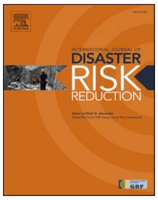
Citation (APA)
Cheshomi, A., Aydin, N. Y., & Comes, T. (2026). A scoping review and stress testing framework for High Impact Low Probability (HILP) events. *International Journal of Disaster Risk Reduction*, *142*, Article 106194.
https://doi.org/10.1016/j.ijdrr.2026.106194

International Journal of Disaster Risk Reduction 142 (2026) 106194

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

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



A scoping review and stress testing framework for High Impact Low Probability (HILP) events

Ali Cheshomi [*], Nazli Yonca Aydin, Tina Comes

Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX, Delft, the Netherlands

ABSTRACT

Societies are increasingly confronted with High-Impact Low-Probability (HILP) events. These events pose important challenges to societies as they disrupt critical infrastructures (CIs), the backbone of modern societies, leading to cascading and systemic disruptions across interconnected systems. Stress testing has emerged as a prominent approach for assessing system performance under adverse conditions. However, its suitability for addressing characteristics of HILP events, such as uncertainty, urgency, and complexity, remains unclear. This paper presents a scoping review of stress testing methodologies developed to date for terrestrial CIs, with the aim of identifying key methodological elements, with particular attention to the context of HILP events. The review reveals that existing stress testing approaches remain largely sector-specific and domain-specific, rely predominantly on hazard-centric scenarios, and insufficiently account for multi-sectoral interdependencies, dynamic system behavior, and the recovery and adaptation phases of resilience. Moreover, current methodologies tend to emphasize quantitative modeling, involve limited stakeholder participation, and lack mechanisms for iterative learning and adaptation, thereby constraining their relevance in rapidly evolving HILP contexts. In response to these gaps, this study proposes a conceptual framework for stress testing structured around three main stages of pre-assessment, assessment, and treatment. The framework emphasizes cross-sectoral and multi-domain analysis, stakeholder-inclusive and participatory approaches, and explicit consideration of recovery and adaptation processes. This study provides a foundation for advancing stress testing practices that are specifically tailored to HILP events and fosters the resilience of CIs.

1. Introduction

High-Impact Low-Probability (HILP) events are rare but potentially catastrophic, characterized by limited historical precedent and high uncertainty about their predictability and potential combinations of effects. These events often emerge as unexpected shocks that can severely disrupt infrastructure, services, and societal functioning [1]. From a hazard magnitude and exposure perspective, global phenomena such as climate change and geopolitical instability are increasing the probability and frequency of high-magnitude events [2]. They are also causing such events to occur in locations where they have not been experienced before. From a vulnerability perspective, the growing complexity and interdependence of critical infrastructures (CIs), driven by the demands of modern life, and reliance on aging yet connected infrastructures, further contribute to the emergence of HILP events through cascading failures. The interdependencies of CIs increase the likelihood that localized disruptions propagate across infrastructure sectors and escalate into large-scale systemic failures [3], i.e., a disruption in one sector can trigger cascading failures across others, potentially amplifying the initial impact and creating a "perfect storm" [4,5]. As such, HILP events may arise from both high-magnitude hazards with very low probabilities of occurrence (e.g., the 2023 Maui, Hawaii, wildfires) or from cascading effects that amplify the consequences of one or multiple triggers (e.g., the 2021 Suez Canal blockage, where a localized disruption triggered global supply chain and energy market impacts).

---

[*] Corresponding author.
E-mail address: a.cheshomi@tudelft.nl (A. Cheshomi).

https://doi.org/10.1016/j.ijdrr.2026.106194
Received 23 February 2026; Received in revised form 6 May 2026; Accepted 8 May 2026

The evolving nature of risks has revealed the limitations of traditional risk assessment approaches, prompting practitioners, policymakers, and researchers to rethink how infrastructure vulnerability and resilience are evaluated [6–9]. However, significant gaps remain in existing knowledge, limiting our ability to adequately assess the performance of CIs under emerging threats and to enhance their resilience [10,11]. In this respect, more advanced and standardized tools for assessing risk and resilience in CIs are required [11–13].

Among the methodologies developed to assess CIs against the changing hazard landscape, Stress Testing has emerged as a promising method [12–15]. Stress testing was originally developed in medicine and materials science to test physical strength and biological responses to stimuli [16–19] and was later adopted in finance, nuclear energy, and CI [17,20,21]. As such, stress testing has evolved into a prominent approach for assessing the vulnerability and resilience of infrastructure systems [11]. Its scenario-based nature allows for the exploration of system performance under extreme conditions designed to expose vulnerabilities and failure thresholds that are often neglected in conventional risk assessment [12,22].

Despite its potential and growing applications for diverse CIs the current stress testing landscape is highly fragmented, with substantial differences in objectives of application, sectoral focuses, modeling approaches, and their underlying modeling assumptions. Surprisingly, there is no review yet that scopes out the emerging field of stress testing research for Critical Infrastructures. In addition, for stress testing to live up to the ambition to reveal system vulnerabilities against HILPs, specific conditions such as extreme uncertainty and evolving disruption dynamics need to be at the heart of a stress test. Yet, because of the fragmented stress testing landscape, it is unclear how and to what extent the different ST paradigms are adequate for HILPs. What is missing is thus a cross-sectoral review of stress testing methodologies, along with an understanding of how they can be adequately tailored to prepare and evaluate our infrastructure for future HILP events.

To address this gap, this study is guided by four research questions designed to capture the core methodological dimensions of stress testing. We split the questions into a narrative and a methodological part. The first three questions are inspired by the scenario planning literature [23,24], especially for low-likelihood situations [25], focusing on the scenario narratives and outcomes that are underlying the ST: (1) *what* systems are tested and for which hazards? ("What is tested?"); (2) *how* are the impacts of a hazard and the related uncertainties modeled? ("How is it modeled?"); and (3) *what* outcome is measured, risk or resilience? ("What is measured?"). The methodological question focuses on the process. Wright and Goodwin [25] have shown that under conditions of low predictability, cognitive biases play an important role. Therefore, we investigate (4) *how* is the stress testing workflow designed? ("What is the stress testing workflow?"), especially if it invites diversity and broad participation to avoid groupthink. These questions from the scenario planning literature address the key challenges by capturing system scope and interdependencies, representing uncertainty and dynamics, evaluating performance beyond failure, and examining how stress testing is implemented. Together, these questions constitute the framework for analyzing stress testing methodologies, allowing us to identify what is needed to tailor CI stress tests specifically for HILP events.

Building on this analytical framework, this study provides a scoping review of the emerging body of academic research on stress testing methods applied across various CIs. Our study focuses on terrestrial CIs, i.e., chemical and nuclear sectors, transportation, water, food supply chains, ICT, energy, and general infrastructure, if a specific infrastructure is not mentioned. From our review, we identify gaps and analyze how current stress testing methodologies need to be extended for HILP events. In response to the identified gaps, we design a conceptual framework that synthesizes existing stress testing approaches into a unified structure that can be adapted to different contexts; moreover, it provides guidance on how stress testing processes can be systematically adjusted to address the specific challenges of HILP events.

This article is organized as follows. The next section provides a detailed description of the research methodology. Section 3 presents the results derived from the literature analysis. Section 4 discusses the key findings, elaborates on the proposed conceptual framework, identifies research gaps, and suggests avenues for future investigation. Finally, Section 5 concludes the study by summarizing the main insights, limitations, and implications for future stress testing methodologies for HILP events.

2. Methodology

In this study, we adopted a scoping literature review, as scoping reviews are a form of knowledge synthesis that enables the systematic mapping of evidence on a given topic and the identification of key concepts, theories, sources, and knowledge gaps [26]. This type of review is particularly suitable for examining emerging or heterogeneous bodies of literature [27,28].

Our review focuses on stress testing methodologies developed for terrestrial CIs. In the following parts of this section, we first describe the steps of the scoping review, from developing the search strategy to data extraction. We then explain how the extracted insights were analyzed using thematic analysis to propose a conceptual framework for stress testing of CIs in the context of HILP events. Fig. 1 summarizes the key data extraction points from the literature and the phases of the thematic analysis that result in the development of the conceptual framework, which is discussed in further detail in the following subsections.

2.1. *Scoping review*

This scoping review was conducted and reported in line with the PRISMA Extension for Scoping Reviews (PRISMA-ScR) guidelines to ensure a transparent and systematic identification, selection, and synthesis of relevant studies [26]. In the following sub-sections, details for each step of the scoping review are provided.

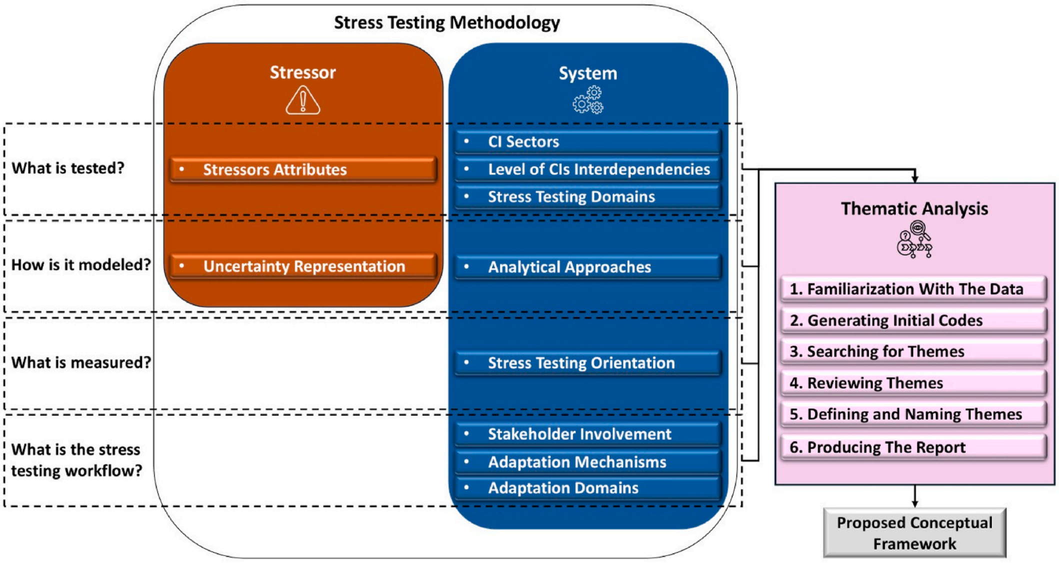


Fig. 1. Theoretical framework to guide the literature review.

2.1.1. *Database search strategy*

The search was developed and carried out in the Scopus database, selected for its extensive coverage across engineering, environmental, and social science disciplines [29,30]. Complementary searches were performed on the Web of Science (WoS) and Google Scholar. The WoS search returned fewer records and no additional unique studies beyond those already retrieved from Scopus, indicating substantial overlap between the two databases. Similarly, verification of the most highly cited publications in Google Scholar confirmed that these were also indexed in Scopus. To further enhance completeness, backward and forward citation tracking was performed to identify additional relevant studies not retrieved in the initial search [31]. Although reliance on a single database is acknowledged as a limitation, the choice of Scopus as the primary data source ensured transparency, reproducibility, and manageability of the review process [32,33].

In this study, we focused on terrestrial CIs, as they constitute the backbone of societal functioning and are particularly characterized by strong interdependency. Using the indicative list of CI sectors and services identified by EU Member States [34], the terrestrial sectors considered in this study include chemical and nuclear, transport and mobility, water, food (particularly its supply chain), information and communication technologies (ICT), and energy. Moreover, we distinguished whether each study is primarily oriented toward Risk stress testing or Resilience stress testing. While Risk stress testing focuses on identifying hazards, vulnerabilities, and failure points to prevent infrastructure collapse, Resilience stress testing examines the ability of infrastructure systems to absorb shocks, sustain critical services, and recover or adapt in the face of disruptions.

Based on this scope and analytical framework, the search strategy was designed to capture studies that address both risk- and resilience-oriented stress testing methods across terrestrial CI sectors. Accordingly, the search strategy used in this study was structured as follows: (a) risk OR resilience; (b) AND stress test; (c) AND (chemical OR nuclear OR transport OR water OR food OR ICT OR energy). Regarding the CI sectors, relevant terms for each sector were incorporated into the search string to capture the most suitable keywords for this study. The final search string was then used to query the Scopus database across titles, abstracts, and author-provided keyword fields. The Appendix provides the complete search string used for the bibliographic search, as well as the filters and limits applied to the database, which are summarized in Table 3.

2.1.2. *Eligibility criteria*

To support a transparent and systematic selection process, a set of inclusion and exclusion criteria was developed. A detailed description of inclusion and exclusion criteria is provided in the Appendix (Table 4), while the main elements are summarized below.

- **Domain Focus:** Studies that focus on terrestrial CIs, either generally or within specific sectors (e.g., chemical, nuclear, transport, water, food supply chain, ICT, or energy). Studies from non-terrestrial CI domains (e.g., finance, banking, clinical medicine, biological sciences, or humanities) were excluded.
- **Stress Testing Methodological Contribution:** Studies that develop, propose, apply, examine, or conceptualize stress testing methodologies for our CIs of interest. Studies using the term “stress testing” in unrelated contexts (e.g., physiological or biomechanical stress testing on humans, or in-situ geological or material stress testing) were excluded.
- **Event Context Scope:** Studies addressing disruptive events, including human-made incidents (e.g., human error or malicious attacks), natural hazards, technological failures, pandemics, contaminations, and other relevant events.

- **Study Type:** Only peer-reviewed journal articles and conference proceedings presenting methodological or conceptual contributions relevant to CI stress testing were included. Other sources, such as theses, dissertations, book chapters, reports, working papers, etc., were excluded.
- **Date and language:** No publication date restrictions were applied. To ensure consistency and comparability, only studies published in English were included.

2.1.3. *Study selection*

The Scopus database search was conducted in January 2026, and the references were imported into Rayyan [35]. After removing duplicates, the remaining articles were screened by at least two reviewers who independently assessed titles and abstracts. Any discrepancies were discussed and resolved with input from a third reviewer. After this step, in the final full-text screening stage, a set of studies met the eligibility criteria and were included for content analysis, as shown in the PRISMA diagram (Fig. 2).

2.1.4. *Data extraction*

To explore and extract data from the included studies, we developed a theoretical framework (Fig. 1). The framework distinguishes two dimensions in stress testing: the **Stressor**, applied to infrastructure systems, and the **System** that is stressed. Guided by these dimensions, we analyzed the literature through four main questions introduced in the introduction section: **(1) What is tested? (2) How is it modeled? (3) What is measured? (4) What is the stress testing workflow?**

Below, we further elaborate on the data extraction within each question. A full list of extraction points is provided in the Appendix (Table 5). These extraction points are explicitly aligned with core features of HILP complexity and uncertainty.

(1) What is tested?

To understand which stresses or shocks the stress testing tests against, we extracted **Stressor Attributes,** distinguishing *Primary Stressors*, i.e., the main stressor types that each study focuses on (e.g., earthquakes, human-made attacks, etc.), from *Secondary Stressors,* i.e., any additional stressors that either result from the primary stressor or occur alongside it. For example, a primary stressor, such as an earthquake, may lead to a secondary stressor, such as a tsunami or a landslide.

To characterize the CI system scope, the interdependencies, and the domain of stress testing examined, we extracted three classes of attributes from each article. The first relates to the **CI Sectors** (e.g., energy, transport, water supply, etc.) that are the focus of the study. The second concerns the **level of CI Interdependencies** considered, describing whether the analysis is conducted at a cross-sectoral level or focused on a specific sector.

Finally, in the **Domain of stress testing**, we assessed whether the studies extend beyond the physical domain to include others, such as organizational, economic, and cyber domains.

(2) How is it modeled?

For the Stressor, we investigated the **Uncertainty Modeling Approach** by identifying whether the study employs a deterministic, probabilistic, or a hybrid approach combining both in the scenario design.

Following Linkov and Kott [36], risk and resilience assessment methods were classified into metric-based and model-based analytical approaches. Accordingly, we analyzed the **stress testing approach**, based on whether the study uses metric-based

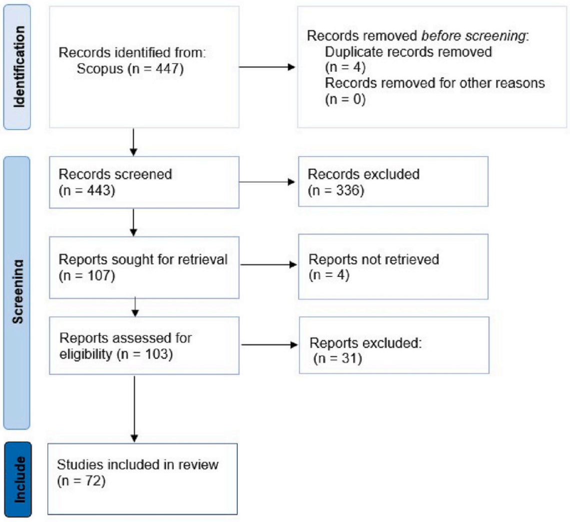


Fig. 2. PRISMA diagram—scoping review [26].

approaches (Individual metrics, indices, dashboards, and decision analytics), model-based approaches (statistical models, network-based methods, simulation-based analysis, scenario analysis, and spatial data analysis), or a combination of both. In addition, all the analytical approaches were further examined with respect to whether they adopt static representations of system performance or explicitly incorporate dynamic system modeling to represent temporal system evolution, cascading effects, or recovery processes.

