Review | Received 3 July 2026; Revised 5 September 2026; Accepted 15 September 2026; Published 23 September 2026
https://doi.org/10.55092/sc20260018

# Monitoring-data-driven reliability assessment of bridge infrastructure: model updating, uncertainty quantification and maintenance decision-making

**Yulei Bai[1], Bingxu Duan[1], Pinghe Ni[1,2,*], Qiang Han[1], Jun Li[2], Xuanyi Zhang[3], Qiang Li[4] and Wangji Yan[5]**

1 State Key Laboratory of Bridge Engineering Safety and Resilience, Beijing University of Technology, Beijing, China
2 Centre for Infrastructural Monitoring and Protection, School of Civil and Mechanical Engineering, Curtin University, Perth, Australia
3 Chair for Reliability Engineering, TU Dortmund University, Dortmund, Germany
4 Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada
5 State Key Laboratory of Internet of Things for Smart City, Department of Civil and Environmental Engineering, University of Macau, Macau, China

* Corresponding author; Email: Pinghe.ni@connect.polyu.hk.

## Highlights:

- A focused Scopus-based critical review synthesizes 56 core bridge studies and 58 supplementary method-transfer or comparison studies published from 2020 through 30 June 2026.
- Monitoring data are linked to reliability through a data-state-variable-limit-state chain.
- Model updating methods are compared by applicability, failure scenarios and reliability contribution.

**Abstract:** Monitoring-data-driven reliability assessment of bridge infrastructure uses structural health monitoring, inspection, load testing, bridge weigh-in-motion, visual/non-destructive testing (NDT) data and digital twins to characterize evolving condition and failure risk. This review asks how monitoring data are transformed into reliability inputs, how model updating methods complement one another, how limited data support time-dependent reliability and rare-event estimation, and how updated reliability supports maintenance decisions. A Scopus-based search of English journal papers published from 2020 through the 30 June 2026 search cutoff was conducted using keywords on bridges, monitoring data, model updating, reliability/risk assessment and maintenance. After thematic screening, 114 papers were coded and critically synthesized, comprising 56 core bridge studies and 58 supplementary studies retained only for method transfer or comparison. The review establishes a technical chain from monitoring data to structural state characterization, model updating, uncertainty quantification, reliability/risk assessment and maintenance decision-making, and formulates a probabilistic framework based on observation models, state/parameter updating, posterior predictive reliability and risk- and loss-based decision-making. The findings show that finite-element updating contributes physical interpretability, Bayesian updating supports evidence fusion and posterior uncertainty, filtering enables online recursion, surrogate modelling and machine learning reduce computational burden, and digital twins are meaningful only when data assimilation, uncertainty propagation, reliability calculation and decision feedback are integrated. Future work should strengthen traceable data-to-limit-state mappings, explicit treatment of model-form and measurement errors, tail failure-probability estimation under sparse monitoring data, and closed-loop reliability-informed bridge maintenance.

**Keywords:** structural health monitoring; bridge weigh-in-motion; load testing; model updating; reliability assessment; maintenance decision-making; digital twin

## 1. Introduction

Bridges and bridge networks are critical components of transportation infrastructure, supporting mobility, logistics and emergency access. With increasing service time, bridges are continuously affected by traffic growth, environmental corrosion, material ageing, fatigue accumulation and extreme hazards, and their actual performance may gradually deviate from the design state [1–3]. Conventional bridge reliability assessment often relies on design models, code-based parameters or periodic inspection results, which are insufficient for capturing the current service condition of existing bridges. The development of long-term structural health monitoring, load testing, bridge weigh-in-motion, visual measurement, non-destructive testing and bridge digital twins has created opportunities to update structural reliability using real-time or staged information [4–6].

Monitoring-data-driven bridge reliability assessment should not be interpreted as substituting sensor readings directly into reliability equations [7–9]. It is better regarded as a closed-loop process involving data, models, probabilistic inference and engineering decisions. Monitoring data first have to be translated into physically meaningful structural state descriptors, such as vehicle load effects, modal properties, deflection, strain, crack width, corrosion degree, support or boundary conditions, damage locations or model parameters. These quantities are then used to update bridge models through finite-element model updating, Bayesian inversion [8,9], Kalman filtering [10,11], particle filtering, machine learning or digital-twin frameworks [3]. The updated model must quantify parameter uncertainty, model error, measurement noise and degradation uncertainty before calculating failure probability, reliability indices, fragility or risk indicators [12].

