) were introduced besides those found in the literature (Table 1). The first two groups included the availability of a standard scale to accurately identify the severity of a defect (i.e., three parameters distributed between “delamination”, “concrete degradation”, and “corrosion” sub-criteria) and confidence in defect detection results (i.e., three parameters distributed between “delamination”, “concrete degradation”, and “corrosion” sub-criteria). The next two were the efficiency of NDT

technologies under different finishes (four parameters) and the ability to mount an NDT technology on a cart/car/drone (i.e., one parameter in “simplicity of data collection”). The two remaining groups comprised the ability to store and compare the current data with old data (one parameter in “simplicity of data analysis”) and traffic delay cost (one parameter in “data collection cost”).

In the first group, the parameters were proposed to examine the availability of a standard scale and its quality to provide useful guidance in identifying the severity of the existing defects. For example, the corrosion threshold levels reported in the ASTM standard for Half-cell Potential test method (ASTM 2013). In the second group, the “confidence in defect detection results” parameter was applied to investigate the degree of uncertainty in the test findings as subsurface defects are not visible to inspectors. Therefore, detecting these types of defects carries a degree of uncertainty, especially in the early stage. The first and second groups are essential in accurately identifying the current condition of a bridge deck.

The “Efficiency of NDT technologies under different finishes” parameter examined the capability of NDT technologies to provide stable performance under different finishes of concrete bridge deck (e.g., asphalt overlay and coated rebar). This is an important factor for using NDT technologies in inspection applications. On the other hand, the “ability to mount an NDT technology on a cart/car/drone” parameter was proposed to investigate to which extent using these pieces of equipment (e.g., cart and drone) facilitates the operation of the device to ease the task of data collection. Similarly, the “ability to store and compare the current data with old data” parameter assessed the simplicity in analysing and interpreting the data due to this privilege (e.g., identifying the progress rate of deterioration). Finally, the “traffic delay cost” parameter was proposed to quantify the impact of inspection activities on traffic flow (i.e., cost of lane closure and bridge users’ delay cost).

2.2.2. *PAM Network*

The new parameters introduced in the previous sub-section and parameters found in the literature (Table 1) were put together in one model. A total of forty parameters were used in this study to measure the strengths and weaknesses of using different NDT technologies. The developed model is shown in Figure 2. The PAM network was divided into three levels: criteria, sub-criteria, and parameters. Starting from the first criterion, “capability” assessed the ability of an NDT technology to detect different types of defects, which included cracks, scaling, spalling, delamination, concrete degradation, and corrosion. Concrete degradation is a term used to describe the reduction in concrete characteristics, such as strength and modulus.

Two aspects should be considered in identifying the capability of NDT technologies: accuracy and uncertainty. Defect detection accuracy reflects the closeness of the inspection findings to the real situation. It is a very important indicator in using a specific device in the inspection. This indicator was considered in the capability criterion by using several parameters, such as “ability to detect defects with different sizes”, “ability to detect defects at varying depths”, and “availability of a standard scale to accurately identify defect severity”. On the other hand, detecting subsurface defects using different technologies contains some sort of uncertainty since these types of defects are hidden beneath the bridge deck surface. This aspect was considered by using the “confidence in defect detection results” parameter.

“Performance under different environment” criterion included three sub-criteria: efficiency under different finishes, consistency, and sensitivity. “Efficiency under different finishes” examined the performance of NDT technologies in the case of different types of overlay layers and coated rebars. On the other hand, “consistency” was used to describe the degree of changes in test results when repeating the test. This criterion has a strong relationship with the sensitivity of NDT technologies to environmental conditions and traffic noise since a highly sensitive device provides low consistency in test results when the test is repeated.

The “Ease of use” criterion was applied to measure the degree of difficulty in conducting an inspection using different technologies. It was divided into two sub-criteria: simplicity of data collection and simplicity of data analysis. Each sub-criterion involved four parameters (Figure 2). Similarly, “speed” and “cost” criteria were divided into two parts: one for data collection and the other for data analysis. As shown in Figure 2, data collection speed depended on the simplicity of data collection since the light-weight device, small-size device, or devices mounted on a cart/car/drone provides a fast inspection process. Similarly, data analysis speed, data collection cost, and data analysis cost depended on the simplicity of data analysis, data collection speed, and data analysis speed, respectively. The aforementioned relationships between the parameters made using the ANP approach instead of the AHP approach mandatory.

[Figure 2 near here]

2.2.3. Analysis Software

Super Decisions, Python, and Microsoft Excel were used for data analysis (Figure 3). First, the PAM model was built in Super Decisions to identify all required data to design the questionnaire. Super Decisions is a free educational decision support software that implements the AHP and the ANP approaches. Further details about the software and the two approaches can be found on Super Decisions website (<http://www.superdecisions.com>). Second, all the collected data were stored in an Excel file (as an input file). Then, the supermatrix for each participant was generated and stored in a separate sheet in the same Excel file. Afterwards, a Python code was written using the PYANP library to perform the ANP calculations for each participant and capture the importance weights of the performance criteria, sub-criteria, and parameters. Supermatrix and limit matrix were the input and the output data of the Python code, respectively. The output of the Python code was verified using Super Decisions results for two participants. The limit matrices for all participants were stored in another Excel file (i.e., output

file). Finally, Microsoft Excel was used to perform TOPSIS calculations (PI), statistical analysis (the mean for the importance weight of each parameter), and output visualization (overall performance scores of NDT technologies).

[Figure 3 near here]

2.3.Questionnaire Design

This stage followed two main steps: identifying the required data and developing the questionnaire format. This section expatiates on the details of these steps.

2.3.1. Required Data

As mentioned before, the relative importance weights of the performance criteria and parameters were determined using the ANP technique. On the other hand, the performance of different technologies regarding these criteria and parameters were evaluated using the TOPSIS technique. Accordingly, it was required from each participant to provide two types of information. The first one was his/her judgment on the relative importance of the performance criteria, sub-criteria, and parameters (i.e., using pairwise comparison matrices). This information was analysed using the ANP approach and the outputs were the relative importance weights of the performance criteria, sub-criteria, and parameters. The second type of information was applied to assess the performance of different technologies regarding the proposed parameters. These data, in addition to the weights calculated from the ANP, were used to calculate the performance index for each technology using the TOPSIS approach.

To accurately identify all the information needed for the analysis stage, the PAM network was built in Super Decisions. Then, all input data needed for the pairwise comparison matrices of the performance criteria, sub-criteria, and parameters were checked and listed. On the other hand, the required performance scores for each technology considering all proposed parameters were identified. Furthermore, the personal information of each participant was also considered, such as name, company, and years of experience.

2.3.2. Questionnaire Format

In this step, questionnaire questions were formulated and ordered accordingly. To achieve this, questionnaires in previous studies (Moufti 2013, Iason 2014, Mahmoud 2017, Omar *et al.* 2017) were revised to identify the most appropriate design (e.g., structure, design of tables, and clarity of instructions for each part). Accordingly, the questionnaire in the current study was divided into five sections. The first section was devoted to collecting personal information from the participants (e.g., name, company, and years of experience). The second section covered the details of the performance assessment model (i.e., the PAM network shown in Figure 2). The third and fourth sections were dedicated to collecting the required data (i.e., degree of importance) for pairwise comparison matrices of the criteria, sub-criteria, and parameters for assessing NDT technologies. In the last section, the performance scores of NDT technologies were measured according to the proposed parameters.

Figure 4 shows the first part of section 3 in the questionnaire, which was dedicated to collecting the data for pairwise comparison matrices of criteria and sub-criteria. This section started with some instructions on how to complete the section. Then, the pairwise comparison matrices of criteria and sub-criteria were provided. Here, each criterion was compared with the other criteria located in the same group to identify the degree of importance for all criteria. A scale of 1 to 9, which was proposed by (Saaty 2004a), was adopted to measure the degree of importance of each criterion. on this scale, 1 implies that criteria A and B have equal importance, while 9 indicates that criterion A is extremely important compared to B. Further details about this scale are shown in Table 2 (Saaty 2004a).

[Figure 4 near here]

[Table 2 near here]

Figure 5 shows a snapshot from section 5 of the questionnaire. In this part, participants were asked to evaluate NDT technologies according to the proposed parameters in the model. Nine technologies were selected in this study for the assessment (Figure 2). A scale of 0 to 10 was used to evaluate the performance of NDT technologies. On this scale, 10 denotes that an NDT technology offers an extremely high performance regarding this parameter, while 1 indicates a very low performance. On the other hand, 0 implies that an NDT technology cannot provide any information regarding this parameter (e.g., the camera cannot detect delamination).

[Figure 5 near here]

2.4.Data Collection

The required data for the model were collected using a survey questionnaire. The data collection stage involved two steps: conducting a pilot study and conducting a full-scale study. The next sub-sections elaborate on the details of these steps.

2.4.1. Pilot Study

The questionnaire format should be tested before conducting a full-scale study to ensure that the instructions needed to fill the questionnaire are adequate and the questions can easily be understood by participants. Furthermore, testing ensures that the wording of the questions is clear, the questions are placed in the best order, and there are no mistakes in formulating the questionnaire. In addition, suggestions about adding some major parameters that have not been considered can be collected at this stage. This allows making the needed corrections before conducting the full-scale study. Therefore, a pilot study is typically conducted on a small number of participants and the questionnaire is edited based on their feedback.

In this regard, five participants were chosen to fulfil these purposes. The participants for this step were chosen to represent all the participants' categories (e.g., government agencies, consultancy agencies, research centres and universities). Their feedback did not indicate any mistakes, disorder, or obscurity in the questionnaire questions. Furthermore, the participants

noted that the different aspects of implementing NDT technologies in the inspection were well covered. The only negative comment received from the participants was that: “questionnaire is too long and too much information is required”. Indeed, the main reason for this comment was that all the key factors (i.e., 5 criteria, 13 sub-criteria, and 40 parameters) that influence the implementation of NDT technologies in the inspection process were considered. This was done to realize a high accuracy of the assessment, which necessitated a long questionnaire to encompass the required information. Based on the findings of this stage, the development of the questionnaire moved to the next stage (i.e., full-scale study) without any changes. The data collected from this stage were analysed with the data collected from the next stage.

2.4.2. Full-scale Study

After ensuring the integrity of the questionnaire format, the final design of the questionnaire was distributed to the experts in this field. They included those who work in government agencies, consultancy agencies, or research centres and universities. As a sampling frame for this study could not be identified, the sample was taken to be a nonprobability sample (Zhao *et al.* 2015). The nonprobability sample method does not depend on a random selection of participants. Instead, participants are chosen based on their willingness to participate in the research study (Wilkins 2011). In this regard, 65 experts were invited to participate in this study. Only 24 participants completed the questionnaire, which resulted in an approximate response rate of 37%.

According to Baruch (1999), the response rate is expected to be in the range (55.6±19.7)%. Clearly, the response rate of our study fell within the specified benchmark. Nevertheless, some researchers also opined that statistical analysis could be conducted if the sample size is greater than 30 (Zhao *et al.* 2016). Furthermore, other studies stated that a survey based on the AHP approach does not need a large sample size for data analysis (Baby 2013, Kil *et al.* 2016, Tarei Pradeep *et al.* 2018, Waris *et al.* 2019). Therefore, several studies in the

literature conducted AHP surveys with a number of participants less than 30 (Sambasivan and Fei 2008, Tarei Pradeep *et al.* 2018, Waris *et al.* 2019, Demirkesen and Bayhan 2020, Yeo *et al.* 2020).

Several methods were adopted to ensure the reliability of the data collected and the validity of our findings. First, the inconsistency ratio for all pairwise comparison matrices used in the ANP calculation was checked. The calculated values fell between 0.024 and 0.0784, which satisfied the proposed limit (i.e. smaller than 0.1 (Saaty 2004a)). Second, following the reasoning proposed by Saaty (1980), Chebyshev's theorem statistical test was applied to all collected data (i.e., the data required to calculate the importance weights and technologies rating scores). According to the theorem, at least 75% of the data must lie within the range of ± 2 standard deviations for acceptability. Our results satisfied the limit of this test. Between 88% and 100% of our dataset (an average of 95.3%) fell within ± 2 standard deviations ranges. Besides, this method has been used by several studies to check the reliability of the collected data from a small sample size (Salman *et al.* 2007, Doloi 2008, Akroush and El-adaway 2017, David and Varghese 2021). Finally, the findings from the data analysis were compared with the findings of two studies in the literature. This is discussed in detail in the model testing section.