(3) What is measured?

Under **Stress Testing Orientation,** we determined whether the methodology of each study is primarily oriented toward Risk stress testing or Resilience stress testing. In this review, Risk stress testing refers to methodologies that evaluate the likelihood and consequences of predefined hazard scenarios, typically focusing on failure probabilities, damage estimation, or threshold exceedance. In contrast, resilience stress testing frames stress testing around different stages of resilience rather than solely around hazard likelihood and evaluates how CI systems maintain their functionality under disruptions and how they recover and adapt over time. Therefore, if a study was classified as Resilience stress testing, we additionally recorded the specific resilience stages under investigation. In the literature, researchers have converged around the National Academy of Science (NAS) definition of disaster resilience as "*the ability to (1) prepare and plan for, (2) absorb, (3) recover from, and more successfully (4) adapt to adverse events*" [37,38]. The subdivision of resilience stress testing into these four stages allowed us to analyze the assumed dynamics and timing of the stress testing.

(4) What is the stress testing workflow?

We analyzed the selected papers to determine whether they proposed a **Framework or Workflow for stress testing** and extract these frameworks. This enabled us to identify the common steps and elements in different workflows developed for stress testing methodologies.

Since stress testing should ultimately lead to improved preparedness, we examined whether these frameworks **Involved Stakeholders** in the design or execution of stress testing, for example, in scenario development, or vulnerability assessment. In addition, we explored whether the frameworks include an **Iterative Adaptation** approach. Iterative adaptation in stress testing involves continuously refining strategies, models, or decisions based on new information, feedback, or evolving system conditions. We assessed whether this is explicitly incorporated, as iterative adaptation plays an important role when designing a methodology that takes into account the changing risk landscape. This is particularly important in contexts characterized by deep uncertainty, such as climate change-related events. The increasing complexity of modern society means that these events and their impacts are constantly changing and evolving. Beyond the general domain areas of stress testing, we also examined whether the studies address **Adaptive Strategies** more broadly. Specifically, we assessed whether stress testing methodologies focus solely on technical (physical) system adaptations (e.g., infrastructure hardening, redundancy, engineering upgrades) or whether they also consider non-technical strategies, such as organizational learning, institutional flexibility, policy change, behavioral adjustments, and governance innovation.

2.2. Thematic analysis of data for framework development

To analyze the diverse insights derived from the extracted data and to support the development of a conceptual framework for stress testing of CIs in the context of HILP events, we applied thematic analysis, adapting the approach proposed by Braun and Clarke [39] to the context of our study. The phases of thematic analysis enabled the systematic identification of methodological steps, recurring patterns, and gaps in the current literature regarding the design and application of stress testing methodologies. These insights directly informed the development of a new conceptual framework tailored to HILP events. Accordingly, after data extraction, we applied our structured thematic analysis, comprising six analytical phases, as discussed in more detail below.

Familiarization with the data: We began by familiarizing ourselves with the extracted data to better understand the scope and focus of each study. This involved reviewing the extracted data points and each extracted workflow or framework for stress testing of CIs for the studies that developed one.

Generating initial codes: Based on the results of extracted data, we systematically identified and organized common features across studies. In other words, for each extraction point (e.g., level of interdependencies), we noted common focuses. For instance, several studies focused on single-sector analyses. These recurring patterns were coded as initial analytical observations. In addition, we mapped and compared the procedural steps described in the identified workflows and conceptual frameworks, and when key elements were unclear, we reviewed the accompanying text to support more accurate categorization.

Searching for themes: Next, we grouped similar steps and organized them into preliminary phases that represent common methodological stages or processes in existing stress testing approaches. These phases include, for example, pre-assessment (e.g., scoping, problem structuring, and scenario definition), assessment (e.g., scenario simulation), and treatment (e.g., monitoring and adaptation). In addition, we mapped the recurring patterns of extracted data points onto these same stages. For example, a focus on single-sector analysis relates to the pre-assessment phase, where decisions about the scope of the problem are made.

Reviewing themes: We refined these preliminary themes through an iterative process, assessing whether the proposed stages and components of our emerging conceptual framework aligned with the frameworks and workflows consistently applied across the reviewed studies.

Defining and naming themes: Once refined, each theme was clearly defined and named to represent a distinct stage or component of the stress test process. These included, for example, Pre-assessment, Assessment, and Treatment, along with their corresponding subthemes. Moreover, we developed Table 1, in which, for each extraction point, we summarized the requirements associated with HILP

Table 1
Summary of HILP requirements across the extraction points.

Extraction Point	Aspect	Requirement for HILP
- CI Sectors - Level of CI Interdependencies	Sectoral Scope	HILP events require a systemic, cross-sectoral stress testing approach that explicitly maps and analyzes interdependencies among CIs, as failures propagate through feedback loops and cascading effects that cannot be captured through single-sector assessments alone [3,11,40–42].
- Domain of stress testing	Domain Coverage	HILP events require stress testing methodologies that explicitly incorporate and analyze interactions across the physical, informational, cognitive, and social domains, as disruptions fracture the integrity of these interconnected domains [11,38,42,43].
- Stressor Attributes - Uncertainty Modeling Approach	Stressor Definition	HILP events require threat-agnostic rather than hazard-specific modeling, as predefined scenario-based approaches fail to capture the complexity of cascading failures, emergent threats, and environmental interdependencies. As the threat landscape facing CI is continuously expanding, such events are characterized by nonlinear escalation, deep uncertainty, and unpredictable cascading dynamics [1,11,42].
- Stress test Analytical Approach	System Modeling	HILP events require an approach that, rather than focusing on individual hazards, emphasizes the system's inherent qualities and capabilities, thereby enabling adaptability and robustness in the face of unforeseen challenges [42,44,45].
- Stress test Orientation	Modeling Orientation	HILP events require a resilience-based stress testing approach that enables threat-agnostic and scalable analysis beyond component-level risk assessments, focusing on system structure, interdependencies, and performance under degraded conditions to strengthen the capacity to absorb, recover, and adapt amid increasing complexity and cascading disruptions [11].
- Specific Resilience Stages	Resilience Stages	HILP events require stress testing frameworks that evaluate how systems maintain critical functions during disruptions, recover from degraded states, and adapt over time, independent of specific hazard scenarios [11,42].
- Involvement of Stakeholders	Stakeholder Engagement	HILP events require active stakeholder involvement in stress testing processes to assess technical and governance interdependencies, address information gaps and misinformation, and evaluate soft functions and coordination dynamics that may act as single points of failure in complex, multi-scalar systems [42,46].
- Iterative Adaptation	Iterative Learning	HILP events require iterative and adaptive stress testing processes, as resilience is a dynamic capability that must continuously evolve with changing threat conditions and shifting system interdependencies [11,44].
- Adaptive Strategies	Adaptation Strategies	HILP events require multi-domain adaptive strategies that go beyond physical protection measures to strengthen social coordination, cyber resilience, and financial robustness, given that disruptions propagate across these interconnected domains [42,43].

events, citing the relevant academic literature.

We then compared the recurring patterns in the extracted data points (e.g., a predominant focus on single-sector analyses) with the requirements of stress testing for HILP events (e.g., the need to account for cross-sector interdependencies and cascading effects). This comparison enabled us to identify gaps and limitations in existing methodologies and to assess whether adjustments were required to better address HILP events in stress testing.

Producing the report: Finally, we consolidated the framework and discussed how the identified focuses and methodological patterns observed across studies could be adapted to HILP events. This step involved reflecting on how existing approaches could be extended to better capture uncertainty, interdependencies, and systemic resilience under extreme conditions.

3. Results

3.1. Overview of the stress testing literature

The 107 studies identified span publications from 2007 through the mid-2020s, with relatively limited activity prior to 2014 and a marked increase over the last decade, indicating growing scholarly interest in stress testing for CIs (Fig. 3). This rise coincides with greater attention to infrastructure resilience in policy and research [34,47].

Fig. 4 distinguishes case studies (n = 62) and conceptual papers (n = 10) and maps this methodological scope onto the geographic distribution of the included studies. Among the case-study applications, the majority focus on high- and upper middle-income contexts (e.g., [14,48,49]), with a strong concentration in Europe (n = 19), followed by North America (n = 9), Asia (n = 8), and Australia (n = 5). A smaller number of studies adopt a synthesized or generic case-study setting (n = 10) that is not tied to a specific geographic location.

Studies explicitly addressing low- and lower-middle-income contexts remain limited (e.g., Lachaut et al. [50]; Ray et al. [51]; Sunkara et al. [52]), with case studies concentrated primarily in Asia (n = 4) and Africa (n = 3). Overall, the distribution highlights a

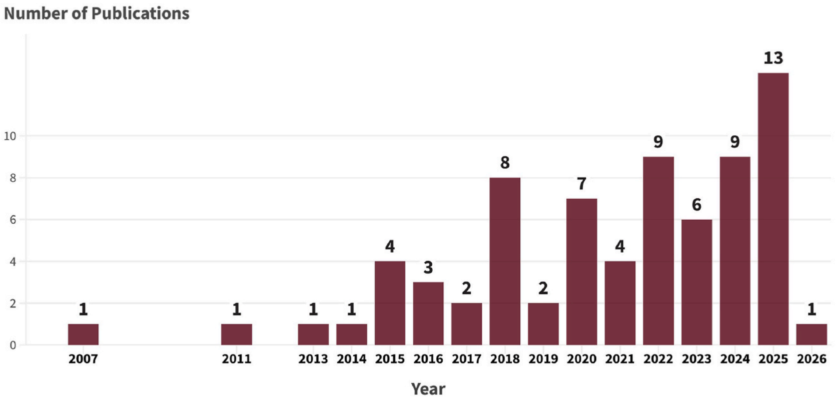


Fig. 3. Number of publications per year.

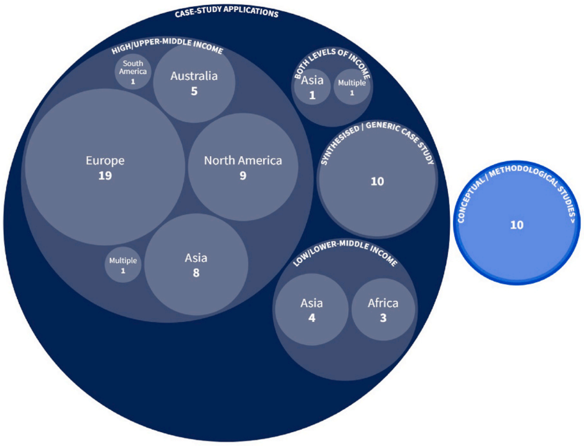


Fig. 4. Methodological scope, income level & geographic scope.

clear imbalance in the empirical application of stress testing methodologies, with substantially greater coverage of high- and upper-middle income regions as compared to middle- and low-income contexts. This imbalance may reflect differences in data availability, funding, and analytical capacity across regions.

3.2. What is tested? Sector, level of interdependencies, domain and stressor

Single-Sector vs. Cross-Sectoral Focus and Sectoral Coverage: The majority of the reviewed studies focus on a single infrastructure sector in isolation, with 66 out of 72 studies (around 92%) adopting a sector-specific perspective (Fig. 5). Within these single-sector studies, water systems dominate the literature (n = 31), followed by transportation (n = 20) and energy systems (n = 17). These studies commonly examine water supply systems (e.g., Becher et al. [53]; Verbist et al. [54]), transportation networks such as road and port systems (e.g., Aydin et al. [22], Clarke and Obrien [14], Pitilakis et al. [55]), and energy infrastructures including hydropower generation and electric power grids (e.g., [56]).

In contrast, cross-sectoral studies remain limited (n = 6 out of 72) and mostly adopt a general or conceptual perspective (e.g., Comes and Bertsch [57], Linkov et al. [11]), rather than modeling interdependencies across specific CI sectors. Only a small number of studies explicitly examine interactions between sectors, such as combined water–energy systems (e.g., Bhave et al. [58]) or broader multi-sector configurations including transportation, water, and energy (e.g., Jovanović et al. [59]). Other critical sectors, such as chemical infrastructure, are largely absent from both single-sector and cross-sector stress testing applications.

Overall, the evidence indicates that most stress testing studies focus on a narrow subset of infrastructure sectors and adopt predominantly sector-specific perspectives, with limited development of approaches capable of capturing interdependencies across CI systems. Given the increasing interdependence of critical infrastructures, there is a broad consensus that their performance cannot be evaluated in isolation from their wider environment or from other infrastructures upon which they depend [60]. This gap limits the ability of current studies to capture risks that emerge from cross-sector interactions. Such a gap in understanding can lead to ineffective coordination and response during emergencies [61] or insufficient preparedness as risks are not adequately evaluated [62].

Domain: Across the reviewed studies, Fig. 6 illustrates that the majority focus exclusively on the physical domain, with 42 out of 72 studies adopting a purely technical perspective (e.g., Bourgin and Le-Cler [63], Choine et al. [64], Lam et al. [65], and Spence and Brown [66]). The remaining 30 studies incorporate at least one additional domain, most commonly economic (n = 13) (e.g., Lam et al. [65], Rehman et al. [67], and L. Xiao et al. [68]) and organizational (n = 9) dimensions (e.g., Jeuken et al. [69], Schätter et al. [70], and Zhang et al. [71]), while social (n = 7) (e.g., Bhave et al. [58] and Lachaut et al. [50]) and cyber (n = 6) domains (e.g., Kilic [72],

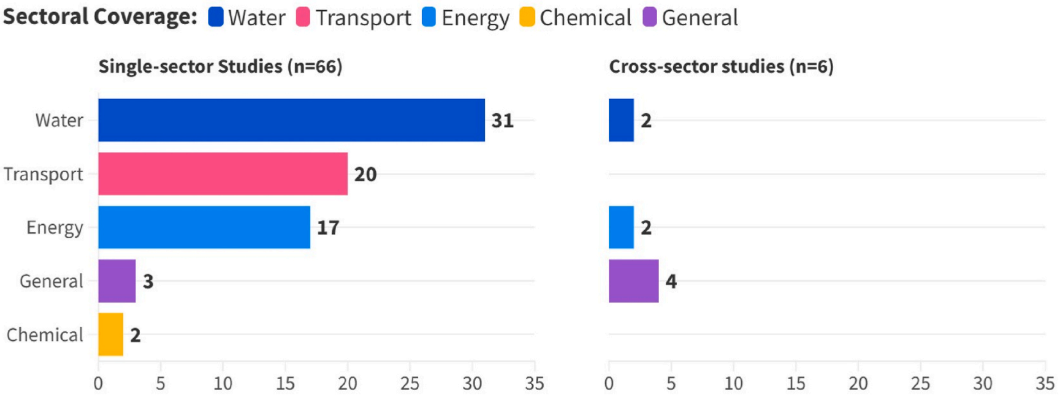


Fig. 5. Comparison of sectoral coverage in single-sector and cross-sectoral studies.

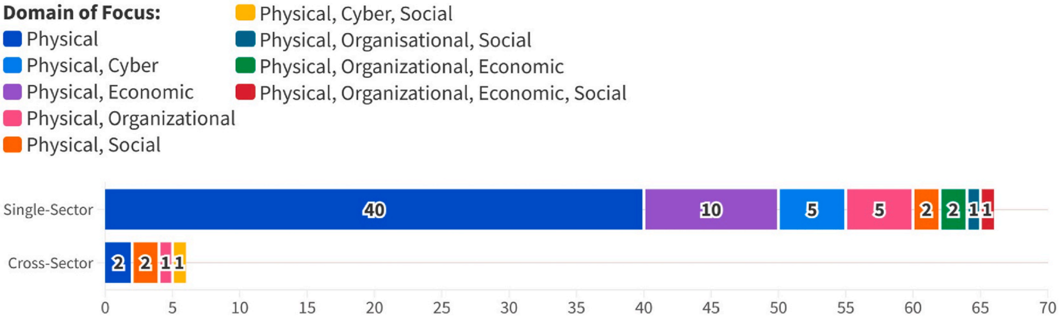


Fig. 6. Sector level and domain of focus.

