

Decision-oriented resilience for sustainable infrastructure across the life cycle

Corresponding Author: Professor Stergios-Aristoteles Mitoulis

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This manuscript presents a multi-criteria decision-making methodology that integrates resilience trajectories LCA, LCC, and expected service life to evaluate six retrofit strategies for an aging prestressed concrete bridge in Greece. Although the concept is good, many issues need to be clearly discussed. I recommend a major revision.

Major:

1. The novelty and contribution are not very clear. Page 2, Lines 67–72. The authors claim that practical decision frameworks rarely integrate probabilistic recovery trajectories, ISO-compliant Life Cycle Assessment and Life Cycle Cost, and expert-informed multi-criteria decision analysis within a single methodological concept. While this may be technically true as a conjunction, each of these components has been individually well-developed, but why is an integration framework needed? I do think the expert-based parameters are essential for the resilience assessment, especially for achieving the recovery data, the authors could re-organize and strengthen the novelty of the manuscript.
2. Although I support the expert judgments, the sample size and bias are not convincing. Page 5, Lines 146–160, the entire resilience parameters and significant LCA/LCC inputs depend on expert judgement from only seven respondents. The authors (Lines 154–157) argue that depth compensates for breadth, citing consistency with “best practice in expert elicitation studies [35]”. However, Ref. [35] (Mitoulis et al., 2021) is the authors’ own work, which weakens this justification as independent evidence. With seven experts, the statistical descriptors used throughout (minimum, Q1, median, Q3, maximum) essentially correspond to individual data points rather than robust distributional estimates. The median is literally the fourth-ranked response out of seven. The authors should distribute at least more than 50 responses (since you mentioned it covers more than six countries) to make the data reliable.
3. For the resilience index, the authors adopt a resilience framework (Equation 1) originally formulated for recovery from hazard events, and reinterpret it for progressive deterioration - a conceptual extension that is, in principle, acceptable. However, the absence of any fragility or structural vulnerability modelling means that the functionality levels are not derived from structural vulnerability analysis or probabilistic damage-to-functionality mapping, but from operational assumptions and expert-based (only seven) construction schedules. As a result, the resilience index reduces to a schedule efficiency metric rather than a structural resilience metric in the sense that Equation 1 was originally intended to capture. This represents a significant simplification that should be explicitly discussed. Additionally, the framework does not account for the possibility of additional hazard events (e.g., seismic events) occurring during the multi-year intervention window, especially a consideration that is particularly relevant for a bridge located in seismically active Greece.
4. To link resilience to LCA and LCC, there is no mechanistic description of the realistic definition, but based on a performance index in Equation 11. However, the resilience assessment is stochastic, yet this variability is collapsed into a single composite score (via the $0.7 \times Q2 + 0.3 \times Q3$ weighting) before entering the MCDA. Meanwhile, the LCA and LCC are treated as fully deterministic, despite carrying their own unquantified uncertainties (e.g., EPD variability, cost estimation error, discount rate sensitivity). Consequently, the final scenario rankings are presented without confidence intervals, making it impossible to assess whether the differences between competing scenarios are statistically meaningful or fall within the uncertainty range. Is it reasonable? I am not convinced by the current version.

Minor:

1. Page 8, Min-max normalisation (Equation 9) is sensitive to outliers. For instance, a single extreme value on any criterion can compress all other alternatives into a narrow band. Given the only 6 retrofitting alternatives, this effect can be substantial. Did the authors consider alternative normalisation schemes, e.g., target-based or z-score?

2. There is a Scenario #0 visible in Fig. 7a that is not discussed anywhere in the text. What does this represent? If it is the existing (pre-intervention) condition, this should be explained. If it is an error, it should be corrected.

Reviewer #2

(Remarks to the Author)

This manuscript proposes an integrated decision-support methodology for retrofit and adaptation of ageing bridge infrastructure, combining probabilistic resilience modelling, life cycle assessment, life cycle costing, and multi-criteria decision analysis. The framework is demonstrated through application to six intervention scenarios for the Polyfytos prestressed concrete viaduct in Greece, a bridge with documented tendon corrosion and serviceability degradation. The stated contribution is a unified, policy-sensitive optimization structure that makes trade-offs between recovery speed, embodied carbon, cost, and service life extension explicit and quantifiable.

The methodological assembly is coherent and the case study is well-documented, but the manuscript does not meet the novelty and rigor standards expected for Nature Communications in its present form. The core concern is that the framework is a structured integration of existing tools (Bruneau resilience metric, ISO-compliant LCA/LCC, TOPSIS/weighted-sum MCDA) rather than a conceptual or methodological advance in any of them.

Before its acceptance, the authors must clearly articulate what is genuinely new at the methodological level, address the unresolved issues in resilience computation under TLC violation, correct the LCC scope limitation and indirect cost model, reconcile the abstract claims with the scenario-dependent results in Table 2, and provide the missing inter-expert variability analysis. Only critical inputs and assumptions must be presented in the main text.

Specific comments:

The central claim—that the methodology "bridges a critical gap" between engineering analysis and strategic infrastructure governance—requires direct comparison with existing integrated frameworks (e.g., Bocchini & Frangopol, 2012, *Computer-Aided Civil and Infrastructure Engineering*; Dong & Frangopol, 2016, *ASCE Journal of Structural Engineering*; Tapia & Padgett, 2016, already cited) to establish that the present formulation offers capabilities absent from those predecessors. The authors must identify, with technical precision, which components of the proposed pipeline are genuinely new versus assembled from established methods.

The resilience index in Equation (1) follows Bruneau et al. (2003) without modification. The authors must clarify what methodological advancement this constitutes beyond parametrizing an existing formula with expert-elicited duration distributions. If the novelty resides in the probabilistic treatment of TRE, this must be stated explicitly with the governing probability distribution, its parameters, and the fitting procedure documented in the main text rather than deferred entirely to supplementary material.

The control time $TLC = 30$ months is adopted based on empirical benchmarks from three bridge cases (Ponte San Giorgio, West Seattle Bridge, Arlington Memorial Bridge). This selection is critical because it directly governs which scenarios achieve resilience compliance. The authors must provide a formal sensitivity analysis of the ranking results to TLC, demonstrating whether the conclusions change if TLC is set to 24 or 36 months, as is plausible for different asset contexts. Table 1 shows that for scenarios #1, #2, #4, #5, and #6, recovery time exceeds TLC at Q3 and Maximum conditions. This means the majority of scenarios fail resilience compliance under upper-quartile uncertainty. The authors interpret this as a robustness concern but do not explain how the MCDA framework handles scenarios that formally violate the control-time constraint. If $TRE > TLC$, the resilience integral in Equation (1) is undefined within the specified bounds. The authors must state how R is computed or bounded in these cases, and whether scenarios violating TLC should be disqualified rather than penalized by weight.

The functionality curves in Fig. 3 appear as step functions with abrupt transitions at discrete construction phases. The authors must justify the piecewise-linear/step form of $Q(t)$ against alternative continuous degradation-recovery models, and clarify whether the area under the curve (i.e., the resilience integral) is computed analytically or numerically, and what discretization error arises from the step representation.

Direct costs in Fig. 7a are described as "cost of materials only (excluding labor and machinery)." This is an unusual and restrictive definition of LCC. Labor and machinery can represent 40–60% of total construction cost for bridge interventions of this type. The authors must either include these components or explicitly justify their exclusion and quantify the sensitivity of the LCC ranking to this omission.

The indirect cost formula (Equation 8) calculates detour-related costs as a function of transport-type distance and impact factor only. There is no vehicle operating cost model, value-of-time component, or traffic volume input. For a bridge of motorway class, indirect costs depend critically on annual average daily traffic (AADT) and detour distance. The authors must provide these input values and their sources, or acknowledge that the indirect cost estimates are incomplete proxies. The LCA system boundary excludes module B (use phase: maintenance, repair, inspection) for all scenarios. For scenarios with shorter expected service lives (e.g., Scenario #3 with localized strengthening), use-phase maintenance interventions could substantially alter total life cycle GWP. The authors must either include module B or demonstrate quantitatively that its exclusion does not reverse the environmental ranking.

The LCA weight coefficients attributed to the Joint Research Centre (JRC) for aggregation of GWP (0.46), ADPM (0.17), ADPF (0.18), and WDP (0.19) are referenced to Sala et al. (2017). The authors must confirm that these weights are drawn from the Environmental Footprint weighting set and that their application to a construction-sector LCA is methodologically appropriate. The JRC EF weights were developed for product-level footprinting and their transfer to infrastructure-level LCA requires explicit justification.

The MCDA sensitivity analysis explores five weighting schemes (T1–T5), but the inter-criteria weight ranges are not reported in the main text. The authors must provide the full weight vectors for T1–T5 explicitly. Without these, the reader cannot assess whether the ranking reversals between scenarios #3 and #6 occur under small or large weight perturbations, which is the essential test of robustness.

The UPI formulation (Equation 11) employs min–max normalization, which is sensitive to the composition of the scenario set: adding or removing a scenario changes the normalized values of all others and therefore the rankings. The authors must acknowledge this dependency explicitly and discuss whether the ranking of any scenario would change if the scenario set were extended or reduced.

The paper claims that "strategies prioritising system-level transformation with minimal structural intervention consistently outperform conventional approaches" in the abstract. Table 2 shows that under equal weighting (T1), Scenario #6 (full prefabricated replacement) ranks first for both UPI and TOPSIS, not a minimal-intervention strategy. Scenario #3 (external prestressing, minimal intervention) leads only under environmental and resilience prioritization (T3, T4). The abstract must be revised to accurately represent the conditional nature of this finding.

Expert elicitation involved seven respondents across six countries for a site-specific Greek motorway bridge. The authors must report the inter-expert variability quantitatively (e.g., coefficient of variation for recovery duration estimates by scenario) and discuss whether this sample size is sufficient to characterize the tails of the recovery time distribution, particularly given that Q3 and maximum values drive resilience non-compliance for most scenarios.

Reference 12 (Mitoulis et al., 2025) is cited as a Zenodo preprint. Nature Communications does not permit citation of non-peer-reviewed preprints in this manner unless clearly identified. The authors must replace this citation with a peer-reviewed source or explicitly flag it as a preprint and ensure the concept of "sustained resilience" it introduces is not used as a foundational premise without independent validation.

References 46 and 47 are cited in support of TOPSIS as "the most appropriate method for the problem at hand," yet Reference 46 concerns environmental impacts of building reuse and Reference 47 addresses failure assessment model evaluation—neither directly supports TOPSIS selection for MCDA of retrofit strategies. The authors must cite methodologically relevant literature on TOPSIS applicability (e.g., Behzadian et al., 2012, *Expert Systems with Applications*, 39(17), 13051–13069) and provide a reasoned justification for preferring TOPSIS over alternatives such as VIKOR or ELECTRE given the decision problem structure.

The paper does not address the transferability constraints of the framework. The six scenarios are specific to a balanced-cantilever prestressed concrete bridge with a defined structural failure mode (tendon corrosion). The methods section implies scalability to "ageing bridge networks in Europe," but no discussion is provided of what additional data, models, or calibration would be required for different bridge typologies, hazard regimes, or regulatory contexts. This limitation must be explicitly acknowledged.

Figure captions for Figs. 4b and 4c contain a numerical inconsistency: the ADP-fossil range for scenarios #3 and #6 is stated as " $4.5\text{--}6.5 \times 10^7$ MJ/km" in the caption, while the figure axis shows values in the range of $4\text{--}18 \times 10^6$ MJ/km. The authors must verify and correct all numerical values reported in figure captions against the plotted data.

Reviewer #3

(Remarks to the Author)

The manuscript is well written and has great insights in the case study. However, some questions remain:

1. Equation 8- symbols are not defined. Also why is the detour considered only for trucks using ltr?
2. What is the necessity to divide by bridge length in Equation 7?
3. The retrofit strategies are an important part of the results and all the discussions are around which strategy is the best- hence the manuscript needs to have a discussion also on what the retrofit strategies are and at the bare minimum, add the table from supplementary material to main article.
4. It is unclear where the uncertainty in the resilience quantification is coming from- is it purely from the survey results or are uncertainties assumed in some other part of the method as well?
5. If the TLC were changed, would we expect any of the rankings in Table 2 to materially change?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have successfully improved the manuscript based on the previous feedback, and I recommend its acceptance subject to a minor comment.

Regarding the response to Comment 4, while I understand the limitations regarding data availability and the necessity of conservative assumptions, this is precisely why the manuscript should explicitly document these constraints and provide clear suggestions for future development. My primary remaining concern is associated with Equation 11. Technically, it represents a way to combine different sources of input into a single unified performance index. Theoretically, however, aggregating incommensurable dimensions, specifically combining the resilience metric with life-cycle costs and environmental impacts, presents major conceptual and weighting challenges, as these metrics represent fundamentally incomparable items. I suggest the authors explicitly acknowledge this theoretical limitation within the text and indicate that alternative integration methodologies should be further explored in future work.

Reviewer #2

(Remarks to the Author)

The authors have addressed most of my previous comments satisfactorily. However, one important issue remains and requires further revision.

In the text preceding Equation 2, the authors state:

"Therefore, LCA and LCC are here performed with a deterministic approach. For LCC, temporal variabilities are considered. Specifically, costs have been adjusted to ensure temporal consistency over the observation time of the analysis using an appropriate inflation rate (2.8%)."

This statement raises two concerns.

First, the use of a fully deterministic approach for life-cycle cost analysis is difficult to justify. LCC deals with future costs, system performance, maintenance needs, interventions, and service conditions, all of which are inherently uncertain. Even when empirical data are limited, uncertainty should not simply be ignored; this is precisely where modelling assumptions, sensitivity analysis, scenario analysis, or probabilistic treatment become important. The authors should revise this part of the analysis and explain how uncertainty in future costs, performance, deterioration, maintenance, and replacement assumptions is treated. If the authors decide to retain a deterministic formulation, they should provide a clear and convincing justification and, at minimum, include a sensitivity analysis of the main cost drivers.

Second, the reference to an "inflation rate" in the context of LCC is conceptually confusing. Life-cycle cost analysis normally requires discounting future costs to present value using an appropriate discount rate or discount function. Inflation and discounting are not equivalent. Inflation adjusts nominal costs over time, whereas discounting converts future costs into present values. The authors should clarify whether Equation 2 uses nominal or real costs, whether inflation has been removed or included, and which discount rate or discount function is used to compute present values. As currently written, the description suggests a possible misunderstanding of a basic distinction in LCC modelling.

These two issues weaken confidence in the LCC component of the manuscript. I therefore recommend that the authors revise this section thoroughly, clarify the economic assumptions, and ensure that the formulation is consistent with standard LCC practice.

Finally, Equation 2 itself is difficult to interpret in its current form. The authors should rewrite the equation, define all variables clearly, distinguish between cost escalation and discounting, and explain in plain terms what the equation computes and how it is used in the analysis.

Reviewer #3

(Remarks to the Author)

A few questions/comments on the revised manuscript:

1. The added section in the revised manuscript to address my previous question on uncertainty needs more careful revision. The variable TRe is defined as stochastic variable- but the uncertainty around is not discussed. How are you modeling the variation around it? Also, from what I understood, the expert elicitation included just 7 complete responses- is that sample size enough to build uncertainty into the framework? Also, how are the mean, std deviation from these 7 responses propagated through to the final resilience and LCA/LCC targets?

2. I appreciate the authors including sensitivity in the TLC variable, however, the option identified in the main manuscript using a TLC of 30 months yields somewhat different results from either 24 or 36 months. There is a 1-2 line discussion in the manuscript on penalizing longer trajectory interventions- however conclusion is "resilience-oriented strategies remain favourable under strict recovery targets, whereas replacement-oriented solutions continue to perform comparatively well when long-term service-life extension is prioritized" - I am not sure this comes through in the sensitivity. Scenario #3 doesn't show up as particularly favorable in either case which is the scenario with maximum resilience score. What explains this?

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors have satisfactorily addressed my suggestions.

Reviewer #3

(Remarks to the Author)

The authors have satisfactorily addressed all of my comments and I recommend no more revisions.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Answers to Reviewers' comments

Reviewer #1 (Remarks to the Author):

Comment: This manuscript presents a multi-criteria decision-making methodology that integrates resilience trajectories LCA, LCC, and expected service life to evaluate six retrofit strategies for an aging prestressed concrete bridge in Greece. Although the concept is good, many issues need to be clearly discussed. I recommend a major revision.

Answer: The authors thank the Reviewer for the constructive comments and for recognising the potential of the proposed approach. We have carefully addressed all the points raised and revised the manuscript accordingly, as detailed below.

Major Comments:

Comment #1. The novelty and contribution are not very clear. Page 2, Lines 67–72. The authors claim that practical decision frameworks rarely integrate probabilistic recovery trajectories, ISO-compliant Life Cycle Assessment and Life Cycle Cost, and expert-informed multi-criteria decision analysis within a single methodological concept. While this may be technically true as a conjunction, each of these components has been individually well-developed, but why is an integration framework needed? I do think the expert-based parameters are essential for the resilience assessment, especially for achieving the recovery data, the authors could re-organize and strengthen the novelty of the manuscript.