Figures (6-8) summarize the participants' information that was provided in the first section of the questionnaire considering the following aspects: year of experience, participant's affiliation, participants' countries, participants' experience regarding different devices, and frequency of using these devices. As shown in Figure 6, 50 % of the participants had more than 10 years of experience and about 34 % of them had above 20 years of experience. According to the findings in Figure 7 (a), 63 % of the participants were working in the industry (42 % in government agencies and 21 % are consultants), while 37 % were working in academic organisations, such as universities. Figure 7 (b) shows participants' countries. Most participants

were working in the United States and Canada (i.e., 16 participants were equally divided between these two countries).

Figure 8 shows the experience of participants regarding different devices and frequency of use. From the figure, GPR was the most used technology, followed by the camera, then IRT. This trend matched the interest of the research community in using these technologies (Abdelkhalek and Zayed 2020a). On the other hand, PR was the least used NDT technology. The trend in participants' experience regarding different devices was quite similar to the frequency of using these devices. About 80 % of participants had experience in multiple technologies (i.e., more than three technologies), with an average of 5.26 technologies per participant. In addition, around 65 % of the participants used at least three technologies, with an average of 3.65 technologies per participant. Some participants utilized other technologies that were not incorporated in this study, such as hammer sounding, chain drag, covermeter test, and carbonation test.

[Figure 6 near here]

[Figure 7 near here]

[Figure 8 near here]

2.5.Data Analysis

Data analysis involved two main steps: calculating the average importance weights for the criteria, sub-criteria, and parameters proposed to assess the performance of NDT technologies and calculating the overall PI for these technologies. These steps are explained in the next sub-sections.

2.5.1. Importance Weights using the ANP

Two steps were adopted to get the average importance weights of the performance criteria, sub-criteria, and parameters (Forman and Peniwati 1998, El-Abbasy *et al.* 2015). The first step preserved the participants' identities to get the relative importance weights of the

criteria, sub-criteria, and parameters. The second step involved aggregating the participants' judgment on the prospective value of the relative importance weights of the criteria, sub-criteria, and parameters (Forman and Peniwati 1998). In the first step, the ANP calculations were performed for each participant to get the limit matrix (i.e., limit matrix carried the importance weights of the criteria, sub-criteria, and parameters). In the ANP calculations, three matrices were computed as in the following sequence: unweighted supermatrix, weighted supermatrix, and finally, limit matrix. The unweighted supermatrix and weighted supermatrix were calculated using Microsoft Excel, while the limit matrix was calculated using the PYANP library in Python. The size of these matrices in this study was (59×59), which was too large to display here.

Second, the average importance weights were calculated by aggregating the importance weights obtained from all participants using an arithmetic mean equation as follows:

$$W_f = \frac{\sum_{g=1}^q W_{fg}}{q} \quad [1]$$

where W_f = arithmetic mean of the importance weight of factor f (factor is used here to represent criterion, sub-criterion, or parameter), W_{fg} = importance weight for factor f given by participants g , q = number of participants.

2.5.2. Performance Index using TOPSIS

Two main steps were needed to get the overall performance index (PI). First, the average performance scores for each technology regarding assessment parameters were calculated using the arithmetic mean as follows:

$$R_{tf} = \frac{\sum_{g=1}^q R_{tfg}}{q} \quad [2]$$

where R_{tf} = arithmetic mean of the performance score of technology t according to parameter f , R_{tfg} = performance score for technology t according to parameter f given by participants g .

Second, the PI for a specific technology was computed using the TOPSIS technique. The average performance scores obtained from Equation (2), and the parameters' weight obtained from Equation (1) were used to perform TOPSIS calculations. The steps to calculate PI are as follows:

Step 1: Normalize the performance scores for each parameter

$$N_{tf} = R_{tf} / \sqrt{\sum_{t=1}^d R_{tf}^2} \quad [3]$$

where d = number of technologies (i.e., 9 technologies in the current study).

Step 2: Calculate the weighted normalized performance scores for each parameter

$$V_{tf} = W_f * N_{tf} \quad [4]$$

Step 3: Identify the positive ideal (I_f^+) and negative ideal (I_f^-) solutions for each parameter

$$I_f^+ = \max_t (V_{tf}) \quad \text{if } f \in \text{benefit criteria} \quad [5]$$

$$I_f^+ = \min_t (V_{tf}) \quad \text{if } f \in \text{cost criteria} \quad [6]$$

$$I_f^- = \min_t (V_{tf}) \quad \text{if } f \in \text{benefit criteria} \quad [7]$$

$$I_f^- = \max_t (V_{tf}) \quad \text{if } f \in \text{cost criteria} \quad [8]$$

Step 4: Calculate separation measures (S_t^+ , S_t^-) for each alternative (i.e., technology)

$$S_t^+ = \sqrt{\sum_{f=1}^r (V_{tf} - I_f^+)^2} \quad [9]$$

$$S_t^- = \sqrt{\sum_{f=1}^r (V_{tf} - I_f^-)^2} \quad [10]$$

where r = number of parameters in the developed model (i.e., 40 parameters).

Separation measures (S_t^+ , S_t^-) are the average distances to positive ideal and negative ideal solutions, respectively.

Step 5: Calculate Performance Index (PI) for each alternative (i.e., technology)

$$PI_t = S_t^- / (S_t^+ + S_t^-) \quad [11]$$

3. Findings and Discussion

3.1. Importance Weights of Criteria, Sub-criteria, and Parameters

Figures (9-11) show the importance weights of the criteria, sub-criteria, and parameters for assessing the NDT technologies. As shown in Figure 9, the highest importance weight was assigned to the “defect detection capability” criterion (0.3636), followed by “inspection cost” (0.2171). There is only a slight difference between the importance weights of “inspection speed” (0.1537) and “performance under different environment” (0.1534). Lastly, the least importance weight was assigned to the “ease of use” criterion (0.1121).

[Figure 9 near here]

Figure 10 shows that “data collection speed” attained the highest importance weight among all sub-criteria (0.132), followed by “simplicity of data collection” (0.1225), and then “delamination” (0.1046). On the other hand, the lowest importance weight was assigned to the “ability to detect surface defects” sub-criterion (0.0439). Despite “defect detection capability” criterion obtained the highest importance weight at the criteria level, its sub-criteria were ranked considerably lower. In fact, they were not ranked in either the first or the second place at the sub-criteria level (e.g., “delamination” and “corrosion” sub-criteria came in the third and fifth places, respectively).

Two reasons made the weights of “data collection speed” and “simplicity of data collection” sub-criteria surpass the weights of the sub-criteria in the “defect detection capability” criterion. First, the weight of the “defect detection capability” criterion was divided between four sub-criteria, while the weights of “ease of use” and “speed” criteria were divided between two sub-criteria. Second, the dependence between the sub-criteria in “ease of use”,

“speed”, and “cost” increased the importance weights of “data collection speed” and “simplicity of data collection” sub-criteria. In general, delamination (i.e., “defect detection capability” criterion), sensitivity (i.e., “performance under different environment” criterion), simplicity of data collection (i.e., “ease of use” criterion), data collection speed (i.e., “speed” criterion), and data collection cost (i.e., “cost” criterion) got the highest importance weights compared to other sub-criteria located in the same main criterion.

[Figure 10 near here]

Figure 11 shows the importance weights of the proposed parameters. The importance weights of the parameters were categorized into three groups: high (red, ≥ 0.04), medium (orange, ≥ 0.02 and < 0.04), and low (blue, < 0.02). “Confidence in delamination detection results” parameter obtained the highest importance weight among all parameters (i.e., 0.0601). Three parameters were proposed to reflect the confidence in the test results: one for delamination and the others for concrete degradation and corrosion. The total weight for these three parameters was 0.1254 (12.54% of the total weight), which proved their impact on the overall performance of NDT technologies.

The “potential for automation in data collection” parameter obtained the second highest importance weight (0.0552), followed by “level of experience of the operator” (0.0524), then “time needed for testing” (0.0518). On the contrary, the “performance of using NDT technology in the case of coated rebar” parameter got the least importance weight (i.e., 0.0078). Parameters with the highest importance weights in each sub-criterion were as follows: ability to detect cracks under “surface defects”, confidence in defect detection results under “delamination, concrete degradation, and corrosion” sub-criteria, performance of asphalt overlay under “efficiency under different finishes”, sensitivity to environmental conditions under “sensitivity”, the potential for automation under “simplicity of data collection and analysis”, the time needed for data collection and analysis under “data collection and analysis

speeds”, traffic delay cost under “data collection cost”, and software cost under “data analysis cost”.

[Figure 11 near here]

3.2. Performance Index for NDT Technologies

The current model evaluates the performance of different NDT technologies regarding the proposed criteria, sub-criteria, and parameters. This is discussed in detail in this section. The model also assembles all performance scores regarding different parameters into one performance index. A technology that has a high-performance index should provide reasonable performance for all or majority of the parameters. Decision-maker should consider the detailed performance of a specific technology besides its overall performance to satisfy his\her needs.

Figure 12 shows the overall PI for different NDT technologies. According to the findings presented in Figure 12, IRT was ranked first ($PI = 0.586$). GPR came in the second place ($PI = 0.58$), followed by IE ($PI = 0.575$), and then UPE ($PI = 0.529$). The differences in the PIs of the first three technologies (i.e., IRT, GPR, and IE) are small. On the other hand, the least overall performance score was assigned to PR ($PI = 0.333$).

[Figure 12 near here]

To have a better understanding of the performance of these technologies, their performance based on different criteria was investigated. Figure 13 shows the PI of NDT technologies according to the five main criteria (i.e., defect detection capability, performance under different environments, ease of use, speed, and cost). Regarding the “defect detection capability” criterion, IE attained the highest PI, followed by GPR, then IRT and UPE (it should be noted that the differences between the PI values of IE, GPR, IRT and UPE are very small). IE provided the best performance in detecting delamination, which had the highest importance weights among all defect types. In addition, IE has multi-defect detection capabilities. These reasons put IE in the first place in terms of the capability criterion.

On the other hand, GPR, IRT, and UPE provided the near-best performance in detecting some defects. In addition, the multi-defect detection capabilities of GPR, IRT, and UPE raised the overall scores of these technologies regarding defect detection capability criterion (e.g., GPR can detect delamination, concrete degradation, and corrosion). On the other hand, the camera, as an NDT technique, came in the last place (with slight differences from ER and PR) since it can only detect surface defects. These types of defects got the least importance weight among other defects under the “capability” criterion. In general, there is a large gap between the PI values of the different NDT technologies at the head of the list and those at the bottom of the list.

Comparing the PIs in the second criterion (i.e., performance under different environment) shows that GPR obtained the highest PI, followed by PR as these technologies provide a stable performance under various bridge deck finishes and different working conditions. There are perceptible differences between the PIs of the aforesaid technologies (i.e., GPR and PR) and the remaining technologies. On the contrary, IRT got the least PI under this criterion since it is very sensitive to working conditions. Under the “ease of use” criterion, the camera attained the highest PI, followed by IRT, and then GPR. Similarly, camera obtained the highest PI under the criterion “speed”, which is markedly higher than the PIs of other technologies. Next in rank were IRT, and then GPR. In fact, the regular camera and infrared camera are portable, easy to use, and can be mounted on a car or a drone, which facilitate collecting data using these devices. These privileges made these technologies surpass other technologies under the “ease of use” and “speed” criteria. Finally, under the “cost” criterion, the highest PI was assigned to the camera, followed by IRT, and then PR.

[Figure 13 near here]

Based on the findings in Figure 13, despite IRT did not obtain the highest score under any criteria, it was ranked first in the overall PI. In fact, IRT provided near-best performance

in most criteria (i.e., the third position in the “capability” criterion and the second position in “ease of use”, “speed”, and “cost” criteria), which made the PI of IRT transcends those of the remaining technologies. On the other hand, GPR provided the best performance with respect to the second criterion (i.e., performance under different environment). Its PI in this criterion is considerably larger than those of other technologies. In addition, it provided near-best performance under the remaining criteria (i.e., the second position in “capability” criterion with a small difference from IE and the third position in “ease of use” and “speed” criteria). These results put GPR in second place.

Three interesting points can be observed in these findings. First, IE obtained the highest PI under only one criterion (i.e., capability) and was not ranked in the first three positions under the remaining criteria. Nevertheless, it attained the third highest overall PI. The main reason behind this is that the “capability” criterion had the highest importance weight among all other criteria, which made this criterion dominates the overall performance index of the NDT technologies. Second, the camera obtained the highest PI under three criteria (i.e., ease of use, speed, and cost), which significantly outweighs the PIs of the other technologies. However, receiving the least score under the “capability” criterion put the camera in the sixth position regarding the overall PI. Third, despite HCP, ER, and PR are efficient methods to detect corrosion, the overall PIs of these technologies were low. That is mainly because other technologies provide multi-defect detection capability (e.g., IE and UPE) and/or provide higher performance under several criteria (e.g., IRT and GPR).