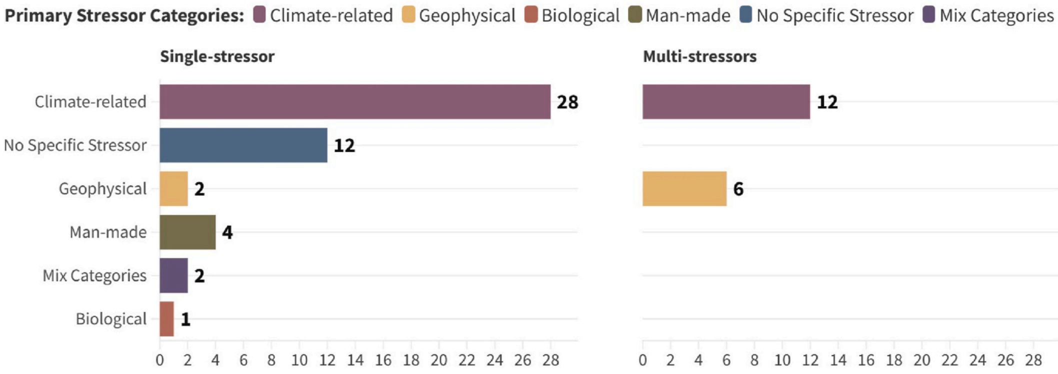


Fig. 7. Overview of primary stressors categories and use of single vs. multi-stressor scenarios in studies.

Nikolopoulos et al. [73], and Nikolopoulos et al. [74]) appear less frequently. Within single-sector studies, multi-domain approaches most often combine the physical domain with economic (n = 10), cyber (n = 5), or organizational (n = 5) aspects, whereas more complex configurations involving three or more domains remain rare. Cross-sector studies show a similar pattern, with two studies focusing only on the physical domain, while the remainder incorporate only one or two additional dimensions.

Despite the inherently multi-domain nature of CI systems, spanning physical, organizational, cyber, economic, and social dimensions, explicit integration across multiple domains remains limited. While some contributions explore cyber–physical interdependencies or discuss multi-domain stress testing from a conceptual perspective, the literature overall continues to prioritize technical and physical analyses over broader cross-domain interdependencies. This results in comparatively limited attention to the cyber, organizational, economic, and social dimensions, as well as to the interdependencies that connect these domains. This limited perspective is concerning because the non-technical domains can also shape how disruptions evolve, how cascading effects propagate, and how effectively systems and communities respond to disastrous events [3,60,75,76].

Stressor Characterization: Across the reviewed studies, environmental shocks emerged as the dominant class of primary stressors, with climate-related events emerging as the most frequently considered class of primary stressors (Fig. 7). Climate related stressors such as floods and droughts are predominantly examined in single-stressor settings (n = 28) [48,77,78], but also feature prominently in multi-stressor applications (n = 12). Within studies adopting climate-related multi-stressor approaches, the additional stressors most commonly incorporated alongside the primary hazard are demand-related changes (e.g., [50,52,54,79]).

Geophysical hazards, primarily earthquakes, constitute the next most frequently addressed class of environmental stressors, although they appear less often overall, with only two single-stressor [22,80] and six multi-stressor studies focusing on such events. Among the multi-stressor earthquake-focused studies, studies mostly consider cascading hazards, including landslides, tsunami impacts, and soil liquefaction (e.g., Argyroudis et al. [12], Choine et al. [64], Clarke and Obrien [14], and Pitilakis et al. [55]).

Human-made stressors, including cyber-physical attacks, remain sparsely represented (n = 4) [72–74,81], while biological stressors are almost absent, appearing as a primary stressor in only one study [70]. Notably, a substantial number of studies (n = 12) adopt a nonspecific threat approach or do not define a particular triggering hazard (e.g., Gauthier et al. [82], Jovanović [83], Kaulbars et al. [84], Liao and Ilic [85], and Z. Xiao and Bai [86]).

An important distinction across the reviewed stress testing studies is whether they adopt a nonspecific or threat-agnostic approach to modeling stressors. As the results show, however, most studies rely on specific stressors, such as earthquakes [55], cyber-physical attacks [81], or climate change-induced droughts [51]. While such specificity is valuable for hazard-focused risk analysis and can support stakeholder engagement by making scenarios more tangible for them to process, it may constrain the generalizability of the methodology when addressing unknown or unpredictable events. In contrast, studies such as Jovanović et al. [59] and Xu et al. [87] do not focus on a specific type of threat and discuss their frameworks without predefining any type of hazard. These approaches may enable the assessment of the system independent of any special trigger event, meaning the methodology can be adapted to different contexts.

3.3. How is it modeled? Uncertainty and analytical approach

Uncertainty Modeling Approach: Fig. 8 illustrates the distribution of uncertainty modeling approaches across infrastructure sectors. Overall, hybrid approaches, combining deterministic and probabilistic elements, are most frequently employed, particularly in single-sector studies (e.g., Argyroudis et al. [12], Gauthier et al. [82], John et al. [88], and Z. Yang et al. [49]).

This approach has strong representation in the water (n = 17), transport (n = 8), and energy (n = 7) sectors, reflecting an effort to balance defined disruption scenarios with explicit representation of uncertainty. In comparison, deterministic and probabilistic approaches appear less frequently and are largely confined to the same sectors (water, transport, and energy). Despite the presence of cascading effects and compounding uncertainties that could benefit from more flexible modeling strategies, cross-sector studies remain limited in the modeling of stress scenarios, with only a few examples of hybrid deterministic and probabilistic approaches across different sectors [59,71,83].

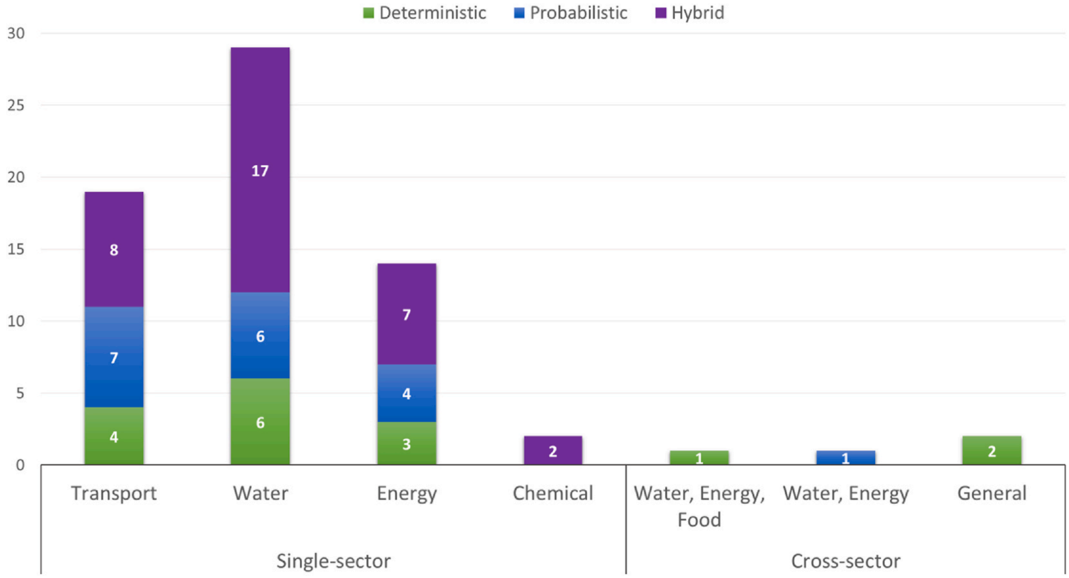


Fig. 8. Scenario modeling approaches across infrastructure sectors.

Across the studies, exploratory scenario approaches are commonly used to address uncertainty. Many studies incorporated probabilistic components, such as stochastic modeling, Monte Carlo simulation, or scenario randomization, to explicitly represent uncertainty [51,89]. Hybrid approaches, which combine deterministic modeling with probabilistic approaches, were frequently employed to broaden the range of potential disruptions and reduce dependence on fixed assumptions [13,55,65,90]. These hybrid frameworks enable the examination of systems under multiple plausible futures, enhancing the robustness of evaluations under uncertain and evolving conditions. By contrast, only a limited number of studies relied exclusively on deterministic approaches [91]. Such approaches can facilitate the analysis of events or hazards that cannot be incorporated into probabilistic risk analyses due to methodological limitations [12]. Consequently, the limited use of deterministic approaches constrains the ability of existing analyses to adequately capture rare, highly uncertain events, such as HILP events.

Overall, hybrid modeling dominates in single-sector studies, likely due to its flexibility in integrating both defined disruption scenarios and probabilistic variability. However, hybrid approaches are not commonly observed in cross-sector contexts, where the complexity and cascading uncertainties of events could benefit from more flexible modeling strategies. Moreover, although deterministic approaches have proven to be useful to take into account highly uncertain and rare events in the analysis, they have been less commonly used compared to probabilistic approaches in general. Therefore, this can limit the ability of existing stress testing methodologies to adequately capture uncertainty, unprecedented disruptions, and complex cascading effects.

Analytical (Model-Based and Metric-Based) Approaches: We categorized the reviewed studies according to their primary analytical approach, distinguishing between model-based, metric-based, and hybrid (model-based and metric-based) frameworks. Fig. 9 highlights a clear contrast between single-sector and cross-sector applications. Studies focusing on a single infrastructure sector predominantly rely on model-based approaches, with 56 out of 65 single-sector studies using exclusively model-based methods (e.g., Bouziotas et al. [92], Lam et al. [65], and Nasrazadani et al. [93]), while eight studies combine model-based and metric-based approaches (e.g., Goodwill et al. [94], Ray et al. [51], and Verbist et al. [54]). Among cross-sector studies, although their overall number is low, most employ model-based and hybrid approaches, and no study relies solely on a metric-based approach. Of the five cross-sector studies, three adopted hybrid assessment, integrating decision analysis, resilience indices, and scenario-based simulation to capture a broader range of system attributes. Jovanović et al. [59] and Linkov et al. [11] further extend these hybrid approaches by proposing the incorporation of network analysis to explicitly represent interdependencies across sectors. The remaining two studies, Bhave et al. [58] and Zhang et al. [71] rely primarily on model-based approaches, using mainly scenario analysis and simulation to evaluate system performance.

Fig. 10 further shows that, when disaggregated by sector, the dominance of model-based approaches remains consistent across infrastructure domains, including water, energy, and transport systems. Most studies rely on scenario-driven simulation [63,77,86],

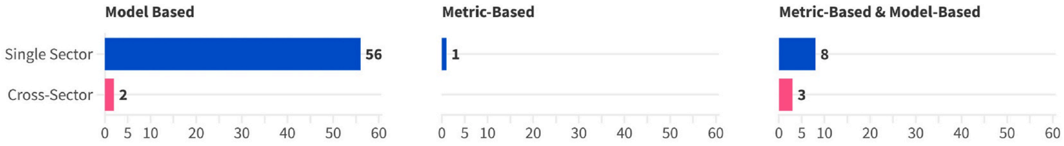


Fig. 9. Analysis approaches and the level of analysis.

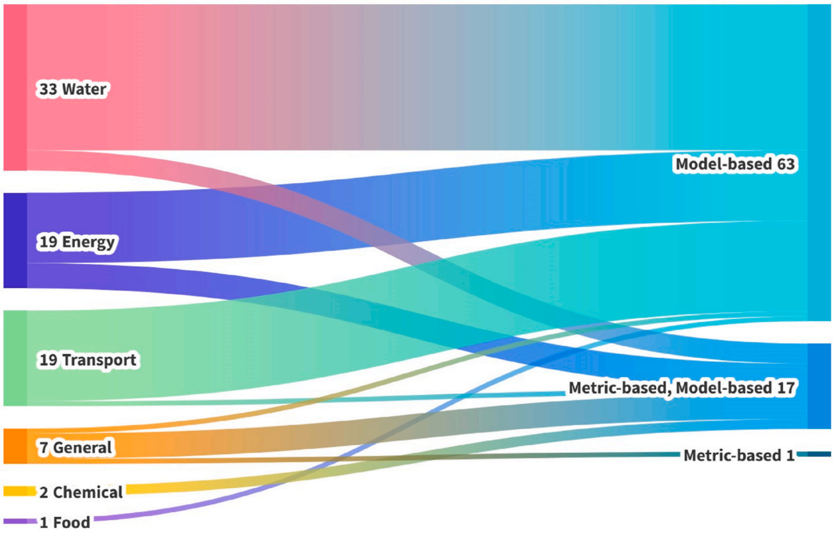


Fig. 10. Sector and analysis approaches.

often complemented by spatial [22,66], network-based [65,95], or statistical modeling approaches [78].

The methodological differences observed between single-sector and cross-sector studies may be linked to variations in data availability and system complexity. Single-sector systems are generally easier to represent in detail, which supports the use of model-based approaches. By contrast, cross-sector or more generic system representations are harder to fully parameterize, leading studies to rely more on hybrid approaches rather than solely model-based approaches. The limited use of metric-based approaches represents an additional shortcoming, particularly in complex systems where such methods can enhance understanding, especially in contexts with limited data availability, for example, through stakeholder engagement.

Furthermore, most modeling approaches adopt static representations of infrastructure performance, with only a limited subset of studies implementing dynamic system modeling to capture temporal system evolution and cascading effects (e.g., [56,85,89]). This pattern may reflect a combination of data limitations, computational complexity, and the analytical scope of existing studies.

Overall, these results indicate that while analytical approaches vary with system scope, current stress testing studies rely mainly on quantitative methods, with qualitative approaches rarely implemented. This reliance reflects a strong preference for formal, data-driven techniques, but may limit the ability of stress testing methodologies to capture contextual, organizational, and decision-making dimensions that are critical in complex disruption scenarios, especially in situations where extensive data are not available [96]. Furthermore, quantitative approaches, whether model-based or hybrid, are most often applied in a static manner, suggesting that current methodologies are not well equipped to represent evolving system dynamics, cascading effects, or recovery processes over time. Although hybrid analytical frameworks are proposed in some cross-sector studies, they are frequently presented at a conceptual level and remain limited in both number and practical implementation, indicating a gap between methodological development and real-world application.

3.4. What is measured? Risk vs. resilience stress testing

Stress Testing Orientation: Fig. 11 illustrates the resilience versus risk orientation of stress testing approaches across the reviewed studies. Of the 72 studies included in the review, the majority (n = 47) adopt a resilience-oriented stress testing framework (e.g., Aydin et al. [22], Bhave et al. [58], Jovanović et al. [59], and Rokstad et al. [97]), while 25 studies are primarily risk-oriented (e.g., Argyroudis et al. [12], Clarke and Obrien [14], Esposito et al. [13], and Z. Xiao and Bai [86]). Notably, all cross-sector studies (n = 6) fall within the resilience-oriented category [11,57–59,71,83], whereas risk-oriented approaches are observed exclusively in single-sector applications. Among single-sector studies (n = 66), both risk- and resilience-oriented frameworks are represented, although resilience-based approaches are more prevalent overall (n = 41) (e.g., Amestoy et al. [98], Kilic [72], L. Xiao et al. [68], and Yin et al. [99]).

Resilience Stages: A closer examination of resilience-oriented studies shows that most analyses address combinations of resilience stages rather than focusing on a single dimension. Robustness (or absorption) emerges as the central component across all resilience-oriented applications (n = 47). The most common combinations pair robustness with adaptation (n = 13) (e.g., Lachaut et al. [50], Ray et al. [51], and Z. Yang et al. [49]) or recovery (n = 11) (e.g., Aguila Téllez et al. [100], Bai et al. [101], and Rokstad et al. [97]). By contrast, only a small number of studies (n = 3) comprehensively address all four resilience stages, preparedness, robustness, recovery, and adaptation, within a single framework [59,83,94]. These findings indicate that although robustness is consistently incorporated, comparatively less attention is given to the equally important dimensions of adaptation and recovery, and comprehensive stress testing approaches that systematically cover all resilience stages remain limited. This trend may partly reflect methodological and practical

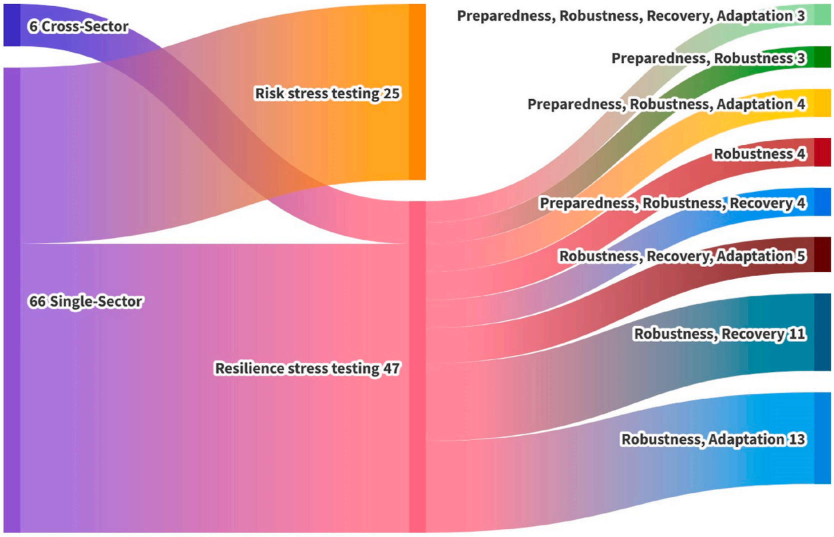


Fig. 11. Sector level and risk or resilience orientation of approaches and resilience stages of interest.

considerations, since robustness and absorption are often more straightforward to operationalize through simulation-based performance assessment, whereas preparedness, recovery, and long-term adaptation require broader institutional, organizational, and governance considerations that are more difficult to quantify and model within traditional stress testing frameworks.