1) Synopsis of the response

We thank the Reviewer for their comment. We agree that the individual components of the framework—probabilistic resilience assessment, ISO-compliant LCA/LCC, and MCDA—have each been previously developed in the literature. However, retrofit and adaptation decisions for ageing infrastructure inherently require balancing short-term recovery performance with long-term environmental, economic, and durability objectives. Resilience and sustainability operate on fundamentally different temporal domains: resilience captures short-term recovery dynamics and service continuity, while sustainability evaluates long-term impacts across the life cycle. In practical retrofits, however, these dimensions are inherently interconnected. Existing approaches commonly evaluate these dimensions separately, making it difficult to consistently assess trade-offs between rapid recovery, service continuity, embodied impacts, costs, and service-life extension within a single decision process.

The novelty of our paper therefore lies in establishing a unified decision-oriented framework that integrates short-term recovery dynamics, representing resilience, with long-term life-cycle performance, expressing sustainability, within a common optimization structure seamlessly, from a temporal perspective, which is the first in the international literature.

In strategy above, resilience is explicitly treated as a recovery-based decision variable, derived from functionality–time trajectories, and directly embedded into the decision process for retrofit selection. The framework incorporates multi-criteria decision-making approaches, including weighted aggregation and TOPSIS, enabling both transparent scalarization and robustness assessment of alternative solutions. Unlike existing approaches, which address resilience and sustainability separately due to temporal incompatibilities, or within hazard-specific formulations, the proposed methodology provides a consistent basis for evaluating trade-offs across recovery, environmental impact, cost, and service life. Our framework is applied on a real-world case study, to demonstrate its applicability to complex infrastructure decision-making under realistic operational constraints.

2) Detailed response

The contribution does not lie in integration per se, but in establishing a unified, decision-oriented framework in which resilience is formulated as a recovery-based decision variable associated with retrofit implementation, and explicitly linked with life-cycle environmental and economic performance (evaluated through the LCA and LCC methodologies) within a consistent decision structure. Resilience and sustainability operate on fundamentally different temporal domains: resilience captures short-term recovery dynamics and service continuity, while sustainability evaluates long-term impacts across the life cycle. In practical retrofit decision-making, however, these dimensions are inherently interconnected. Strategies that minimise recovery time may not minimise environmental impacts or maximise service life, while interventions that optimise long-term durability may require prolonged closures and increased short-term disruption. Existing approaches commonly evaluate these dimensions separately, making it difficult to consistently assess such trade-offs within a single decision process. The proposed framework addresses this gap by enabling the joint evaluation of recovery performance, environmental impact, cost, and expected service life within the same decision structure.

Existing multi-criteria decision-making (MCDM) approaches are widely used, and several integrated frameworks have been developed, particularly within the sustainability domain (e.g. Tapia & Padgett, 2016) and the structural resilience domain (e.g. Bocchini & Frangopol, 2012; Dong & Frangopol, 2016). However, these studies are generally formulated either within hazard-driven or performance-based contexts, where resilience is derived from probabilistic damage models, or within sustainability-oriented approaches focused on life-cycle environmental and economic indicators. In most cases, resilience is treated as a consequence of hazard response, rather than as a recovery-driven decision variable explicitly linked to intervention implementation processes and operational recovery trajectories. The proposed methodology differs in that it explicitly formulates resilience through functionality–time recovery trajectories under uncertainty and directly embeds these trajectories into the retrofit selection process alongside LCA, LCC, and service-life considerations. This enables explicit quantification of trade-offs between rapid recovery, service continuity, embodied impacts, economic expenditure, and long-term durability within a single optimization and decision-support framework.

Methodologically, the framework coordinates complementary MCDM approaches, including weighted aggregation and TOPSIS, enabling both transparent scalarization and robustness assessment of alternative solutions under varying stakeholder and policy priorities. The inclusion of expert-informed recovery parameters is particularly important in this context, as empirical recovery datasets for ageing and landmark infrastructure undergoing complex retrofit interventions are rarely available. For unique bridges such as the Polyfyto bridge, recovery trajectories, implementation durations, governance delays, and operational constraints are highly case-specific and cannot typically be derived directly from historical observations or standardized databases alone. Expert elicitation therefore plays a critical role in parameterizing the resilience assessment and enabling the construction of realistic functionality–time recovery trajectories under uncertainty. The need for more integrated resilience-sustainability decision frameworks has also been recognized in the literature. Ayyub (2020a,b), for instance, emphasizes that understanding the relationship between resilience and sustainability is essential for effective and cost-efficient infrastructure decision-making, while also highlighting the multidimensional nature of their quantification. More recently, Sánchez-Silva et al. (2025) further highlighted that resilience and sustainability involve potentially conflicting objectives operating across different temporal and decision domains, and emphasized the importance of systems-thinking, multicriteria life-cycle evaluation, and integrated decision-making approaches for evolving infrastructure systems. Similar priorities are reflected in international governance and policy frameworks, which increasingly advocate the joint consideration of resilience, sustainability, and economic performance in infrastructure planning and adaptation (e.g. OECD, 2025; Hallegatte et al., 2019; EU Adaptation Strategy, 2021).

References:

- Ayyub, B. M. (2020a). Disaster resilience and sustainability of infrastructures: Relationships and quantification methods. In *Gulf Conference on Sustainable Built Environment* (pp. 289-308). Cham: Springer International Publishing.
- Ayyub, B. M. (2020b). Infrastructure resilience and sustainability: Definitions and relationships. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 6(3), 02520001.
- Bocchini, P., & Frangopol, D. M. (2012). Restoration of bridge networks after an earthquake: Multicriteria intervention optimization. *Earthquake spectra*, 28(2), 427-455.
- Dong, Y., & Frangopol, D. M. (2016). Performance-based seismic assessment of conventional and base-isolated steel buildings including environmental impact and resilience. *Earthquake Engineering & Structural Dynamics*, 45(5), 739-756
- EU Adaptation Strategy (2021). <https://climate-adapt.eea.europa.eu/en/eu-adaptation-policy/strategy>
- Hallegatte, S., Rentschler, J., & Rozenberg, J. (2019). *Lifelines: The resilient infrastructure opportunity*. World Bank Publications. <https://openknowledge.worldbank.org/entities/publication/c3a753a6-2310-501b-a37e-5dcab3e96a0b>
- OECD (2025). Recommendation of the Council on the Governance of Infrastructure, OECD/LEGAL/0460
- Sánchez-Silva, M., Gardoni, P., Val, D. V., Yang, D. Y., Frangopol, D. M., Limongelli, M. P., ... & Straub, D. (2025). Moving toward resilience and sustainability in the built environment. *Structural Safety*, 113, 102449.

3) Amendments to the manuscript to address the comment

For all the amendments please refer to the revised manuscript, in the text highlighted with yellow. The authors do not provide amended text here because of the extensive revisions.

The manuscript **Abstract, Introduction and Conclusions** have been revised to better clarify the research gap, the novelty of the proposed framework, and its positioning with respect to existing literature.

In the **Abstract**, the gap statement was sharpened by specifying that current adaptation decisions often evaluate recovery dynamics, life cycle environmental impacts, economic costs, structural performance, service continuity, and service life separately. The innovation statement was also revised to clarify that the proposed methodology formulates resilience as a recovery-based variable linked to retrofit implementation and integrates it with Life Cycle Assessment, Life Cycle Cost, and multi criteria decision analysis.

In the **Introduction**, the literature review was expanded to include existing integrated decision-making frameworks in both the sustainability domain and the structural resilience domain. The gap statement was then revised to clarify that existing frameworks are generally formulated either within hazard driven or performance-based contexts, where resilience is derived from probabilistic damage and recovery models, or within sustainability-oriented approaches focused on life cycle environmental and economic indicators. Additional text was added to explain that resilience is often treated as an outcome of hazard response rather than as a recovery-based decision variable explicitly linked to retrofit implementation processes and operational recovery trajectories.

In the **Conclusions**, the contribution was refined to emphasize that the core innovation lies in operationalising resilience as a recovery based, time bounded, and uncertainty aware decision variable linked to retrofit implementation, which can be directly weighed against carbon, cost, and longevity. The conclusions were also revised to clarify the policy relevance of the framework, its ability to support context dependent prioritisation of retrofit strategies, and its role in making resilience a measurable and governable property of infrastructure systems.

Figure 1 has also been substantially revised to improve the clarity of the proposed workflow, explicitly illustrate the governance and policy inputs, distinguish the integrated assessment and modelling core, clarify the role of expert elicitation and recovery-oriented resilience modelling, and better represent the interaction between resilience, LCA, LCC, service-life assessment, and policy-sensitive MCDA optimisation.

Comment #2. Although I support the expert judgments, the sample size and bias are not convincing. Page 5, Lines 146–160, the entire resilience parameters and significant LCA/LCC inputs depend on expert judgement from only seven respondents. The authors (Lines 154–157) argue that depth compensates for breadth, citing consistency with “best practice in expert elicitation studies [35]”. However, Ref. [35] (Mitoulis et al., 2021) is the authors’ own work, which weakens this justification as independent evidence. With seven experts, the statistical descriptors used throughout (minimum, Q1, median, Q3, maximum) essentially correspond to individual data points rather than robust distributional estimates. The median is literally the fourth-ranked response out of seven. The authors should distribute at least more than 50 responses (since you mentioned it covers more than six countries) to make the data reliable.

1) Synopsis of the response

The authors thank the Reviewer for their comment. This component of the study should be interpreted as a targeted expert elicitation and consultation exercise rather than as a conventional questionnaire-based survey. The objective was not to obtain statistically representative population-level estimates, but to support the demonstration and application of the proposed methodology through high-quality, case-specific expert input. The elicitation process required substantial time commitment (approximately one day per expert to respond and follow-up communications) and highly specialized expertise, including detailed assessment of retrofit strategies, recovery trajectories, operational constraints, and implementation procedures for a landmark ageing bridge crossing a major water body. Hence, our approach was to prioritise depth, technical relevance, and contextual understanding over sample size. The resulting data are therefore intended to inform a decision-oriented methodological framework under uncertainty, rather than to provide statistically generalisable distributions. Variability and uncertainty are explicitly reported through robust descriptive statistics and detailed in Supplementary Information, allowing the expert input to be interpreted as informed expert judgement supporting the framework application.

2) Detailed response

This component of the study is intended as a targeted expert elicitation and consultation exercise rather than as a conventional questionnaire-based survey aimed at statistical inference. The objective was not to derive population-level estimates or statistically representative distributions, but to support the demonstration and application of the proposed methodology through high-quality, context-specific expert input under conditions of limited empirical data availability. More than 50 international experts were initially invited to participate. However, the elicitation process required substantial time commitment and highly specialized expertise, including detailed evaluation of retrofit strategies, recovery trajectories, implementation procedures, operational continuity constraints, and life-cycle implications (see Supplementary Information). The questionnaire therefore differed fundamentally from a short-form survey and instead functioned as a structured technical consultation exercise requiring several hours of engagement and, in some cases, follow-up clarification and supporting documentation. Consequently, only complete and sufficiently detailed responses were retained for analysis.

The expertise required for this exercise is also highly specialized and interdisciplinary, involving bridge assessment, retrofit engineering, resilience analysis, and life-cycle evaluation. Furthermore, the case study concerns a landmark ageing bridge crossing a major water body, characterised by unique structural, operational, and constructability constraints. As a result, the relevance and quality of expert input were considered more important than the size of the respondent pool. In this context, increasing the number of responses would not necessarily improve the methodological contribution or reliability of the framework application, as the purpose of the study is methodological demonstration rather than large-scale statistical generalisation.

For this reason, the expert input should be interpreted as informed expert judgement supporting decision-making under uncertainty, consistent with structured expert elicitation approaches reported in the literature

(e.g. O'Hagan et al., 2006; Aspinall & Cooke, 2011; Hemming et al., 2018). The role of the elicitation is to define plausible recovery durations, implementation uncertainties, and operational constraints for the analysed retrofit strategies, particularly where empirical recovery datasets are unavailable or insufficient.

We acknowledge the Reviewer's observation regarding the interpretation of statistical descriptors for a limited number of responses. Accordingly, the manuscript has been revised to clarify that the reported minimum, quartiles, median, and maximum values are intended as robust descriptive indicators of the spread and clustering of expert judgement, rather than as statistically representative distributional estimates. Their purpose is to transparently capture variability and uncertainty within the elicited responses and to support comparative evaluation of alternative retrofit scenarios within the proposed framework.

Finally, we acknowledge the Reviewer's point regarding the paper cited in the original submission. The revised manuscript includes additional references on structured expert elicitation practices and decision-making under uncertainty, justifying the expert elicitation approaches followed in our manuscript.

Aspinall, W P., & Cooke, R.M. (2011). Quantifying scientific uncertainty from expert judgement elicitation. In Risk and uncertainty assessment for natural hazards (pp. 64-99). Cambridge University Press.

French, S., Hanea, A.M., Bedford, T., & Nane, G. F. (2021). Introduction and overview of structured expert judgement. In *Expert Judgement in Risk and Decision Analysis* (pp. 1-16). Cham: Springer International Publishing.

Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F., & Wintle, B.C. (2018). A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution*, 9(1), 169-180.

3) Amendments to the manuscript to address the comment:

For all the amendments please refer to the revised manuscript, in the text highlighted with yellow. The authors do not provide amended text here because of the extensive revisions.

The manuscript has been revised in the **Methods, Results and Discussion, Conclusions**, and **Supporting Information**. All corresponding changes are highlighted in yellow in the revised manuscript. In the **Methods**, we reframed the data collection as a structured expert elicitation and consultation exercise, rather than a conventional questionnaire survey. We clarified that the objective was to obtain high quality, case specific input for methodological demonstration, not statistically representative population level estimates. We also clarified that more than 50 international experts were invited, but only complete and sufficiently detailed submissions were retained because the elicitation required specialised expertise and substantial technical engagement in bridge assessment, retrofit engineering, resilience analysis, operational constraints, and life cycle evaluation.

The **role of expert judgement** was clarified by explaining that empirical recovery datasets for landmark bridge retrofit interventions are rarely available, and that expert informed parameterisation is therefore needed to construct realistic functionality time recovery trajectories under uncertainty. The manuscript was revised to avoid over interpreting the elicited values. Minimum, first quartile, median, third quartile, and maximum values are now described as indicators of variation within the expert judgements, rather than as statistically representative distributional estimates. Additional independent references on structured expert elicitation and decision making under uncertainty were added to support the methodological approach and reduce reliance on self-citation.

In the **Results and Discussion**, the interpretation of the recovery trajectories was justified and linked explicitly to the elicited expert judgement and the analysed case study conditions. In **Conclusions**, we clarified that the framework is intended as a methodological decision support tool under uncertainty, rather than a statistically generalisable predictive model.

Comment #3. For the resilience index, the authors adopt a resilience framework (Equation 1) originally formulated for recovery from hazard events, and reinterpret it for progressive deterioration - a conceptual extension that is, in principle, acceptable. However, the absence of any fragility or structural vulnerability modelling means that the functionality levels are not derived from structural vulnerability analysis or probabilistic damage-to-functionality mapping, but from operational assumptions and expert-based (only seven) construction schedules. As a result, the resilience index reduces to a schedule efficiency metric rather than a structural resilience metric in the sense that Equation 1 was originally intended to capture. This represents a significant simplification that should be explicitly discussed. Additionally, the framework does not account for the possibility of additional hazard events (e.g., seismic events) occurring during the multi-year intervention window, especially a consideration that is particularly relevant for a bridge located in seismically active Greece.

1) Synopsis of the response

We agree with the Reviewer that the proposed application represents a conceptual extension of classical hazard-based resilience formulations. In this study, resilience is interpreted within a decision-oriented framework focused on recovery associated with progressive deterioration and planned retrofit interventions rather than on probabilistic hazard-induced damage process. Accordingly, the functionality–time trajectories are derived from intervention sequencing, operational continuity, implementation constraints, and expert-informed recovery estimates reflecting realistic retrofit conditions for the analysed landmark bridge. The objective of the framework is therefore not to perform fragility or vulnerability analysis for acute hazard events (e.g. earthquakes), but to support comparative decision-making across retrofit strategies by integrating recovery dynamics with environmental, economic, and service-life considerations for chronic stressors. Within this context, the resilience metric captures not only intervention duration, but also functionality loss and restoration, partial service continuity, sequencing of recovery phases, and uncertainty in implementation. Additional hazard events during the intervention phase are not considered in the present study, as the analysis intentionally isolates the performance of the retrofit strategies themselves and does not include concurrent external hazard occurrence during the intervention period. The above assumptions have been discussed in the revised version of the manuscript.

2) Detailed response

We agree with the Reviewer that Equation (1) originates from hazard-based resilience formulations and that the present application represents a conceptual extension from acute hazard recovery to recovery associated with progressive deterioration and planned intervention processes. In this study, resilience is interpreted within a decision-oriented context, where the focus is on the ability of infrastructure systems to maintain and recover functionality during and after retrofit implementation, rather than on probabilistic damage evolution resulting from external hazards.