Three important points are worth noticing: (i) the performance of NDT technologies in detecting different defect types (e.g., surface defects, delamination, concrete degradation, and corrosion), (ii) the performance of NDT technologies under different bridge deck finishes, and (iii) the impact of test spacing on the defect detection capabilities of NDT technologies. As for the first point, as shown in Figure 14, the camera got the highest PI regarding surface defects

detection capability, which is clearly higher than that of IRT in the second place. IE was ranked first in detecting delamination, followed by IRT, and then UPE. Regarding concrete degradation, USW placed first in detecting this type of defect, while GPR and UPE provided almost the same performance after USW. IE was ranked third in detecting this type of defect. In corrosion detection capability, HCP provided the highest performance, followed by ER, then GPR, and finally PR.

[Figure 14 near here]

Referring to the second point, Figure 15 shows that GPR provided a stable performance under different bridge deck finishes. This is also observed in Figure 13 since GPR obtained the highest PI in “performance under different environment” criterion. PR provided almost the same performance in the first three cases, while in the last case (i.e., coated rebar), the defect detection capability of PR was significantly reduced. On the other hand, the performances of IE, USW, UPE, and ER were greatly affected by the presence of overlay layers (i.e., cases 2 and 3), especially in the case of the asphalt layer. Similarly, the delamination detection capability of IRT was significantly affected by the presence of overlay layers. Nevertheless, the overall defect detection capability of IRT was not drastically affected since it can be used in detecting surface defects, such as cracking in the pavement layer. This characteristic enhanced the performance of IRT under different overlay layers. Finally, the performance of HCP was considerably affected by the presence of overlay layers and coated rebars (i.e., cases 2, 3, and 4).

[Figure 15 near here]

The last point is related to the defect detection capabilities of NDT technologies when test spacing is increased from 1 foot to 2 or 3 feet. In this part, participants were asked to provide their feedback on the impact of increasing the test space from 1 foot to 2 feet and 3 feet on the defect detection capability of NDT technologies. Many participants did not respond

to this question. Others who responded stated that there is a negative impact if the test spacing is increased, but they could not quantify this impact. Therefore, there is a need to investigate further.

Figure 16 summarizes the responses received from the remaining participants. It is obviously clear that increasing the test spacing from 1 foot to 2 and 3 feet negatively impacted the defect detection capabilities of NDT technologies. This concept is considered an intuitive phenomenon, but the purpose of this question was to quantify this impact.

[Figure 16 near here]

4. Model Testing

Two studies from the literature were used to check the validity of the findings obtained in the current study (Gucunski *et al.* 2013, Omar *et al.* 2017). These two studies were chosen based on the amount of similarity found in the proposed parameters, evaluated technologies, and categorization structure. For example, the first study (Gucunski *et al.* 2013) examined the performance of NDT technologies according to twenty-five parameters, while the second study (Omar *et al.* 2017) considered fifteen parameters. Most of the parameters proposed in these two studies were assessed in the current study. In addition, new parameters were also proposed to enhance the precision of the assessment.

Table 3 shows the comparison results of nine items in the three studies. The first item focused on the methods of data collection, which included expert opinion, field tests, and laboratory tests. As shown in Table 3, Gucunski *et al.* (2013) relied on several sources for data collection (i.e., expert opinion, field, and laboratory tests), whereas the second study (Omar *et al.* 2017) and our current work adopted only one approach (i.e., expert opinion). In the second item, the current study evaluated all the technologies included in the other two studies. In

addition, the regular camera was considered since it is an efficient tool to identify surface defects (Abdelkhalek and Zayed 2020a).

Regarding delamination detection capability, IE was ranked first, followed by IRT in the three studies. However, the technology in the third rank differs in the three studies. Furthermore, in the first study (Gucunski *et al.* 2013), USW got the same PI score as IE. In concrete degradation detection capability, the first study recommended only one technology (i.e. USW) to detect this type of defect, while the second study (Omar *et al.* 2017) excluded this criterion. In the current study, USW was ranked first in detecting this type of defect similar to the first study, followed by GPR and UPE, and then IE.

Regarding corrosion detection capability, HCP occupied the first position in the three studies. Next in rank was PR in the first study and GPR in the second study. ER got the same score as HCP in the first study, while it came in the second place in the current study (ER and PR were not considered in the second study). In ease of use, IRT obtained the highest score in the first and second studies, while it was categorized in the second position in the current study. This is because the camera occupied the first position in the current study, which was not considered in the first and second studies. The first study rated ER in the second place, whereas the second study put HCP in the second place, though the first study ranked it in the third place. GPR occupied third place in the second and current studies.

[Table 3 near here]

In terms of the “speed” criterion, IRT and GPR were ranked first and second, respectively, as reported in the first and second studies. But the current study placed the camera in the first position, followed by IRT, and then GPR. The first place under the “cost” criterion was occupied by IRT, HCP, ER, and PR (in the first study), IRT (second study), and camera (current study). The second place was occupied by GPR, IE and USW (first study), GPR

(second study), and IRT (current study). For the overall performance, GPR came in the first place in the first study, third in the second study, and second in the current study. IE placed third in the first and current studies, whereas it placed first in the second study. In the current study, the highest overall PI was assigned to IRT. In general, despite the similarities found in several aspects of the three studies (e.g., number of technologies, data collection techniques, parameters considered), there are still some disparities. Therefore, some differences in the findings are observed.

5. Sensitivity Analysis

A sensitivity analysis was conducted to check how changing the importance weights of the proposed parameters affects the PI of the NDT technologies and their priorities. The weight of each parameter was changed in the interval [-50%, +50%] from its value with a step of 5%. After each change, TOPSIS calculations were conducted to get the overall performance index of different technologies and their rankings. Table 4 shows the findings of the sensitivity analysis. Figure 17 shows the impact of changing the importance weights of five parameters on the priority of different technologies with respect to overall PI. The findings show that twelve parameters exerted some influence in prioritizing NDT technologies (Table 4). This impact was categorized into four groups: low, medium, high, and very high. Most of the parameters was categorized under the low impact (i.e., seven parameters). Two parameters were sorted under the medium impact, while the remaining parameters (i.e., three) provided high to very high impact.

For example, the “ability to detect cracks” parameter significantly impacted (i.e., high impact) the priority of five technologies (Figure 17[a]). It reversed the ranking of GPR and IRT in the range -50 to -15, IE and GPR in the range +15 to +50, and HCP and Camera in the range -50 to -30. The “ability to detect delamination with different sizes” parameter provided a low

impact on the priority of two technologies. The rankings of IE and GPR were reversed in the range +35 to +50, as shown in Figure 17(b). “Confidence in delamination detection results” parameter provided the highest impact among all the parameters. It impacted the priority of six technologies, as shown in Figure 17(c). This parameter had the highest importance weight among all parameters (Figure 11).

Eleven parameters out of twelve are located under the “defect detection capability” criterion. The main reason for this is the imperfect defect detection capability of different technologies. Therefore, increasing the importance weight of any parameter under the “defect detection capability” criterion raises the PI of the technologies that can detect this defect. Furthermore, it reduces the PI of the technologies that cannot detect this type of defect and vice versa. For example, in Table 4, reducing the weight of the “confidence in corrosion detection results” parameter made the PI of IE (i.e., cannot detect corrosion) exceeds the PI of GPR (i.e., can detect corrosion) (Figure 17 [e]). On the other hand, increasing the weight of “confidence in delamination detection results” parameter made the PI of IE surpass those of GPR and IRT (Figure 17 [c]). Despite the ability of the three technologies (i.e., IE, GPR, and IRT) to detect delamination, IE provided the best performance regarding this defect (Figure 14). Figure 17(a) shows that changing the importance weight of “ability to detect cracks” parameter significantly increased the PI of the camera.

The most impacted technologies due to changing the value of importance weights of different parameters were IRT, GPR, and IE. The overall PIs of these technologies are too close (Figure 12), which makes the priority of these technologies very sensitive to any change in the importance weights.

6. A Selection Model for Inspection System Components

It is crucial to carefully choose the devices that will be involved in the inspection process to ensure successful outcomes. In this regard, this section provides some guidelines to aid the selection of the most effective NDT techniques for inspection purposes (i.e., hereinafter referred to as the selection of the inspection system components). Figure 18 shows three main filters required to select the components of the inspection system. First, the decision-maker should select technologies that can detect all targeted defects (filter 1). Second, the performance of these technologies should be checked to ensure their efficiency under the type of finishing in the bridge that requires inspection (filter 2). Findings in Figure 15 can be used to guide decision-making in this stage. Finally, if there are various alternatives for the components of the inspection system, the different combinations can be evaluated by calculating the performance index for each combination (filter 3).

To calculate the performance index for the proposed system, the decision-maker should assign importance weights for different performance parameters and rate the components of the system (i.e., NDT technologies) regarding these parameters (Filter 3: steps 1 and 2 in Figure 18). Decision-maker can use the importance weights and scores that are provided in this study or propose values according to his/her preferences. After that, system defect detection capability score regarding each defect type can be determined using Equation (12) as follows (Filter 3: step 3):

$$R_{scapf} = \max \{R_{1capf}, R_{2capf}, \dots, R_{mcapf}\} \quad for f = 1 to 4 \quad [12]$$

where R_{scapf} = performance score for the inspection system according to parameter f in defect detection capability (i.e., there are four parameters in defect detection capability: surface defects, delamination, concrete degradation, and corrosion), R_{1capf} = performance score for the

technology 1 according to parameter f in defect detection capability, m = number of technologies involved in the inspection system.

The overall score for defect detection capability of the inspection system can be calculated using the following equation:

$$R_{scap} = \sum_{f=1}^4 W_f * R_{scap_f} \quad [13]$$

The performance score of the inspection system (R_{sf}) regarding the “ease of use” criterion can be determined using Equation (14) as follows (Filter 3: step 4):

$$R_{sf} = \min \{R_{1f}, R_{2f}, \dots, R_{mf}\} \quad \text{for } f = \text{ease of use} \quad [14]$$

The performance score of the inspection system (R_{sf}) regarding “speed and cost” criteria can be determined using Equation (15) as follows (Filter 3: step 4):

$$R_{sf} = \frac{\sum_{t=1}^m R_{tf}}{m^2} \quad \text{for } f = \text{speed and cost} \quad [15]$$

The overall performance index (PI_s) for the system can be calculated using the calculated scores (R_{sf}) from Equations (13) to (15) and the proposed importance weights (W_f) as follows (Filter 3: step 5):

$$PI_s = \sum_{f=1}^r W_f * R_{sf} \quad [16]$$

where r = performance criteria (i.e., defect detection capability, ease of use, speed, and cost criteria).

The above steps were implemented on two types of bridges: one without an overlay layer and the other with an asphalt overlay. The inspection scope was to detect delamination and corrosion in the two bridges. According to filter 1 in Figure 18, three systems were

proposed to be used in the inspection: GPR, IE+HCP, and IRT+HCP. Two sets of importance weights were used to evaluate the alternatives as shown in Table 5. Scores in Figure 18 were used to get the overall performance index.

In the first case (i.e., without an overlay layer), all the proposed systems were efficient (filter 2). Applying the proposed steps in filter 3 to get the overall performance indices of the proposed systems, Equation (12) was used to get delamination and corrosion scores (Table 5). Equations (13) was then, used to calculate the defect detection capability score. Scores regarding ease of use, speed, and cost were determined using Equations (14) and (15). The overall performance index was calculated using Equation (16). In the first set of importance weights, the system that incorporates GPR got the highest PI (0.486), while, in the second set of importance weights, the system that incorporates IE and HCP surpassed the other systems (PI = 0.593). The difference in the weight of defect detection capability between the two sets of weights (i.e., the weight of defect detection capability in the second set is higher than the first one) is the main reason that affects the decision.

In the second case (i.e., with asphalt overlay), after checking the finishing constraints (Figure 15) on using some technologies in filter 2 (Figure 18), it was found that only one system was suitable to do the inspection (i.e., GPR). In this case, there is no need to go to filter 3 as there is only one alternative to use in the inspection.

7. Conclusions

This paper discusses the Performance Assessment Model (PAM) that was developed to evaluate the performance of NDT technologies incorporated in concrete bridge deck inspection. The model utilized forty parameters to precisely assess the performance of NDT technologies from diverse perspectives. Three levels were proposed to organize these parameters: criteria, sub-criteria, and parameters. A survey questionnaire was used to collect

the data needed for the proposed model. In data analysis, the ANP technique was used to calculate the importance weights of the proposed criteria, sub-criteria, and parameters, whereas TOPSIS was utilized to prioritize NDT technologies according to the proposed parameters (assign a performance index for each technology).