Cross-sector studies demonstrate a stronger tendency to integrate multiple resilience dimensions, including adaptation and recovery, whereas single-sector applications more commonly concentrate on robustness-related system performance under stress. This may be due to the inherently systemic perspective of cross-sector analyses, which more explicitly address interdependencies and cascading effects, thereby necessitating consideration of recovery and adaptive capacity beyond the immediate system performance.

Among the cross-sector studies, Comes and Bertsch [57] propose a dynamic resilience-based approach aimed at improving the robustness of CIs against unrecognized and unexpected hazards. Linkov et al. [11] present a tiered resilience-based stress testing framework that focuses on absorption, recovery, and adaptation, though their approach remains largely conceptual. Bhave et al. [58] implement a resilience-based stress test in the Lake Malawi Shire River Basin, evaluating system robustness under extreme conditions while also assessing adaptation strategies such as revised operational rules and contingency planning. Jovanović et al. [59] propose and implement a framework that explicitly incorporates all resilience stages, preparedness, absorption (robustness), recovery, and adaptation.

Overall, these findings suggest that although resilience has become a dominant conceptual framing for stress testing, empirical applications remain concentrated on robustness. The implementation of stress testing across multiple sectors and across the full spectrum of resilience stages remains comparatively limited. This narrow emphasis resulted in stress tests characterized by static, single-system performance analysis and a focus on component failures, rather than simulating the evolving states of interconnected systems and exploring cascading failures over time. As a consequence, far less attention is devoted to recovery and adaptation, stages that determine the long-term resilience of CI systems under conditions of extreme uncertainty. As Esposito and Stojadinović [15] argue, assessing only the risk profile of a civil infrastructure system does not provide sufficient insight into its capacity to function and recover after a disaster. Traditional risk-based approaches typically capture immediate losses following a disaster but overlook how community needs and service demands change over time and how infrastructure systems, such as water, gas, and electricity, gradually restore their capacity to meet those needs [11,102,103]. The temporal dimension is therefore essential [104,105], as the dynamics of service restoration, recovery trajectories, and the adaptive behavior of communities are more accurately represented through resilience-informed modeling than through probabilistic risk assessments.

3.5. What is the stress testing workflow? Research frameworks, stakeholder involvement, iterative adaptation, adaptive strategies

Among the 72 reviewed studies, 28 studies present some form of a conceptual framework or workflow for the stress testing of CIs. As illustrated in Fig. 12, the majority of these frameworks are sector-specific, with most developed for the water sector (n = 10) (e.g., He et al. [91], Koh et al. [78], and Verbist et al. [54]) and the transport sector (n = 10) (e.g., Aydin et al. [22], Nasrazadani et al. [106], and Yin et al. [99]), followed by the energy sector (n = 4) (e.g., Bourgin and Le-Cler [63], Liao and Ilic [85], and Xu et al. [87]).

The literature also includes a smaller number of generic frameworks (n = 4) without focusing on a specific sector. For example, Esposito et al. [13] propose a structured stress testing workflow consisting of four phases of Pre-assessment, Assessment, Decision, and Reporting, and explicitly describe interactions among key actors during expert consultation and input integration processes. Although the approach has been applied to different sectors separately, it does not explicitly account for interdependencies among CI. Comes and Bertsch [57] introduce an iterative stress testing framework that couples stress test design with participatory approaches and

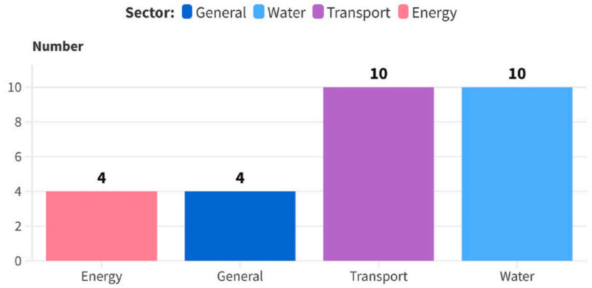


Fig. 12. Conceptual frameworks and workflows identified in the literature for stress testing by sector.

incorporates dynamic elements, such as evolving external trends, changing preferences, risk aversion, and societal goals. Their methodology emphasizes system-inherent vulnerabilities and adopts a generic perspective that is not tied to specific initiating hazard events. Similarly, Linkov et al. [11] develop a tiered framework for integrated risk and resilience stress testing of CI systems, structured to move from preliminary screening assessments to increasingly detailed quantitative analyses. More recently, Jovanović [83] presents a resilience indicator-based methodology for stress testing infrastructures exposed to adverse events and polycrises. Their framework is structured into four phases of preparation, testing, closing, and post-testing, and considers both threat-side and asset-side dimensions of resilience. These frameworks integrate threat identification, indicator-based resilience assessment, scenario development, interdependency analysis, and standardized reporting within a coordinated multi-actor framework. However, despite these advances, a comprehensive framework that synthesizes the different existing approaches into an adaptable structure applicable across different contexts is still lacking. Such a framework could help guide the future development of stress testing methodologies for CIs across varying infrastructural, organizational, and geographic settings.

As mentioned in the methodology section and illustrated in the theoretical framework guiding this review, in addition to analyzing existing frameworks and workflows for aspects such as their procedural steps, we also examined whether stakeholders are explicitly included as part of the framework or workflow due to the high importance of their perspectives and requirements in the stress testing analysis. Furthermore, we assessed whether the proposed frameworks adopt an iterative and adaptive approach and whether adaptation is limited to technical aspects or extends beyond them. These points are elaborated on in more detail below.

Stakeholder Engagement: Stakeholder engagement is not addressed consistently across the reviewed stress testing frameworks. Fewer than half of the studies (n = 13 out of n = 28) explicitly include stakeholder perspectives in the design or execution of stress tests (e.g., Albano et al. [90], Comes and Bertsch [57], Esposito et al. [13], Jovanović [83], Linkov et al. [11], and Pitilakis et al. [55]), while the remaining studies either exclude stakeholder involvement or do not explicitly mention it. Among the frameworks that do engage stakeholders, the level and form of involvement vary considerably, ranging from expert-centered interaction workflows [13,55] to approaches that either directly involve stakeholders in the stress testing process or explicitly incorporate stakeholder preferences and priorities into the analysis [57,78,90]. Taken together, these findings indicate that stakeholder engagement is not systematically embedded within stress test frameworks and lacks a consistent mode of implementation. This may stem from the predominant focus on model-based and quantitative approaches, in which stakeholders are typically involved only in a consultative role rather than as active participants in the analysis. This narrow engagement constrains the ability of stress tests to incorporate contextual knowledge, behavioral dynamics, organizational cultures, social norms, and indigenous knowledge, all of which strongly influence how systems respond and adapt during real crises [1,107–110]. While certain operational constraints can be represented and tested through models, these deeper socio-cultural and behavioral dimensions are often underrepresented, despite their critical role in shaping system resilience.

Iterative adaptation: Iterative adaptation in stress testing refers to the explicit use of feedback and model updating to refine strategies, models, or decisions as system conditions and information evolve. Among the reviewed studies, only a limited number explicitly incorporate an iterative process (n = 9 out of n = 28). These studies mostly implement the stress test iterations through dynamic stress testing frameworks that allow for progressive refinement of analysis, scenario updating, and strategy adjustment over time. Iterative elements are commonly supported by mechanisms such as stakeholder feedback, adaptive decision-making processes, or time-sequenced simulations [11,52,57,85,90].

Overall, the predominance of non-iterative approaches highlights a significant gap in the ability of current frameworks to represent dynamic risk evolution and support adaptive decision-making. This limitation is particularly critical given that CI systems are complex adaptive systems composed of numerous interacting components whose behavior evolves over time [41,111]. As noted by Comes and Bertsch [57], designing such systems requires approaches that account not only for external trends but also for changing societal preferences, risk perceptions, and goals. This, in turn, implies that stress testing methodologies must be capable of continuous adjustment as new information emerges and conditions evolve. However, most existing frameworks rely on one-off evaluations and fixed models, offering only limited capacity for iterative adaptation.

Adaptive strategies beyond technical: This dimension captures the extent to which stress testing methodologies address not only technical system adaptations (e.g., infrastructure hardening, redundancy, or engineering upgrades) but also non-technical strategies such as organizational learning, institutional flexibility, policy change, behavioral adjustments, and governance innovation. Among the reviewed frameworks, most studies limited their focus to adaptation to the technical domain, often overlooking broader dimensions. Only a small number of studies (n = 6 out of n = 27), such as Albano et al. [90], Bourginand and Le-Cler [63], Bouziotas et al. [92], Linkov et al. [11], and Schätter et al. [70], considered targeted adaptation strategies beyond the technical level. These studies incorporated a variety of intervention options beyond physical aspects, such as organizational coordination, policies, business continuity planning, and operational flexibility, thus offering a more holistic approach to resilience. Overall, focusing primarily on technical adaptation measures leads many studies to overlook organizational, policy, and coordination strategies. These measures often enable adaptive behavior and may, in some cases, offer more cost-effective improvements to resilience than purely technical solutions.

4. Discussion

In this section, we synthesize the results of the scoping review by reflecting on the major gaps identified across the reviewed studies and discussing the limitations these gaps introduce for the analysis in the context of HILP events. Based on our findings, we present a new conceptual framework to align existing stress testing methodologies with the requirements of HILP events.

4.1. Synthesis

Building on the analytical structure developed in Table 1, which defined the key aspects of stress testing analysis considering the extraction points and articulated their associated HILP-oriented requirements, we now examine how gaps in existing stress testing approaches perform against these requirements. We begin by discussing six major categories of knowledge gaps we identified across the reviewed studies, with each subcategory (gap) representing a methodological or conceptual constraint in current stress testing practices. For each gap, we explain its limitations with respect to the requirements of HILP events and propose directions for advancing research and practice.

At the end of this synthesis, Table 2 summarizes the dominant themes in the current stress testing literature, highlights their gaps, and links them to the limitations of these approaches when examined through the lens of HILP events. In addition, Table 6 in the Appendix provides a complementary overview of how studies proposing explicit stress testing frameworks or workflows perform against the identified HILP-oriented requirements, highlighting their key strengths and limitations.

4.1.1. Lack of cross-sector stress testing

Scholars studying infrastructure interdependency have long noted that failures do not remain confined to the originating system. Cascading risks are recognized as a major form of failure within interdependent CI networks [41]. These cascading effects arise from interdependent infrastructure networks [112,113], where feedback loops can significantly amplify even minor disruptions [114]. Therefore, given the growing interdependencies of the aging, tightly coupled, and complex infrastructure systems, it is crucial to

Table 2

Synthesis of main focuses, gaps in the current literature on stress testing approaches for CIs, and their limitations in addressing HILP events.

Gap No.	Aspect	Major focus of traditional approaches	Shortcoming for HILP events
1. Lack of Cross-Sector Stress Testing			
	Sectoral Scope	Single-sector stress tests (e.g., water, transport, energy)	Fails to capture cross-sector interdependencies and cascading failures; HILP disruptions propagate across multiple interconnected systems.
2. Limited Multi-Domain Perspectives and Non-Technical Adaptation Strategies			
2.1	**Domain Coverage**	Predominant focus on the physical domain of infrastructures	Overlooks non-technical domains, particularly organizational dimensions and interdependencies across different domains, which are central to shaping HILP impacts and response capacity.
2.2	**Adaptation Strategies**	Emphasis on technical measures (e.g., infrastructure hardening)	Overlooks organizational, policy, and coordination measures that enable adaptive response under deep uncertainty.
3. Limited Threat-Agnostic Modeling and Insufficient System Focus			
3.1	**Stressor Definition**	Hazard-specific scenarios (e.g., floods, droughts, earthquakes) and hazard-centric modeling	Cannot address unknown and unpredictable threats; lacks the threat-agnostic flexibility required under HILP uncertainty.
3.2	**Modeling Orientation**	Hazard-centric and risk-based stress testing	Shifts attention away from system vulnerabilities and interdependencies, limiting identification of systemic bottlenecks.
4. Underrepresentation of Recovery and Adaptation Stages and Lack of Dynamic System Modeling			
4.1	**Resilience Orientation**	Focus on preparedness and absorption (robustness)	Ignores recovery and adaptation processes that determine long-term resilience under extreme and prolonged disruption.
4.2	**System Modeling**	Static system performance analysis; component level failures	Does not capture cascading failures over time or evolving system states, missing critical HILP dynamics.
5. Lack of Broader Stakeholder Engagement and Participatory Approaches			
5.1	**Stakeholder Engagement**	Expert-driven, analyst-led, and predominantly quantitative approaches	Misses contextual knowledge, decision constraints, information flows, coordination needs, and behavioral responses critical during unprecedented crises.
6. Limited Iterative Adaptation			
6.1	**Iterative Learning**	One-off evaluations using fixed models	Cannot evolve with changing risks, feedback, or learning, limiting relevance for evolving HILP conditions.

account for cross-sector interdependencies. Doing so enables better identification of vulnerabilities that may result in cascading effects across CI systems and helps map single points of failure that can escalate even small triggers into HILP events [1,3,11].

Our review shows that current stress testing research focuses predominantly on CI sectors in isolation, with a particular emphasis on the assets within each sector. Little has been explicitly documented regarding the interdependencies and interactions between multiple sectors [40]. A systemic perspective is therefore essential for understanding the vulnerabilities that arise within networked and interdependent systems. Stress testing methodologies must move beyond single-sector assessments to explicitly map and analyze cross-sector interdependencies.

4.1.2. Limited multi-domain perspectives and non-technical adaptation strategies

According to the U.S. Army's Network-Centric Warfare doctrine, any complex system, including CI, operates across four management domains: physical, information, cognitive, and social [38,115]. These domains are the constructs that form the system, and diverse and unpredictable threats, such as HILP events, can fracture infrastructural integrity across multiple domains, affecting physical assets, social response systems, cyber components, and financial stability [43].

Our findings indicate that current stress testing methodologies concentrate predominantly on the physical and technical dimensions of CI systems. Overemphasizing the physical infrastructure and asset-based performance in stress testing limits our understanding of how disruptions propagate across interdependent systems and hinders the role of institutions and governance dimensions, such as coordination strategies, organizational arrangements, and policy frameworks, which are critical for enabling effective and even more cost-efficient responses to HILP events.

These findings highlight the need for stress testing methodologies that adopt a systems-thinking perspective and explicitly incorporate different domains and interconnectedness across these domains. Since resilience emerges from interactions among physical, informational, cognitive, and social elements, stress testing must move beyond the single-domain analysis and better capture the full range of adaptive capacities required to manage HILP disruptions. Incorporating multi-domain perspectives also supports the identification of organizational and governance-based interventions, which are essential for enabling coordinated and adaptive responses in complex socio-technical systems.

4.1.3. Over-reliance on hazard-specific scenarios, hazard-centric stress testing and lack of system focus

Recent research argues that hazard-specific resilience approaches are inherently limited because their reliance on predefined scenarios prevents them from capturing the full complexity of cascading failures, novel threat combinations, and the dynamic evolution of risks over time, particularly in systems where environmental interactions play a critical role [42]. Similar concerns are echoed in other studies, which emphasize the need to prepare for complex scenarios and acknowledge that predicting all details of specific threats, the extent of disruption, or the likelihood of a scenario is often nearly impossible ([36]).

Our review indicates that the current stress testing literature increasingly adopts hybrid scenario approaches that combine deterministic and probabilistic elements to better represent uncertain and evolving disruptions. However, the findings also show that most stress testing studies continue to rely predominantly on hazard-specific scenarios, particularly climate-related hazards (e.g., floods, droughts) and geophysical events (e.g., earthquakes). While such events are undeniably significant due to their global impacts, this emphasis reinforces a hazard-centric perspective in which resilience is assessed relative to predefined threat categories, with greater attention to the hazard itself than to systemic behavior and structural vulnerabilities.

However, the consequences of a disaster are not only a function of the triggering hazard but are also shaped by how hazards interact with exposure and the vulnerability of systemic interdependencies. As Helbing et al. [116] note, disasters, regardless of their cause, tend to generate comparable systemic consequences such as transportation breakdowns, supply chain disruptions, and impacts on trade and communication systems. These cascading outcomes arise from interdependencies within infrastructure networks, not exclusively from the magnitude or specific type of the initiating event. Therefore, approaches that focus on a particular hazard type, as well as those that prioritize the hazard rather than the system, may overlook this dimension of how disruptions propagate through interconnected systems, which can substantially amplify impacts beyond those directly caused by initiating hazards. Addressing such shortcomings requires contextual, system-level knowledge that supports decision-making across multiple infrastructures and governance levels and over different temporal and spatial scales.