Accordingly, the objective of the study is not to model structural vulnerability or derive damage-to-functionality relationships, but to support comparative decision-making across alternative retrofit strategies, by integrating recovery dynamics with environmental, economic and service-life consideration. The functionality–time trajectories are therefore constructed based on intervention sequencing, operational constraints, and expert-informed recovery estimates, reflecting realistic retrofit conditions for the analysed bridge.

In this context, it is important to distinguish between acute hazard-driven (e.g. seismic events) resilience and resilience associated with long-term infrastructure adaptation under chronic deterioration processes (e.g. fatigue, environmental degradation, corrosion). The latter is governed primarily by intervention planning, constructability, operational continuity, and recovery management rather than by immediate hazard-induced damage states. The proposed framework is intended to address this second class of problems.

We acknowledge the Reviewer's observation that the resulting resilience metric differs from classical hazard-based resilience formulations relying on fragility analysis. However, within the present framework, the resilience index does not reduce to a simple schedule-efficiency metric. Beyond intervention duration, it explicitly incorporates functionality loss and restoration trajectories, partial service continuity, sequencing of recovery phases, uncertainty in implementation duration, and the likelihood of exceeding recovery targets. These aspects are directly relevant for infrastructure adaptation decisions under progressive deterioration conditions, where maintaining operational functionality during intervention is itself a critical resilience objective.

The inclusion of fragility or vulnerability models would correspond to a different modelling framework focused on hazard-induced damage simulation and cascading failure analysis. While such extensions are important, they would introduce fundamentally different modelling assumptions and objectives from those considered in the present decision-oriented adaptation framework.

Regarding the possibility of additional hazard events during the intervention period (e.g. seismic events), we agree that this is an important aspect, particularly for infrastructure located in seismically active regions such as Greece. In the present work, additional hazard occurrence during the intervention period is not considered, as the analysis is intended to isolate and compare the retrofit strategies themselves rather than to couple retrofit decision-making with stochastic multi-hazard simulation. Future developments of the framework could incorporate time-dependent hazard occurrence and multi-hazard interaction during intervention periods, enabling assessment of resilience under concurrent deterioration and external hazard scenarios.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised in the **Introduction, Methods, Results and Discussion, and Conclusions**. All corresponding changes are highlighted in yellow in the revised manuscript.

In the **Introduction**, we clarified that the proposed framework addresses **adaptation and retrofit decision making under progressive deterioration**, rather than classical hazard driven resilience based on fragility or vulnerability modelling. In the **Methods**, before the resilience formulation, we added text to explicitly distinguish between hazard driven resilience and deterioration driven, recovery-based resilience. We clarified that the functionality time trajectories are not derived from fragility analysis, but from intervention sequencing, operational continuity assumptions, implementation constraints, and expert informed recovery estimates. We also clarified that, within the proposed framework, the resilience index is not intended as a full structural vulnerability metric. Rather, it captures recovery performance during retrofit implementation, including functionality loss and restoration, partial service continuity, sequencing of intervention phases, implementation uncertainty, and compliance with recovery targets.

In the **Results and Discussion**, the terminology was further clarified to avoid implying comprehensive hazard based structural resilience modelling. We now refer more explicitly to recovery oriented and decision-oriented resilience assessment for retrofit and adaptation planning. In **Conclusions**, we added a limitation and future development statement clarifying that the present framework does not explicitly model additional hazard events during the intervention period, nor hazard specific fragility or vulnerability relationships. We identify this as a future extension, particularly for multi hazard settings and infrastructure located in seismic regions.

Comment #4. To link resilience to LCA and LCC, there is no mechanistic description of the realistic definition, but based on a performance index in Equation 11. However, the resilience assessment is stochastic, yet this variability is collapsed into a single composite score (via the $0.7 \times Q2 + 0.3 \times Q3$ weighting) before entering the

MCDA. Meanwhile, the LCA and LCC are treated as fully deterministic, despite carrying their own unquantified uncertainties (e.g., EPD variability, cost estimation error, discount rate sensitivity). Consequently, the final scenario rankings are presented without confidence intervals, making it impossible to assess whether the differences between competing scenarios are statistically meaningful or fall within the uncertainty range. Is it reasonable? I am not convinced by the current version.

1) Synopsis of the response

We acknowledge the Reviewer's observation regarding the different treatment of uncertainty across resilience, LCA, and LCC components. In the present study, LCA and LCC are evaluated through a deterministic, scenario-based approach consistent with standard engineering practice and with the type and availability of data for the specific geographical and temporal context of the case study. In particular, the available environmental and cost datasets do not provide sufficient information to support a robust and context-consistent probabilistic uncertainty propagation. By contrast, resilience explicitly incorporates variability because recovery trajectories are directly derived from expert-informed implementation estimates and exhibit significant dispersion, which is represented through descriptive statistical indicators. Within the MCDA framework, all criteria are ultimately synthesized into normalized representative indicators in order to enable consistent comparison across alternatives. Accordingly, the objective of the framework is not to produce probabilistic rankings or statistically significant differences between scenarios, but to support comparative decision-making under a transparent and coherent set of assumptions.

2) Detailed response

We would like to clarify that LCA and LCC are not treated as probabilistic variables in this study, but are evaluated using a deterministic, scenario-based approach, consistent with data availability and standard practice. In particular, LCA results are derived from Environmental Product Declarations (EPDs) selected based on geographical and temporal representativeness. Where multiple datasets exist, conservative assumptions (i.e. higher-impact values) are adopted. This did not occur for most of the construction materials considered in the designed rehabilitation options. Also, for some construction works (e.g. dismantling and transport to waste processing site), information relevant for LCA and LCC analyses was derived through literature and experts' consultation. Literature did not provide confidence intervals to perform an uncertainty analysis. Regarding LCC, cost estimates are based on available statistics for the specific regional and temporal context, which are limited in number and do not support a statistically meaningful uncertainty analysis. All costs are discounted using a rate of 2.8% (see Supplementary information, p.17), reflecting the average base rate in Greece [24, 25] over the last two years. The authors decided also to not admit carbon (or environmental impact) discounting. The authors agree with literature deeming environmental impact discounting fundamentally an ethical stance. Carbon (or environmental impact) discount rates can reduce in fact the present value of future climate issues (climate change, acidification, etc.). If environmental damage is discounted at a high rate, it might appear almost negligible, ignoring the potential for catastrophic, non-linear tipping points. Also, because emissions persist in the biosphere for years (e.g. CO₂ persists in the atmosphere for centuries), the true cost of emissions is felt long after the economic benefit of producing them has passed. Finally, unlike financial investments, environmental destruction such as biodiversity loss or ecosystem collapse cannot be "reversed" by future technological innovation, making it unethical to discount these impacts.

As a result, while uncertainties are inherently present (e.g. due to variability in EPD datasets and cost assumptions), the available information does not allow for a robust and context-consistent propagation of uncertainty. Performing a formal uncertainty analysis would therefore require the use of additional data that are not representative of the specific geographical or temporal context of the case study. In contrast, the resilience assessment explicitly incorporates variability, as recovery times are directly elicited from experts

and exhibit significant dispersion. This justifies the use of statistical descriptors (e.g. Q2 and Q3) to represent uncertainty in recovery dynamics.

It is also important to note that, within the MCDA framework, all criteria are ultimately represented by single normalized values, in order to enable consistent comparison across alternatives. Therefore, even when variability is initially characterized (e.g. for resilience), it is intentionally synthesized into a representative indicator prior to aggregation. The MCDA is thus intended as a decision-support tool, rather than a probabilistic ranking method. The aim is not to assess statistical significance between alternatives, but to identify consistent trends and relative performance across scenarios under a common set of assumptions. A full uncertainty propagation (e.g. confidence intervals on rankings) would require a fundamentally different data structure and modelling approach and is therefore beyond the scope of the present decision-oriented framework.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised to clarify the assumptions underlying the treatment of LCA, LCC, and resilience, and to explicitly state the limitations related to data availability and uncertainty representation. Specifically, in section “Integration of sustainability”, we clarify why uncertainties were not propagated in LCA and LCC analyses. Also, we mentioned explicitly that we considered costs variabilities and increases with an appropriate inflation rate.

“Because of a lack of context-consistent data, it was not possible to perform a formal uncertainty analysis. Therefore, LCA and LCC are here performed with a deterministic approach. For LCC, temporal variabilities are considered. Specifically, costs have been adjusted to ensure temporal consistency over the observation time of the analysis using an appropriate inflation rate (2.8%)”.

Minor comments:

Comment #5. Page 8, Min-max normalisation (Equation 9) is sensitive to outliers. For instance, a single extreme value on any criterion can compress all other alternatives into a narrow band. Given the only 6 retrofitting alternatives, this effect can be substantial. Did the authors consider alternative normalisation schemes, e.g., target-based or z-score?

1) Answer: The authors thank the Reviewer for this comment. Alternative normalisation schemes such as z-score or target-based normalization were considered during the development of the framework. However, these approaches were not considered appropriate for the present study. Z-score normalization relies on statistically meaningful estimates of mean and standard deviation, which are not robust for a limited number of alternatives, and may introduce instability and reduced interpretability in the MCDA context. Moreover, the objective of the present framework is comparative decision support rather than statistical inference, making distribution-based normalization less suitable.

Similarly, target-based normalization requires the definition of reference or benchmark values for each criterion. In this study, consistent and non-arbitrary target values are not available across the resilience, environmental, economic, and service-life indicators considered. Introducing such targets would therefore require additional assumptions that are not sufficiently supported by the available data. In contrast, min–max normalization was selected because it preserves the relative ordering and proportional separation among alternatives while maintaining transparency and interpretability within the decision-oriented framework adopted in this study. Furthermore, no extreme outlier behaviour was observed among the analysed retrofit alternatives that would lead to severe compression effects within the normalized criteria space.

2) Amendments to the manuscript to address the comment:

A clarification has been added in the revised manuscript at Equation (9):

“Min–max normalization was adopted to ensure transparent and consistent relative comparison among retrofit alternatives within the MCDA framework. Alternative normalization schemes, including z-score and target-based normalization, were considered; however, they were not deemed appropriate for the present application. Z-score normalization relies on statistically robust estimates of mean and standard deviation, which are not reliable for the limited number of analysed alternatives, while target-based normalization would require the introduction of external benchmark values that are not consistently available across the considered resilience, environmental, economic, and service-life criteria. Furthermore, no extreme outlier behaviour was observed among the analysed alternatives that would lead to severe compression effects within the normalized criteria space.”

Comment #6. There is a Scenario #0 visible in Fig. 7a that is not discussed anywhere in the text. What does this represent? If it is the existing (pre-intervention) condition, this should be explained. If it is an error, it should be corrected.

Answer: The authors thank the Reviewer for this comment. Scenario #0 was included in error from a preliminary version and has now been removed from the figure.

Reviewer #2 (Remarks to the Author):

This manuscript proposes an integrated decision-support methodology for retrofit and adaptation of ageing bridge infrastructure, combining probabilistic resilience modelling, life cycle assessment, life cycle costing, and multi-criteria decision analysis. The framework is demonstrated through application to six intervention scenarios for the Polyfyto prestressed concrete viaduct in Greece, a bridge with documented tendon corrosion and serviceability degradation. The stated contribution is a unified, policy-sensitive optimization structure that makes trade-offs between recovery speed, embodied carbon, cost, and service life extension explicit and quantifiable.

Comment #1. The methodological assembly is coherent and the case study is well-documented, but the manuscript does not meet the novelty and rigor standards expected for Nature Communications in its present form. The core concern is that the framework is a structured integration of existing tools (Bruneau resilience metric, ISO-compliant LCA/LCC, TOPSIS/weighted-sum MCDA) rather than a conceptual or methodological advance in any of them.

Before its acceptance, the authors must clearly articulate what is genuinely new at the methodological level, address the unresolved issues in resilience computation under TLC violation, correct the LCC scope limitation and indirect cost model, reconcile the abstract claims with the scenario-dependent results in Table 2, and provide the missing inter-expert variability analysis. Only critical inputs and assumptions must be presented in the main text.

1) Synopsis of the response

We agree that the novelty of the work needed to be articulated more clearly and that several methodological assumptions required clarification. Current adaptation decisions remain fragmented because resilience and sustainability operate on fundamentally different temporal domains: resilience captures short-term recovery dynamics and service continuity, while sustainability evaluates long-term impacts across the life cycle. In this paper, we depart from this temporally disjointed approach of resilience and sustainability. Instead, of analyzing the tradeoffs between resilience and sustainability metrics we establish a unified decision-oriented framework that integrates short-term recovery dynamics, representing resilience, with long-term life-cycle performance, expressing sustainability, within a common optimization structure seamlessly, from a temporal perspective, which is the first in the international literature. The manuscript has therefore been substantially revised to better define the contribution as a decision-oriented framework that operationalises resilience as a recovery-based variable linked to retrofit implementation, and integrates it with LCA, LCC, service-life assessment, and MCDA within a single retrofit decision structure.

We have also addressed the specific methodological concerns raised by the Reviewer by clarifying the treatment of T_{LC} exceedance, revising the LCC scope and indirect cost assumptions, aligning the Abstract with the scenario-dependent results in Table 2, strengthening the presentation of inter-expert variability, and streamlining the main text so that only the essential methodological inputs and assumptions are retained. These revisions improve the clarity, rigour, and positioning of the manuscript while making the scope and limitations of the proposed framework more explicit.

2) Detailed response: The authors thank the Reviewer for the detailed and constructive assessment. We have carefully addressed all the points raised and substantially revised the manuscript accordingly.

In particular, we have clarified the methodological contribution and positioning of the work more explicitly in the Introduction, Methods and Discussion sections. We agree that the individual methodological components employed in the study—including resilience formulations, ISO-compliant LCA/LCC approaches, and MCDA techniques—are themselves established in the literature. The contribution of the present work

does not lie in proposing a new standalone resilience metric, environmental assessment method, or optimisation algorithm, but in establishing a unified decision-oriented formulation that explicitly integrates recovery-based resilience with long-term environmental, economic and service-life performance within a common retrofit decision structure.

More specifically, the framework formulates resilience as a recovery-oriented decision variable derived from functionality–time trajectories associated with intervention implementation and operational continuity under progressive deterioration conditions. This departs from conventional hazard-driven resilience approaches based on fragility, vulnerability, and probabilistic damage from external events. Instead, it focuses on infrastructure adaptation and retrofit decision-making, where the key challenge is to evaluate trade-offs between recovery performance, environmental impact, cost, operational disruption, and long-term durability within a consistent comparative framework. This distinction is important because the decision problem addressed in the present study is not purely technical or hazard-driven. Retrofit strategies must be evaluated under multiple and potentially conflicting constraints, including recovery performance, environmental impact, economic cost, operational disruption, and service life. Accordingly, the proposed framework is intended as a decision-support methodology under uncertainty rather than as a probabilistic predictive model. The objective is therefore not statistical inference or probabilistic ranking, but transparent comparative evaluation of retrofit alternatives under a coherent set of assumptions and stakeholder priorities.

In response to the Reviewer’s additional comments, we have further revised the manuscript to:

- clarify the interpretation of resilience under TLC exceedance conditions;
- improve the discussion of LCC scope and indirect cost assumptions;
- reconcile the abstract statements with the scenario-dependent ranking results;
- strengthen the presentation and interpretation of inter-expert variability;
- and streamline the manuscript so that only the most critical assumptions and methodological inputs are retained in the main text, with supplementary details appropriately transferred to the Supporting Information.

We believe these revisions substantially improve the clarity, rigor, transparency, and positioning of the manuscript.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised throughout the **Abstract, Introduction, Methods, Results and Discussion, Conclusions, and Supporting Information**. For all the amendments please refer to the revised manuscript, in the text highlighted with yellow. The authors do not provide amended text here because of the extensive revisions.

In the **Abstract**, we clarified the research gap and revised the novelty statement to specify that the framework treats resilience as a recovery-based variable linked to retrofit implementation and integrates it with LCA, LCC, and MCDA. In the **Introduction**, we expanded the positioning of the work with respect to existing sustainability and structural resilience decision frameworks, and clarified that the contribution lies in linking recovery-based resilience, life cycle environmental and economic performance, and service life within a single retrofit decision structure. In the **Methods**, we clarified the methodological contribution, the treatment of TLC exceedance, and the role of TLC as a policy relevant threshold rather than a strict exclusion criterion. Figure 1 has also been substantially revised to improve the clarity of the proposed workflow, explicitly illustrate the governance and policy inputs, distinguish the integrated assessment and modelling core, clarify the role of expert elicitation and recovery-oriented resilience modelling, and better represent the interaction between resilience, LCA, LCC, service-life assessment, and policy-sensitive MCDA optimisation

In the **Integration of sustainability** subsection, we revised the LCA and LCC scope, clarified the functional unit, observation period, system boundaries, replacement stage, and deterministic treatment of LCA/LCC inputs. The LCC scope was corrected to include direct intervention and life cycle costs together with indirect user costs associated with traffic disruption and detours. In the **Results and Discussion**, we strengthened the presentation of inter expert variability by reporting recovery statistics through minimum, quartiles, median, maximum values, and compliance characteristics. The interpretation of the ranking results is also being aligned with the scenario dependent outcomes in Table 2. In the **Conclusions**, we revised the contribution statement to emphasize that the core innovation is the operationalisation of resilience as a recovery based, time bounded, uncertainty aware decision variable linked to retrofit implementation and weighed against carbon, cost, and longevity.