Existing studies indicate that monitoring-data-driven bridge reliability assessment is driven jointly by structural health monitoring, structural reliability, Bayesian statistics, computational mechanics, machine learning and bridge asset management. Applications include foundation identification, bridge weigh-in-motion [5,6], fatigue-life assessment [13], corrosion degradation [14], seismic fragility [15], proof-load reliability updating [16], and bridge-network resilience [17]. Therefore, a review of this topic should not be limited to monitoring technologies or reliability algorithms alone. It should explain how bridge monitoring data enter reliability assessment through model updating and uncertainty quantification, and how reliability results further support inspection, load restriction, repair, strengthening and bridge-network risk management.

This review focuses on four research questions. First, how can bridge monitoring, inspection and load-testing data be transformed from raw responses into state variables, model parameters or probabilistic inputs required for reliability analysis? Second, what roles do finite-element updating, Bayesian updating, filtering, surrogate modelling and digital twins play in the bridge reliability chain, and how should their boundaries and complementarities be defined? Third, what useful information can limited-monitoring data provide for time-dependent reliability, small failure probabilities, fatigue life, degradation reliability and bridge-network risk? Fourth, how do existing studies translate updated reliability results into inspection, maintenance, load restriction, strengthening and bridge asset-management decisions?

The main contributions of this review are fourfold. First, it constructs an integrated chain linking bridge monitoring data, structural state characterization, model updating, uncertainty quantification, reliability/risk assessment and maintenance decision-making, avoiding a fragmented discussion of monitoring technologies or reliability algorithms. Second, it performs statistical coding and thematic synthesis of a 114-paper corpus published from 2020 through the 30 June 2026 search cutoff, comprising 56 core bridge studies and 58 supplementary method-transfer or comparison studies. Third, it compares different model updating methods in terms of applicable conditions, limitations and complementary roles, emphasizing that monitoring data must be propagated through uncertainty models before they can affect reliability. Fourth, it summarizes bridge component, bridge system and bridge-network scenarios, and discusses the boundary conditions under which bridge experience may or may not be transferred to other infrastructure systems.

## 2. Literature search and screening methodology

This review addresses monitoring-data-driven reliability assessment of bridge infrastructure using Scopus as the unified search source. The statistical scope was restricted to English journal papers published from 2020 through the search cutoff of 30 June 2026. The search strategy combined bridge-object terms with data-source, model-analysis and condition, reliability, risk, and decision-analysis terms; method-specific terms were used only as supplementary expansion terms. Records available by the cutoff were screened and coded, so the 2026 value in Figure 1 represents a partial publication year.

After keyword searching, topical screening, full-text relevance assessment and functional coding, 114 papers were retained in two evidence strata. The core bridge stratum contains 56 records whose primary infrastructure code is bridges or bridge networks and supports bridge-specific quantitative statements and conclusions. The supplementary stratum contains the remaining 58 records, including non-bridge applications and records without a defensible primary infrastructure category. These records are used only to compare transferable methods and are not treated as bridge-specific evidence. The combined corpus is used to describe methodological coverage and is not an exhaustive bibliometric census of bridge monitoring or structural reliability research.

The search strategy was designed to identify studies addressing the transformation of bridge monitoring evidence into reliability-related information. It was therefore organized according to three levels. The first level described data sources, including structural health monitoring, monitoring data, bridge weigh-in-motion, weigh-in-motion, load testing, proof-load testing, bridge inspection, non-destructive testing, traffic-load data, and environmental monitoring. The second level described model-analysis processes, including model updating, data assimilation, state estimation, and uncertainty quantification. The third level described the resulting condition-, reliability-, risk-, and decision-analysis objectives, including condition assessment, reliability assessment, time-dependent reliability, failure probability, risk assessment, maintenance decision-making, inspection planning, digital twin, and life-cycle performance. Bridge-related object terms, such as bridge, bridge component, bridge network, girder bridge, railway bridge, reinforced concrete (RC) bridge, steel bridge, and bridge infrastructure, were combined with these three levels of terms during the search.