These insights underline the need for stress testing methodologies that move beyond hazard-specific modeling. Given the defining features of HILP events, such as low recurrence, limited historical data, nonlinear escalation, and unpredictable consequences [1], stress testing cannot rely on models designed around specific hazards. Instead, it should prioritize systemic fragility and adaptive capacity irrespective of the initiating event. Therefore, stress testing methodologies should evolve from hazard-centric, scenario-bound approaches to threat-agnostic, system-focused frameworks that explore how infrastructures fail under deep uncertainty. Such approaches would emphasize identifying critical vulnerabilities, mapping cascading pathways, and strengthening adaptive capacities across interconnected systems, ultimately aligning stress testing practice with the resilience requirements of HILP events.

4.1.4. Under-representation of recovery and adaptation stages, in resilience stress testing, and lack of dynamic system modeling

Our analysis shows that most existing stress testing approaches concentrate primarily on risk assessment, and that even studies framed within resilience assessments tend to focus largely on the preparedness and absorption stages. This narrow emphasis leads to static, single-system, and component-focused stress tests that overlook system evolution, cascading effects, and the critical roles of recovery and adaptation in long-term resilience. Risk-based approaches capture immediate losses but fail to represent how infrastructure systems restore functionality over time, thereby limiting the ability to capture the temporal and nonlinear dynamics of interconnected CI systems that are essential for understanding their behavior under disruption. These findings point out a major mismatch between existing stress testing approaches and the dynamic nature of HILP events where both the nature and intensity of threats evolve over time, requiring system resilience characteristics to evolve accordingly [42]. In this context, while risk-oriented stress testing may identify vulnerabilities or thresholds of failure, it provides limited insight into how systems recover or adapt in the aftermath of disruption, especially under conditions of deep uncertainty. A resilience-oriented perspective reframes stress testing from being merely a “strength test” to a dynamic assessment of system performance not only at the moment of impact but also in the minutes, hours, days, and weeks that follow. This broader view examines whether and how the system continues functioning, the speed at which essential services are restored, and the extent to which the system adapts to changing post-disruption conditions. This implies that stress testing methodologies for civil infrastructure systems need to explicitly incorporate the modeling of evolving system states and account for recovery and adaptation processes. Such approaches should capture how infrastructures restore and deliver essential services over time, as well as how communities gradually regain their capacity to use these services. Only by integrating recovery and adaptation within dynamic, system-level stress testing can we obtain a comprehensive understanding of infrastructure resilience under deep uncertainty and generate actionable insights for long-term resilience planning.

4.1.5. *Lack of stakeholder engagement and participatory approaches in stress testing*

Our review indicates that although some stress testing studies involve stakeholders in their design, engagement is typically limited to decision-makers and technical experts. Most approaches remain model-based and analyst-driven and focus predominantly on quantitative modeling, with little to no inclusion of a broader range of stakeholders. These findings align with previous work by Comes and Bertsch [57], Linkov et al. [11], Trump et al. [42], and Pescaroli et al. [1], emphasizing that participatory cross-sectoral and decision-maker-oriented approaches to stress testing and resilience assessment are necessary. This limited engagement restricts the inclusion of contextual, behavioral, and socio-cultural factors, such as organizational dynamics and local knowledge, that are critical for understanding how systems respond and adapt during real crises.

Therefore, participatory approaches that integrate stakeholder values, missions, operational knowledge, and governance realities into stress testing are essential, enabling the development of adaptive methodologies that both reduce over-reliance on highly quantitative, bottom-up modeling and ensure that stress tests dynamically incorporate diverse perspectives, contextual insights, and evolving system conditions. Such participatory approaches can inform qualitative and semi-quantitative models that require less data yet more accurately represent real-world system behavior at a higher level. They also help identify coordination bottlenecks and vulnerabilities that emerge not only from technical failures but also from the actions, decisions, and interactions of the people and organizations responsible for managing crises. Participatory approaches can also strengthen fully quantitative models by grounding model structure, parameters, assumptions, and validation processes in stakeholder knowledge and real-world system dynamics. In this way, participatory stress testing can enhance the robustness, credibility, and contextual relevance of quantitative analyses, while also identifying coordination bottlenecks and vulnerabilities that emerge not only from technical failures but from the actions, decisions, interactions, and governance arrangements that shape crisis response.

4.1.6. *Limited iterative adaptation*

Iterative adaptation is only marginally represented across existing stress testing frameworks, revealing a significant limitation in addressing evolving risks and adaptive decision-making. The lack of iterative adaptation may especially constrain the ability of methodologies in the context of HILP events. Evolving hazard patterns may lead to extreme events occurring in regions with no prior exposure or to increases in the magnitude and frequency of known hazards, such as those induced by climate change. These changing external stressors interact with emerging internal system vulnerabilities, raising the likelihood of cascading failures in which even small triggers can escalate into large-scale disruptions. In this context, the evolving nature of hazards and the deep uncertainty associated with HILP events underscore the need for stress testing methodologies that iteratively integrate new information on interdependencies, system performance, and environmental stressors.

To address these challenges, adaptive methodologies are essential for ensuring that stress tests remain relevant in fast-changing environments and accurately reflect the evolving risk landscape [57]. Therefore, stress testing for CIs must be designed as a comprehensive, efficient, and dynamic process. Such an approach should enable the continuous updating of scenarios, model parameters, system representations, and performance expectations as new data, shifting priorities, or changing environmental conditions emerge.

4.2. *Conceptual framework for stress testing of CIs in the context of HILP events*

While several stress testing frameworks have been proposed in the literature, our review highlights two closely related issues. First, as summarized in Table 2, existing approaches exhibit important methodological limitations, such as a narrow focus on specific sectors or hazards and limited consideration of interdependencies and system dynamics. Second, as a result, the current frameworks remain fragmented and application-specific, lacking a comprehensive and flexible structure that can serve as a general basis for developing stress testing methodologies, especially in the context of HILP events.

In response, this study proposes a conceptual framework (Fig. 13) that builds directly on the identified gaps and advances the state of the art by explicitly incorporating the requirements of HILP events into the stress testing process. Compared to existing frameworks which are typically sector specific and focused on static risk assessment, our new framework is designed to be an adaptable guiding structure that supports the development of different stress testing methodologies across contexts.

The proposed framework introduces five key advancements. First, it adopts an integrative perspective by explicitly linking the model world, workshop world, and real world, thereby bridging analytical modeling, participatory processes, and real-world

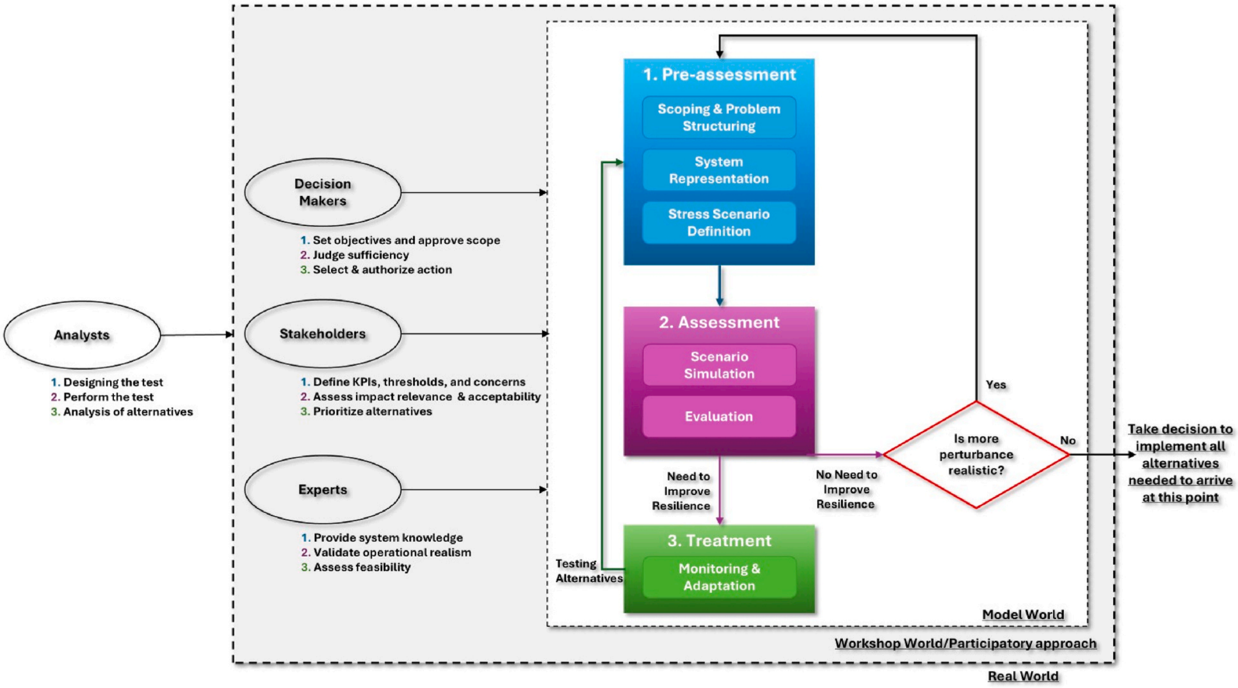


Fig. 13. Conceptual framework for stress testing of critical infrastructures in the context of high impact low probability events.

implementation, an integration that is largely absent in existing frameworks. Second, it embeds stakeholder engagement as a core component across all stages (pre-assessment, assessment, and treatment), ensuring that diverse perspectives and contextual knowledge are systematically incorporated. Third, it integrates iterative adaptation as a fundamental mechanism, enabling continuous learning, feedback, and refinement under evolving system conditions. Fourth, rather than prescribing fixed solutions, the framework systematically indicates how stress testing processes should be adjusted at each stage to better account for HILP-specific requirements. Fifth, by integrating these elements into a unified structure and synthesizing the frameworks from the literature, the proposed framework moves beyond existing approaches and offers a more comprehensive and flexible foundation for developing stress testing methodologies that better address the challenges of HILP events.

The proposed framework, illustrated in Fig. 13, distinguishes between the model world, the workshop world, and the real world. The model world represents the analytical and simulation environment in which system representations, stress scenarios, and performance evaluations are developed. The workshop world captures the participatory setting in which analysts, experts, and decision-makers collaboratively structure the problem, interpret results, and define interventions. The real world refers to the socio-technical system in which decisions are implemented and where cascading effects ultimately occur. The framework is further structured around three main stages, pre-assessment, assessment, and treatment, through which these elements are integrated. The following sections elaborate on these stages and demonstrate how the framework aligns stress testing practices with the requirements of HILP-oriented analysis. Fig. 14 visually summarizes this alignment within the three stages of the proposed framework.

4.2.1. Pre-assessment

The pre-assessment stage establishes the foundation of the stress test. In this stage, analysts, with input from stakeholders, define the problem to be examined, including the infrastructure network of interest, relevant hazards or stressors, and the overarching objectives of the analysis. A clear definition of system boundaries is also essential. Spatial boundaries refer to the geographical extent of the network, possible hazard sources, and locations where consequences may manifest. Temporal boundaries, as outlined by Clarke and Obrien [14], involve determining the time horizon of the analysis, the number and size of time intervals, and whether the system representation is static or dynamic.

Once the scope is defined, analysts model the system of interest and collaboratively develop scenarios. Scenario development should not aim to enumerate all possible worst-case events. Instead, credibility and relevance are central to effective scenario thinking [117]. Drawing on expertise from multiple disciplines ensures that scenarios remain scientifically credible while reflecting diverse perspectives and stakeholder concerns. At this stage, it is critical to shift beyond sector-specific framings and incorporate cross-sectoral integration and multi-domain interdependencies. Equally important is the inclusion of broader stakeholder groups, beyond technical experts and decision-makers, to ensure that the defined problem, system structure, and plausible scenarios reflect contextual knowledge and operational realities. We therefore propose including scenarios that enable the participation of a wide range of stakeholders, including even those beyond the "usual suspects," such as frontline operators, community representatives, and

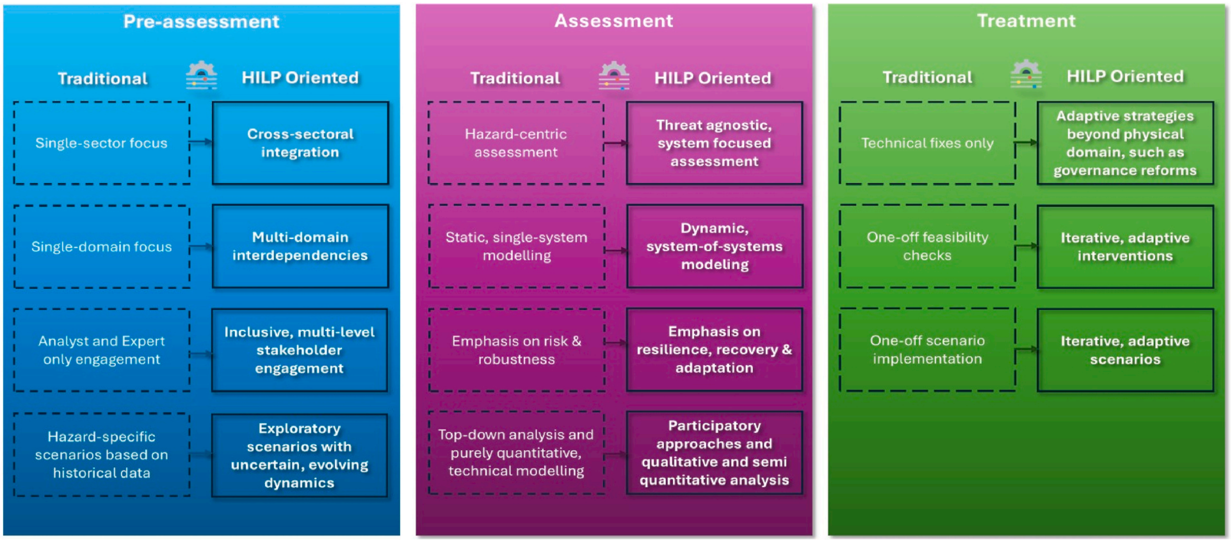


Fig. 14. Adjustments needed for stress testing methodologies to adapt to high impact low probability events.

non-traditional actors whose insights may not be systematically captured in current stress testing practices. Engaging such diverse perspectives helps ensure that scenarios reflect the full spectrum of system interdependencies and decision constraints, enabling us to better map the bottlenecks that may trigger HILP events. In addition, we recommend constructing scenarios that combine deterministic and stochastic elements and support the exploration of compound, cascading, and emerging or unknown threats. This approach not only broadens stakeholder engagement but also better captures the uncertainty and complexity that characterize HILP events.

4.2.2. *Assessment*

In the assessment stage, analysts determine the consequences associated with each scenario. Depending on the scale, purpose, and data availability of the assessment, analyses may be qualitative, semi-quantitative, or quantitative. Scenarios are simulated through topological analysis, system models, and computational tools to identify vulnerabilities, cascading pathways, and components whose failure would generate severe consequences.

However, for HILP events, assessment must move beyond traditional hazard-centric analysis. Rather than evaluating vulnerabilities solely with respect to specific threats, stress testing should focus on system vulnerabilities, independent of the initiating event. Moreover, instead of relying on static models, a dynamic assessment is required—one that evaluates system behavior at the moment of impact and throughout the subsequent minutes, hours, days, and weeks. This requires integrating models of evolving system states and explicitly incorporating recovery and adaptation processes.

Methodologies must also move beyond purely quantitative models, including simulation-based models (such as agent-based models), network-based analyses, probabilistic models, or data-driven statistical models, which often require extensive datasets that may be unavailable for HILP events. Participatory approaches and qualitative or semi-quantitative analyses are essential for capturing perspectives that would otherwise remain unrepresented. Coupling participatory methods with technical modeling ensures that stress tests reflect external trends as well as changing priorities, stakeholder needs, and contextual constraints.

Such an approach supports the development of adaptive methodologies that reduce reliance on highly quantitative, bottom-up modeling and enable stress tests to integrate diverse perspectives, contextual knowledge, governance realities, and evolving system conditions. By combining model-based outputs with expert and stakeholder assessments, stress testing becomes more reflective of real-world system behavior and more capable of identifying vulnerabilities across technical, organizational, and social dimensions.

4.2.3. *Treatment*

After the assessment stage and the generation of results, any unacceptable outcomes must be addressed through the design and implementation of appropriate interventions. Stress testing for CIs must therefore be conceived as a comprehensive, adaptive, and dynamic process, enabling the continuous updating of scenarios, model parameters, system representations, and performance expectations as new data, shifting priorities, or changing environmental conditions arise.

In this stage, adaptation measures aimed at improving system resilience and reducing the likelihood of CI failure should be identified with the support of experts and their domain-specific knowledge. Importantly, the set of potential interventions should not be limited to technical solutions. Rather, it should also include organizational, governance, and coordination-based measures, which are often essential for enabling effective and adaptive responses in complex socio-technical systems. Once a preliminary set of alternatives has been identified, these interventions should be implemented within the modeled system so that their effects can be evaluated. Re-running the analysis with these candidate interventions allows analysts to assess how each measure influences system performance, cascading effects, and recovery trajectories.