Specific comments:

Comment #2. The central claim—that the methodology "bridges a critical gap" between engineering analysis and strategic infrastructure governance—requires direct comparison with existing integrated frameworks (e.g., Bocchini & Frangopol, 2012, *Computer-Aided Civil and Infrastructure Engineering*; Dong & Frangopol, 2016, *ASCE Journal of Structural Engineering*; Tapia & Padgett, 2016, already cited) to establish that the present formulation offers capabilities absent from those predecessors. The authors must identify, with technical precision, which components of the proposed pipeline are genuinely new versus assembled from established methods.

1) Synopsis of the response

The authors thank the Reviewer for this constructive comment. We agree that the claim of bridging engineering analysis and strategic infrastructure governance required clearer positioning relative to existing integrated frameworks, and the manuscript has been revised accordingly. In particular, proposed methodology is now compared more explicitly with prior resilience, sustainability, and life cycle decision frameworks, including Bocchini and Frangopol (2012), Dong and Frangopol (2016), and Tapia and Padgett (2016).

We clarify that the novelty of the present study does not lie in introducing new standalone metrics for resilience, LCA, LCC, or MCDA, but in establishing a unified decision-oriented retrofit selection framework in which recovery based resilience is explicitly formulated as a decision variable and evaluated jointly with environmental, economic, operational and service life performance. Unlike existing hazard-driven frameworks based primarily on fragility, vulnerability, and probabilistic damage modelling, the proposed methodology focuses on intervention-driven recovery under progressive deterioration conditions associated with ageing infrastructure adaptation. The framework is therefore intended as a transparent decision-support methodology under uncertainty, enabling comparative evaluation of retrofit alternatives under multiple and potentially conflicting constraints, including recovery performance, operational disruption, environmental impact, cost, and long-term durability.

2) Detailed response

We agree that several integrated frameworks already exist in the literature, and we have revised the manuscript to provide a clearer and more technically explicit comparison with these approaches.

The methodological contribution of the present study does not lie in proposing a new standalone resilience metric, LCA or LCC methodology, or MCDA algorithm in isolation. Rather, the novelty lies in introducing a unified decision-oriented formulation for retrofit selection in which recovery-based resilience is explicitly incorporated as a decision variable, and is evaluated jointly with life-cycle environmental, economic and service-life performance within a single consistent framework. More specifically, the framework treats

resilience through functionality–time recovery trajectories associated with planned interventions and operational continuity under progressive deterioration conditions. This differs from conventional hazard-driven approaches, where resilience is typically derived from fragility relationships, probabilistic damage states, and post-event recovery following acute external hazards. In the present study, the focus is instead on intervention-driven recovery associated with ageing infrastructure adaptation and retrofit implementation. This distinction is important because the decision problem addressed here is not limited to structural performance under external hazards. Retrofit strategies must be evaluated under multiple and potentially conflicting constraints, including recovery performance, environmental impact, operational disruption, economic cost, and long-term service life. Accordingly, the proposed framework is intended as a decision-support methodology under uncertainty rather than as a probabilistic predictive model. The objective is therefore not statistical inference or probabilistic ranking, but transparent comparative evaluation of retrofit alternatives under a coherent set of assumptions and stakeholder priorities.

With respect to the works cited by the Reviewer, the methodological distinctions are as follows.

In Bocchini and Frangopol (2012), the focus is on post-disaster restoration of bridge networks under seismic hazard, where resilience is formulated within a hazard-driven context and network-oriented optimisation framework concerned with restoration timing and resource allocation following an acute event. The framework does not address life-cycle environmental impacts or deterioration-driven retrofit decision-making.

Dong and Frangopol (2016), intergate resilience and sustainability within a performance-based seismic framework, using fragility functions and probabilistic damage modelling. Here, resilience remains hazard-driven and damage-based, and the methodology is primarily oriented toward structural performance assessment under seismic scenarios rather than intervention-driven retrofit selection.

Tapia and Padgett (2016) proposed a multi-objective optimisation framework for retrofit and repair decisions under natural hazards, incorporating life-cycle sustainability indicators. However, resilience remains represented through hazard-related damage and loss metrics derived from probabilistic fragility modelling, rather than recovery trajectories associated with planned intervention implementation and operational continuity.

Accordingly, the present work advances beyond these studies in three methodological respects:

- (i) resilience is formulated as a recovery-oriented decision (hazard-agnostic approach) variable associated with retrofit implementation, rather than hazard-induced damage.
- (ii) short-term recovery dynamics are explicitly linked with long-term environmental, economic and service-life performance.
- (iii) these dimensions are integrated within a unified decision-support framework for comparing retrofit strategies under operational and policy-relevant constraints.

Furthermore, the need for more integrated resilience–sustainability decision frameworks have been explicitly recognised in the literature, which emphasises both the multidimensional nature of these concepts and the importance of considering them jointly in infrastructure planning, maintenance, adaptation and upgrading (Ayyub, 2020a,b; Faber et al., 2020). More recently, Sánchez-Silva et al. (2025) further highlighted that resilience and sustainability involve potentially conflicting objectives operating across different temporal and decision domains, reinforcing the need for integrated and decision-oriented evaluation frameworks for infrastructure systems. Hence, the present study contributes a more explicit and operational decision

structure for retrofit selection. The manuscript has been revised accordingly to clarify these distinctions more explicitly, identify which methodological components are adopted from prior work, and better articulate the specific contribution of the present study.

References:

- Ayyub, B.M. (2020a). Disaster resilience and sustainability of infrastructures: Relationships and quantification methods. In *Gulf Conference on Sustainable Built Environment* (pp. 289-308). Cham: Springer International Publishing.
- Ayyub, B.M. (2020b). Infrastructure resilience and sustainability: Definitions and relationships. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 6(3), 02520001.
- Bocchini, P., & Frangopol, D. M. (2012). Restoration of bridge networks after an earthquake: Multicriteria intervention optimization. *Earthquake spectra*, 28(2), 427-455.
- Dong, Y., & Frangopol, D. M. (2016). Performance-based seismic assessment of conventional and base-isolated steel buildings including environmental impact and resilience. *Earthquake Engineering & Structural Dynamics*, 45(5), 739-756.
- Faber, M.H., Miraglia, S., Qin, J., & Stewart, M. G. (2020). Bridging resilience and sustainability - decision analysis for design and management of infrastructure systems. *Sustainable and Resilient Infrastructure*, 5(1-2), 102-124.
- Sánchez-Silva, M., Gardoni, P., Val, D. V., Yang, D. Y., Frangopol, D. M., Limongelli, M. P., ... & Straub, D. (2025). Moving toward resilience and sustainability in the built environment. *Structural Safety*, 113, 102449.
- Tapia, C., & Padgett, J.E. (2016). Multi-objective optimisation of bridge retrofit and post-event repair selection to enhance sustainability. *Structure and Infrastructure Engineering*, 12(1), 93-107.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised in the **Abstract, Introduction, Methods, and Conclusions**. For all the amendments please refer to the revised manuscript, in the text highlighted with yellow. The authors do not provide amended text here because of the extensive revisions.

In the **Abstract**, we sharpened the gap and novelty statements by clarifying that the framework addresses the separate treatment of recovery dynamics, life cycle environmental impacts, economic costs, structural performance, service continuity, and service life. In the **Introduction**, we added a direct comparison with existing integrated frameworks, including Bocchini and Frangopol, Dong and Frangopol, and Tapia and Padgett. We clarified that these studies provide important methodological foundations, but are mainly formulated within hazard driven, post disaster, performance based, or sustainability oriented decision contexts. We then clarified how the present framework differs from those predecessors: it addresses retrofit planning under progressive deterioration, formulates resilience as an implementation dependent recovery variable, and jointly evaluates recovery performance with life cycle environmental impact, cost, and expected service life.

In the **Methods**, we further specified that the individual components build on established resilience, LCA, LCC, and MCDA methods, while the methodological contribution lies in their decision-oriented coupling around recovery-based resilience for retrofit and adaptation planning. The revised Figure 1 also improves the positioning of the framework relative to existing resilience–sustainability decision workflows by explicitly highlighting the policy-sensitive and decision-oriented structure of the methodology. In the **Conclusions**, we refined the contribution statement to emphasize that the framework operationalises resilience as a recovery based, time bounded, and uncertainty aware decision variable linked to retrofit implementation and weighed against carbon, cost, and longevity.

Comment # 3. The resilience index in Equation (1) follows Bruneau et al. (2003) without modification. The authors must clarify what methodological advancement this constitutes beyond parametrizing an existing formula with expert-elicited duration distributions. If the novelty resides in the probabilistic treatment of TRE, this must be stated explicitly with the governing probability distribution, its parameters, and the fitting procedure documented in the main text rather than deferred entirely to supplementary material.

The control time $T_{LC} = 30$ months is adopted based on empirical benchmarks from three bridge cases (Ponte San Giorgio, West Seattle Bridge, Arlington Memorial Bridge). This selection is critical because it directly governs which scenarios achieve resilience compliance. The authors must provide a formal sensitivity analysis of the ranking results to T_{LC} , demonstrating whether the conclusions change if T_{LC} is set to 24 or 36 months, as is plausible for different asset contexts.

1) Synopsis of the response

The authors thank the Reviewer for the insightful comment. We clarify that Equation (1) adopts the resilience index of Bruneau et al. (2003) without modification, and that the methodological contribution lies in its operational use within a decision oriented and uncertainty aware framework for infrastructure adaptation and retrofit selection. In the revised manuscript, resilience is explicitly framed as a recovery driven metric derived from intervention sequencing, operational constraints, and expert informed recovery trajectories. Recovery time uncertainty is represented through non-parametric descriptive indicators rather than fitted probability distributions, thereby avoiding unsupported parametric assumptions while maintaining transparency and interpretability within the decision framework. The manuscript has been revised accordingly to clarify the contribution, the treatment of uncertainty, and the role of recovery time in the resilience assessment.

In addition, we performed a formal sensitivity analysis for alternative control times ($T_{LC} = 24, 30$ and 36 months). The results confirm that the framework responds systematically and transparently to different recovery-policy assumptions, while preserving the overall strategic trade-off structure across the considered retrofit scenarios. In particular, resilience-oriented strategies remain favourable under strict recovery targets, whereas long-service-life solutions become increasingly attractive under extended recovery horizons. These analyses and the corresponding rankings are now reported in Supplementary Note 5.

Due to space limitations, the detailed sensitivity analysis is provided in the present response to the Reviewer, while the manuscript has been revised to clarify the policy- and asset-dependent nature of the selected control time.

2) Detailed response

We agree that Equation (1) follows the original formulation of Michel Bruneau et al. (2003) and does not constitute a modification of the resilience metric itself. The contribution of this work does not lie in redefining resilience metric, but in its operationalization within a decision-oriented and uncertainty-aware framework for infrastructure adaptation and retrofit selection.

Specifically, resilience is reformulated from a hazard-centric perspective to a recovery-driven, implementation-based metric, where functionality–time trajectories $Q(t)$ are explicitly from intervention sequencing, operational continuity constraints, implementation processes, and expert-informed recovery durations associated with planned retrofit strategies under progressive deterioration conditions. This enables direct integration of resilience with life-cycle environmental, economic, and service-life indicators within a unified decision framework.

Regarding uncertainty treatment, recovery time is considered an uncertain variable derived from structured expert elicitation. However, no parametric probability distribution is assumed. Given the limited sample size and the nature of the survey, fitting a formal probability distribution would introduce unsupported assumptions and potentially artificial statistical precision. Instead, uncertainty is represented using non-parametric descriptive indicators (minimum, quartiles, median, and maximum), which transparently capture both central tendency and dispersion in expert-informed recovery estimates. These descriptors are

subsequently used to construct alternative recovery trajectories and evaluate resilience under both typical and upper-tail recovery conditions relevant to operational and policy constraints.

The manuscript has been revised accordingly to clarify explicitly in the main text: (i) that Equation (1) is adopted from Bruneau et al. (2003) without modification; (ii) that the methodological contribution lies in the recovery-oriented operationalization of resilience within the retrofit decision framework; and (iii) that uncertainty in recovery duration is represented through non-parametric expert-informed descriptors rather than fitted probabilistic distributions.

Regarding the selection of the control time (T_{LC}), we agree that this parameter is important because it directly affects the interpretation of recovery compliance and resilience acceptability across retrofit scenarios. The selected baseline value of 30 months was derived from empirical benchmarks associated with major bridge intervention and restoration projects of comparable scale and operational importance (e.g. Ponte San Giorgio, West Seattle Bridge, and Arlington Memorial Bridge).

In addition, following the Reviewer's suggestion, we performed a formal sensitivity analysis to evaluate the influence of the control time (T_{LC}) on resilience-driven prioritisation by considering alternative recovery horizons of 24 and 36 months in addition to the baseline 30-month assumption. The resulting rankings and decision spaces are now reported in Supplementary Note 5. The analysis confirms that the framework responds systematically and transparently to different recovery-policy assumptions, as expected, because T_{LC} defines the acceptable recovery horizon against which resilience compliance is assessed.

Importantly, despite changes in absolute rankings, the overall strategic trade-off structure remains broadly consistent across T_{LC} assumptions. Resilience-oriented strategies continue to perform favourably under stricter recovery requirements, whereas scenarios emphasizing extended service life remain comparatively advantageous under longer recovery horizons. The sensitivity analysis further shows that changes are concentrated primarily within the resilience dimension, while the economic, environmental and service-life indicators remain comparatively stable. This indicates that the framework captures policy-dependent recovery priorities in a coherent and interpretable manner rather than exhibiting methodological instability.

The manuscript and Supplementary Information have been revised accordingly.

Given the already extensive length of the manuscript and the supporting nature of this additional analysis, the full sensitivity-analysis tables are included in the present Reviewer response rather than in the main manuscript. Nevertheless, the revised manuscript now explicitly clarifies that (T_{LC}) should not be interpreted as a universal recovery threshold, but rather as a decision- and asset-specific parameter reflecting infrastructure criticality, acceptable disruption duration, network redundancy, and stakeholder priorities.

$T_{LC} = 30$ months (912.5 days)

Scenario	Min	Q1	Median (Q2)	Q3	Max
#1	0.83	0.73	0.48	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#2	0.80	0.70	0.44	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#3	0.89	0.83	0.66	0.54	$T_{RE} > T_{LC}$
#4	0.78	0.58	0.34	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#5	0.78	0.64	0.34	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#6	0.81	0.63	0.32	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$

$T_{LC} = 24$ months (730 days)

Scenario	Min	Q1	Median (Q2)	Q3	Max
----------	-----	----	-------------	----	-----

#1	0.79	0.67	0.35	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#2	0.75	0.63	0.30	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#3	0.86	0.78	0.57	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#4	0.73	0.47	0.00	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#5	0.73	0.55	0.00	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#6	0.77	0.53	0.00	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$

$T_{LC} = 36 \text{ months (1095 days)}$

Scenario	Min	Q1	Median (Q2)	Q3	Max
#1	0.86	0.78	0.57	0.36	$T_{RE} > T_{LC}$
#2	0.83	0.75	0.53	0.35	$T_{RE} > T_{LC}$
#3	0.91	0.86	0.72	0.62	$T_{RE} > T_{LC}$
#4	0.82	0.65	0.45	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#5	0.82	0.70	0.45	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$
#6	0.84	0.69	0.43	0.34	$T_{RE} > T_{LC}$

3) Amendments to the manuscript to address the comment:

The manuscript has been revised in the **Methods, Results and Discussion**, and **Supplementary Note 5**. All corresponding changes are highlighted in yellow in the revised manuscript.

In the **Methods**, we clarified that Equation (1) is adopted from Bruneau et al. (2003) without modification, and that the methodological contribution does not consist in redefining the resilience index itself. Rather, the contribution lies in its recovery-oriented operationalisation within a retrofit decision framework, where functionality time trajectories are derived from intervention sequencing, operational continuity, implementation constraints, and expert informed recovery estimates. We also clarified that no fitted probability distribution is assumed for the recovery time. Recovery uncertainty is represented through expert elicited recovery time estimates under different recovery conditions, rather than through parametric distribution fitting. This clarification was added to avoid implying unsupported probabilistic inference and to better reflect the decision support nature of the framework. The treatment of T_{LC} exceedance was also revised. The manuscript now explains that when $T_{RE} > T_{LC}$, the functionality trajectory is evaluated over the fixed control horizon and the scenario is interpreted as not satisfying the recovery target, rather than producing an undefined resilience value. A compliance indicator was added to make explicit whether each retrofit strategy satisfies the target recovery time under uncertainty.