The findings of this study showed that the “capability” criterion was the most important factor in implementing NDT technologies in the inspection process (importance weight = 0.3636), followed by the “inspection cost” criterion (0.2171), and then the “inspection speed” criterion (0.1537). On the other hand, the “data collection speed” sub-criterion and “confidence in delamination detection results” parameter got the highest importance weight at the sub-criteria and parameter levels, respectively. In measuring the performance scores for using NDT technologies in the inspection process, IE and GPR were ranked first under “capability” and “performance under the different environment” criteria, respectively, whereas the camera placed first under the remaining main criteria (ease of use, speed, and cost). Overall, IRT got the highest overall performance index (0.586), followed by GPR (0.58), and then IE (0.575).

To validate the findings of this study, a comparison was conducted with two studies from the literature. Furthermore, a sensitivity analysis was conducted to assess the impact of changing the importance weights of the proposed parameters on the priority of NDT technologies. The results of the sensitivity analysis showed that: 1) due to the imperfect defect detection capability of NDT technologies, most of the parameters (i.e., eleven) located in the “defect detection capability” criterion exhibited some level of impact on the priority of NDT technologies, 2) “confidence in delamination detection results” parameter had the highest impact among all the parameters, and 3) the most impacted technologies were IRT, GPR, and IE. Finally, a model was developed to facilitate the selection of the most efficient technologies for inspection purposes. The model applies three filters to satisfy the decision-making needs and ensure successful outcomes.

Two main limitations were identified in the current study. First, the sample size for the survey was not large (24 responses). The participants mentioned that different aspects of implementing NDT technologies in the inspection were well covered, but they criticized the length of the questionnaire. This may represent the main challenge to get more responses. Second, this study utilized one method for data collection (i.e., survey questionnaire). Despite this method reflects the expert opinion based on conducting non-destructive tests several times under different circumstances, incorporating other approaches with the current one (e.g., field and laboratory tests) could address more confidence in the assessment findings. Therefore, this aspect represents the future direction for the current study.

8. Disclosure Statement

No potential conflict of interest was reported by the authors.

9. References

- Abdelkhalek, S. & Zayed, T. (2019). State-of-the-art review of defect detection in concrete bridge deck. *CIB World Building Congress 2019, Hong Kong SAR, China, June 17-21*.
- Abdelkhalek, S. & Zayed, T. (2020a). Comprehensive inspection system for concrete bridge deck application: Current situation and future needs. *Journal of Performance of Constructed Facilities*, 34 (5), 03120001.
- Abdelkhalek, S. & Zayed, T. (2020b). Simulation-based planning of concrete bridge deck inspection with non-destructive technologies. *Automation in Construction*, 119, 103337.
- Agdas, D., Rice, J.A., Martinez, J.R. & Lasa, I.R. (2016). Comparison of visual inspection and structural-health monitoring as bridge condition assessment methods. *Journal of Performance of Constructed Facilities*, 30 (3), 04015049.
- Akroush, N.S. & El-Adaway, I.H. (2017). Utilizing construction leading safety indicators: Case study of tennessee. *Journal of Management in Engineering*, 33 (5), 06017002.
- Alsharqawi, M., Zayed, T. & Dabous, S.A. (2020). Integrated condition-based rating model for sustainable bridge management. *Journal of Performance of Constructed Facilities*, 34 (5), 04020091.
- Astm (2013). Standard test method for detecting delaminations in bridge decks using infrared thermography. ASTM International, West Conshohocken, PA, USA.
- Azari, H. & Lin, S. (2019). Evaluation of the impact echo method for concrete bridge decks with asphalt overlays. *Transportation Research Record*, 2673 (2), 436-444.
- Azari, H., Nazarian, S. & Yuan, D. (2014). Assessing sensitivity of impact echo and ultrasonic surface waves methods for nondestructive evaluation of concrete structures. *Construction and Building Materials*, 71, 384-391.
- Baby, S. (2013). AHP modeling for multicriteria decision-making and to optimise strategies for protecting coastal landscape resources. *International Journal of Innovation, Management Technology*, 4 (2), 218-227.

- Baruch, Y. (1999). Response rate in academic studies-a comparative analysis. *Human Relations*, 52 (4), 421-438.
- David, A. & Varghese, S. (2021). AHP model for performance improvement in LSGD projects. *Proceedings of SECON 2020*. Cham: Springer International Publishing, 835-850.
- Demirkesen, S. & Bayhan, H.G. (2020). A lean implementation success model for the construction industry. *Engineering Management Journal*, 32 (3), 219-239.
- Doloi, H. (2008). Application of AHP in improving construction productivity from a management perspective. *Construction Management and Economics*, 26 (8), 841-854.
- Duque, L., Seo, J. & Wacker, J. (2018). Synthesis of unmanned aerial vehicle applications for infrastructures. *Journal of Performance of Constructed Facilities*, 32 (4), 04018046.
- El-Abbasy, M.S., Senouci, A., Zayed, T. & Mosleh, F. (2015). A condition assessment model for oil and gas pipelines using integrated simulation and analytic network process. *Structure and Infrastructure Engineering*, 11 (3), 263-281.
- El-Abbasy, M.S., Zayed, T., El Chanati, H., Mosleh, F., Senouci, A. & Al-Derham, H. (2019). Simulation-based deterioration patterns of water pipelines. *Structure and Infrastructure Engineering*, 15 (7), 965-982.
- Forman, E. & Peniwati, K. (1998). Aggregating individual judgments and priorities with the analytic hierarchy process. *European Journal of Operational Research*, 108 (1), 165-169.
- Gucunski, N., Imani, A., Romero, F., Nazarian, S., Yuan, D., Wiggengerhauser, H., Shokouhi, P., Taffe, A. & Kutrubes, D. (2013). Nondestructive testing to identify concrete bridge deck deterioration. SHRP 2 Report (S2-R06A-RR-1), Washington, D.C., USA: The National Academies Press: Available at Transportation Research Board, <http://www.trb.org/Main/Blurbs/167278.aspx>.
- Hesse, A.A., Atadero, R.A. & Ozbek, M.E. (2017). Using expert opinion to quantify uncertainty in and cost of using nondestructive evaluation on bridges. *Advances in Civil Engineering*, 7925193, 1-12.
- Huston, D., Cui, J., Burns, D. & Hurley, D. (2011). Concrete bridge deck condition assessment with automated multisensor techniques. *Structure and Infrastructure Engineering*, 7 (7-8), 613-623.
- Hwang, C.L. & Yoon, K.P. (1981). *Multiple attribute decision making : Methods and applications : A state-of-the-art survey (lecture notes in economics and mathematical systems ; 186)*: Springer-Verlag, Berlin, Heidelberg.
- Iason, G. (2014). *Infrastructure performance assessment of subway networks*. (M.Sc. Thesis). Available at Concordia University, <https://spectrum.library.concordia.ca/978397/>.
- Janků, M., Cikrle, P., Grošek, J., Anton, O. & Stryk, J. (2019). Comparison of infrared thermography, ground-penetrating radar and ultrasonic pulse echo for detecting delaminations in concrete bridges. *Construction and Building Materials*, 225, 1098-1111.
- Kil, S.H., Lee, D.K., Kim, J.H., Li, M.H. & Newman, G. (2016). Utilizing the analytic hierarchy process to establish weighted values for evaluating the stability of slope revegetation based on hydroseeding applications in south korea. *Sustainability (Switzerland)*, 8 (1), 1-17.
- Lin, S., Meng, D., Choi, H., Shams, S. & Azari, H. (2018). Laboratory assessment of nine methods for nondestructive evaluation of concrete bridge decks with overlays. *Construction and Building Materials*, 188, 966-982.
- Mahmoud, S.a.I. (2017). *Integrated sustainability assessment and rehabilitation framework for existing buildings*. (Ph.D. Thesis). Available at Concordia University, <https://spectrum.library.concordia.ca/982747/>.
- Moufti, S. (2013). *A defect-based approach for detailed condition assessment of concrete bridges*. (M.Sc. Thesis). Available at Concordia University, <https://spectrum.library.concordia.ca/978014/>.
- Oh, T., Kee, S.-H., Arndt Ralf, W., Popovics John, S. & Zhu, J. (2013). Comparison of NDT methods for assessment of a concrete bridge deck. *Journal of Engineering Mechanics*, 139 (3), 305-314.

- Omar, T., Nehdi, M. & Zayed, T. (2017). Performance of NDT techniques in appraising condition of reinforced concrete bridge decks. *Journal of Performance of Constructed Facilities*, 31 (6), 04017104.
- Pailes, B.M. & Gucunski, N. (2015). Understanding multi-modal non-destructive testing data through the evaluation of twelve deteriorating reinforced concrete bridge decks. *Journal of Nondestructive Evaluation*, 34 (4), 30.
- Popescu, C., Täljsten, B., Blanksvärd, T. & Elfgrén, L. (2019). 3D reconstruction of existing concrete bridges using optical methods. *Structure and Infrastructure Engineering*, 15 (7), 912-924.
- Rens Kevin, L., Nogueira Carnot, L. & Transue David, J. (2005). Bridge management and nondestructive evaluation. *Journal of Performance of Constructed Facilities*, 19 (1), 3-16.
- Rong, X., Zhang, S., Bakhtawar, B., Hussein, M., Abdelmageed, S., Tariq, S. & Zayed, T. (2020). Performance assessment for modular integrated construction projects in hong kong. *International Journal of Architecture, Engineering and Construction*, 9 (3), 1-3.
- Saaty, T.L. (1980). *The analytical hierarchy process*: McGraw-Hill, New York, USA.
- Saaty, T.L. (2004a). Decision making — the analytic hierarchy and network processes (AHP/ANP). *Journal of Systems Science and Systems Engineering*, 13 (1), 1-35.
- Saaty, T.L. (2004b). Fundamentals of the analytic network process — dependence and feedback in decision-making with a single network. *Journal of Systems Science and Systems Engineering*, 13 (2), 129-157.
- Salehi, S., Jalili Ghazizadeh, M. & Tabesh, M. (2018). A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems. *Structure and Infrastructure Engineering*, 14 (6), 743-765.
- Salman, A.F.M., Skibniewski, M.J. & Basha, I. (2007). BOT viability model for large-scale infrastructure projects. *Journal of Construction Engineering and Management* 133 (1), 50-63.
- Sambasivan, M. & Fei, N.Y. (2008). Evaluation of critical success factors of implementation of ISO 14001 using analytic hierarchy process (AHP): A case study from malaysia. *Journal of Cleaner Production*, 16 (13), 1424-1433.
- Scott, M., Rezaizadeh, A., Delahaza, A., Santos, C.G., Moore, M., Graybeal, B. & Washer, G. (2003). A comparison of nondestructive evaluation methods for bridge deck assessment. *NDT & E International*, 36 (4), 245-255.
- Seo, J., Duque, L. & Wacker, J. (2018a). Drone-enabled bridge inspection methodology and application. *Automation in Construction*, 94, 112-126.
- Seo, J., Duque, L. & Wacker, J.P. (2018b). Field application of UAS-based bridge inspection. *Transportation Research Record*, 2672 (12), 72-81.
- Sultan, A.A. & Washer, G.A. (2018). Comparison of two nondestructive evaluation technologies for the condition assessment of bridge decks. *Transportation Research Record*, 2672 (41), 113-122.
- Tarei Pradeep, K., Thakkar Jitesh, J. & Nag, B. (2018). A hybrid approach for quantifying supply chain risk and prioritizing the risk drivers: A case of indian petroleum supply chain. *Journal of Manufacturing Technology Management*, 29 (3), 533-569.
- Tinkey, Y. & Olson, L.D. (2010). Vehicle-mounted bridge deck scanner. Highway IDEA Program (Final Report for Highway IDEA Project 132) , Washington, D.C., USA: Transportation Research Board: Available at: <https://trid.trb.org/view/933394>.
- Vaghefi, K., Oats Renee, C., Harris Devin, K., Ahlborn Theresa, M., Brooks Colin, N., Endsley, K.A., Roussi, C., Shuchman, R., Burns Joseph, W. & Dobson, R. (2012). Evaluation of commercially available remote sensors for highway bridge condition assessment. *Journal of Bridge Engineering*, 17 (6), 886-895.
- Vemuri, S.H. & Atadero, R.A. (2017). Case study on rapid scanning techniques for concrete bridge decks with asphalt overlay: Ground-penetrating radar and infrared thermography. *Practice Periodical on Structural Design and Construction*, 22 (2), 05016004.