Finally, decision-makers must be engaged in selecting the final set of validated interventions for implementation. Their involvement ensures that the chosen measures align with strategic priorities, operational realities, and community needs, thereby supporting informed and context-sensitive resilience planning.

5. Conclusion

In this study, we conducted a scoping review of stress-testing methodologies for terrestrial CI to identify key methodological elements and shortcomings, particularly in the context of HILP events. The review demonstrates that existing gaps in stress testing methodologies developed to date constrain their capacity to address HILP events, which are characterized by a combination of features, including deep uncertainty, cascading and nonlinear dynamics, and prolonged, evolving system disruption.

Building on these findings, the primary contribution of our study lies in the development of a conceptual framework structured around three main stages of pre-assessment, assessment, and treatment. It discusses how the necessary adjustments can be incorporated at each stage to enable future stress testing methodologies to better adapt to and be fine-tuned for HILP events. Rather than proposing a single prescriptive method, the framework provides a flexible and integrative structure within which future stress testing methodologies can be designed, refined, and aligned with the complex and uncertain nature of HILP events.

The proposed framework contributes to both academic research and practice in the fields of stress testing and CI resilience. From an academic standpoint, it synthesizes existing stress testing methodological approaches for terrestrial CIs into a coherent conceptual structure explicitly tailored to the challenges posed by HILP events. Therefore, it provides a structured foundation for the development of methodologies that better account for challenges such as extreme uncertainty and non-linear system behavior. From a practical perspective, the framework also offers guidance for practitioners seeking to design, implement, or refine stress testing exercises. Embedding stress testing into training activities and an iterative cycle of reflection can transform training from a one-off analytical exercise into a continuous capacity-building process. This dual role can enable decision-makers to identify critical bottlenecks that may escalate disturbances into HILP events, while simultaneously fostering learning, adaptation, and improved coordination for real-world crisis response.

Several limitations of this study should be acknowledged. The review was limited to peer-reviewed journal articles and conference papers and did not include gray literature, which may have excluded relevant practitioner and policy perspectives. In addition, no formal quality appraisal of the included studies was undertaken beyond the scoping review approach. The conceptual framework used to guide the analysis may also have influenced the interpretation of findings, and the restriction to English-language publications may have resulted in the omission of relevant work from other regions. Despite these limitations, this study provides a coherent synthesis of existing stress testing practices and offers a structured base for the development of future stress testing methodologies that are more adaptive, systemic, and aligned with the challenges posed by HILP events.

CRediT authorship contribution statement

Ali Cheshomi: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nazli Yonca Aydin:** Writing – review & editing, Validation, Supervision, Investigation, Conceptualization. **Tina Comes:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of competing interest

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

Acknowledgements

The work of the authors has been possible thanks to the support of HE AGILE (AGnostic risk management for high Impact Low probability Events), funded by HORIZON EUROPE Research and Innovation Action (Project number: 101121356; Call: HORIZON-CL32022- DRS-01; Topic: HORIZON-CL3-2022-DRS-01-02), and the UK government's Horizon Europe funding guarantee n∘10062626. The authors gratefully acknowledge the contribution of Dr. Arka Bhattacharyya. The authors used AI-based tools for language editing and text refinement. All content was reviewed and approved by the authors.

Appendix

Search string

TITLE-ABS-KEY ((infrastructure* OR (energ* OR electricit* OR grid* OR “fossil fuels”) OR (telecom* OR ICT* OR “communica*" OR “Information and Communication Technologies” OR IT*) OR water* OR (transport* OR mobility OR railway* OR road* OR bridge*) OR (nuclear OR “radiological risk” OR “radiation leak”) OR food OR (chemical* OR “hazardous material*" OR hazmat OR “toxic release”)) AND (resilience* OR risk*) AND (“stress test*”) AND NOT disease* AND NOT medical* AND NOT cardio* AND NOT bank*)AND (EXCLUDE (SUBJAREA, “MEDI”) OR EXCLUDE (SUBJAREA, “PHAR”) OR EXCLUDE (SUBJAREA, “NURS”) OR EXCLUDE (SUBJAREA, “ARTS”) OR EXCLUDE (SUBJAREA, “IMMU”) OR EXCLUDE (SUBJAREA, “HEAL”) OR EXCLUDE (SUBJAREA, “VETE”)

OR EXCLUDE (SUBJAREA, "DENT") OR EXCLUDE (SUBJAREA, "PSYC") OR EXCLUDE (SUBJAREA, "ECON") OR EXCLUDE (SUBJAREA, "MATE") OR EXCLUDE (SUBJAREA, "BIOC")) AND (LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re") OR LIMIT-TO (DOCTYPE, "cr")) AND (LIMIT-TO (LANGUAGE, "English")) AND (EXCLUDE (EXACTKEYWORD, "Physiology") OR EXCLUDE (EXACTKEYWORD, "Animals") OR EXCLUDE (EXACTKEYWORD, "Mental Stress") OR EXCLUDE (EXACTKEYWORD, "Hydrocortisone") OR EXCLUDE (EXACTKEYWORD, "Electrocardiography") OR EXCLUDE (EXACTKEYWORD, "Pathophysiology") OR EXCLUDE (EXACTKEYWORD, "Financial Markets") OR EXCLUDE (EXACTKEYWORD, "Financial Risk") OR EXCLUDE (EXACTKEYWORD, "Financial System"))

Table 3
Filters/limits that are applied to the database.

Criteria Type	Name	Description
Inclusion	Infrastructure Scope	Mentions of physical Critical Infrastructures (e.g., energy, electricity, water, ICT, telecom, transport, food, chemical, nuclear) in title, abstract, or keywords.
	Conceptual Focus	Contains terms like *resilience*, *risk*, and *"stress test*"* in the search fields.
	Document Type	Limited to articles (ar), reviews (re), conference papers (cp), and conference reviews (cr).
	Language	Limited to English-language publications.
Exclusion	Irrelevant Terms	Excludes terms unrelated to CI (e.g., *disease*, *medical*, *cardio*, *bank*).
	Subject Areas	Excludes subject areas: MEDI, PHAR, NURS, ARTS, IMMU, HEAL, VETE, DENT, PSYC, ECON, MATE, BIOC.
	Exact Keywords	Excludes keywords: *Physiology*, *Animals*, *Mental Stress*, *Hydrocortisone*, *Electrocardiography*, *Pathophysiology*, *Financial Markets*, *Financial Risk*, *Financial System*.

Inclusion & exclusion criteria

Table 4
Eligibility Criteria

Inclusion (Eligible)	Exclusion (Not Eligible)
Studies focusing on physical Critical Infrastructures (CI) within sectors such as chemical and nuclear, transport, water, food, ICT, and energy or mentioning CIs in general.	Studies from other domains such as Finance, Banking, Medicine, Pharmacology, Toxicology, Pharmaceutics, Nursing, Arts & Humanities, Immunology & Microbiology, Health Professions, Veterinary, Dentistry, Psychology, Economics, Econometrics & Finance, Materials Science, Biochemistry, Genetics, Molecular Biology. Studies focusing on non-physical infrastructure or specific infrastructures outside the mentioned CI sectors.
Studies that develop, propose, apply, examine, or conceptualize Stress Testing (ST) methodologies for CIs. Includes: (1) development or application of ST methodologies, (2) conceptual frameworks for ST methodology development, (3) contributions to ST under uncertainty, (4) comparative or analytical studies of ST methodologies.	Non-relevant use of "stress testing" terminology (e.g., biomechanics/ physiological stress testing on humans; in-situ geological stress testing of rocks/materials).
Disruptive events, including human-made incidents (e.g., human error or malicious attacks), natural hazards, technological failures, pandemics, contaminations, and other relevant events.	
Peer-reviewed journal articles and conference proceedings presenting methodological/conceptual contributions relevant to CI Stress Testing.	Not peer-reviewed (e.g., theses, dissertations, preprints, reports, book chapters, white papers, project deliverables, working papers, government documents)
Articles published in English. No date restrictions were applied.	Articles in languages other than English.

Full text extraction

Table 5
Full text extraction elements

Question	Category	Extraction Point	Sub-Extraction Point	Explanation	Examples
What is tested?	Stressor	Stressor Attributes	Primary Stressors	The stressor type that is the main focus of the study.	Earthquake, flood, terrorist attack, etc.
			Secondary Stressors	It is either caused by the primary stressor, for example, when the primary stressor is an earthquake and the secondary one is a tsunami. Alternatively, it may occur alongside the primary stressor, such as when the primary stressor is drought and the secondary one is population growth.	Tsunami, landslide, population growth, etc.

(*continued on next page*)

Table 5 (*continued*)

Question	Category	Extraction Point	Sub-Extraction Point	Explanation	Examples
	System	CI Sectors	Sectors of focus	Identifying the sector or sectors that the study focuses on.	Energy, transport, water supply, etc.
	System	Level of CI Interdependencies	One sector	The study examines only one CI sector in isolation, without accounting for cross-sector effects.	Energy sector only, water sector only.
			Cross-sectoral	The study explicitly considers interdependencies between two or more CI sectors.	Energy–water–transport interdependencies.
	System	Domain of ST	Physical	Focuses on physical infrastructure components, assets, or networks and their functional failures.	Power lines, substations, pipelines, roads, bridges.
			Cyber	Focuses on digital systems, cyber assets, cybersecurity vulnerabilities, and disruptions affecting ICT components that support CI operations.	Cyberattacks on SCADA; communication network outages.
			Social	Addresses human behavior, population dynamics, community response, or societal vulnerability.	Evacuation behavior, population exposure, social preparedness.
			Organizational	Examines coordination, decision-making, emergency procedures, and institutional interactions between CI operators.	Inter-agency communication, command and control structures.
			Economic	Focuses on financial impacts, resource allocation, economic losses, or cost–benefit analysis under stress.	Direct/indirect economic losses, business interruption costs.
How is it modeled?	Stressor	Uncertainty Modeling Approach	Deterministic scenario design	Uses predefined, fixed scenarios without explicitly modeling uncertainty or variability.	Single earthquake scenario with set magnitude; predefined flood depth.
			Probabilistic scenario design	Represents uncertainty explicitly using probability distributions, stochastic simulations, or hazard likelihoods.	Probabilistic flood frequency curves; Monte Carlo simulations.
			Hybrid	Combines deterministic elements with probabilistic components (e.g., deterministic hazard paired with probabilistic fragility).	Deterministic and probabilistic earthquake scenarios.
	System	ST approach	Metric-based approach	Uses individual metrics, indices, dashboards, or decision analytics to assess system performance under stress without simulating full system dynamics.	Resilience indices, vulnerability metrics, performance scores.
			Model-based approaches	Use computational, analytical, or simulation models to represent system behavior under stress.	Network flow models, agent-based models, system dynamics, cascading failure simulations.
			Hybrid	Combine metric-based and model-based elements for a more comprehensive assessment.	—
What is measured?	System	ST Orientation	Risk based ST	Tests system performance under predefined severe or extreme hazard scenarios. It remains hazard-specific and often uses probabilistic or scenario-based assumptions.	Risk assessment under 1-in-100-year flood; failure probability of substations under magnitude 7 earthquake.
			Resilience based ST	Assesses how the system prepares for, absorbs, recovers from, and adapts to	Recovery curve analysis; cascading failure modelling

(*continued on next page*)

Table 5 (*continued*)

Question	Category	Extraction Point	Sub-Extraction Point	Explanation	Examples
				disruptions. Focuses on system functionality and performance over time and is suitable for HILP or hazard-agnostic scenarios.	across interdependent networks.
What ST framework/workflow is proposed?	—	ST framework/workflow	—	Describes whether the study proposes an explicit framework or workflow including steps such as scenario development, modeling, and assessment.	—
	System	Stakeholders involvement	Stakeholders involvement in methodology	Identifies whether stakeholders are engaged in the design, implementation, or interpretation of the analysis.	Workshops with CI operators; expert elicitation; codesign of scenarios; joint model validation.
	System	Adaptation	Adaptation mechanism	Identifies whether the frameworks include an iterative adaptation approach. Iterative adaptation in ST refers to the continuous refinement of strategies, models, or decisions based on new information, feedback, or evolving system conditions.	—
	System	Adaptation	Adaptation domain	Identifies whether ST methodologies focus solely on technical (physical) system adaptations (e.g., infrastructure hardening, redundancy, engineering upgrades) or whether they also consider non-technical strategies, such as organizational learning, institutional flexibility, policy change, behavioral adjustments, and governance innovation.	—

Synthesis of strengths and limitations of stress testing frameworks and workflows in the context of HILP events

Table 6
Summary of strengths and limitations of stress testing frameworks and workflows in the context of HILP events

Reference	Framework	Infrastructure	Cross-Sector Stress Testing	Analysis Beyond The Physical Domain (Multi-Domain)	Adaptive Strategies Beyond Technical	Non-Specific Stressor	Exploratory Scenario Development	Resilience Oriented Stress Testing	Dynamic System Modeling	Stakeholder Engagement in Stress Testing	Iterative Adaptation Analysis Approach
[13]	Stress test workflow and actor interaction process	General	×	×	×	✓	✓	×	×	✓	×
[11]	Tiered approach to integrate risk and resilience stress test for CIs	General	✓	✓	✓	✓	Not Determined	✓	✓	✓	✓
[57]	Framework for dynamic stress tests for CI resilience	General	✓	✓	✓	✓	Not Determined	✓	×	✓	✓
[83]	A resilience indicator-based methodology for stress testing	General	✓	✓	×	✓	×	✓	×	✓	✓

(*continued on next page*)

Table 6 (*continued*)

Reference	Framework	Infrastructure	Cross-Sector Stress Testing	Analysis Beyond The Physical Domain (Multi-Domain)	Adaptive Strategies Beyond Technical	Non-Specific Stressor	Exploratory Scenario Development	Resilience Oriented Stress Testing	Dynamic System Modeling	Stakeholder Engagement in Stress Testing	Iterative Adaptation Analysis Approach
	infrastructures exposed to adverse events and polycrises.										
[78]	A modified CRIDA framework	Water	×	×	×	✓	✓	×	×	✓	✓
[90]	A framework for climate adaptation planning based on stress testing scenarios	Water	×	✓	✓	×	✓	✓	×	✓	✓
[51]	Multidimensional stress testing of hydropower under various uncertainties	Water	×	✓	×	×	✓	✓	×	✓	×
[66]	Integrated modeling chain for evaluating adaptation performance	Water	×	×	×	×	✓	✓	×	×	×
[89]	Workflow to assess climate risks to water supply systems	Water	×	✓	×	×	✓	×	✓	✓	×
[91]	A stress testing framework for evaluating urban system flooding	Water	×	×	×	×	×	✓	✓	×	×
[54]	A stress testing method to assess water security under climate stress	Water	×	×	×	×	✓	✓	✓	✓	×
[53]	Multi-hazard stress testing for water supply systems to climate-related disruptions	Water	×	×	×	×	✓	×	✓	×	×
[92]	A simulation-based framework to (re-) design and stress test water systems	Water	×	×	✓	×	✓	✓	×	✓	×
[118]	A framework to examine the climate change impacts on rainfall–runoff dynamics and regional water allocation.	Water	×	×	×	×	✓	✓	✓	×	×
[55]	Port stress test framework (from Esposito et al., 2017)	Transport	×	×	×	×	✓	×	×	✓	×
[65]	Stress test for the road network under hydro hazards	Transport	×	✓	×	×	✓	✓	×	×	×
[22]	Graph-based urban road network stress testing under seismic risks	Transport	×	×	×	×	✓	✓	×	×	×
[14]	Multi-hazard stress testing framework for transport networks	Transport	×	×	×	×	✓	×	×	✓	×

(*continued on next page*)

Table 6 (*continued*)

Reference	Framework	Infrastructure	Cross-Sector Stress Testing	Analysis Beyond The Physical Domain (Multi-Domain)	Adaptive Strategies Beyond Technical	Non-Specific Stressor	Exploratory Scenario Development	Resilience Oriented Stress Testing	Dynamic System Modeling	Stakeholder Engagement in Stress Testing	Iterative Adaptation Analysis Approach
[95]	Workflow for the INFRARISK decision support tool	Transport	×	×	×	×	✓	×	×	×	✓
[99]	A framework for assessing flood resilience in mountain communities using hydrodynamic modeling and stress testing.	Transport	×	×	×	×	✓	✓	✓	×	×
[106]	A multi-layer framework linking source events to network and societal impacts through cascading model interactions	Transport	×	✓	×	×	✓	✓	✓	×	×
[93]	A framework linking hazards to infrastructure damage, network performance, recovery, and societal impacts.	Transport	×	✓	×	×	✓	✓	✓	×	×
[119]	A network-based simulation framework assessing flood impacts and recovery dynamics in urban rail transit systems.	Transport	×	✓	×	×	×	✓	✓	×	×
[70]	Business continuity model against food supply chain disruptions	Transport	×	✓	✓	×	✓	✓	×	✓	✓
[85]	A stress testing model for cascading failures in power networks	Energy	×	×	×	×	×	×	✓	×	✓
[87]	Workflow for extreme scenario analysis of the energy system	Energy	×	×	×	✓	✓	✓	✓	×	×
[63]	Integrated model chain for stress testing hydropower under climate risks	Energy	×	✓	✓	×	✓	✓	×	×	✓

Data availability

Data will be made available on request.