In the **Results and Discussion**, we revised the interpretation of the resilience results to emphasize recovery-oriented performance, implementation uncertainty, and compliance with the recovery target, rather than presenting the resilience index as a full hazard based structural resilience metric. Finally, in response to the Reviewer's request, we added a **sensitivity analysis on the control time** T_{LC} . In addition to the baseline value of 30 months, alternative control times of 24 and 36 months were analysed. The revised manuscript now states that the ranking outcomes respond coherently to different recovery policy assumptions: stricter recovery horizons penalize strategies with longer recovery trajectories, while longer horizons increase the relative attractiveness of strategies with greater long term service life benefits. Full results are reported in **Supplementary Note 5**.

Comment # 4. Table 1 shows that for scenarios #1, #2, #4, #5, and #6, recovery time exceeds TLC at Q3 and Maximum conditions. This means the majority of scenarios fail resilience compliance under upper-quartile uncertainty. The authors interpret this as a robustness concern but do not explain how the MCDA framework handles scenarios that formally violate the control-time constraint. If $T_{RE} > T_{LC}$, the resilience integral in Equation (1) is undefined within the specified bounds. The authors must state how R is computed or bounded in these cases, and whether scenarios violating TLC should be disqualified rather than penalized by weight.

1) Synopsis of the response

The authors thank the Reviewer for this comment. We have clarified how resilience is computed when recovery exceeds the control time (T_{LC}), defined such cases as non-compliant rather than undefined, and explicitly explained how they are treated within the decision framework. In addition, we introduced compliance characteristics to quantify the likelihood of meeting recovery-time targets under uncertainty.

2) Detailed response

We have clarified the formulation of the resilience metric and its interpretation when recovery time exceeds the control time (T_{LC}). In the revised manuscript (Methods section), we explicitly state that resilience is consistently evaluated over a fixed control horizon $[T_0, T_{LC}]$. In cases where the recovery time (T_{RE}) exceeds T_{LC} , the functionality–time curve is truncated at T_{LC} , such that $Q(T_{LC}) < 1$. This ensures that the resilience integral remains well-defined and comparable across all scenarios, while reflecting incomplete recovery within the acceptable time window.

We further clarify that such cases represent non-compliance with resilience targets, rather than undefined behaviour. This interpretation is now explicitly reflected in both the Methods and Results sections, including the discussion of upper-quartile and maximum conditions in Table 1.

To better capture this aspect, we have additionally introduced compliance characteristics, defined as the probability of satisfying the recovery-time constraint, $P(T_{RE} \leq T_{LC})$. Given the available data, compliance is expressed in semi-quantitative terms based on quartile analysis (High >75%, Moderate 25–75%, Low <25%), and is now reported alongside resilience values in Table 1. This provides an explicit indication of the likelihood of meeting policy-relevant recovery thresholds under uncertainty.

Regarding the decision framework, we have revised the manuscript to clarify that T_{LC} is treated as a policy-relevant threshold rather than a strict feasibility constraint. Scenarios exceeding T_{LC} are therefore not excluded a priori but are retained within the decision space and penalized through their resilience scores and compliance characteristics. This allows the framework to explicitly capture trade-offs between short-term recovery performance and longer-term benefits such as extended service life or reduced environmental impact, which are central to infrastructure decision-making in practice.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised as follows:

Methods (Resilience definition):

T_{LC} represents the maximum acceptable time for the complete restoration of the bridge's functionality level and is selected based on strategic, operational, and societal requirements [37]. The recovery time T_{RE} is a stochastic variable significant uncertainty, reflecting variability in intervention duration and administrative delays. While resilience targets aim to achieve $T_{RE} \leq T_{LC}$, in practice T_{RE} may exceed T_{LC} under adverse conditions. In such cases, the functionality–time curve is truncated at T_{LC} , such that $Q(T_{LC}) < 1$. This ensures that the resilience integral remains well-defined over the fixed control horizon $[T_0, T_{LC}]$ and reflects incomplete recovery within the acceptable time window. Scenarios with $T_{RE} > T_{LC}$ are therefore considered non-compliant with resilience targets [38-42].

Methods (Uncertainty treatment):

To complement the resilience metric, a compliance ratio defined as $P(T_{RE} \leq T_{LC})$ is introduced to quantify the probability that a given intervention satisfies the recovery-time constraint under uncertainty.

Within the decision framework, T_{LC} is treated as a soft constraint. Scenarios that exceed the control time are not excluded a priori but are penalized through reduced resilience scores and increased dispersion in recovery trajectories, allowing trade-offs between short-term recovery performance and long-term benefits (e.g. service life extension or reduced environmental impact) to be explicitly evaluated.

While all strategies satisfy the control time under favourable and median conditions, several scenarios exceed the 30-month threshold under upper-quartile and maximum uncertainty conditions, indicating potential non-compliance with resilience targets.

This behaviour reflects differing levels of resilience compliance across scenarios. In particular, upper-tail outcomes (Q3 and maximum) indicate that several intervention strategies are likely to exceed acceptable recovery durations under realistic implementation uncertainty.

Results (Resilience trajectories and Table 1):

Specifically, while all retrofit strategies remain compliant with the control time under average conditions, upper-tail behaviour reveals systematic non-compliance with resilience targets for most scenarios. Full deck replacement strategies (scenarios #4–#6) exhibit both lower median resilience and frequent exceedance of the control time under uncertainty, indicating increased risk of unacceptable recovery durations. In contrast, scenario #3 demonstrates higher resilience and improved compliance across statistical conditions, reflecting both faster recovery and reduced sensitivity to implementation delays. These differences highlight that resilience performance depends not only on expected recovery time but also on the likelihood of meeting policy-relevant recovery thresholds.

Table 1 | Resilience index (R) and compliance characteristics $P(T_{RE} \leq T_{LC})$ (High >75%; Moderate 25–75%; Low <25%) for the six retrofit scenarios across different statistical assumptions on recovery time. Values are shown for Minimum, First Quartile (Q1), Median (Q2), Third Quartile (Q3), and Maximum conditions*

Scenario	Min	Q1	Median (Q2)	Q3	Max	P
#1	0.83	0.73	0.48	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$	Moderate
#2	0.80	0.70	0.44	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$	Moderate
#3	0.89	0.83	0.66	0.54	$T_{RE} > T_{LC}$	High
#4	0.78	0.58	0.34	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$	Low
#5	0.78	0.64	0.34	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$	Low
#6	0.81	0.63	0.32	$T_{RE} > T_{LC}$	$T_{RE} > T_{LC}$	Low

*While favourable conditions (Min, Q1) yield high R-indices (typically $R > 0.70$, with Scenario #3 reaching 0.89 at Min and 0.83 at Q1), median values are notably lower (ranging from 0.32 for #6 to 0.66 for #3). Critically, for upper quartiles and maximum conditions, recovery time (T_{RE}) often exceeds the control period (T_{LC}), particularly for all scenarios at Max values and most at Q3 (exceptions: Scenario #3 maintains $R = 0.54$ at Q3 before exceeding T_{LC} at Max). This behaviour reflects the probabilistic nature of recovery processes and highlights differences in compliance likelihood across scenarios.

Decision optimization section:

In this formulation, the control time acts as a policy-relevant threshold rather than a strict feasibility constraint. Scenarios exceeding T_{LC} are retained within the decision space but are penalized through their resilience scores and compliance characteristics, enabling explicit evaluation of trade-offs between short-term recovery performance and long-term system benefits.

Comment # 5. The functionality curves in Fig. 3 appear as step functions with abrupt transitions at discrete construction phases. The authors must justify the piecewise-linear/step form of $Q(t)$ against alternative

continuous degradation-recovery models, and clarify whether the area under the curve (i.e., the resilience integral) is computed analytically or numerically, and what discretization error arises from the step representation.

1) Synopsis of the response

The authors thank the Reviewer for this comment. We clarify that the stepwise functionality curves are intentionally used to represent the discrete nature of construction and retrofit interventions, where changes in service functionality occur at identifiable phases such as closure, partial reopening, traffic management changes, and completion of works. The resilience integral is computed numerically from the discretised functionality time trajectories.

2) Detailed response

The stepwise representation of the functionality curves reflects the discrete nature of construction and intervention processes, where changes in functionality occur at specific phases (e.g. closure, partial reopening, completion of works), rather than as continuous processes. For this reason, $Q(t)$ is modelled as a piecewise function associated with intervention sequencing and operational constraints. While continuous degradation–recovery models could be adopted in other contexts, they would require assumptions on intermediate behaviour that are not supported by the available data. In contrast, the adopted representation is directly grounded in the structure of the intervention process and expert-informed schedules. The resilience integral is computed numerically based on the discretised $Q(t)$ trajectories. Given the limited number of phases and the piecewise-constant formulation, the discretisation error is negligible relative to the scale of the analysis and does not affect the comparative ranking of scenarios.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised in the **Results and Discussion** section, close to **Fig. 3**. The corresponding changes are highlighted in yellow in the revised manuscript. We clarified that the stepwise functionality curves reflect the discrete nature of retrofit implementation, where service functionality changes at identifiable intervention phases, such as closure, partial reopening, traffic management changes, and completion of works. We also clarified that the resilience integral was computed numerically from the discretised functionality time trajectories.

Comment #6. Direct costs in Fig. 7a are described as "cost of materials only (excluding labor and machinery)." This is an unusual and restrictive definition of LCC. Labor and machinery can represent 40–60% of total construction cost for bridge interventions of this type. The authors must either include these components or explicitly justify their exclusion and quantify the sensitivity of the LCC ranking to this omission.

1) Synopsis of the response

The authors thank the Reviewer for this observation. We agree that the previous wording in Fig. 7a could be misleading because it suggested that direct costs included materials only. The manuscript has been revised to clarify that the cost data used for the LCC are derived from Greek construction price lists, where unit costs already include construction expenditures such as labour, machinery, dismantling, demolition, and transport to waste processing facilities, in addition to material costs.

Accordingly, labour and machinery were not excluded from the LCC analysis. Rather, they are embedded in the adopted unit cost data. We have clarified this point in the Integration of sustainability section and corrected the caption of Fig. 7 to avoid ambiguity. The revised text now specifies that, for LCC, the cost data

merge material and construction expenditures into one value, while the figure caption no longer describes direct costs as “materials only.”

2) Detailed response

The cost statistics used are derived from a Greek price list. This price list includes the costs of construction work such as dismantling and deconstruction, as well as transport to a waste processing site. These costs also include labour and machinery and are added to the cost of the materials used. We have clarified this aspect in the manuscript to avoid any possible misinterpretation.

Please see the example from the price list below. The full list is unavailable for confidentiality reasons.

			Works €1.50 Million	<Works € Million	<Works €10 Million	<Works €10 Million
DEMOLITIONS						
Demolition of structures with load-bearing elements made of reinforced concrete						
For height up to 4.0 m	OIK-2227	m ³	19.20*	17.50*	16.70*	15.80*
For height beyond 4.0 m	OIK-2227	m ³	27.40*	25.00*	23.80*	22.60*
CONSTRUCTION WORKS:						
Prestressed slabs and slab-beams from concrete C30/37						
At height from the ground H ≤ 7.00 m	OΔO-2565	m ³	173.00	158.00	150.00	143.00
At height from the ground 7.00 < H ≤ 15.00 m	OΔO-2566	m ³	196.00	179.00	170.00	162.00
At height from the ground H > 15.00 m	OΔO-2566	m ³	219.00	200.00	190.00	181.00
Prestressed bridge structures from concrete C30/37 using suspended movable formwork	OΔO-2566	m ³	242.00	221.00	210.00	200.00

The provided table illustrates the integration of labour and machinery in the price list. To ensure a nuanced analysis, project costs are categorized by asset value (e.g., above or below €10 million). As indicated in the data, projects with higher total asset values tend to exhibit lower costs per m². This inverse relationship is primarily driven by **economies of scale**:

- **Fixed Cost Absorption:** Larger projects allow for fixed expenses—such as mobilization of machinery and site management—to be distributed across a greater surface area, thereby reducing the unit cost.
- **Non-Proportional Scaling:** Many high-cost construction elements do not scale linearly with project size. Smaller builds often face disproportionately high costs because certain essential processes cannot be "scaled down" without incurring significant logistical inefficiencies.
- **Operational Efficiency:** In contrast to smaller projects, larger developments often benefit from more optimized labor and machinery utilization, minimizing the impact of idle time and fixed overheads on the final cost per m².

3) Amendments to the manuscript to address the comment:

We clarified this issue in the manuscript (see Section Integration of Sustainability)

“...for LCC, the used cost data merges material and construction expenditures into one value.”

Caption of Fig 7 has been also corrected. “Direct costs (a) represent the cost of materials, labor and machinery.”

Comment #7. The indirect cost formula (Equation 8) calculates detour-related costs as a function of transport-type distance and impact factor only. There is no vehicle operating cost model, value-of-time component, or traffic volume input. For a bridge of motorcostsclass, indirect costs depend critically on annual average daily traffic (AADT) and detour distance. The authors must provide these input values and their sources, or acknowledge that the indirect cost estimates are incomplete proxies.

1) Synopsis of the response

We thank the Reviewer for this comment. We have clarified the assumptions underlying the indirect cost estimation, including the approximate traffic demand and detour distance adopted for the case study bridge. We also explicitly acknowledge that the indirect costs should be interpreted as comparative planning-level estimates rather than detailed traffic-engineering forecasts. Importantly, the same traffic and detour assumptions were applied consistently across all retrofit scenarios; therefore, they do not affect the relative ranking trends within the MCDA.

2) Detailed response

The indirect cost component in Equation (8) was intended to provide a comparative estimate of user-related disruption costs associated with bridge closure and detour routing, rather than a detailed transportation-demand or traffic-assignment model. We agree that indirect costs are strongly influenced by traffic volume, detour distance, and vehicle operating assumptions, and that these inputs should be stated explicitly.

For the Polyfytos bridge case study, approximate daily traffic demand was estimated based on local traffic conditions and the population of the two nearest urban centres (Servia and Kozani), assuming that the average daily traffic corresponds approximately to 15% of the combined local population, resulting in an estimated traffic volume of approximately 10,000 vehicles/day. The detour distance associated with bridge closure was approximated as an additional 25 km based on route mapping using OpenStreetMap and Google Maps road-network information.

The resulting indirect costs should therefore be interpreted as planning-level comparative estimates intended to capture the order of magnitude of disruption effects rather than precise operational traffic forecasts. Importantly, the same traffic-demand and detour assumptions were applied consistently across all retrofit scenarios. Consequently, while these assumptions influence the absolute magnitude of indirect costs, they do not materially affect the relative ranking patterns or comparative trends observed in the MCDA results.

3) Amendments to the manuscript to address the comment:

The following has been amended in the revised manuscript.

“For the Polyfytos bridge case study, indirect user costs were estimated using approximate traffic-demand and detour assumptions representative of local operating conditions. Average daily traffic was approximated as 10,000 vehicles/day based on local population and traffic characteristics of the nearby urban centres of Servia and Kozani, assuming traffic demand corresponding to approximately 15% of the combined population. The detour associated with bridge closure was estimated as an additional travel distance of approximately 25 km based on route mapping using OpenStreetMap and Google Maps. These estimates are intended to provide comparative planning-level indicators of disruption costs rather than detailed traffic-engineering forecasts. Since identical traffic-demand and detour assumptions were adopted across all retrofit scenarios, these parameters do not affect the relative ranking trends within the MCDA framework.”

Comment #8. The LCA system boundary excludes module B (use phase: maintenance, repair, inspection) for all scenarios. For scenarios with shorter expected service lives (e.g., Scenario #3 with localized strengthening), use-phase maintenance interventions could substantially alter total life cycle GWP. The authors must either

include module B or demonstrate quantitatively that its exclusion does not reverse the environmental ranking.

1) Synopsis of the response

The authors thank the Reviewer for this important comment and agree that use phase processes should be clearly addressed, particularly for scenarios with different service lives. The manuscript has been revised to clarify the treatment of Module B in accordance with EN 15804, EN 15643, and EN 15978. Maintenance and inspection activities under Module B2 are expected to be comparable across the considered options and to have limited influence on the comparative ranking, while repair activities under Module B3 are excluded because they relate to unexpected events outside the study scope.

In response to the comment, the analysis has been updated to include replacement activities under Module B4 over the 50-year observation period. These include replacement of the draining wearing layer, binder course, primer, and waterproofing membrane at defined intervals. Because these non-structural elements are similar across all six retrofit options examined here, their inclusion increases the absolute environmental impacts and costs but does not alter the results of comparison and thus do not affect the environmental ranking. The revised manuscript and Supplementary Note 4 now report the updated LCA and LCC scope, including Module B4 replacement contributions and updated environmental datasets.

2) Detailed response

We thank the Reviewer for their feedback and we agree with this observation. We recognise that the three options are not functionally equivalent due to their differing service lives and performance levels. We would like to highlight that, however, all scenarios, however, as a minimal requirement, 50 years' service life. We highlighted this aspect in the supplementary note 4.

“The analysis quantifies the environmental impacts and costs associated with the six retrofit scenarios at asset level, using a functional unit of 1 km of bridge. It is essential to note that scenarios are not functionally equivalent in terms of durability performance and service life extension. All designed options, have, as a minimal requirement, a reference service life of 50 years. Scenarios #1, #2, #4 and #5 have 70 years' service life. Where scenarios reach similar operational capacity of 90–100%, differences in longevity and performance are addressed in the MCDA using fit-for-purpose criteria.”