- Waris, M., Panigrahi, S., Mengal, A., Soomro, M.I., Mirjat, N.H., Ullah, M., Azlan, Z. & Khan, A. (2019). An application of analytic hierarchy process (AHP) for sustainable procurement of construction equipment: Multicriteria-based decision framework for malaysia. *Mathematical Problems in Engineering*, 2019, 1-20.
- Wilkins, J.R. (2011). Construction workers' perceptions of health and safety training programmes. *Construction Management and Economics*, 29 (10), 1017-1026.
- Yadollahi, M., Abd Majid, M.Z. & Mohamad Zin, R. (2015). Post-pareto optimality approach to enhance budget allocation process for bridge rehabilitation management. *Structure and Infrastructure Engineering*, 11 (12), 1565-1582.
- Yehia, S., Abudayyeh, O., Nabulsi, S. & Abdelqader, I. (2007). Detection of common defects in concrete bridge decks using nondestructive evaluation techniques. *Journal of Bridge Engineering*, 12 (2), 215-225.
- Yeo, S.Z., How, B.S., Ngan, S.L., Ng, W.P.Q., Leong, W.D., Lim, C.H. & Lam, H.L. (2020). An integrated approach to prioritise parameters for multi-objective optimisation: A case study of biomass network. *Journal of Cleaner Production*, 274, 123053.
- Zhao, X., Hwang, B.-G. & Lee, H.N. (2016). Identifying critical leadership styles of project managers for green building projects. *International Journal of Construction Management*, 16 (2), 150-160.
- Zhao, X., Hwang, B.-G., Pheng Low, S. & Wu, P. (2015). Reducing hindrances to enterprise risk management implementation in construction firms. *Journal of Construction Engineering and Management*, 141 (3), 04014083.

List of Figures:

- Figure 1:** Methodology of the study
- Figure 2:** PAM network
- Figure 3:** Software used and data flow from questionnaire design stage to data analysis stage
- Figure 4:** Section 3 in the questionnaire
- Figure 5:** Section 5 in the questionnaire
- Figure 6:** Participants' experience
- Figure 7:** Participant's affiliation
- Figure 8:** Device's use experience
- Figure 9:** Importance weights of main criteria
- Figure 10:** Importance weights of sub-criteria
- Figure 11:** Importance weights of parameters
- Figure 12:** Overall Performance Index of non-destructive technologies
- Figure 13:** Performance Index of non-destructive technologies according to main criteria
- Figure 14:** Performance scores of non-destructive technologies according to sub-criteria in defect detection capability
- Figure 15:** Impact of bridge deck finishes on the performance of NDT technologies
- Figure 16:** Impact of increasing test spacing

978 **Figure 17:** Sensitivity analysis results for five parameters: a) Ability to detect cracks, b) Ability to
979 detect delamination at varying depths, c) Confidence in delamination detection results, d)
980 Ability to detect depth and size of corrosion, e) Confidence in corrosion detection results
981 **Figure 18:** Inspection system components selection model

982 **List of Tables:**

983 **Table 1:** Proposed parameters to evaluate the performance of non-destructive technologies

984 **Table 2:** Degree of comparative importance scale (Saaty, 2004a)

985 **Table 3:** Comparison of the findings of the present study with other studies

986 **Table 4:** Results of sensitivity analysis

987 **Table 5:** Case study

988

989

1 *Table 1: Proposed parameters in the literature to evaluate the performance of non-*
2 *destructive technologies*

No.	Criteria	Sub-criteria	Parameters
1	Capability	Surface defects	<ul style="list-style-type: none"> • Ability to detect cracks • Ability to detect scaling • Ability to detect spalling
		Delamination	<ul style="list-style-type: none"> • Ability to detect different defect depths
		Concrete degradation	<ul style="list-style-type: none"> • Ability to detect different defect sizes
		Corrosion	
2	Performance under different environments	Efficiency under different finishes	Defect detection capability under different overlay layers
		Consistency	Repeatability
		Sensitivity	Sensitivity to ambient noise and/or environmental conditions
3	Ease of use	Ease of data collection	<ul style="list-style-type: none"> • Portability • Expertise needed for data collection • Potential for automation and improvement
		Ease of data analysis	<ul style="list-style-type: none"> • Complexity of analysis and interpretation • Expertise needed for data analysis • Potential for automation and improvement
4	Cost	Data collection cost	<ul style="list-style-type: none"> • Equipment cost • Maintenance cost • Operator Cost
		Data analysis cost	<ul style="list-style-type: none"> • Software cost • Analyst cost
5	Speed	Data collection speed	<ul style="list-style-type: none"> • Pre-collection preparation • Time needed for data collection
		Data analysis speed	Time needed for data analysis

4 *Table 2: Degree of comparative importance scale (Saaty, 2004a)*

Degree of importance	1	3	5	7	9	2,4,6,8
Definition	Equal importance	Moderate importance	Strong importance	very Strong importance	Extreme importance	Intermediate values

5

6 *Table 3: Comparison of the findings of the present study with other studies*

Item	Gucunski <i>et al.</i> (2013)		Omar <i>et al.</i> (2017)	Current study
Data collection technique	Expert opinion, field and laboratory tests		Expert opinion	Expert opinion
Technologies included	GPR, IE, USW, HCP, ER, PR, IR, chain drag		GPR, IE, UPE, HCP, IRT	GPR, IE, USW, UPE, HCP, ER, PR, IRT, Camera
Delamination detection capability	1st	IE, USW	IE	IE
	2nd	IRT	IRT	IRT
	3rd	GPR	UPE	UPE
Concrete degradation detection capability	USW		N/A	1st USW
				2nd GPR and UPE
				3rd IE
Corrosion detection capability	1st	HCP, ER	HCP	HCP
	2nd	PR	GPR	ER
	3rd	GPR	N/A	GPR and PR
Ease of use	1st	IRT	IRT	Camera
	2nd	ER	HCP	IRT
	3rd	HCP	GPR	GPR
Speed	1st	IRT	IRT	Camera
	2nd	GPR	GPR	IRT
	3rd	ER, HCP	HCP	GPR
Cost	1st	IRT, HCP, ER, PR	IRT	Camera
	2nd	GPR, IE, USW	GPR	IRT
	3rd	N/A	HCP	PR
Overall performance	1st	GPR	IE	IRT
	2nd	USW	HCP	GPR
	3rd	IE	GPR	IE
	4th	ER	UPE	UPE

20 *Table 4: Results of sensitivity analysis*

Parameter	Impact level	Impact range	Impacted technologies
Ability to detect cracks	high	-50 to -15	GPR and IRT
		+15 to +50	IE and GPR
		-50 to -30	HCP and Camera
Ability to detect delamination with different sizes	Low	+35 to +50	IE and GPR
Ability to detect delamination at varying depths	Low	+20 to +50	IE and GPR
Confidence in delamination detection results	very high	-50 to -10	GPR and IRT
		-50	Camera and USW
		+35 to +50	IE and IRT
		+5 to +50	IE and GPR
Standard scale to identify deteriorated concrete severity	low	+45 to +50	UPE and GPR
		+30 to +50	GPR and IRT
		+35 to +50	GPR and IRT
Confidence in concrete degradation results	low	+40 to +50	IE and IRT
Ability to identify active corrosion	low	+35 to +50	GPR and IRT
Ability to detect corrosive environment	low	+35 to +50	GPR and IRT
standard scale to identify rebar corrosion severity	medium	-50 to -20	IE and GPR
		+20 to +50	GPR and IRT
Ability to detect depth and size of corrosion	medium	-50 to -15	IE and GPR
		+15 to +50	GPR and IRT
Confidence in corrosion detection results	high	-50 to -15	IE and GPR
		+15 to +50	GPR and IRT
		+25 to +50	HCP and Camera
Level of the experience of the operator	low	-50 to -30	GPR and IRT
		+35 to +50	IE and GPR

21

22

23

24

25

26

27

28

29

30 *Table 5: Case study*

<u>Weight set 1</u>					
Parameter		weight	Scores		
			<i>GPR</i>	<i>IE+HCP</i>	<i>IRT+HCP</i>
Delamination		0.55	0.563	1.000	0.672
Corrosion		0.45	0.802	0.992	0.992
Defect detection capability		0.3	0.671	0.996	0.816
Ease of use		0.2	0.479	0.364	0.367
Speed		0.25	0.406	0.156	0.233
Cost		0.25	0.352	0.137	0.240
Performance index			0.486	0.445	0.436
<u>Weight set 2</u>					
Parameter		weight	Scores		
			<i>GPR</i>	<i>IE+HCP</i>	<i>IRT+HCP</i>
Delamination		0.55	0.563	1.000	0.672
Corrosion		0.45	0.802	0.992	0.992
Defect detection capability		0.5	0.671	0.996	0.816
Ease of use		0.1	0.479	0.364	0.367
Speed		0.2	0.406	0.156	0.233
Cost		0.2	0.352	0.137	0.240
Performance index			0.535	0.593	0.539