References

[1] G. Pescaroli, L. McMillan, M. Gordon, N.Y. Aydin, T. Comes, M. Maraschini, J. Oliveria-Palmas, S. Torresan, B. Trump, M. Pelling, I. Linkov, Definitions and taxonomy for high impact low probability (Hilp) and outlier events. https://doi.org/10.2139/ssrn.5113134, 2025.

[2] Y. Wang, O. Zhao, L. Zhang, Multiplex networks in resilience modeling of critical infrastructure systems: a systematic review, Reliab. Eng. Syst. Saf. 250 (2024) 110300, https://doi.org/10.1016/j.ress.2024.110300.

[3] G. Pescaroli, D. Alexander, Critical infrastructure, panarchies and the vulnerability paths of cascading disasters, Nat. Hazards 82 (1) (2016) 175–192, https://doi.org/10.1007/s11069-016-2186-3.

[4] S. Hasan, G. Foliente, Modeling infrastructure system interdependencies and socioeconomic impacts of failure in extreme events: emerging R&D challenges, Nat. Hazards 78 (3) (2015) 2143–2168, https://doi.org/10.1007/s11069-0151814-7.

[5] R. Hoff, R. Sparks, M. Chester, A. Mustafa, N. Johnson, A. Birchfield, T. McPhearson, R. Li, N. Ahmad, I. Searles, Cascading failure propagation and perfect storms in interdependent infrastructures. https://doi.org/10.1061/AOMJAH.AOENG-0045, 2025.

[6] G. Giannopoulos, M. Theocharidou, Risk Assessment Methodologies for Critical Infrastructure Protection. Part II, A New Approach, Publications Office, 2015, https://doi.org/10.2788/621843.

[7] G. Giannopoulos, R. Filippini, M. Schimmer, Risk Assessment Methodologies for Critical Infrastructure Protection.Part I, a State of the Art, Publications Office, 2012. Retrieved October 28, 2024, from, https://data.europa.eu/doi/10.2788/22260.

[8] A. Opdyke, A. Javernick-Will, M. Koschmann, Infrastructure hazard resilience trends: an analysis of 25 years of research, Nat. Hazards 87 (2) (2017) 773–789, https://doi.org/10.1007/s11069-017-2792-8.

[9] UNISDR, Disaster resilience scorecard for cities - toolkit - beta version: campaign, from, Sendai Framework for Disaster Risk Reduction 2015 - (2021) 2030, https://www.unisdr.org/campaign/resilientcities/toolkit/article/disaster-resilience-scorecard-for-cities.htmlUnitedNationsOfficeforDisasterRiskReduction. (Accessed 16 February 2026), 2015.

[10] S.A. Argyroudis, S. Mitoulis, M.G. Winter, A.M. Kaynia, Fragility of transport assets exposed to multiple hazards: State-of-the-art review toward infrastructural resilience, Reliab. Eng. Syst. Saf. 191 (2019) 106567, https://doi.org/10.1016/j.ress.2019.106567.

[11] I. Linkov, B.D. Trump, J. Trump, G. Pescaroli, W. Hynes, A. Mavrodieva, A. Panda, Resilience stress testing for critical infrastructure, Int. J. Disaster Risk Reduct. 82 (2022) 103323, https://doi.org/10.1016/j.ijdrr.2022.103323.

[12] S.A. Argyroudis, S. Fotopoulou, S. Karafagka, K. Pitilakis, J. Selva, E. Salzano, A. Basco, H. Crowley, D. Rodrigues, J.P. Matos, A.J. Schleiss, W. Courage, J. Reinders, Y. Cheng, S. Akkar, E. Uçkan, M. Erdik, D. Giardini, A. Mignan, A risk-based multi-level stress test methodology: application to six critical non-nuclear infrastructures in Europe, Nat. Hazards 100 (2) (2020) 595–633, https://doi.org/10.1007/s11069-019-03828-5.

[13] S. Esposito, B. Stojadinović, A. Babič, M. Dolšek, S. Iqbal, J. Selva, M. Broccardo, A. Mignan, D. Giardini, Risk-based multilevel methodology to stress test critical infrastructure systems, J. Infrastruct. Syst. 26 (1) (2020), https://doi.org/10.1061/(ASCE)IS.1943-555X.0000520.

[14] J. Clarke, E. Obrien, A multi-hazard risk assessment methodology, stress test framework and decision support tool for transport infrastructure networks, Transp. Res. Procedia 14 (2016) 1355–1363, https://doi.org/10.1016/j.trpro.2016.05.208.

[15] S. Esposito, B. Stojadinovic, *Report on Strategies for Stress Test Implementation at Community Level and Strategies to Enhance Societal Resilience Using Stress Tests* (Tech. Rep.), ETH Zurich, 2016, July, https://doi.org/10.3929/ETHZ-B-000234348 (Artwork Size: 25 pp. Medium: application/pdf.

[16] A. Atrens, C.C. Brosnan, S. Ramamurthy, A. Oehlert, I.O. Smith, Linearly increasing stress test (LIST) for SCC research, Meas. Sci. Technol. 4 (11) (1993) 1281, https://doi.org/10.1088/0957-0233/4/11/017.

[17] M.H. Ellestad (Ed.), Stress Testing: Principles and Practice, fifth ed., Oxford University Press, 2003.

[18] R. Hoffman, A generalized concept of resilience, Textil. Res. J. 18 (3) (1948) 141–148, https://doi.org/10.1177/004051754801800301.

[19] N. Winzer, A. Atrens, W. Dietzel, G. Song, K.U. Kainer, Comparison of the linearly increasing stress test and the constant extension rate test in the evaluation of transgranular stress corrosion cracking of magnesium, Mater. Sci. Eng., A 472 (1) (2008) 97–106, https://doi.org/10.1016/j.msea.2007.03.021.

[20] H. Jeong, B.-Y. Yun, Developing the Stress Test Gap Analysis Methodology for Nuclear Power Plants, 2018.

[21] J.L. Yellen, Financial Stability a Decade After the Onset of the Crisis, 2017.

[22] N.Y. Aydin, H.S. Duzgun, F. Wenzel, H.R. Heinimann, Integration of stress testing with graph theory to assess the resilienceofurbanroadnetworksunderseismichazards, Nat. Hazards 91 (1) (2018) 37–68, https://doi.org/10.1007/s11069017-3112-z.

[23] K. Cordova-Pozo, E.A.J.A. Rouwette, Types of scenario planning and their effectiveness: a review of reviews, Futures 149 (2023) 103153, https://doi.org/10.1016/j.futures.2023.103153.

[24] P.J.H. Schoemaker, When and how to use scenario planning: a heuristic approach with illustration, J. Forecast. 10 (6) (1991) 549–564, https://doi.org/10.1002/for.3980100602.

[25] G. Wright, P. Goodwin, Decision making and planning under low levels of predictability: enhancing the scenario method, Int. J. Forecast. 25 (4) (2009) 813–825, https://doi.org/10.1016/j.ijforecast.2009.05.019.

[26] A.C. Tricco, E. Lillie, W. Zarin, K.K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M.D. Peters, T. Horsley, L. Weeks, S. Hempel, E.A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M.G. Wilson, C. Garritty, S.E. Straus, PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation, Ann. Intern. Med. 169 (7) (2018) 467–473, https://doi.org/10.7326/M18-0850.

[27] J. Magana, S. Hinrichs-Krapels, W. Bramer, T. Comes, Patient flow logistics from disaster to care: a scoping review of actors, transport modes and decision problems, J. Humanit. Logist. Supply Chain Manag. (2025) 1–25, https://doi.org/10.1108/JHLSCM-10-2024-0147.

[28] Z. Munn, M.D.J. Peters, C. Stern, C. Tufanaru, A. McArthur, E. Aromataris, Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach, BMC Med. Res. Methodol. 18 (1) (2018) 143, https://doi.org/10.1186/s12874-018-0611-x.

[29] J.F. Burnham, Scopus database: a review, Biomed. Digit. Libr. 3 (1) (2006) 1, https://doi.org/10.1186/1742-5581-3-1.

[30] M. Gusenbauer, N.R. Haddaway, Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources, Res. Synth. Methods 11 (2) (2020) 181–217, https://doi.org/10.1002/jrsm.1378.

[31] J.K. Owens, Systematic reviews: brief overview of methods, limitations, and resources, Nurse Author & Ed. 31 (3–4) (2021) 69–72, https://doi.org/10.1111/nae2.28.

[32] J. Helbach, D. Pieper, T. Mathes, T. Rombey, H. Zeeb, K. Allers, F. Hoffmann, Restrictions and their reporting in systematic reviews of effectiveness: an observational study, BMC Med. Res. Methodol. 22 (1) (2022) 230, https://doi.org/10.1186/s12874-022-01710-w.

[33] N. Shaheen, A. Shaheen, A. Ramadan, M.T. Hefnawy, A. Ramadan, I.A. Ibrahim, M.E. Hassanein, M.E. Ashour, O. Flouty, Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability, Frontiers in Research Metrics and Analytics 8 (2023), https://doi.org/10.3389/frma.2023.1268045.

[34] P.O.o. t.E. Union, /* COM/2005/0576 final */, Green Paper on a European programme for critical infrastructure protection, CELEX1. Retrieved October 15, 2025, from, https://op.europa.eu/en/publication-detail/-/publication/4e3f9be0-ce1c-4f5c-9fdc-07bdd441fb88/language-en?utm_source=chatgpt.com, 2005, November.

[35] M. Ouzzani, H. Hammady, Z. Fedorowicz, A. Elmagarmid, Rayyan—A web and mobile app for systematic reviews, Syst. Rev. 5 (1) (2016) 210, https://doi.org/10.1186/s13643-016-0384-4.

[36] I. Linkov, A. Kott, Fundamental concepts of Cyber resilience: introduction and overview, in: A. Kott, I. Linkov (Eds.), Cyber Resilience of Systems and Networks, Springer International Publishing, 2019, pp. 1–25. Retrieved October 20, 2024, from, http://link.springer.com/10.1007/978-3-319-77492-3_1.

[37] E.B. Connelly, C.R. Allen, K. Hatfield, J.M. Palma-Oliveira, D.D. Woods, I. Linkov, Features of resilience, Environ. Syst. Decis. 37 (1) (2017) 46–50, https://doi.org/10.1007/s10669-017-9634-9.

[38] C. Fox-Lent, M.E. Bates, I. Linkov, A matrix approach to community resilience assessment: an illustrative case at Rockaway Peninsula, Environ. Syst. Decis. 35 (2) (2015) 209–218, https://doi.org/10.1007/s10669-015-9555-4.

[39] V. Braun, V. Clarke, Using thematic analysis in psychology, Qual. Res. Psychol. 3 (2) (2006) 77–101, https://doi.org/10.1191/1478088706qp063oa.

[40] M. Hogan, L.R. Team, Anytown: Final Report, 2013.

[41] S. Rinaldi, J. Peerenboom, T. Kelly, Identifying,understanding,andanalyzingcriticalinfrastructureinterdependencies, Control Systems, IEEE 21 (2002) 11–25, https://doi.org/10.1109/37.969131.

[42] B.D. Trump, S.-A. Mitoulis, S. Argyroudis, G. Kiker, J. Palma-Oliveira, R. Horton, G. Pescaroli, E. Pinigina, J. Trump, I. Linkov, Threat-agnostic resilience: framing and applications, Int. J. Disaster Risk Reduct. 124 (2025) 105535, https://doi.org/10.1016/j.ijdrr.2025.105535.

[43] B.D. Trump, K. Poinsatte-Jones, M. Elran, C. Allen, B. Srdjevic, M. Merad, D.M. Vasovic, J.M. Palma-Oliveira, Social resilience and critical infrastructure systems, in: I. Linkov, J.M. Palma-Oliveira (Eds.), Resilience and Risk, Springer, Netherlands, 2017, pp. 289–299, https://doi.org/10.1007/978-94-024-1123-2_9.

[44] J. Andersson, V. Grassi, R. Mirandola, D. Perez-Palacin, A conceptual framework for resilience: fundamental definitions, strategies and metrics, Computing 103 (4) (2020) 559–588, https://doi.org/10.1007/s00607-020-00874-x.

[45] I. Linkov, C. Fox-Lent, L. Read, C.R. Allen, J.C. Arnott, E. Bellini, J. Coaffee, M.-V. Florin, K. Hatfield, I. Hyde, W. Hynes, A. Jovanovic, R. Kasperson, J. Katzenberger, P.W. Keys, J.H. Lambert, R. Moss, P.S. Murdoch, J. Palma-Oliveira, D. Woods, Tiered approach to resilience assessment, Risk Anal. 38 (9) (2018) 1772–1780, https://doi.org/10.1111/risa.12991.

[46] R. Horton, B.D. Trump, J. Trump, H.S. Knowles, I. Linkov, P. Jones, G. Kiker, Performance metrics for resilience of airport infrastructure, Transport. Res. Transport Environ. 142 (2025) 104676, https://doi.org/10.1016/j.trd.2025.104676.

[47] B. Rathnayaka, C. Siriwardana, D. Robert, D. Amaratunga, S. Setunge, Improving the resilience of critical infrastructures: Evidence-based insights from a systematic literature review, Int. J. Disaster Risk Reduct. 78 (2022) 103123, https://doi.org/10.1016/j.ijdrr.2022.103123.

[48] P. Ray, S. Wi, A. Schwarz, M. Correa, M. He, C. Brown, Vulnerability and risk: climate change and water supply from California's Central Valley water system, Clim. Change 161 (1) (2020) 177–199, https://doi.org/10.1007/s10584-02002655-z.

[49] Z. Yang, P. Liu, L. Cheng, L. Zhang, B. Ming, M. Xiong, Q. Xia, The complementary management of large-scale hydro-photovoltaic hybrid power systems reinforces resilience to climate change, J. Hydrol. 612 (2022) 128214, https://doi.org/10.1016/j.jhydrol.2022.128214.

[50] T. Lachaut, J. Yoon, C. Klassert, A. Tilmant, Aggregation in bottom-up vulnerability assessments and equity implications: the case of Jordanian households' water supply, Adv. Water Resour. 169 (2022) 104311, https://doi.org/10.1016/j.advwatres.2022.104311.

[51] P. Ray, L. Bonzanigo, S. Wi, Y.-C.E. Yang, P. Karki, L.E. García, D.J. Rodriguez, C.M. Brown, Multidimensional stress test for hydropower investments facing climate, geophysical and financial uncertainty, Glob. Environ. Change 48 (2018) 168–181, https://doi.org/10.1016/j.gloenvcha.2017.11.013.

[52] S. Sunkara, R. Singh, D. Gold, P. Reed, A. Bhave, How should diverse stakeholder interests shape evaluations of complex water resources systems robustness when confronting deeply uncertain changes? Earths Future 11 (8) (2023) https://doi.org/10.1029/2022EF003469.

[53] O. Becher, R. Pant, J. Verschuur, A. Mandal, H. Paltan, M. Lawless, E. Raven, J. Hall, A multi-hazard risk framework to stress-test water supply systems to climate-related disruptions, Earths Future 11 (1) (2023), https://doi.org/10.1029/2022EF002946.

[54] K. Verbist, H. Maureira-Cortés, P. Rojas, S. Vicuña, A stress test for climate change impacts on water security: a CRIDA case study, Clim. Risk Manag. 28 (2020), https://doi.org/10.1016/j.crm.2020.100222.

[55] K. Pitilakis, S. Argyroudis, S. Fotopoulou, S. Karafagka, K. Kakderi, J. Selva, Application of stress test concepts for port infrastructures against natural hazards. The case of Thessaloniki port in Greece, Reliab. Eng. Syst. Saf. 184 (2019) 240–257, https://doi.org/10.1016/j.ress.2018.07.005.

[56] J. Yang, S. Liu, Z. Lu, Z. Yan, X. Xu, Source-grid-load combined security assessment of PV-Penetrated distribution network. 2020 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), 2020, pp. 1–5, https://doi.org/10.1109/APPEEC48164.2020.9220690.

[57] T. Comes, V. Bertsch, Designing Dynamic Stress Tests for Improved Critical Infrastructure Resilience, 2013.

[58] A.G. Bhave, D. Conway, S. Dessai, A.J. Dougill, D. Mkwambisi, Stress-testing development pathways under a changing climate: Water-energy-food security in the lake Malawi-Shire river system, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 380 (2221) (2022) 20210134, https://doi.org/10.1098/rsta.2021.0134.