We agree on the relevance of the use stage, especially with regard to elements' replacement. Please note that the author refers to the definitions provided in EN 15804, EN 15643 and EN 15978.

- The B2 module, Maintenance, involves planned, regular actions to preserve the functionality and appearance of a building component. This includes cleaning, servicing, and minor replacements, such as painting and applying sealants. For environmental impact assessments (performed with LCA), maintenance and inspection work is usually considered negligible due to its low environmental impact. This assumption is mainly applied to assets constituted by durable materials or those requiring low-intensity maintenance, such as structural elements. However, maintenance could be significant for pavements and finishing layers. These components are identical for all planned options.
Maintenance can cause additional costs, especially because of regular inspections. However, also these activities are expected to be performed identically for all options, with exception of scenarios having external prestressing cables. In these circumstances, specific inspection activities are expected, which will not cause significant additional costs.

- The authors refer to the definition of Repair (Module B3) according to EN15798. Repair activities include all works carried out following unexpected events, such as earthquakes or vandalism. These events are outside the scope of the study and are therefore excluded.
- When entire elements are substituted (e.g. prefabricated slabs), the authors refer to replacement activities (B4 module). This should occur at the end of the element's nominal service life and can determine total environmental impacts and costs.

Given the above, the authors decided to consider the Reviewer's suggestions by performing **Life Cycle Assessment (LCA) and Life Cycle Costing (LCC)** calculations that include replacement activities (**Module B4**). Specifically, we calculated the total environmental impacts and costs over a **50-year** observation period. Within this timeframe:

- **The draining wear layer** is replaced every 8 years. This layer is identical for all **options**.
- **The binder course** is replaced every 16 years. This layer is identical for all **options**.
- **The primer and waterproofing membrane** are replaced every 32 years. These layers are identical for all **options**.
- **Structural elements**, which vary by design option, reach their end-of-life at the end of the period. Consequently, their environmental impacts and costs are accounted for in **Modules C3-C4**.

In summary, Module B4 accounts for the replacement of non-structural elements, which are identical across all six options. For example, we assessed an absolute cost increment of **€650k (costs are discounted considering an inflation rate of 2.8%)**. See the example of the calculation for Scenario #1.

ELEMENTS and MATERIALS		QUANTITY	UNIT	Service life	DIRECT Costs (EUR)						
				years	A1-A3 +A5	B4	C1_old	C1_new	C2_old	C2_new	TOTAL
Structural elements	Prestressed Steel Y1860	3,6	tons	70	25.200,00 €	0,00 €	42.217,50 €	6.130,75 €	6.130,75 €	4.162,66 €	1.118.500,29 €
	Concrete C40-50	4	m³	70	776,00 €	0,00 €					
	Structural Steel	312	tons	70	327.600,00 €	0,00 €					
	Concrete C40-50	260	m³	70	50.440,00 €	0,00 €					
	Steel B450-B500	31,2	tons	70	32.760,00 €	0,00 €					
Pavement	Primer Asphalt	1300	m²	32	1.430,00 €	3.460,32 €					
	Waterproofing Membrane	1300	m²	32	12.090,00 €	29.255,40 €					
	Asphalt Concrete (Binder layout)	1300	m²	16	8.060,00 €	66.698,42 €					
	Asphalt (Draining wear layer)	1300	m²	8	9.360,00 €	530.996,34 €					
	Expansion joints	27	m	16	0,00 €	0,00 €					
	Milling for replacement pavement up to 6 cm	1300	m²	8	/	12.105,04 €					
	Milling for replacement pavement up to 8 cm	1300	m²	16	/	8.268,77 €					
	TOTAL				467.716,00 €	650.784,29 €					

The same occurs with environmental impact assessment. For example, we assessed an absolute GWP increment of **1,95E+05 kgCO₂eq**. See the example of calculation for Scenario #1

ELEMENTS and MATERIALS		QUANTITY	UNIT	Service life	GWP - Direct [kg CO2 eq.]										
				years	A1-A3	A5	B4	C1 (old)	C2_old	C3-C4_old	C1_new	C2_new	C3-C4	D	TOTAL
Structural elements	Prestressed Steel Y1860	3,6	tons	70	4,90E+03	1,01E+02	0,00E+00	1,37E+03	9,58E+03	7,72E+03	4,67E+04	2,40E+04	0,00E+00	-1,77E+03	8,79E+05
	Concrete C40-50	4	m³	70	1,28E+03	0,00E+00	0,00E+00						4,35E+01	-2,01E+01	
	Structural Steel	312	tons	70	4,27E+05	0,00E+00	0,00E+00						1,95E+04	-7,58E+04	
	Concrete C40-50	260	m³	70	8,32E+04	0,00E+00	0,00E+00						2,83E+03	-1,31E+03	
	Steel B450-B500	31,2	tons	70	2,15E+04	0,00E+00	0,00E+00						8,23E+01	4,80E+03	
Pavement	Primer Asphalt	130000	kg	32	7,59E+04	0,00E+00	7,66E+04						6,85E+02	0,00E+00	
	Waterproofing Membrane	1300	m²	32	6,79E+03	2,82E+01	7,95E+03						1,17E+03	1,03E+02	
	Asphalt Concrete (Binder layout)	184600	kg	16	1,75E+04	1,42E+02	5,26E+04						0,00E+00	-4,27E+03	
	Asphalt (Draining wear layer)	184,6	tonne	8	1,29E+03	2,47E+01	7,83E+03						1,62E+01	-3,45E+02	
	Expansion joints	27	m	16	1,63E+04	0,00E+00	4,99E+04						2,78E+02	-7,29E+03	
	TOTAL				6,56E+05	2,96E+02	1,95E+05	1,37E+03	9,58E+03	7,72E+03	4,67E+04	2,40E+04	2,46E+04	-8,59E+04	8,79E+05

All calculated increments are equal for all scenarios. In light of the above, the authors contend that the inclusion of element replacements, even if relevant because affect the total costs and environmental impacts,

will not affect the comparison of the planned scenarios, nor will it result in a significant change to the final ranking provided through the MCDA.

The recent iteration of the document discloses the environmental impacts and costs associated with the replacement module. It is also important to note that the use of updated environmental databases has enabled the identification of more specific environmental datasets for some materials used in the planned scenarios of intervention. This is the rationale behind the decision to incorporate updated LCA values into the current version.

3) Amendments to the manuscript to address the comment:

In section “Integration of sustainability”, we clearly defined the system boundaries of LCA analyses.

“The system boundaries of the analysis encompass the demolition and waste processing of damaged bridge components (in this study called C1-C4_old), raw material extraction and manufacturing, transportation, construction and installation processes (A1-13), elements replacement (B4), and end-of-life stages of newly installed elements (C1-C4_new), following standards specific for construction works (EN 15804 and EN 15643 [43,44]). Impacts are assessed considering 50 years' observation time.”

We also justified why maintenance is not considered in LCA and LCC

“The environmental impacts caused by maintenance (Module B2) are disregarded, as they are deemed to be negligible and similar across the designed interventions. Consequently, they do not influence the ultimate decision-making process.”

“Similarly to LCA, costs associated with maintenance are excluded on the basis that the activities in question are executed in a consistent manner across all the scenarios that have been designed.”

Updated results are found in the section “LCA and LCC evaluation”. For further information regarding the performed LCA/LCC, please refer to the Supplementary Note 4.

Comment # 9. The LCA weight coefficients attributed to the Joint Research Centre (JRC) for aggregation of GWP (0.46), ADPM (0.17), ADPF (0.18), and WDP (0.19) are referenced to Sala et al. (2017). The authors must confirm that these weights are drawn from the Environmental Footprint weighting set and that their application to a construction-sector LCA is methodologically appropriate. The JRC EF weights were developed for product-level footprinting and their transfer to infrastructure-level LCA requires explicit justification.

1) Synopsis of the response

The authors thank the Reviewer for this clarification. We confirm that the adopted weights are drawn from the JRC Environmental Footprint weighting set reported by Sala et al. (2017). The manuscript has been revised to clarify that these weights are used as a transparent and policy recognised weighting basis for aggregating selected LCA impact categories within the MCDA, rather than as a claim that a bridge specific Product Environmental Footprint method is being applied.

We also clarified that the LCA is performed at the bridge asset level, not at the full infrastructure network level. Although the JRC EF weighting set was originally developed for product and organisation footprinting, its use here is limited to assigning relative importance to environmental impact categories based on environmental damage, policy relevance, planetary boundaries, and category uncertainty. The revised manuscript now explicitly states this methodological interpretation and acknowledges that the weighting factors are used to support comparative decision making across retrofit options.

2) Detailed response

In order to justify the use of the JRC EF weights for the selected case study, two preliminary remarks are necessary.

Remark 1: JRC EF Weighting and its application

In the context of LCA to support policy decisions and multicriteria decision-making, weighting is used to:

- identify the most relevant impact categories;
- guide decision-makers towards the most effective solutions for reducing environmental impacts, e.g. in relation to eco-innovation policies and strategies;
- present results in an aggregated manner (up to a single score) to improve decision-making and communication.

In the European context, the first Environmental Footprint pilot adopted an equal weighting approach (1:1:1, etc.) after characterisation and normalisation. This approach assigns an equal weight to each of the 16 midpoint impact categories captured in the LCA analysis. The JRC weighting system aimed to develop a method for weighting the Environmental Footprint Impact Categories according to their relevance to overall environmental issues (Sala et al. 2017). The weighting consider exclusively:

- The environmental damage and planetary boundaries,
- The importance of impact categories.
- underlying impact categories' uncertainties. (Sala et al. 2017)

As also underlined by the European Platform on LCA, the EU Product Environmental Footprint (PEF) method (and its associated weighting system) can be used for any product (goods or services) or organisation without restrictions. There is not a standalone PEF method in force for more complex systems such as bridges or buildings.

Remark 2: Definition of infrastructure

Infrastructure can be interpreted as singular construction work (ISO 6701-1:2020), here called also **infrastructure asset**, or as the entire infrastructure system, also called infrastructure networks (OECD)

ISO 6701-1:2020

Infrastructures are construction works (3.1.1) comprising a structure (3.1.4), such as a dam (3.2.24), bridge (3.3.19), road (3.3.1), railway (3.3.3), runway, utilities, pipeline (3.2.32), or sewerage system (5.4.40), or the result of operations such as dredging, earthwork (7.1.6), geotechnical processes, but excluding a building (3.1.3) and its associated site (3.1.6) works.

Organisation for Economic Co-operation and Development, Defining Infrastructure

Infrastructure has been defined as “the physical components of interrelated systems providing commodities and services essential to enable, sustain, or enhance societal living conditions and maintain the surrounding environment.

Justification for JRC weights application in bridge asset

We agree that there are possible challenges related to the application of EF to infrastructure systems. However, the scope of the present LCA study is limited to the bridge asset rather than the entire infrastructure system.

Even if there is no standalone PEF method in force for bridges, the envisioned task in the environmental impact analysis is the identification of weighting factors to be used to set priorities among the selected environmental impact categories. Weighting factors of PEF methodology do not strictly depend on the scope of the assessment, but solely on the importance of environmental issues and underlying uncertainties in the environmental impact assessment. For this reason, we decided to consider the weighting factors of JRC EF methodology and apply them to the assessed total environmental impacts.

3) Amendments to the manuscript to address the comment:

We clarified this concept by adding the following in the section “Multi-criteria decision analysis and optimization”.

The weighting factors of this study are based on environmental damage and planetary boundaries, the importance of impact categories in current policies, and the underlying impact categories’ uncertainties. Weighting factors do not strictly depend on the scope of the assessment [45].

Comment #10. The MCDA sensitivity analysis explores five weighting schemes (T1–T5), but the inter-criteria weight ranges are not reported in the main text. The authors must provide the full weight vectors for T1–T5 explicitly. Without these, the reader cannot assess whether the ranking reversals between scenarios #3 and #6 occur under small or large weight perturbations, which is the essential test of robustness.

1) Synopsis of the response

We thank the Reviewer for this comment. We have now explicitly reported the full weight vectors for all sensitivity scenarios (T1–T5) in the Methods section to ensure transparency and enable assessment of ranking robustness.

2) Detailed response

We agree that explicit reporting of weight vectors is essential to assess the robustness of the MCDA results. In the revised manuscript, we now provide the full set of weight vectors for all five scenarios (T1–T5) directly in the Methods section. These weight configurations correspond to distinct policy priorities, including equal weighting (T1) and strong emphasis on individual criteria (T2–T5), where one dimension is assigned a dominant weight (0.55) and the remaining criteria are equally weighted (0.15). This design allows systematic evaluation of ranking stability under substantial, rather than marginal, shifts in decision priorities.

The Results section has been updated to reference these explicitly defined weight vectors, ensuring that the reader can directly interpret ranking changes in the context of the corresponding weighting schemes.

3) Amendments to the manuscript to address the comment:

In Methods (MCDA section), we have added explicit definition of all weight vectors (T1–T5), including numerical values for w_C , w_E , w_R , and w_L .

“The specific weight vectors used in the sensitivity analysis (T1–T5) are defined explicitly as follows: T1 (equal weighting): $w_C = 0.25$, $w_E = 0.25$, $w_R = 0.25$, $w_L = 0.25$; T2 (economic priority): $w_C = 0.55$, $w_E = 0.15$, $w_R = 0.15$, $w_L = 0.15$; T3 (environmental priority): $w_C = 0.15$, $w_E = 0.55$, $w_R = 0.15$, $w_L = 0.15$; T4 (resilience priority): $w_C = 0.15$, $w_E = 0.15$, $w_R = 0.55$, $w_L = 0.15$; T5 (service-life priority): $w_C = 0.15$, $w_E = 0.15$, $w_R = 0.15$, $w_L = 0.55$.”

In Results (ranking section), we have added reference to Methods section for full specification of weight vectors.

“The baseline scenario corresponds to T1, while the four test scenarios T2–T5 prioritize economic (T2), environmental (T3), resilience (T4) and service-life (T5) objectives (see Methods for full weight vectors).”

Comment # 11. The UPI formulation (Equation 11) employs min–max normalization, which is sensitive to the composition of the scenario set: adding or removing a scenario changes the normalized values of all others and therefore the rankings. The authors must acknowledge this dependency explicitly and discuss whether the ranking of any scenario would change if the scenario set were extended or reduced.

1) Answer: The authors thank the Reviewer for this comment. We acknowledge that min–max normalization is dependent on the composition of the scenario set, and that adding or removing alternatives may affect the normalized values and resulting rankings. However, in the present study, the analysis is conducted on a fixed and predefined set of retrofit alternatives, identified through structured expert consultation and representing realistic intervention strategies for the case study. The purpose of the normalization is therefore to enable consistent relative comparison within this defined decision space, rather than to derive rankings that are invariant to changes in the scenario set. In this context, the results should be interpreted as scenario-dependent rankings, reflecting the relative performance of the considered alternatives. The manuscript has therefore been revised to explicitly acknowledge this dependency and to clarify that the reported rankings are relative comparisons within the predefined set of retrofit strategies analysed in this case study.

2) Amendments to the manuscript to address the comment:

A clarification has been added in the **Multi-criteria decision analysis and optimization** section to explicitly acknowledge that min–max normalization depends on the analysed scenario set. The revised text clarifies that the rankings should be interpreted as relative comparisons within the predefined decision space considered in this study.

Comment #12. The paper claims that "strategies prioritising system-level transformation with minimal structural intervention consistently outperform conventional approaches" in the abstract. Table 2 shows that under equal weighting (T1), Scenario #6 (full prefabricated replacement) ranks first for both UPI and TOPSIS, not a minimal-intervention strategy. Scenario #3 (external prestressing, minimal intervention) leads only under environmental and resilience prioritization (T3, T4). The abstract must be revised to accurately represent the conditional nature of this finding.

1) Answer:

The authors thank the Reviewer for this comment. We agree that the original wording in the Abstract overgeneralised the results by implying that one class of strategies consistently outperforms the others. The Abstract has therefore been revised to reflect the conditional and policy dependent nature of the findings. The revised wording no longer states that strategies with limited structural intervention “consistently outperform” conventional approaches. Instead, the Abstract now clarifies that the framework evaluates retrofit strategies under competing resilience, cost, carbon, and longevity objectives, and that system level transformation with limited structural intervention **can** outperform conventional strategies depending on the adopted decision priorities.

2) Amendments to the manuscript to address the comment:

To further avoid overstatement, the **Results and Discussion and Conclusions** have also been revised to clarify that no single retrofit solution is universally preferred across all policy perspectives, and that the preferred strategy depends on whether decision makers prioritise short term recovery, environmental impact, economic cost, or long term service life.

Comment #13. Expert elicitation involved seven respondents across six countries for a site-specific Greek motorway bridge. The authors must report the inter-expert variability quantitatively (e.g., coefficient of variation for recovery duration estimates by scenario) and discuss whether this sample size is sufficient to characterize the tails of the recovery time distribution, particularly given that Q3 and maximum values drive resilience non-compliance for most scenarios.