**Figure 1.** Primary coding characteristics of the combined 114-paper evidence corpus. Each paper has one mutually exclusive primary code in every displayed dimension. The "Unclassified/no code" categories contain 28 records for infrastructure type, 18 for method category, 2 for data type and 5 for assessment objective; every panel therefore sums to 114. The core/supplementary evidence-stratum label is assigned separately, and bridge-specific synthesis uses the 56-record core bridge dataset. The 2026 publication count is partial because the search cutoff was 30 June 2026.

Method-specific terms, including Bayesian updating, Bayesian inversion, Kalman filtering, particle filtering, surrogate modelling, and machine learning, were used as supplementary search terms and method-classification terms. They were not treated as independent research objectives. This distinction reflects the focus of this review, which is not to compare isolated algorithms, but to examine how different model-analysis methods enable monitoring data to enter structural reliability updating and subsequent maintenance decision-making.

Eligibility was defined separately for the two evidence strata. Core bridge studies had to examine a bridge, bridge component or bridge network; use monitoring, inspection, load-testing, condition-rating, traffic-load or environmental data; apply model updating, state estimation, uncertainty quantification, reliability or risk assessment, or maintenance decision-making; and support assessment of an existing bridge. Supplementary studies could examine other civil infrastructure only when they demonstrated a specific method transferable to one or more links in the bridge data-state-limit-state-decision chain. Studies dealing only with new-structure design optimization, material-scale parameter inversion, construction management, or generic algorithms without a relevant civil-infrastructure application were excluded.

The screening procedure comprised three steps. First, titles, abstracts and keywords were screened in the Scopus results to remove studies clearly outside the monitoring-data, model-updating or reliability-assessment chain. Second, full texts or core contents were examined to determine whether a logical chain existed from monitoring or inspection data to structural state descriptors, model or parameter updating, and reliability or risk assessment. Third, retained papers were assigned an evidence-stratum label and functionally coded by publication year, infrastructure type, data type, method category and assessment objective. For Figure 1, each paper received one mutually exclusive primary code in every displayed dimension. When no substantive primary assignment was defensible, the record was coded as “Unclassified/no code.” These categories contain 28 records for infrastructure type, 18 for method category, 2 for data type and 5 for assessment objective; consequently, every panel sums to 114. Secondary characteristics were retained only for qualitative synthesis. The coding counts describe corpus composition and are not used for meta-analysis or evidence weighting.

The evidence-stratum label is separate from the four primary coding dimensions shown in Figure 1. The core bridge dataset comprises 56 records whose primary infrastructure code is bridges or bridge networks; bridge-specific quantitative statements use this denominator. The other 58 records, including non-bridge and infrastructure-unclassified records, are labelled supplementary method-transfer or comparison evidence. Counts describing the combined methodological landscape use the full 114-paper denominator and report all unclassified records explicitly.

The final set covers bridge scour monitoring, railway-bridge monitoring, bridge condition thresholds, bridge weigh-in-motion, influence lines, active-learning reliability, digital-twin maintenance, bridge-network deterioration and maintenance, ground-penetrating-radar assessment of bridge decks, health-monitoring-based strengthening assessment, Bayesian updating, filtering, surrogate modelling and maintenance decision-making. These papers form the empirical basis for the descriptive statistics and methodological critique in this review.

## 3. Overall framework for monitoring-data-driven bridge reliability assessment

As shown in Figure 2, monitoring-data-driven bridge reliability assessment can be summarized as a chain from monitoring data to structural state characterization, model updating, uncertainty quantification, reliability/risk assessment and maintenance decision-making. The framework emphasizes that monitoring data do not directly provide reliability. They must first be interpreted as structural state descriptors, then used for model updating and uncertainty quantification before they can enter reliability or risk assessment and support maintenance decisions.