31

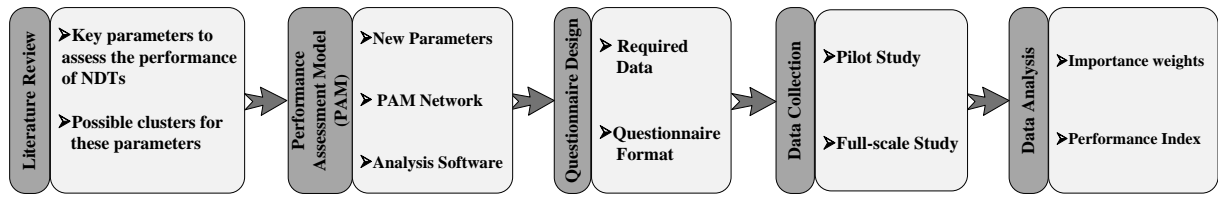


Figure 1: Methodology of the study

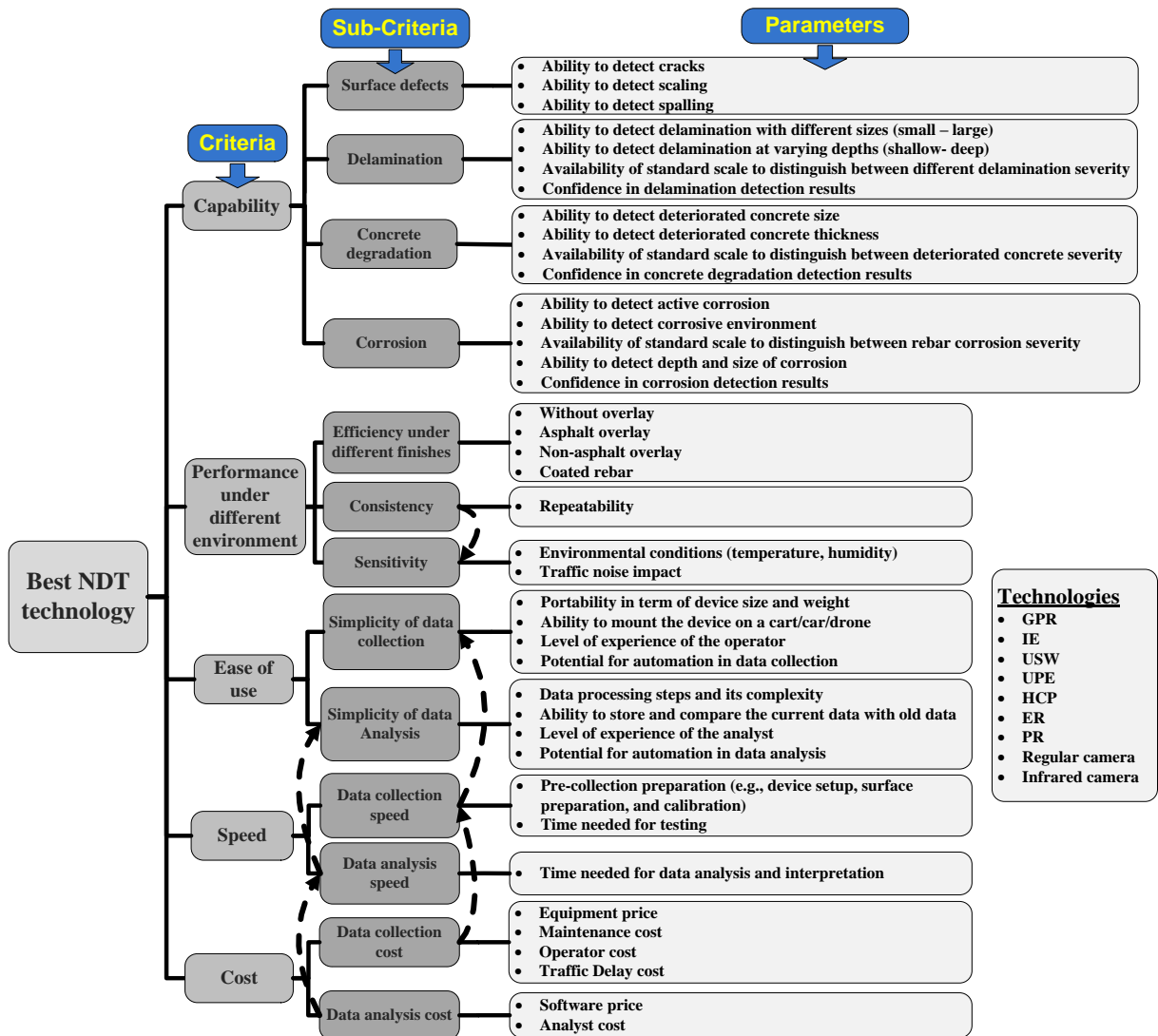


Figure 2: PAM network

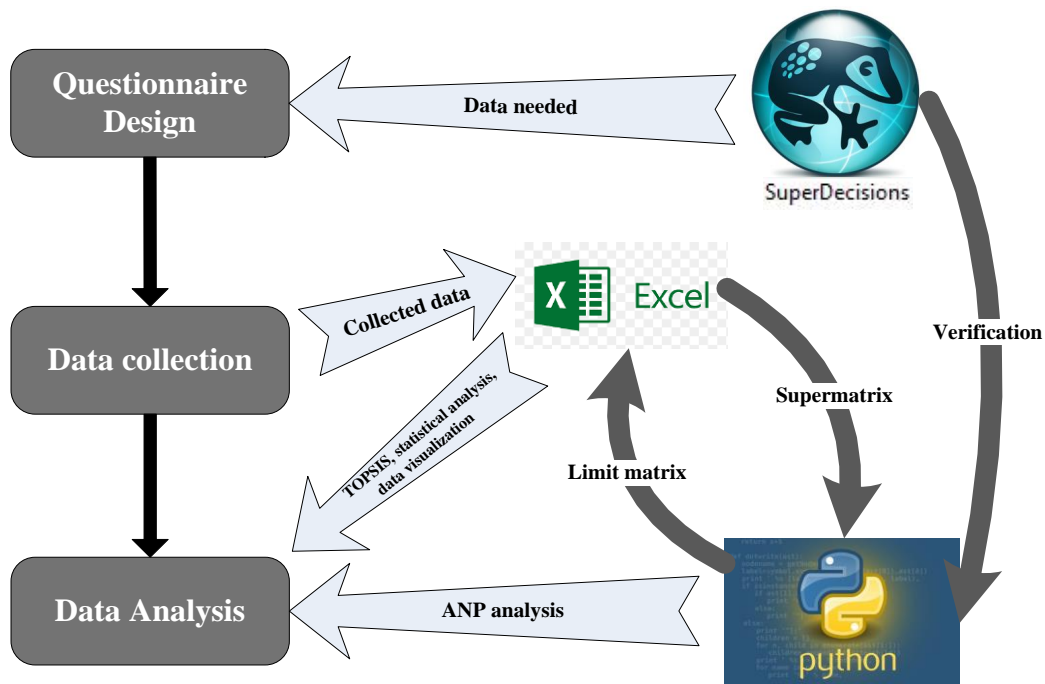


Figure 3: Software used and data flow from questionnaire design stage to data analysis stage

Section 3: Degree of Importance for Main Criteria and Sub-criteria (Pairwise Comparison)

In this section, criteria/sub-criterion in each group in the previous figure will be compared to each other to identify the degree of importance for all of them. Please, fill the table below as in the following example.

6. What is the degree of importance of X criterion when compared to Y criteria?

Example:

Criteria	Absolute		Very strong		Strong		Moderate		Equal		Moderate		Strong		Very strong		Absolute	Criteria
	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
X criterion		✓			X is more important								Y is more important					Y1 criterion
																		Y2 criterion
																		Y3 criterion

This means that the importance of criterion X is 8 times the importance of criterion Y1

This means that the importance of criterion X is equal to the importance of criterion Y2

This means that the importance of criterion X is 1/8 times the importance of criterion Y3

Criteria X	Absolute		Very strong		Strong		Moderate		Equal		Moderate		Strong		Very strong		Absolute	Criteria Y
	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
	Insert the following symbol in this table → ✓																	
Criterion X is more important ← Main Criteria → Criterion Y is more important																		
Capability of NDT technology to detect different defects																		Performance under different environment
																		Ease of use
																		Speed
																		Cost

Figure 4: Section 3 in the questionnaire

Section 5: Rating of NDT Technology

9. Rate NDT technologies (i.e., in the head of the table) according to the parameters in the following table (see this example→)?

- Please use a **scale from 1 to 10** in your rating as in the following example: 10= very high & 1= very low
- Please, use **0** when the device cannot provide this parameter
- The **highlighted** notes in **yellow** indicate that this parameter will use an **inverse scale**: 10 = very low & 1=very high

Example:

GPR = Ground Penetrating Radar, IE = Impact Echo, USW = Ultrasonic Surface Wave, UPE = Ultrasonic Pulse Echo, HCP = Half-cell Potential, ER = Electrical Resistivity, PR = Polarization Resistance, IRT = Infrared Thermography

Parameters	GPR	IE	USW	UPE	HCP	ER	PR	IRT	Camera	Notes
Parameters in Surface Defects										
Ability to detect cracks									9	10: very High & 1: very Low
Parameters in Delamination										
Ability to detect delamination at varying depths (shallow to deep)										10: very High & 1: very Low
Parameters in Sensitivity										
Sensitivity to environment conditions (inverse scale)		8						2		10: very low & 1: very High
Parameters in Portability										
Portability in term of device size and weight									10	very High & 1: very Low

This means that IE has a low sensitivity to environment conditions (inverse scale)

This means that IRT has a high sensitivity to environment conditions (inverse scale)

This means that the camera is an extremely portable device

This means that camera can detect surface cracks (very high capability)

Parameters	GPR	IE	USW	UPE	HCP	ER	PR	IRT	Image	Notes
Parameters in Surface Defects										
Ability to detect cracks										10: very High & 1: very Low
Ability to detect scaling										10: very High & 1: very Low
Ability to detect spalling										10: very High & 1: very Low
Parameters in Delamination										
Ability to detect delamination with different sizes (small – large)										10: very High & 1: very Low
Ability to detect delamination at varying depths (shallow to deep)										10: very High & 1: very Low
Availability of standard scale to accurately identify delamination severity										10: very High & 1: very Low

Figure 5: Section 5 in the questionnaire

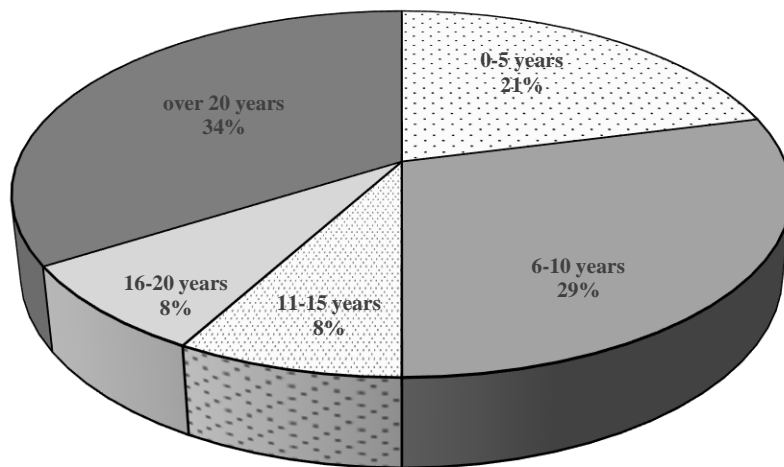
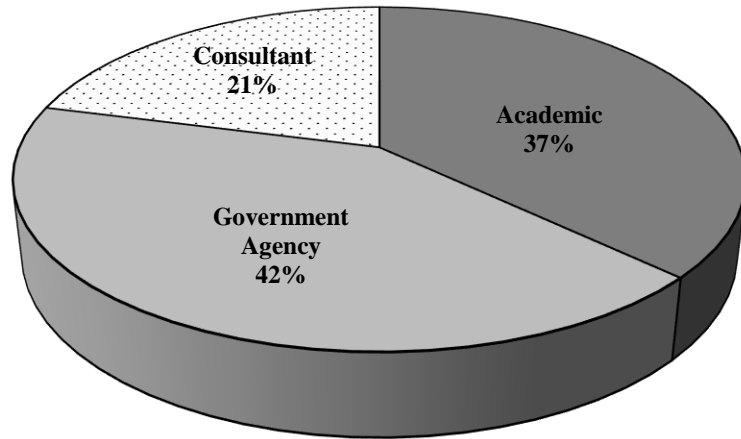
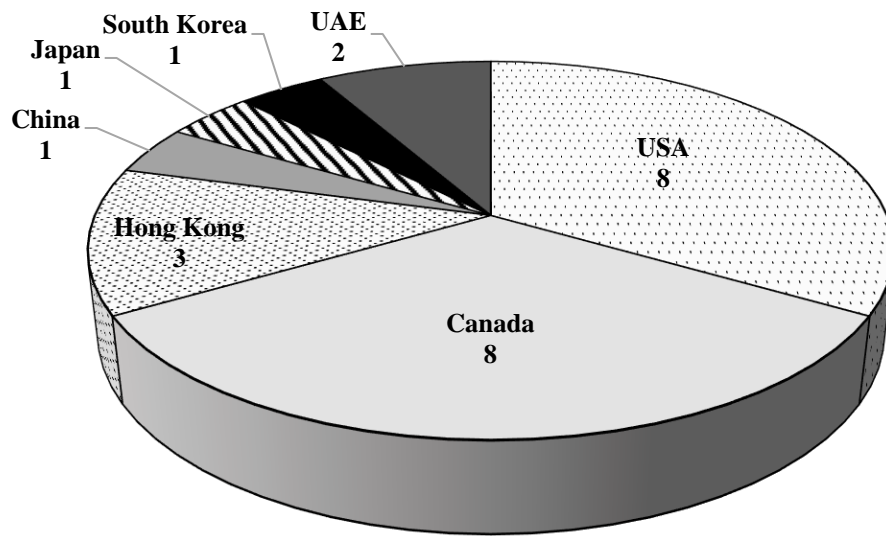


Figure 6: Participants' experience



a) Participant's affiliation



b) Participant's countries

Figure 7: Participant's information: a) Participant's affiliation and b) Participant's countries

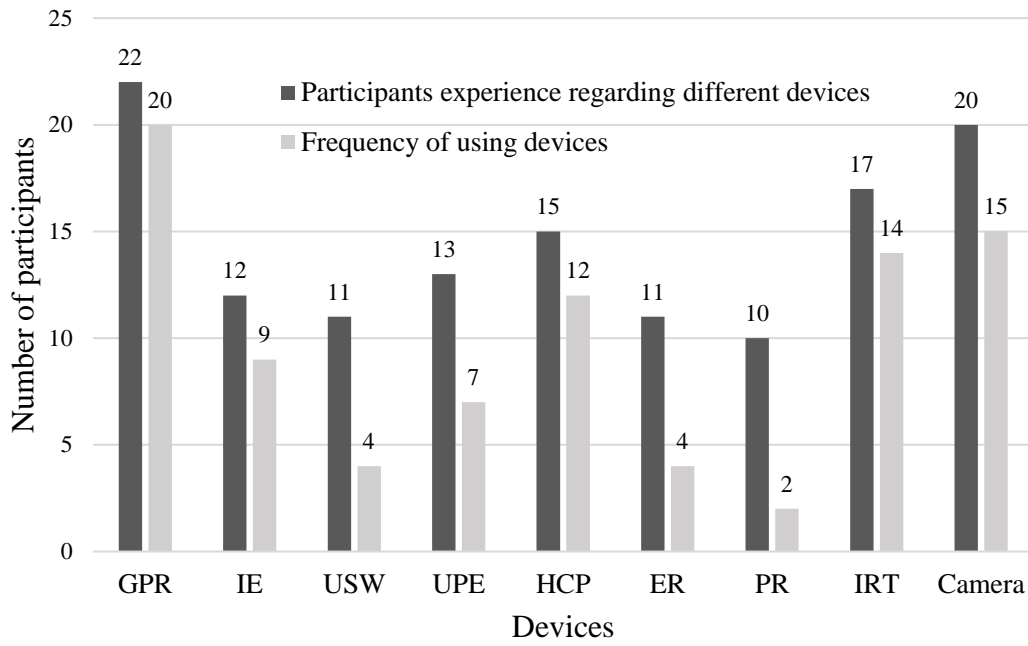


Figure 8: Participants experience regarding different devices and frequency of using devices