1) Synopsis of the response

The authors thank the Reviewer for this comment. We agree that inter-expert variability and the interpretation of upper recovery values needed to be clarified. The revised manuscript and Supplementary Information now report quantitative variability measures for the expert-elicited recovery estimates, while clarifying that the seven responses are not intended to characterise statistical tails of a recovery-time distribution. Instead, Q3 and maximum values are interpreted as adverse expert-elicited recovery conditions used to test the robustness of retrofit strategies within the proposed decision framework.

2) Detailed response

Inter-expert variability is explicitly reported in the **Supplementary Information** through descriptive statistics, including minimum, quartiles, median, maximum, standard deviation, and interquartile range (IQR) for recovery durations and related quantities. These metrics were selected to characterise variability in a robust and transparent manner, with particular emphasis on the IQR to capture clustering of expert-informed estimates while limiting the influence of extreme values.

This study adopts a targeted expert elicitation and consultation approach, where experts are selected to cover multiple dimensions of the problem (bridge assessment, retrofit design, life-cycle analysis, and resilience-oriented decision-making), rather than to form a statistically representative sample. The objective is therefore not to perform statistical inference or to derive population-level distributions, but to obtain high-quality, context-specific input across different domains of expertise.

We did not report coefficient of variation because the elicitation was not designed to estimate statistically representative recovery-time distributions. With a targeted panel of seven experts, CV values would have limited inferential meaning and could suggest a level of statistical precision that is not supported by the data. Instead, inter-expert variability is reported through transparent descriptive indicators, including minimum, first quartile, median, third quartile, maximum, standard deviation, and interquartile range. These values are used to represent the spread of expert judgement and to define favourable, central, and adverse recovery conditions for the decision framework.

The use of Q3 and maximum values is therefore intended to represent upper-bound recovery conditions derived from expert judgement, which are directly relevant for assessing the robustness of retrofit strategies within the decision-oriented framework, rather than to characterise statistical tail behaviour.

In this context, the results are intended to support the demonstration and application of the proposed methodology to a specific, complex case study, rather than to provide generalised statistical estimates.

3) Amendments to the manuscript to address the comment:

The manuscript has been revised in the **Methods, Results and Discussion**, and **Supplementary Information** to clarify that the expert-elicited values are used to support comparative decision making under uncertainty, not to estimate statistically representative recovery-time distributions. We also clarified that Q3 and maximum values are interpreted as adverse expert-elicited recovery conditions rather than as statistically characterised distribution tails.

Comment #14. Reference 12 (Mitoulis et al., 2025) is cited as a Zenodo preprint. Nature Communications does not permit citation of non-peer-reviewed preprints in this manner unless clearly identified. The authors

must replace this citation with a peer-reviewed source or explicitly flag it as a preprint and ensure the concept of "sustained resilience" it introduces is not used as a foundational premise without independent validation.

Answer: This reference has been replaced with an alternative one in the revised manuscript:

Kopiika, N., Di Baria, R., Argyroudis, S., & Mitoulis, S. A. (2026). *Sustained resilience in prioritising critical infrastructure recovery in war-impacted built environments*. In T. Comes, R. M. Dijkman, J. Jörin, M. Mühlhäuser, M. Pregnotato, C. Reuter, & C. Schaffner (Eds.), *Book of Abstracts of the International Conference on Resilient Systems (ICRS) 2026* (p. 172). Eindhoven University of Technology. <https://doi.org/10.6100/qp5f-nb93>

Comment #15. References 46 and 47 are cited in support of TOPSIS as "the most appropriate method for the problem at hand," yet Reference 46 concerns environmental impacts of building reuse and Reference 47 addresses failure assessment model evaluation—neither directly supports TOPSIS selection for MCDA of retrofit strategies. The authors must cite methodologically relevant literature on TOPSIS applicability (e.g., Behzadian et al., 2012, *Expert Systems with Applications*, 39(17), 13051–13069) and provide a reasoned justification for preferring TOPSIS over alternatives such as VIKOR or ELECTRE given the decision problem structure.

1) Answer

The authors thank the Reviewer for this important comment. We acknowledge that the previous references did not provide an appropriate methodological justification for the use of TOPSIS in the present MCDA framework. The manuscript has therefore been revised in the **Methods** section, within the **TOPSIS robustness assessment and verification** subsection, and all corresponding changes are highlighted in yellow.

We replaced the previous generic statement that TOPSIS was “the most appropriate method” with a more specific and qualified rationale. TOPSIS is now presented as a complementary robustness assessment method used to verify the weighted sum UPI ranking, rather than as a universally superior MCDA method. We also added methodologically relevant references on TOPSIS applications and MCDA use, including Behzadian et al. (2012) and Zyoud and Fuchs Hanusch (2017).

2) Amendments to the manuscript to address the comment:

The revised text clarifies that TOPSIS is suitable for the present retrofit selection problem because it provides a transparent complete ranking of alternatives through a closeness coefficient based on distance from ideal and anti-ideal solutions. This is appropriate for comparing retrofit strategies across multiple conflicting criteria, including resilience, life cycle cost, environmental impact, and expected service life.

We also added a brief comparison with ELECTRE and VIKOR. Compared with ELECTRE, TOPSIS avoids the need to define additional preference, indifference, and veto thresholds. Compared with VIKOR, TOPSIS is more consistent with the objective of assessing relative proximity to a multi criteria ideal solution, rather than identifying a regret-based compromise solution. Accordingly, TOPSIS is used as a complementary verification tool rather than as a substitute for stakeholder preference definition.

References

Behzadian, M., Otaghsara, S. K., Yazdani, M., & Ignatius, J. (2012). A state-of-the-art survey of TOPSIS applications. *Expert Systems with applications*, 39(17), 13051-13069.

Zyoud, S. H., & Fuchs-Hanusch, D. (2017). A bibliometric-based survey on AHP and TOPSIS techniques. *Expert systems with applications*, 78, 158-181.

Comment #16. The paper does not address the transferability constraints of the framework. The six scenarios are specific to a balanced-cantilever prestressed concrete bridge with a defined structural failure mode (tendon corrosion). The methods section implies scalability to "ageing bridge networks in Europe," but no discussion is provided of what additional data, models, or calibration would be required for different bridge typologies, hazard regimes, or regulatory contexts. This limitation must be explicitly acknowledged.

1) Answer

The authors thank the Reviewer for this comment. We agree that the distinction between the transferability of the methodological framework and the case-specific nature of the application should be stated more clearly. The manuscript has therefore been revised accordingly. The objective of this study is not to generalise the numerical results obtained for the Polyfytos Bridge to other bridge systems, but to demonstrate a structured decision-support framework through a representative case study. The resulting rankings and performance outcomes are inherently dependent on the characteristics of the examined bridge, including its structural typology, deterioration mechanism, retrofit alternatives, recovery assumptions, life-cycle inventories, cost data, hazard exposure, and applicable regulatory thresholds. The proposed framework, however, is transferable at the methodological level. Its application to other bridge typologies, infrastructure networks, or regional contexts would require appropriate case-specific calibration and input data, including structural configuration, deterioration and hazard models, retrofit strategies, recovery scenarios, environmental inventories, economic parameters, and policy or regulatory constraints. A clarification acknowledging these limitations and transferability requirements has now been added to the manuscript.

2) Amendments to the manuscript to address the comment:

The manuscript has been revised in the **Conclusions** to explicitly acknowledge this transferability limitation. "The proposed framework is transferable and thus can be implemented to diverse critical infrastructure, yet, the resulting rankings are case-specific. Hence, its application to other infrastructure typologies or regional contexts would require appropriate case-specific calibration and input data, including structural configuration, deterioration and hazard models, adaptation strategies, environmental inventories, economic parameters, and regulatory constraints."

Comment # 17. Figure captions for Figs. 4b and 4c contain a numerical inconsistency: the ADP-fossil range for scenarios #3 and #6 is stated as " $4.5\text{--}6.5 \times 10^7$ MJ/km" in the caption, while the figure axis shows values in the range of $4\text{--}18 \times 10^6$ MJ/km. The authors must verify and correct all numerical values reported in figure captions against the plotted data.

Answer:

The authors would like to thank the Reviewer for highlighting the issue with the captions for Figures 4b and 4c. We have checked the values and corrected the typos. It is also important to note that the use of updated environmental databases has enabled the identification of more specific environmental datasets for some materials used in the planned scenarios of intervention. This is the rationale behind the decision to incorporate updated LCA values into the current version. Updated results are found in the section "LCA and LCC evaluation". For further information, please refer to the Supplementary note 4.

Reviewer #3 (Remarks to the Author):

The manuscript is well written and has great insights in the case study. However, some questions remain:

Answer: The authors thank the Reviewer for the positive assessment and constructive comments. The manuscript has been revised to clarify the points raised.

Comment #1. Equation 8- symbols are not defined. Also why is the detour considered only for trucks using I_{tr} ?

Answer: The authors would like to thank the Reviewer for highlighting the issue with Equation 8. We have addressed this issue and clarified the formula and all elements associated with it. We clarified that all means of transportation are included in the calculation. The contributions of each mode of transportation are summed to derive the total indirect costs.

“Indirect costs due to the traffic detour are calculated as product of total distance of each mean of transportation ($D_{detour,tr}$), expressed in kilometer and the cost of a kilometer traveled I_{tr} of each mean of transportation. All contributions of each mean of transportation tr are finally summed to derive the total indirect costs I_{ind} ”.

Comment #2. What is the necessity to divide by bridge length in Equation 7?

Answer: Thank you for this comment regarding the choice of the functional unit (FU). In Life Cycle Assessment (LCA) and Life Cycle Costing (LCC), impacts must be normalized against a functional unit to standardize the assessment based on the service delivered rather than the physical characteristics of the product (ISO 14040). This approach ensures a transparent and fair comparison between different technical solutions that may vary in lifespan, efficiency, or material intensity. FU must be consistent with the goal and scope of the assessment (see ISO 14040). For this study, the goal is to evaluate the sustainability of various rehabilitation strategies for the Polyfyto bridge. Since all proposed options are designed to restore the bridge to full operational capacity, the primary "service" is the continued availability of the asset. Therefore, the scope of the assessment is the whole bridge asset.

LCA studies of bridges found in literature usually refer to the following FUs: 1) the entire bridge structure, 2) the bridge surface area (m^2), 3) the bridge length (m or km) Life cycle carbon assessment of bridges: Integrating design, construction processes, maintenance, and circular end-of-life strategies - ScienceDirect Frontiers | Life cycle assessment of the sustainability of bridges: methodology, literature review and knowledge gaps. We selected one kilometer of bridge length as the functional unit. This choice provides several advantages:

- **Comparability:** It aligns with standard infrastructure benchmarking practices (which would have not been possible by preferring option 1).
- **Scalability:** It allows for easier comparison with other bridge rehabilitation projects of varying lengths (not possible with option 2).
- **Ranking Integrity:** While selecting a different FU (such as option 2) would alter the absolute numerical results, it would do so proportionally across all scenarios. Consequently, the relative performance and final ranking of the rehabilitation options obtained with bridge length as FU remain unchanged.

We have clarified this definition in the revised manuscript (see Section Integration of Sustainability) to ensure the scope and scaling factors are transparent.

“Since the assessment is conducted at asset level and all proposed options are designed to restore the bridge to full operational capacity, the selected functional unit is one kilometer of bridge length. The selected unit allows for comparability of scenarios and scalability of costs and impacts for other case studies.”

Comment #3. The retrofit strategies are an important part of the results and all the discussions are around which strategy is the best- hence the manuscript needs to have a discussion also on what the retrofit strategies are and at the bare minimum, add the table from supplementary material to main article.

Answer: The authors thank the Reviewer for this comment. We agree that the retrofit strategies should be more immediately accessible in the main manuscript, since they are central to the interpretation of the results and ranking discussion. The revised manuscript now includes a concise summary of the six retrofit strategies in the **Infrastructure assessment, retrofit and adaptation strategies** section, immediately before Fig. 2. The new text briefly describes each scenario, from hybrid repair and external prestressing solutions to full deck replacement and prefabricated modular reconstruction. The detailed comparative table remains in Supplementary Table 1, where intervention type, Gerber joint treatment, external prestressing, traffic disruption, and degree of system transformation are reported. This avoids overloading the main manuscript while ensuring that readers can understand the retrofit alternatives directly from the main text.

Comment #4. It is unclear where the uncertainty in the resilience quantification is coming from- is it purely from the survey results or are uncertainties assumed in some other part of the method as well?

Answer: The authors thank the Reviewer for this comment. We clarified in the revised manuscript that uncertainty in the resilience quantification originates from the structured expert elicitation, specifically from expert informed estimates of intervention duration, administrative delays, operational continuity, and implementation constraints used to construct the alternative functionality time trajectories. No additional probabilistic uncertainty model is introduced in the resilience calculation. We also note that LCA and LCC are not treated probabilistically in the present application. As already clarified in the **Integration of sustainability** section, LCA and LCC inputs are treated deterministically because context consistent uncertainty data were not available. Their influence on the final decision is instead explored through MCDA weighting and sensitivity analyses. The manuscript has been revised in the **Fit for purpose resilience assessment** section to clarify the origin and treatment of uncertainty in the resilience formulation.

Comment #5. If the TLC were changed, would we expect any of the rankings in Table 2 to materially change?

Answer: The authors thank the Reviewer for this comment. We agree that the selected control time T_{LC} can influence the resilience values and, consequently, the ranking outcomes when resilience is given substantial weight in the decision process. This issue has now been explicitly addressed in the revised manuscript. In addition to the baseline value of $TLC = 30$ months, we performed a sensitivity analysis using alternative control times of 24 and 36 months. The results show that changing TLC affects the absolute resilience values and the degree of recovery target compliance, as expected. Shorter control times penalise strategies with longer or more uncertain recovery trajectories, whereas longer control times increase the relative attractiveness of strategies offering extended service life and long-term renewal benefits.

Importantly, the sensitivity analysis indicates that the overall strategic interpretation remains broadly consistent. Resilience oriented strategies remain favourable under stricter recovery targets, while replacement-oriented solutions remain competitive when long term service life extension is prioritised. Thus, TLC should not be interpreted as a universal constant, but as a policy and asset dependent parameter reflecting infrastructure criticality, acceptable disruption duration, network redundancy, and stakeholder priorities. The manuscript has been revised in the **Methods** and **Results and Discussion** sections to clarify the

role of T_{LC} , its treatment as a policy relevant recovery threshold, and the results of the sensitivity analysis. The full sensitivity analysis for $TLC = 24, 30$, and 36 months is reported in **Supplementary Note 5**.

ANSWERS TO REVIEWERS' COMMENTS

COMMSENG-26-0068A

Reviewer #1 (Remarks to the Author):

The authors have successfully improved the manuscript based on the previous feedback, and I recommend its acceptance subject to a minor comment.

Regarding the response to Comment 4, while I understand the limitations regarding data availability and the necessity of conservative assumptions, this is precisely why the manuscript should explicitly document these constraints and provide clear suggestions for future development. My primary remaining concern is associated with Equation 11. Technically, it represents a way to combine different sources of input into a single unified performance index. Theoretically, however, aggregating incommensurable dimensions, specifically combining the resilience metric with life-cycle costs and environmental impacts, presents major conceptual and weighting challenges, as these metrics represent fundamentally incomparable items. I suggest the authors explicitly acknowledge this theoretical limitation within the text and indicate that alternative integration methodologies should be further explored in future work.

1) Answer

The authors thank the Reviewer for this constructive comment. We agree that aggregating resilience, life-cycle cost, environmental impact, and expected service life into a single scalar indicator raises important theoretical considerations because these criteria represent different and inherently non-commensurable value domains. The objective of Eq. (11) is therefore not to imply physical equivalence among these dimensions, but rather to provide a transparent preference-based decision model that allows alternative stakeholder priorities to be explored explicitly through weighting assumptions.

In this context, the UPI (Unified Performance Index) should be interpreted as a decision-support tool rather than a universal measure of infrastructure performance. The resulting rankings are conditional upon the adopted weighting scheme and decision context, which is why the manuscript includes systematic sensitivity analyses and complementary evaluation using TOPSIS. These analyses were intended to reduce reliance on a single aggregation approach and to assess the robustness of the ranking outcomes under alternative value preferences.

We agree that alternative approaches for integrating incommensurable criteria deserve further investigation. The manuscript has therefore been revised to explicitly acknowledge this limitation and to identify exploration of alternative decision-integration methodologies as a future research direction.

2) Amendments to the manuscript

The manuscript has been revised in the Methods section, close to Eq. (11) (now Eq. (12)), and in the Conclusions. The revised text explicitly acknowledges that resilience, LCC, LCA, and service life are not physically commensurable dimensions. We clarified that the UPI is a preference-based decision indicator conditional on the adopted weights and decision context. We also added a future work statement recommending exploration of alternative integration methodologies for handling incommensurable criteria.

Amended text in the Methods and Conclusions sections is reported in what follows; it has been highlighted in green in the revised manuscript.