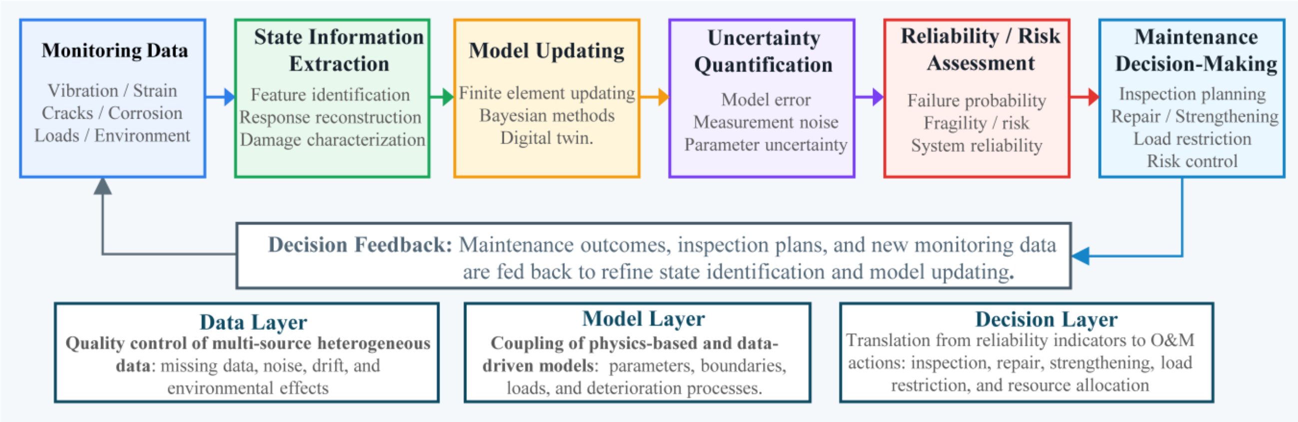


**Figure 2.** Technical framework for monitoring-data-driven reliability assessment of bridge infrastructure.

### *3.1. Characteristics of the literature set*

To improve the systematic nature of the review, the 114 Scopus papers were statistically coded by publication year, infrastructure type, method category, data type, and assessment target. Each paper received one mutually exclusive primary code in each displayed dimension. An unclassified code was used when a paper had no substantive primary assignment in that dimension; consequently, every panel in Figure 1 sums to 114. Secondary codes used in the qualitative synthesis are not counted in Figure 1. The figure therefore describes the composition of the combined evidence corpus, whereas bridge-specific interpretations are based on the 56-record core bridge dataset.

Within the combined evidence corpus, the annual distribution does not by itself support a monotonic publication-growth trend. The counts are used only to describe the temporal composition of the reviewed evidence. The 2026 count is partial because the Scopus search was completed on 30 June 2026 and is therefore not comparable with full-year counts.

The primary infrastructure coding identifies 56 bridge and bridge-network records as the core bridge dataset and 58 records as supplementary method-transfer or comparison evidence. Across the combined corpus, the primary method codes cover Bayesian or probabilistic inversion, filtering or sequential estimation, surrogate modelling or machine learning, physics-based finite-element updating, digital twins or multisource fusion, and model-error or identifiability studies. These counts describe the composition of the selected corpus rather than the publication output of the whole field.

Within the selected evidence corpus, vibration, modal, acceleration, strain, displacement, deflection, crack, visual, and non-destructive testing (NDT) data provide the main evidence for state identification. Bridge weigh-in-motion, traffic loads, temperature, environmental data, long-term monitoring, and digital-twin data represent additional evidence pathways. The assessment targets are centered on component reliability, fatigue reliability, time-dependent failure probability, load-carrying safety, and maintenance decisions. These distributions describe the evidence represented in the selected corpus; they should not be interpreted as a field-wide time trend.