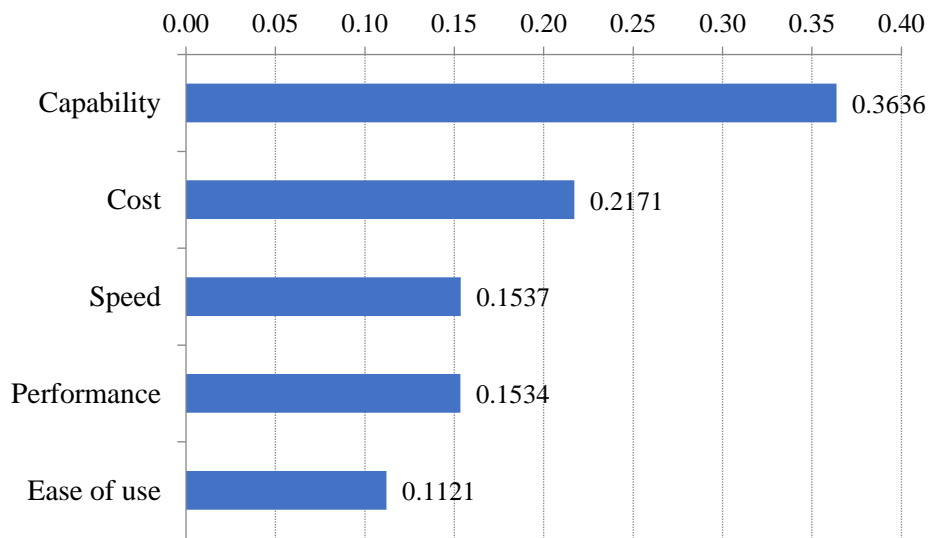
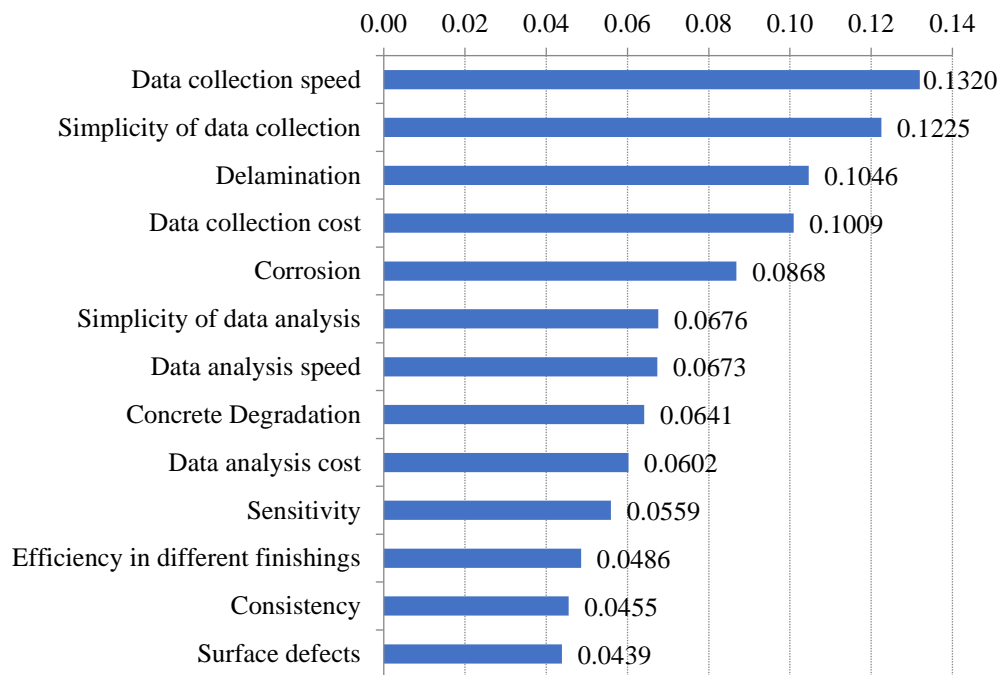


Figure 9: Importance weights of main criteria

35



36

37

Figure 10: Importance weights of sub-criteria

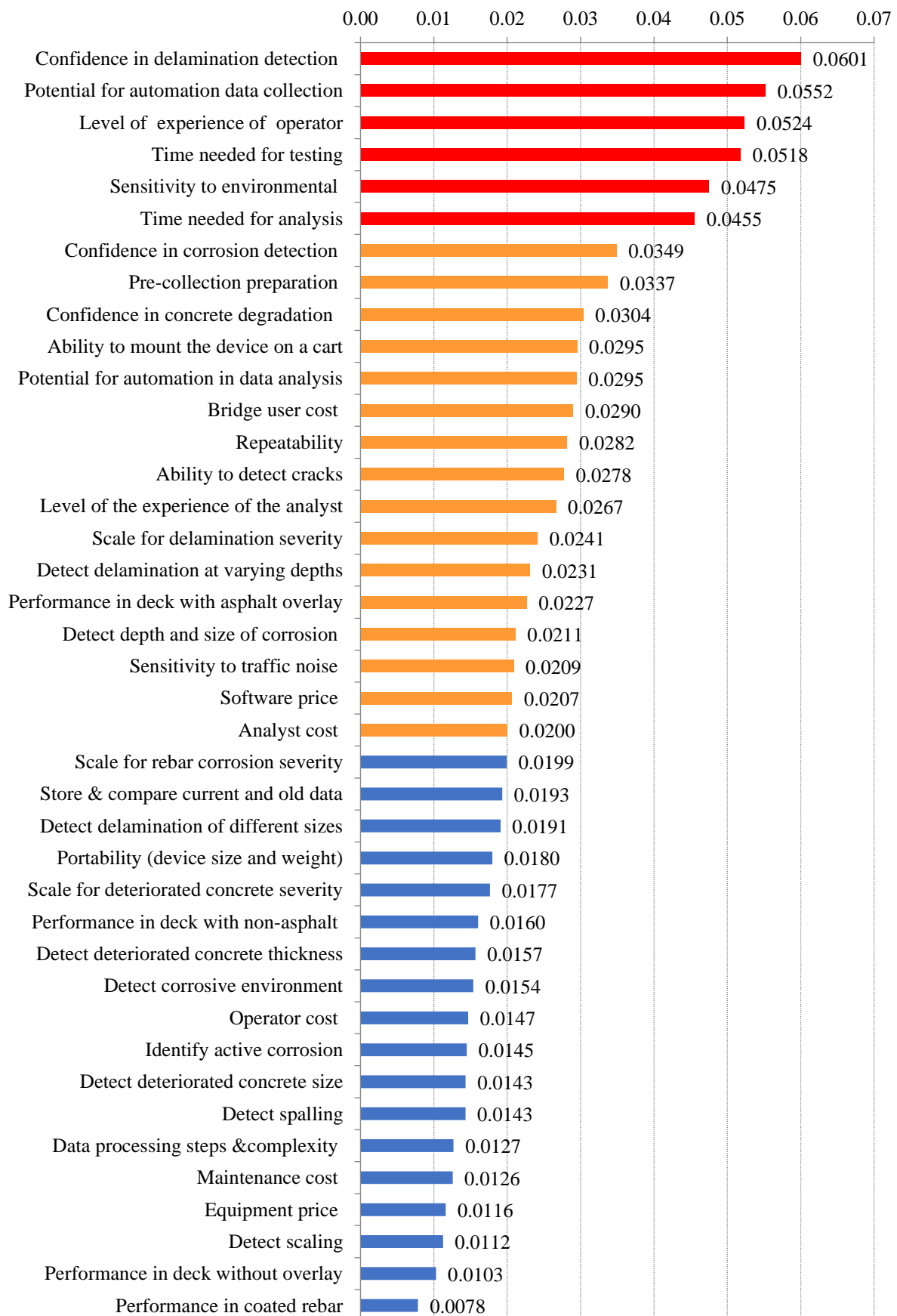


Figure 11: Importance weights of parameters

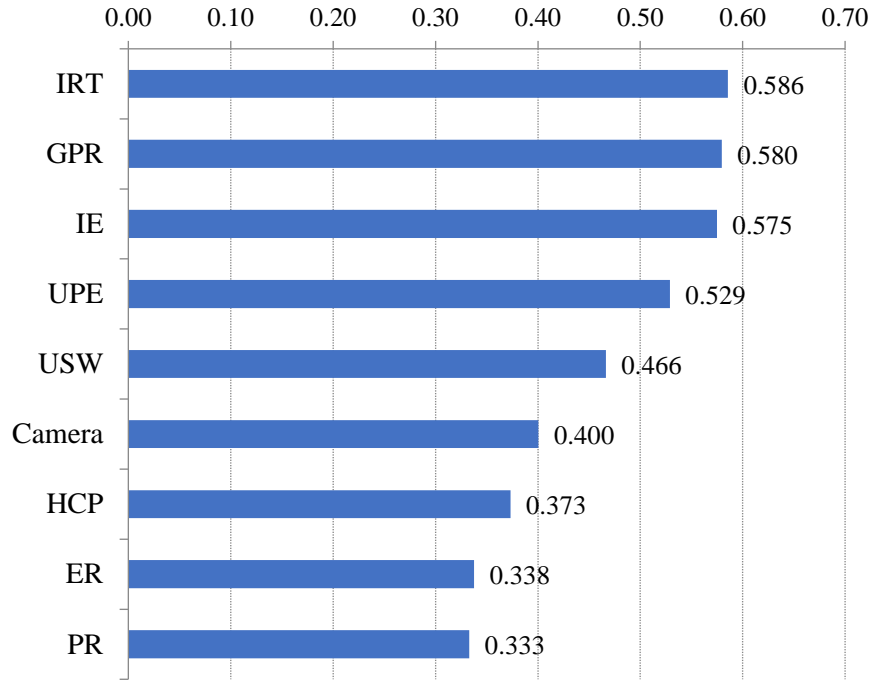


Figure 12: Overall Performance Index of non-destructive technologies

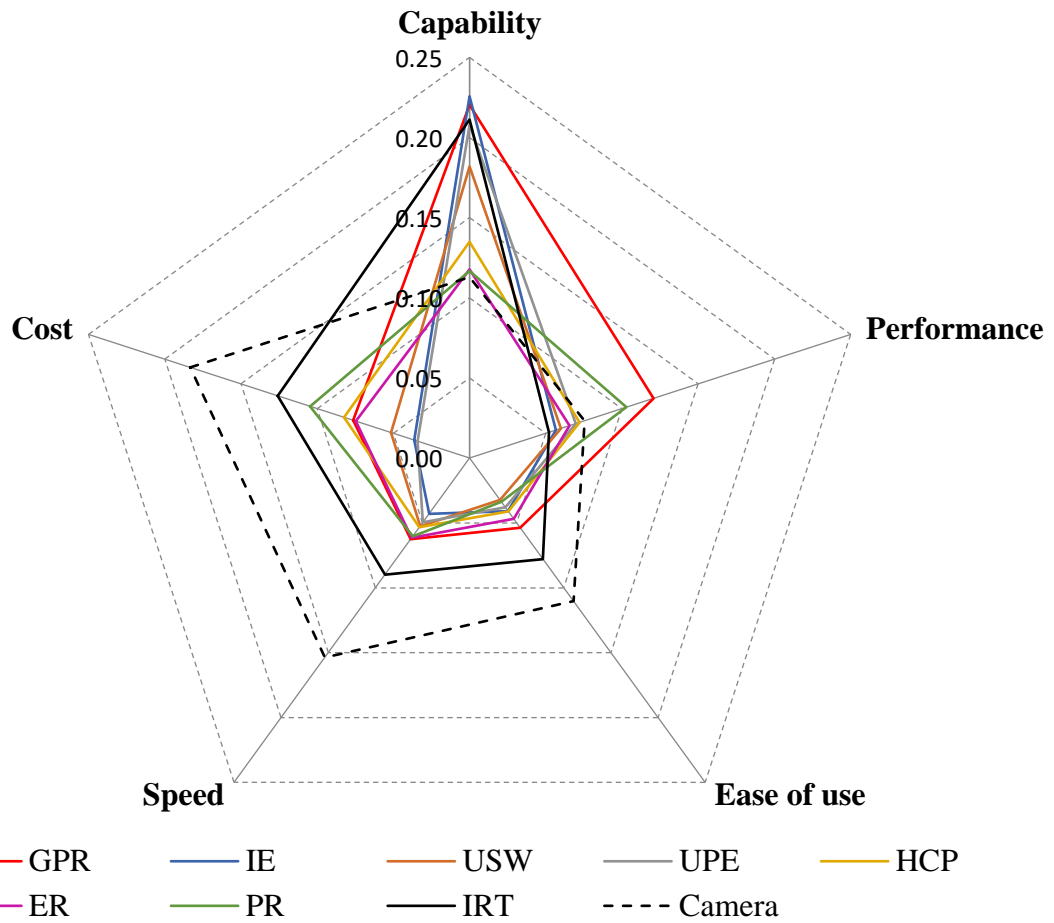


Figure 13: Performance Index of non-destructive technologies according to main criteria