"It should be noted that the UPI is not intended to imply physical commensurability among resilience, life-cycle cost, environmental impact, and expected service life. These dimensions represent different value domains and cannot be reduced to a common physical unit. The scalarization in Eq. (12) should therefore be interpreted as a preference-based decision model that makes weighting assumptions explicit, rather than as a mechanistic equivalence among criteria. Consequently, the resulting rankings are conditional upon the

adopted weighting scheme and decision context. The use of alternative weighting scenarios and complementary TOPSIS analysis was intended to assess the robustness of ranking outcomes under different stakeholder preferences. By making value judgements explicit through weights, control times, and uncertainty thresholds, the method enables transparent alignment between engineering choices and policy objectives within real-world regulatory and budgetary constraints.”

“Future work would also explore alternative integration strategies, including outranking methods, robust decision making, and participatory multi attribute utility approaches, to further address the theoretical challenges associated with aggregating incommensurable resilience, environmental, economic, and durability criteria.”

Reviewer #2 (Remarks to the Author):

The authors have addressed most of my previous comments satisfactorily. However, one important issue remains and requires further revision.

In the text preceding Equation 2, the authors state:

“Therefore, LCA and LCC are here performed with a deterministic approach. For LCC, temporal variabilities are considered. Specifically, costs have been adjusted to ensure temporal consistency over the observation time of the analysis using an appropriate inflation rate (2.8%).”

This statement raises two concerns.

First, the use of a fully deterministic approach for life-cycle cost analysis is difficult to justify. LCC deals with future costs, system performance, maintenance needs, interventions, and service conditions, all of which are inherently uncertain. Even when empirical data are limited, uncertainty should not simply be ignored; this is precisely where modelling assumptions, sensitivity analysis, scenario analysis, or probabilistic treatment become important. The authors should revise this part of the analysis and explain how uncertainty in future costs, performance, deterioration, maintenance, and replacement assumptions is treated. If the authors decide to retain a deterministic formulation, they should provide a clear and convincing justification and, at minimum, include a sensitivity analysis of the main cost drivers.

Second, the reference to an “inflation rate” in the context of LCC is conceptually confusing. Life-cycle cost analysis normally requires discounting future costs to present value using an appropriate discount rate or discount function. Inflation and discounting are not equivalent. Inflation adjusts nominal costs over time, whereas discounting converts future costs into present values. The authors should clarify whether Equation 2 uses nominal or real costs, whether inflation has been removed or included, and which discount rate or discount function is used to compute present values. As currently written, the description suggests a possible misunderstanding of a basic distinction in LCC modelling.

These two issues weaken confidence in the LCC component of the manuscript. I therefore recommend that the authors revise this section thoroughly, clarify the economic assumptions, and ensure that the formulation is consistent with standard LCC practice.

Finally, Equation 2 itself is difficult to interpret in its current form. The authors should rewrite the equation, define all variables clearly, distinguish between cost escalation and discounting, and explain in plain terms what the equation computes and how it is used in the analysis.

1) Answer

We thank the Reviewer for these constructive comments. We provide point-by-point responses detailing how these concerns have been comprehensively addressed in the revised text and supplementary materials.

Concern 1: Cost uncertainty and sensitivity analysis

We agree with the Reviewer that future costs, system performance, and operational conditions are inherently bound to long-term uncertainties.

First, to provide necessary context regarding our modelling choices, we would like to highlight that this study is built upon a specifically established, distinct definition of infrastructure resilience. In our framework, we explicitly distinguish between:

1. **Acute hazard-driven resilience** (e.g., sudden seismic, flooding, or blast events), and
2. **Chronic resilience associated with long-term infrastructure adaptation** under continuous deterioration processes (e.g., structural fatigue, environmental degradation, and chemical corrosion).

The latter class of resilience is governed primarily by strategic intervention planning, constructability, operational continuity, and life-cycle recovery management, rather than by immediate, hazard-induced structural damage states. The proposed framework was intentionally engineered to address this second class of chronic, adaptation-driven problems.

Consequently, in the present work, additional stochastic hazard occurrences during and after the intervention period were excluded from the scope. The explicit objective of this analysis is to isolate, decouple, and benchmark the retrofit strategies themselves, rather than to couple localized retrofit decision-making with a complex, stochastic multi-hazard simulation. While future iterations of this framework will aim to incorporate time-dependent hazard occurrences and concurrent multi-hazard interactions during intervention intervals, the baseline formulation for this study remains decoupled from active performance degradation curves.

However, we fully acknowledge the Reviewer's excellent point that economic and operational parameters remain uncertain even within a fixed structural timeline. Therefore, to validate our baseline and satisfy the Reviewer's recommendation, we have executed a robust **one-way sensitivity analysis** on the two primary cost drivers of our model:

- **Initial Capital Costs** (the direct cost of implementing the rehabilitation scenarios, corresponding to life-cycle modules A1–A3).
- **Replacement and Maintenance Intervals** (the operational lifespans governing module B4).

In strict alignment with standard asset management and engineering cost-estimation protocols (such as the Association for the Advancement of Cost Engineering AACE international guidelines), the baseline values of these primary cost drivers were varied by **20%**.

A 20% variation is a highly defensible standard in techno-economic literature; it is sufficiently broad to capture realistic market volatility, macroeconomic supply-chain fluctuations, and labour rate variances over an extended observation period. Furthermore, because the objective of a standard sensitivity analysis is to identify meaningful economic "switching points" (the thresholds where competing options invert their rankings) rather than to model catastrophic, low-probability market failures (e.g., a 200% cost spike), this 20% bound operates precisely within the envelope of realistic economic probability.

The sensitivity analysis yielded the following critical insights (the comprehensive data and tables have been compiled in **Note 4 of the Supplementary Material**):

- **Impact of Initial Capital Cost Variation ($\pm 20\%$):**

Varying the initial material and construction costs (Modules A1–A3) resulted in a total life-cycle cost variation bounded strictly between **16% and 19%** across all scenarios. Because this variance propagates in a strictly proportional manner across all alternatives, it **does not alter the final preference ranking** of the rehabilitation strategies. Strategy #3 consistently emerges as the optimal choice (lowest overall life-cycle cost), while Strategy #5 consistently remains the least favourable (highest overall cost), completely independent of initial capital variations.

- **Impact of Operational Lifespan/Replacement Interval Variation ($\pm 20\%$):**

To model uncertainty in component performance and service life, we introduced a 20% extension and reduction to the Reference Service Life (RSL) of the components within each strategy. Reducing the service life forced premature asset replacements, triggering a total LCC increase ranging between **$\pm 0.5\%$ and $\pm 2\%$** .

Interestingly, the scenarios split into two distinct sensitivity groups based on their inherent financial structures: the first group comprises Scenarios #1, #2, #4 and #5 with variations (absolute value) $< 1\%$, while

the second group comprises Scenarios #3 and #6 with variations (absolute value) >1%. These distinct sensitivities are directly proportional to the baseline share of Module B4 (Replacement) relative to the total LCC. The higher the initial baseline allocation to Module B4, the more sensitive that specific scenario's total LCC is to variations in component durability.

Crucially, just as with the capital costs, **this variation does not disrupt the final multi-criteria options ranking**. Strategy #3 retains its position as the most cost-effective alternative, and Strategy #5 remains the most expensive, across the entire operational uncertainty envelope.

Ultimately, this sensitivity analysis proves that our baseline economic assumptions and deterministic calculation methods possess a high degree of mathematical robustness. The underlying cost structures are stable enough to reliably feed the subsequent Multi-Criteria Decision Analysis (MCDA) and guarantee that the final option rankings are sound, reproducible, and resilient to market fluctuations. Because the sensitivity analysis confirmed that our findings are highly stable, we didn't deem it critical to add an extensive discussion about it to the main manuscript and we preferred adding this in the Supplementary Material Note 4.

Concern 2: Cost inflation and discounting

We agree with the Reviewer comments, and we revised this part of the methodology to separate cost escalation from temporal discounting. In detail:

- The cash flows are modelled using **nominal costs**.
- The **2.8% rate** previously cited represents the general price inflation rate for the region. This has now been mathematically segregated from the present-value calculations to establish absolute conceptual correctness.

To bring future nominal cash flows back to Present Value (PV), we have formally introduced a **real discount rate of 4.0%**. This specific rate represents the standard macroeconomic benchmark recommended by the European Commission's Cost-Benefit Analysis (CBA) guidelines for public infrastructure, civil engineering, and environmental projects. Given that our case study fits precisely within this public infrastructure domain, a 4.0% real discount rate is highly appropriate. It strictly aligns with established European Union reference benchmarks for member states and accurately reflects the long-term temporal and social preferences for macroeconomic capital investments in Greece.

As a result, the Equation providing the present value for LCC is the following.

$$PV(y) = C_{nom} \frac{(1+e)^y}{(1+r)^y} \quad (2)$$

Where C_{nom} is the nominal cost (based on price statistics of the Greek construction cost catalogue), e is the inflation rate and r is the real discount rate.

To reflect this approach, explanatory texts have been provided rewritten to conform to standard Life-Cycle Costing (LCC) nomenclature, explicitly mapping how nominal price escalation and real discounting interact across the project timeline.

2) Amendments to the manuscript and supplementary material

1. **Manuscript (Section Integration of Sustainability):** We added a new Equation that explicitly illustrates the calculation of the Present Value. We have explicitly defined our baseline cost drivers, introduced the discount rate ($r = 4\%$) and inflation parameters ($e = 2.8\%$). The results have been updated accordingly.

"Moreover, to bring future nominal cash flows back to Present Value (PV), we have formally introduced a real discount rate of 4.0%. This specific rate represents the standard macroeconomic

benchmark recommended by the European Commission's Cost-Benefit Analysis (CBA) guidelines for public infrastructure, civil engineering, and environmental projects. As a result, the Equation providing the present value for LCC is the following:

$$PV(y) = C_{nom} \frac{(1+e)^y}{(1+r)^y} \quad (2)$$

where C_{nom} is the nominal cost (according to price statistics from the Greek construction costs' catalogue), e is the inflation rate and r is the real discount rate."

2. **Manuscript (Section Multi-criteria decision analysis and optimization):** Since the LCC results have been updated, we updated also the results of the MCDA (see Table 2 in the Manuscript).
3. **Supplementary Material (Note 4):** We have added a dedicated section in the Supplementary Material containing complete datasets, breakdown tables, and comparative sensitivity analysis for both the 20% CAPEX and 20% RSL variations. This section explicitly documents the mathematical robustness of the final MCDA ranking.

Reviewer #3 (Remarks to the Author):

A few questions/comments on the revised manuscript:

1. The added section in the revised manuscript to address my previous question on uncertainty needs more careful revision. The variable TRe is defined as stochastic variable- but the uncertainty around is not discussed. How are you modeling the variation around it? Also, from what I understood, the expert elicitation included just 7 complete responses- is that sample size enough to build uncertainty into the framework? Also, how are the mean, std deviation from these 7 responses propagated through to the final resilience and LCA/LCC targets?

1) Answer

The authors thank the Reviewer for the comment. We clarify that TRe is not modelled through a fitted probability distribution. The previous use of the term “stochastic variable” was imprecise and has therefore been revised. In the updated manuscript, TRe is described as an uncertain expert elicited variable reflecting variation in intervention duration, administrative delays, implementation sequencing, and operational continuity assumptions.

The variation around TRe is represented through a non-parametric treatment of the elicited responses. Specifically, the expert elicited values are summarised through descriptive recovery conditions, namely minimum, first quartile, median, third quartile, and maximum values. These conditions are used to construct alternative functionality time trajectories and to evaluate resilience under favourable, central, and adverse implementation assumptions. The seven complete responses are therefore not used as a statistical sample from which to infer a population level recovery time distribution. Rather, they are used as structured expert judgement within a targeted expert elicitation and consultation process. The objective was to obtain high quality, case specific input for a landmark ageing bridge system, where empirical recovery datasets are not available and the evaluation of retrofit implementation requires specialised expertise. The elicitation required detailed technical engagement with recovery trajectories, operational constraints, intervention sequencing, implementation procedures, and expected service life. For this reason, expert relevance, technical depth, and contextual understanding were prioritised over sample size.

Mean and standard deviation are reported in the Supplementary Information only as descriptive indicators of the spread of expert judgements. They are not used to fit probability distributions, compute confidence intervals, or propagate probabilistic uncertainty through the final MCDA. In the baseline decision model, the median recovery trajectory is used to represent central recovery behaviour, while the upper quartile is used to represent adverse recovery conditions. These two descriptors are then incorporated into the resilience score before MCDA aggregation, consistently with the decision support purpose of the framework.

We also clarify that LCA and LCC targets are not obtained by propagating the mean or standard deviation of the seven expert responses. Expert elicitation informs recovery durations, operational continuity assumptions, implementation sequencing, and expected service life. By contrast, material quantities, environmental characterisation factors, unit cost data, and life cycle cost components are derived from engineering inventories, environmental datasets, and cost sources. LCA and LCC are therefore calculated through the inventory and cost model described in the manuscript. In response to the Reviewer’s concern, the LCC component has also been complemented by sensitivity analysis on the main cost drivers.

2) Amendments to the manuscript

The manuscript has been revised in the Methods section, particularly in the expert elicitation paragraph, the Fit for purpose resilience assessment subsection, and the MCDA subsection. Amended text in the Methods section is reported in what follows; it has been highlighted in green in the revised manuscript.

“Expert elicitation informed the definition of retrofit options, recovery durations, operational continuity assumptions, implementation sequencing, and expected service life. It did not define material quantities, environmental characterization factors, or unit cost data, which were derived from engineering inventories, environmental datasets, and cost sources. Accordingly, expert judgement affects the resilience trajectories

and service-life criterion, while LCA and LCC values are calculated from the inventory and cost model described below."

"The recovery time T_{Re} is treated as an uncertain expert elicited variable, reflecting variation in intervention duration, administrative delays, and operational continuity assumptions. Variation is represented through descriptive recovery conditions derived from the elicitation, namely minimum, first quartile, median, third quartile, and maximum values"

"The mean and standard deviation of elicited responses are reported only as descriptive indicators of variation across expert judgements. They are not used to fit a probability distribution or to perform probabilistic propagation through the MCDA. For the baseline decision model, the median recovery trajectory represents central recovery behaviour, while the upper quartile represents adverse recovery conditions."

2. I appreciate the authors including sensitivity in the TLC variable, however, the option identified in the main manuscript using a TLC of 30 months yields somewhat different results from either 24 or 36 months. There is a 1-2 line discussion in the manuscript on penalizing longer trajectory interventions- however conclusion is "resilience-oriented strategies remain favourable under strict recovery targets, whereas replacement-oriented solutions continue to perform comparatively well when long-term service-life extension is prioritized" - I am not sure this comes through in the sensitivity. Scenario #3 doesn't show up as particularly favorable in either case which is the scenario with maximum resilience score. What explains this?

1) Answer

We thank the Reviewer for this observation. We agree that the previous wording could imply that Scenario #3 becomes the preferred overall option under alternative control times, which was not our intention. In response to this comment, we carefully re-examined and verified the sensitivity analysis results to ensure that the reported rankings and their interpretation are correct. The verification confirmed that Scenario #3 consistently exhibits the strongest resilience performance, when the resilience criterion is considered directly, due to its shorter, more predictable recovery trajectory and its ability to maintain partial functionality during implementation. However, the results reported in Supplementary Note 5 correspond to overall MCDA rankings, not resilience-only rankings. Even under the resilience-oriented weighting scheme T4, resilience contributes 55% of the total score, while cost, environmental impact and expected service life jointly account for the remaining 45%. As a result, the scenario with the highest resilience index does not necessarily achieve the highest overall MCDA ranking.

By contrast, replacement-oriented solutions such as Scenario #4 remain competitive under weighting schemes that place greater emphasis on expected service life (T5), illustrating that the preferred intervention depends on the balance between recovery performance and long-term asset renewal objectives rather than on resilience alone. The sensitivity analysis therefore illustrates a policy-dependent trade-off between short-term recovery and long-term asset renewal, rather than a direct one-to-one correspondence between resilience score and overall ranking.

To address this point, we have revised the discussion in the manuscript to explicitly distinguish resilience performance from overall MCDA performance and to clarify that the TLC sensitivity analysis affects the relative importance of recovery performance within the broader multi-criteria evaluation, rather than determining the preferred option solely on the basis of resilience.

2) Amendments to the manuscript

The manuscript has been revised in the Ranking of the retrofitting strategies and the sensitivity analysis subsection. Amended text is reported in what follows; it has been highlighted in green in the revised manuscript.

"The sensitivity analysis shows that the effect of T_{LC} is reflected primarily through the resilience criterion, while the final MCDA rankings remain governed by trade-offs among all evaluation criteria. Scenario #3 consistently exhibits the strongest recovery-resilience performance; however, it does

not always achieve the highest overall MCDA score because, even under the resilience-oriented weighting scheme substantially weight is still assigned to life-cycle cost, environmental impact and expected service life. This distinction explains why replacement-oriented solutions, particularly Scenario #4, may remain competitive or become preferred under certain TLC assumptions, as their superior long-term service-life performance can offset weaker recovery performance.”