[59] A. Jovanović, H. Schernberg, S. Brem, S. Chakravarty, M. Jelić, Indicator-based assessment of extreme threats (XTs) and their impacts on the resilience of geographical areas and critical infrastructures, Environ. Syst. Decis. 43 (4) (2023) 599–624, https://doi.org/10.1007/s10669-023-09950-x.

[60] A. Fekete, Critical infrastructure and flood resilience: cascading effects beyond water, WIREs Water 6 (5) (2019) e1370, https://doi.org/10.1002/wat2.1370.

[61] P. Pederson, D. Dudenhoeffer, S. Hartley, M. Permann. Critical Infrastructure Interdependency Modeling: A Survey of U.Sand International Research, Idaho National Laboratory, INL/EXT-06-11464, Idaho Falls, 2006. ID.

[62] O. Renn, New challenges for risk analysis: systemic risks, J. Risk Res. 24 (1) (2021) 127–133. https://doi.org/10.1080/13669877.2020.1779787.

[63] P.-Y. Bourgin, S. Le-Cler, A pragmatic approach to assess the climate resilience of hydro projects, E3S Web of Conferences 346 (2022), https://doi.org/10.1051/e3sconf/202234604004.

[64] M. Choine, A. O'Connor, P. Gehl, D. D'Ayala, M. García-Fernández, M.-J. Jiménez, K. Gavin, P. Van Gelder, T. Salceda, R. Power, A multi hazard RISK assessment methodology accounting for cascading hazard events. 12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP 2015, 2015. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84978681127&partnerID=40&md5=f5191ddaf41495a22063f43cecbd8500.

[65] J.C. Lam, B.T. Adey, M. Heitzler, J. Hackl, P. Gehl, N. Van Erp, D. D'Ayala, P. Van Gelder, L. Hurni, Stress tests for a road network using fragility functions and functional capacity loss functions, Reliab. Eng. Syst. Saf. 173 (2018) 78–93, https://doi.org/10.1016/j.ress.2018.01.015.

[66] C. Spence, C. Brown, Decision analytic approach to resolving divergent climate assumptions in water resources planning, J. Water Resour. Plann. Manag. 144 (9) (2018), https://doi.org/10.1061/(ASCE)WR.19435452.0000939.

[67] A.U. Rehman, V. Glenis, E. Lewis, C. Kilsby, C. Walsh, Robust blue-green urban flood risk management optimised with a genetic algorithm for multiple rainstorm return periods, J. Flood Risk Manag. 18 (3) (2025) e70118, https://doi.org/10.1111/jfr3.70118.

[68] L. Xiao, A.S. Khoojine, H. Chen, C. Wang, A dynamic information-theoretic network model for systemic risk assessment with an application to China's maritime sector, Mathematics 13 (18) (2025). Retrieved February 20, 2026, from, https://www.mdpi.com/2227-7390/13/18/2959.

[69] A. Jeuken, P. Ray, E. Penning, L. Bouaziz, J. Tracy, S. Wi, S. McEvoy, Ü. Taner, M. Hegnauer, Challenges for upscaling hydrological effectiveness of nature-based solution for adaptation to climate change in watersheds, Aquatic Ecosystem Health & Management 26 (2) (2023) 19–32, https://doi.org/10.14321/aehm.026.02.019.

[70] F. Schätter, O. Hansen, M. Herrmannsdörfer, M. Wiens, F. Schultmann, Conception of a simulation model for business continuity management against food supply chain disruptions, Procedia Eng. 107 (2015) 146–153, https://doi.org/10.1016/j.proeng.2015.06.068.

[71] B. Zhang, Y. Li, C. Zhang, Exploring pathways for sustaining stable transboundary cooperation for the LancangMekong River Basin under changing hydrological conditions, J. Water Resour. Plann. Manag. 151 (8) (2025) 04025027, https://doi.org/10.1061/JWRMD5.WRENG-6690.

[72] A. Kilic, A quantitative framework for physical cybersecurity in public EVSE systems, Comput. Secur. 159 (2025) 104685, https://doi.org/10.1016/j.cose.2025.104685.

[73] D. Nikolopoulos, C. Makropoulos, D. Kalogeras, K. Monokrousou, I. Tsoukalas, Developing a stress-testing platform for cyber-physical water infrastructure, in: Proceedings - 2018 4th International Workshop on Cyber-Physical Systems for Smart Water Networks, CySWater 2018, 2018, pp. 9–11, https://doi.org/10.1109/CySWater.2018.00009.

[74] D. Nikolopoulos, G. Moraitis, D. Bouziotas, A. Lykou, G. Karavokiros, C. Makropoulos, Cyber-physical StressTesting platform for water distribution networks, J. Environ. Eng. 146 (7) (2020) 04020061, https://doi.org/10.1061/(ASCE)EE.1943-7870.0001722.

[75] J. Valinejad, L. Mili, Cyber–physical–social model of community resilience by considering critical infrastructure interdependencies, IEEE Internet Things J. 10 (19) (2023) 17530–17543, https://doi.org/10.1109/JIOT.2023.3277450.

[76] P. Yan, F. Zhang, F. Zhang, L. Geng, A systematic review and conceptual framework of urban infrastructure cascading disasters using scientometric methods, Buildings 15 (7) (2025) 1011, https://doi.org/10.3390/buildings15071011.

[77] S. Culley, S. Westra, B. Bennett, H. Maier, J. Newman, L. van der Linden, M. Irvine, Water security analysis of the middle river supply system: a foreSIGHT application, in: Proceedings of the International Congress on Modelling and Simulation, MODSIM, 2021, pp. 547–553. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85177054545&partnerID= 40&md5=f7130ab2f9190a9981a6b799c8e66e03.

[78] R. Koh, M. Babel, V. Shinde, G. Mendoza, Towards climate resilient municipal water supply in Bangkok: a collaborative risk informed analysis, Clim. Risk Manag. 35 (2022), https://doi.org/10.1016/j.crm.2022.100406.
[79] A. Tariq, R.J. Lempert, J. Riverson, M. Schwartz, N. Berg, A climate stress test of Los Angeles' water quality plans, Clim. Change 144 (4) (2017) 625–639, https://doi.org/10.1007/s10584-017-2062-5.
[80] M. Garcia-Fernandez, K. Assatourians, M.-J. Jimenez, An operational-oriented approach to the assessment of low probability seismic ground motions for critical infrastructures, J. Seismol. 22 (1) (2018) 123–136, https://doi.org/10.1007/s10950-017-9695-8.
[81] D. Nikolopoulos, C. Makropoulos, Stress-testing water distribution networks for cyber-physical attacks on water quality, Urban Water J. 19 (3) (2022) 256–270, https://doi.org/10.1080/1573062X.2021.1995446.
[82] P. Gauthier, A. Furno, N.-E. El Faouzi, Road network resilience: how to identify critical links subject to day-to-day disruptions, Transp. Res. Rec. 2672 (1) (2018) 54–65, https://doi.org/10.1177/0361198118792115.
[83] A. Jovanović, Stress-Testing the resilience of critical infrastructures exposed to polycrises triggered by emerging risks, Int. J. Disaster Risk Sci. 16 (4) (2025) 594–617, https://doi.org/10.1007/s13753-025-00663-0.
[84] D. Kaulbars, F. Schweikowski, C. Wietfeld, Spatially distributed traffic generation for stress testing the robustness of mission-critical smart grid communication, 2015 IEEE Globecom Workshops (GC Wkshps) (2015) 1–6, https://doi.org/10.1109/GLOCOMW.2015.7414168.
[85] H. Liao, M. Ilic, Stress test model of cascading failures in power grids, NAPS 2011 - 43rd North American Power Symposium (2011), https://doi.org/10.1109/NAPS.2011.6025200.
[86] Z. Xiao, X. Bai, Impact of local port disruption on global container trade: an example of stressing testing Chinese ports using a D-vine copula-based quantile regression, Ocean Coast Manag. 228 (2022) 106295, https://doi.org/10.1016/j.ocecoaman.2022.106295.
[87] Z. Xu, G. Lu, J. Zhao, Multi-scenario generation technology considering extreme scenarios in energy system modeling, IOP Conf. Ser. Earth Environ. Sci. 546 (2) (2020) 022057, https://doi.org/10.1088/17551315/546/2/022057.
[88] A. John, A. Horne, R. Nathan, K. Fowler, J. Webb, M. Stewardson, Robust climate change adaptation for environmental flows in the Goulburn River, Australia, Front. Environ. Sci. 9 (2021), https://doi.org/10.3389/fenvs.2021.789206.
[89] S. Whateley, J.D. Walker, C. Brown, A web-based screening model for climate risk to water supply systems in the northeastern United States, Environ. Model. Software 73 (2015) 64–75, https://doi.org/10.1016/j.envsoft.2015.08.001.
[90] C. Albano, M. McCarthy, M. Dettinger, S. McAfee, Techniques for constructing climate scenarios for stress test applications, Clim. Change 164 (3) (2021), https://doi.org/10.1007/s10584-021-02985-6.
[91] J. He, Y. Qiang, H. Luo, S. Zhou, L. Zhang, A stress test of urban system flooding upon extreme rainstorms in Hong Kong, J. Hydrol. 597 (2021) 125713, https://doi.org/10.1016/j.jhydrol.2020.125713.
[92] D. Bouziotas, S. Stofberg, J. Frijns, D. Nikolopoulos, C. Makropoulos, Assessing the resilience of circularity in water management: a modeling framework to redesign and stress-test regional systems under uncertainty, Urban Water J. 20 (5) (2023) 532–549, https://doi.org/10.1080/1573062X.2023.2190030.
[93] H. Nasrazadani, M. Nogal, B.T. Adey, S.A. Mitoulis, Prioritizing simulation-based stress tests to assess the resilience of transport systems: a computation-free methodology, Journal of Infrastructure Preservation and Resilience 6 (1) (2025) 16, https://doi.org/10.1186/s43065-025-00128-0.
[94] J.E. Goodwill, P. Ray, D. Nock, C.M. Miller, Emerging investigator series: moving beyond resilience by considering antifragility in potable water systems, Environ Sci (2022).
[95] Z. Sabeur, P. Melas, K. Meacham, R. Corbally, D. D'Ayala, B. Adey, An integrated decision-support information system on the impact of extreme natural hazards on critical infrastructure, IFIP Adv. Inf. Commun. Technol. 507 (2017) 302–314, https://doi.org/10.1007/978-3-319-89935-0_25.
[96] T. Comes, B. Van de Walle, L. Van Wassenhove, The coordination-information bubble in humanitarian response: theoretical foundations and empirical investigations, Prod. Oper. Manag. 29 (11) (2020) 2484–2507, https://doi.org/10.1111/poms.13236.
[97] E. Rokstad, C. Makropoulos, M. Rokstad, Resilience assessment of water distribution networks exposed to substance intrusion, Urban Water J. (2023), https://doi.org/10.1080/1573062X.2023.2237944.
[98] T. Amestoy, A. Hamilton, P. Reed, Integrated river basin assessment framework combining probabilistic streamflow reconstruction, Bayesian bias correction, and drought storyline analysis, Environ. Model. Software 195 (2026) 106756, https://doi.org/10.1016/j.envsoft.2025.106756.
[99] M. Yin, H. Huang, F. Yu, A. Wu, Y. Tao, X. Sun, Study on stress testing and the evaluation of flood resilience in Mountain communities, Sustainability 17 (16) (2025) 7463, https://doi.org/10.3390/su17167463.
[100] A. Aguila Téllez, N. Krishnan, E. García, D. Carrión, M. Ruiz, A probabilistic framework for reliability assessment of active distribution networks with high renewable penetration under extreme weather conditions, Energies 18 (24) (2025) 6525, https://doi.org/10.3390/en18246525.
[101] Y. Bai, Y. Sui, X. Deng, X. Wang, Quantum-inspired robust optimization for coordinated scheduling of PV-hydrogen microgrids under multi-dimensional uncertainties, Sci. Rep. 15 (1) (2025) 29589, https://doi.org/10.1038/s41598025-12280-4.
[102] S. Marasco, O. Kammouh, G.P. Cimellaro, Disaster resilience quantification of communities: a risk-based approach, Int. J. Disaster Risk Reduct. 70 (2022) 102778, https://doi.org/10.1016/j.ijdrr.2021.102778.
[103] A.A. Shah, A. Ullah, N.A. Khan, M.H. Shah, R. Ahmed, S.T. Hassan, M.A.U.R. Tariq, C. Xu, Identifying obstacles encountered at different stages of the disaster management cycle (DMC) and its implications for rural flooding in Pakistan, Front. Environ. Sci. 11 (2023) 1088126, https://doi.org/10.3389/fenvs.2023.1088126.
[104] S. Krishnan, N.Y. Aydin, T. Comes, TIMEWISE: Temporal Dynamics for Urban Resilience - theoretical insights and empirical reflections from Amsterdam and Mumbai, Npj Urban Sustainability 4 (1) (2024) 4, https://doi.org/10.1038/s42949-024-00140-5.
[105] M. Sirenko, T. Comes, A. Verbraeck, Urban heatwaves reverse vulnerability-resilience relationships throughout the day, Npj Urban Sustainability 6 (1) (2026) 22, https://doi.org/10.1038/s42949-025-00327-4.
[106] H. Nasrazadani, B.T. Adey, L. Dorren, Simulation-based stress tests: integrated simulation approach for assessing resilience of transport systems under multiple stressors, J. Infrastruct. Syst. 30 (3) (2024) 04024010, https://doi.org/10.1061/JITSE4.ISENG-2417.
[107] R.M. Bhaduri, Leveraging culture and leadership in crisis management, Eur. J. Train. Dev. 43 (5–6) (2019) 554–569, https://doi.org/10.1108/EJTD-10-2018-0109.
[108] M.A. Hakim, F. Kirana Anggarani, A. Okvitawanli, T. Hamamura, How culture shapes community resilience: a cultural-psychological case study of coastal communities' responses to climate disaster in Indonesia, Asian J. Soc. Psychol. 28 (3) (2025) e70039, https://doi.org/10.1111/ajsp.70039.
[109] J. Kujala, S. Sachs, H. Leinonen, A. Heikkinen, D. Laude, Stakeholder engagement: past, present, and future, Bus. Soc. 61 (5) (2022) 1136–1196, https://doi.org/10.1177/00076503211066595.
[110] F.H. Norris, S.P. Stevens, B. Pfefferbaum, K.F. Wyche, R.L. Pfefferbaum, Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness, Am. J. Community Psychol. 41 (1–2) (2008) 127–150, https://doi.org/10.1007/s10464-007-9156-6.
[111] J.M. Ottino, Engineering complex systems, Nature 427 (6973) (2004) 399, https://doi.org/10.1038/427399a, 399.
[112] E. Luiijf, A. Nieuwenhuijs, M. Klaver, M. van Eeten, E. Cruz, Empirical findings on critical infrastructure dependencies in Europe, in: R. Setola, S. Geretshuber (Eds.), Critical Information Infrastructure Security, CRITIS 2008, Lecture Notes in Computer Science 5508, 2009, pp. 302–310, https://doi.org/10.1007/978-3-642-03552-4_28. Springer, Berlin, Heidelberg.
[113] E. Luiijf, M. Klaver, Expand the crisis? Neglect critical infrastructure! Insufficient situational awareness about critical infrastructure by emergency management: insights and recommendations, in: Proceedings of CRISE 2013, 2013, pp. 293–304. Weimar, Germany, 27–29 May 2013.
[114] L. Buzna, K. Peters, H. Ammoser, C. Kühnert, D. Helbing, Efficient response to cascading disaster spreading, Physical Review E 75 (5) (2007) 056107, https://doi.org/10.1103/PhysRevE.75.056107.
[115] D.S. Alberts, R.E. Hayes, Power to the Edge: Command, Control in the Information Age, CCRP Publication Series, 2003.
[116] D. Helbing, H. Ammoser, C. Kühnert, Disasters as extreme events and the importance of network interactions for disaster response management, in: S. Albeverio, V. Jentsch, H. Kantz (Eds.), Extreme Events in Nature and Society, The Frontiers Collection, Springer, Berlin, Heidelberg, 2006, pp. 319–348. https://doi.org/10.1007/3-540-28611-X_15.

[117] C. Selin, Trust and the illusive force of scenarios, Futures 38 (1) (2006) 1–14, https://doi.org/10.1016/j.futures.2005.04.001.

[118] A. Horne, A. John, K. Fowler, Z. Zhang, R. Nathan, Stress testing water allocations across large river basins, Climatic Change 178 (2025) 226. https://doi.org/10.1007/s10584-025-04051-x.

[119] W. Bi, J. Schooling, K. MacAskill, Assessing flood resilience of urban rail transit systems: Complex network modelling and stress testing in a case study of London, Trans. Res. Part D Trans Environ 134 (2024) 104263. https://doi.org/10.1016/j.trd.2024.104263.