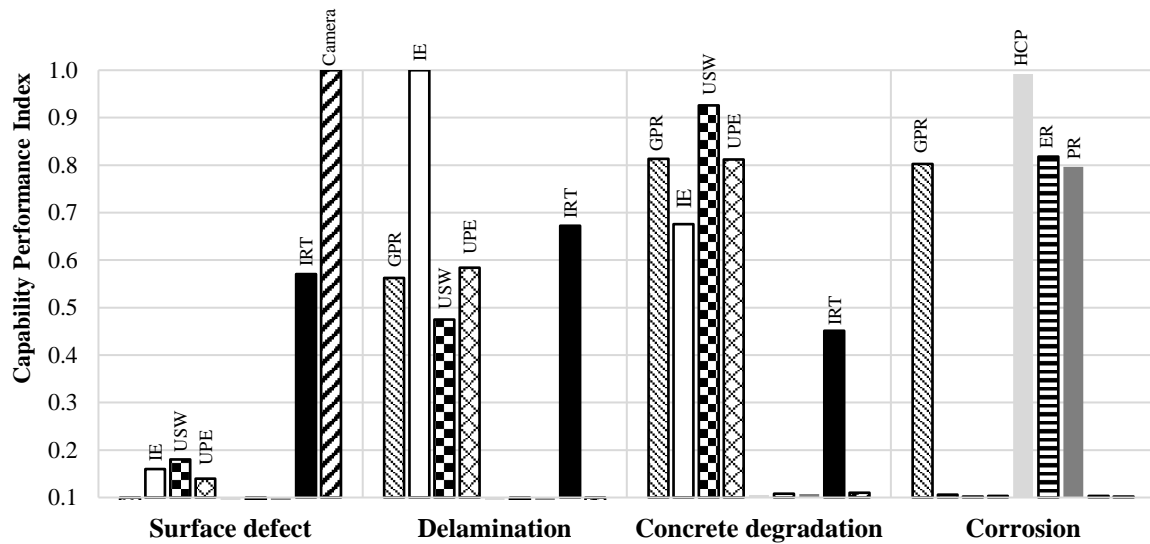


Figure 14: Performance index of non-destructive technologies according to sub-criteria in defect detection capability

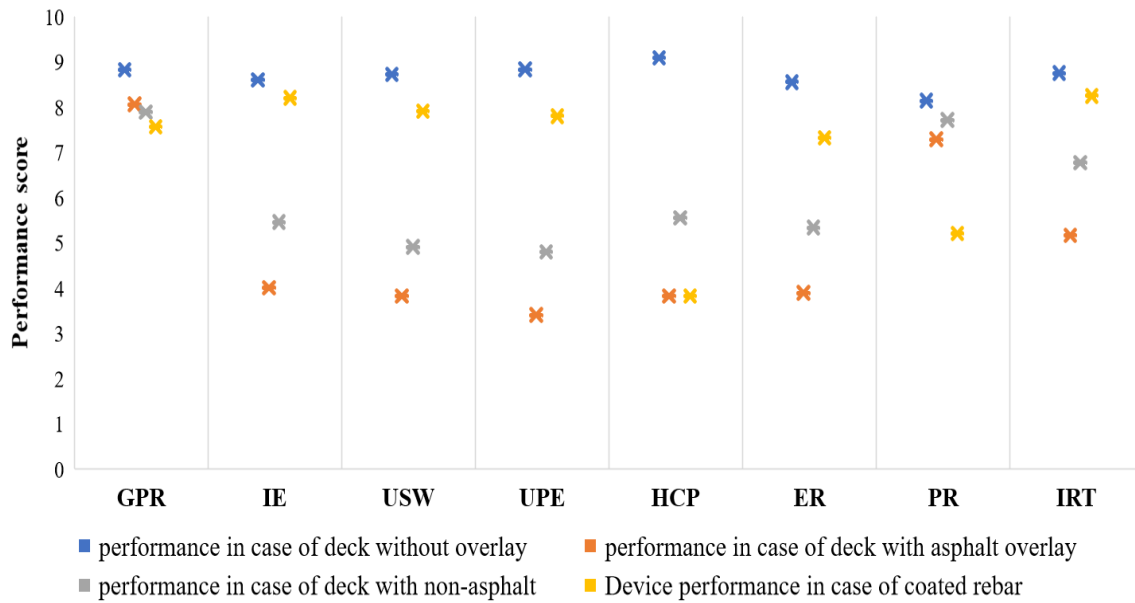


Figure 15: Impact of bridge deck finishes on the performance of NDT technologies

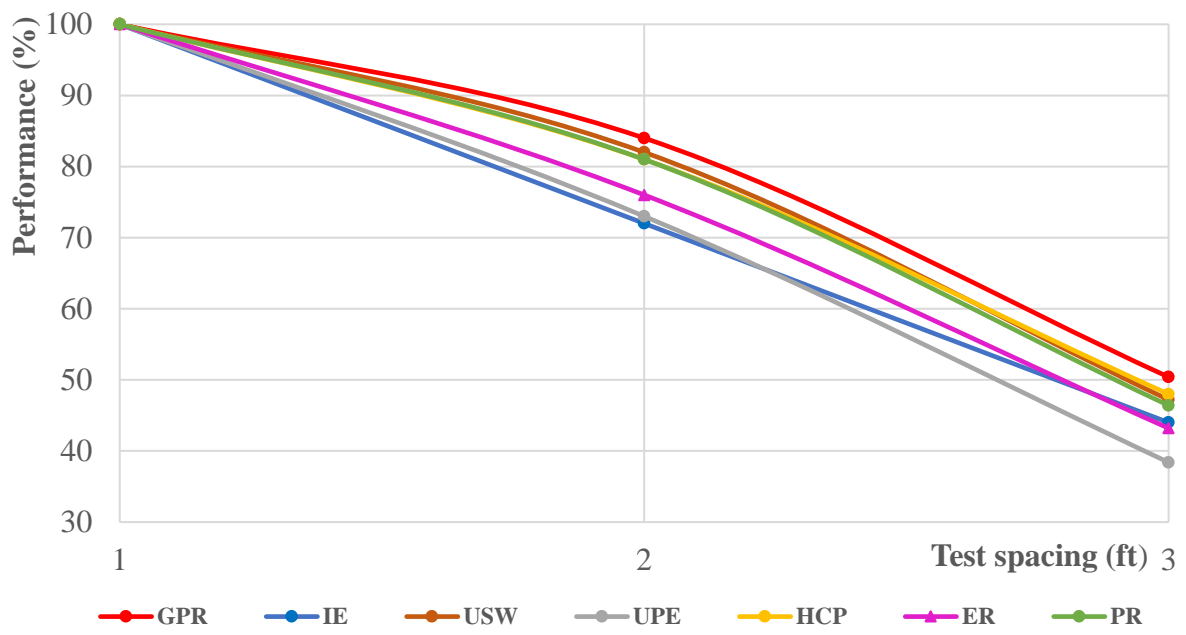


Figure 16: Impact of increasing test spacing

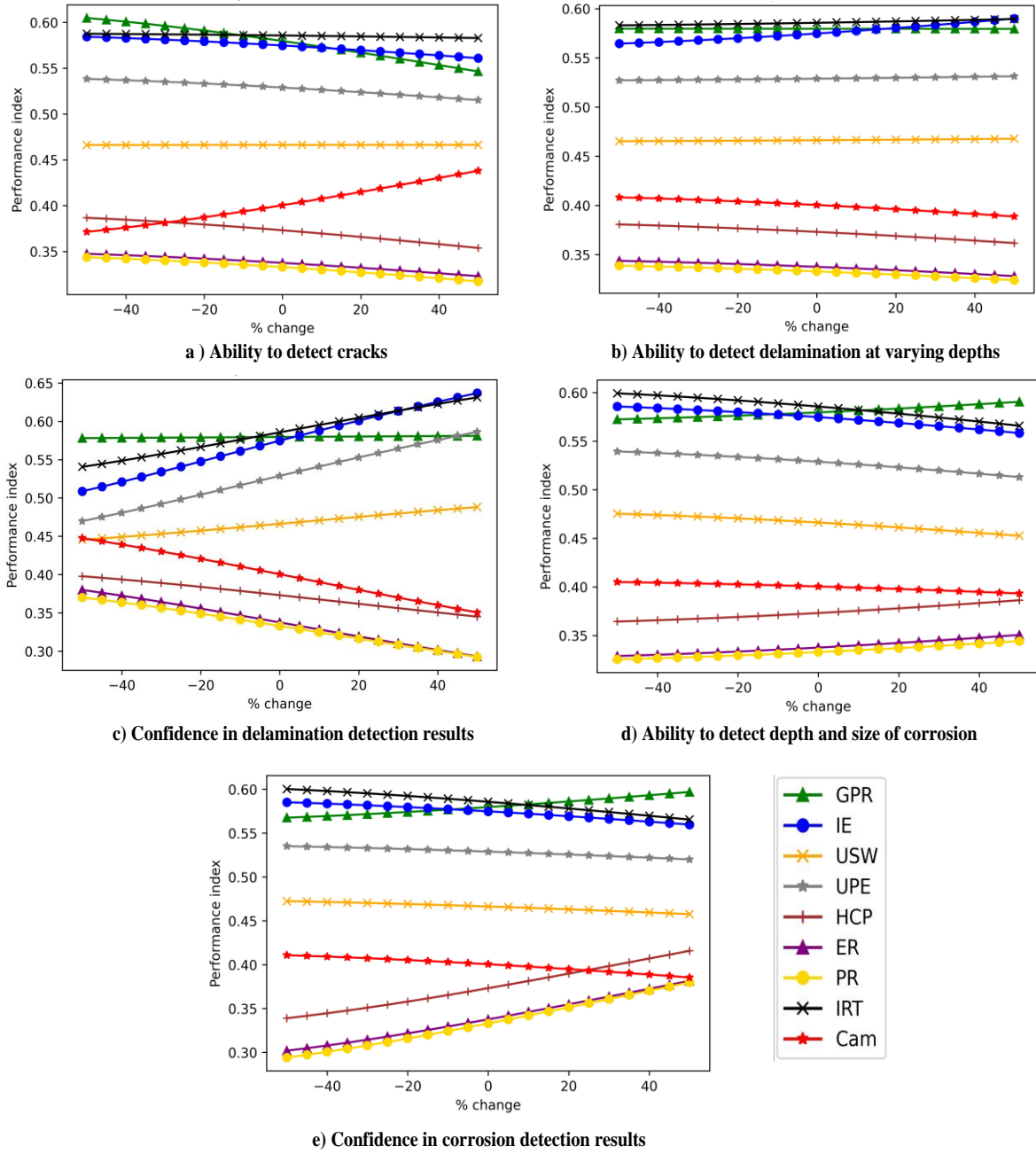


Figure 17: Sensitivity analysis results for five parameters: a) Ability to detect cracks, b) Ability to detect delamination at varying depths, c) Confidence in delamination detection results, d) Ability to detect depth and size of corrosion, e) Confidence in corrosion detection results

Performance index of NDT technologies regarding different criteria

Defect detection capability

Surface defects

Camera (1.0)
IRT (0.570)

Delamination

IE (1.0)
IRT (0.672)
UPE (0.584)
GPR (0.563)
USW (0.475)

Concrete degradation

USW (0.926)
GPR (0.813)
UPE (0.812)
IE (0.675)

Corrosion

HCP (0.992)
ER (0.818)
GPR (0.802)
PR (0.796)

Ease of use

Camera (0.983)
IRT (0.694)
GPR (0.479)
ER (0.417)
HCP (0.367)
IE (0.364)
UPE (0.338)
PR (0.30)
USW (0.288)

Speed

Camera (1.0)
IRT (0.584)
GPR (0.406)
ER (0.399)
PR (0.392)
HCP (0.346)
USW (0.341)
UPE (0.322)
IE (0.279)

Cost

Camera (0.844)
IRT (0.579)
PR (0.482)
HCP (0.379)
GPR (0.352)
ER (0.344)
USW (0.237)
IE (0.167)
UPE (0.157)

Filter 1: Select technologies for targeting defects

Surface defects

1. Camera
2. IRT

Delamination

1. IE
2. IRT
3. UPE
4. GPR
5. USW

Concrete degradation

1. USW
2. GPR
3. UPE
4. IE

Corrosion

1. HCP
2. ER
3. GPR
4. PR

If the chosen technology cannot work with this type of finishes, choose another technology

Filter 2: Consider efficiency under different finishes

Without overlay

Asphalt overlay

Non-asphalt overlay

Coated rebar

Filter 3: Calculate overall system performance index

1. Assign importance weights for different parameters

2. Evaluate NDT technologies regarding different parameters

3. Calculate defect detection capability score

4. Calculate ease of use, speed, and cost average score

5. Calculate overall system score

Figure 18: Selection model for inspection system components