### *3.2. Basic reliability theory and monitoring-data entry points*

The core of bridge reliability assessment is to quantify the probability that a structure crosses a limit state under traffic loads, environmental actions and degradation at a given time. For a bridge component, bridge system or bridge network, a generic limit-state function can be written as

$$g(\mathbf{X},t) = R(\mathbf{X}_R,t) - S(\mathbf{X}_S,t) \tag{1}$$

where $R(\mathbf{X}_R,t)$ denotes resistance or load-carrying capacity, $S(\mathbf{X}_S,t)$ denotes load effect or demand, $\mathbf{X}_R$ and $\mathbf{X}_S$ are uncertain variables affecting resistance and demand, respectively, and $t$ is service time. The time-dependent failure probability $P_f(t)$ and reliability index $\beta(t)$ can then be expressed as

$$P_f(t) = P[g(\mathbf{X},t) \le 0] \tag{2}$$

$$\beta(t) = -\Phi^{-1}[P_f(t)] \tag{3}$$

The reliability index $\beta(t)$ in Equation (3) describes the time-dependent reliability level of the assessed bridge. When this value is compared with a target reliability level, the target should be interpreted as an assessment-specific target reliability index, $\beta_{\mathrm{t,assess}}$, rather than the design target reliability index for a new bridge. Existing bridges differ from newly designed structures in age, deterioration state, available inspection or monitoring information, reference period, remaining service life, redundancy, failure consequence, and maintenance feasibility. Therefore, the design target reliability index may provide a useful background benchmark, but it should not be directly used as the acceptance criterion for condition assessment of existing bridges. The assessment-specific distinction is consistent with risk-informed target-reliability frameworks for bridge networks and existing structures [18,19].

These expressions show that reliability assessment does not ask whether a single monitored value exceeds a threshold. It asks for the probability that the limit-state function is non-positive under the combined uncertainty of loads, resistance, model error and degradation state.

Monitoring data $D$ enter reliability analysis primarily by updating uncertain variables or model parameters. Let $\theta$ denote material parameters, boundary conditions, damage variables, degradation rates, model bias or load-model parameters. Bayesian updating can be written as

$$p(\theta \mid D) = \frac{p(D \mid \theta)p(\theta)}{\int p(D \mid \theta)p(\theta)d\theta} \tag{4}$$

The updated bridge failure probability is then

$$P_f(t \mid D) = \int I[g(x,\theta,t) \le 0]p(x,\theta \mid D)dxd\theta \tag{5}$$

where $I[.]$ is the failure indicator and $p(x,\theta \mid D)$ contains monitoring-updated parameter, state and model uncertainties. Therefore, the key issue is not the size of the monitoring data set, but whether the data identify variables relevant to the limit state and whether the updated uncertainty is properly propagated to failure probability, risk or maintenance decisions.

Here, $\mathbf{X}$ denotes the vector of basic random variables entering the limit-state function, and $x$ denotes a realization of $\mathbf{X}$. Accordingly, $p(x,\theta \mid D)$ in Equation (5) represents the joint posterior density of $\mathbf{X}$ and $\theta$ conditional on the monitoring data $D$; integration over $x$ and $\theta$ propagates their posterior uncertainty into the updated failure probability.

When reliability results are used for operation and maintenance, failure probability must be combined with consequences, inspection and repair costs, and operational impacts. A simplified life-cycle risk objective can be written as

$$a^* = \arg\min_a E\left[C_I(a) + C_M(a) + C_F P_f(t \mid D,a) + C_U(a)\right] \tag{6}$$

where a denotes an inspection, repair, strengthening, load-restriction or monitoring-layout decision; $C_I$, $C_M$, $C_U$ and $C_F$ denote inspection cost, maintenance cost, operational-impact cost and failure consequence, respectively. This expression emphasizes that reliability updating has asset-management value only when it is embedded in a decision objective.

### *3.3. A unified probabilistic framework*

To avoid treating monitoring-data-driven bridge reliability assessment as a set of unrelated algorithms, it can be expressed as a probabilistic loop composed of an observation model, state/parameter updating, posterior predictive reliability and risk- and loss-based decision-making. The core objects are monitoring data *D*, latent bridge state $z_t$, model parameters $\theta$, and the limit-state function g(.). Monitoring data constrain bridge states and model parameters through observation models or likelihood functions, forming conditional posterior distributions. Reliability assessment then uses these posterior distributions to compute conditional failure probabilities, posterior predictive failure probabilities, system reliability and time-dependent reliability. Finally, reliability results enter maintenance decisions through risk and loss functions.

After the *k*th observation, recursive state and parameter updating can be summarized as

$$p\left(z_t,\theta \middle| D_{1:k}\right) \propto p\left(y_k \middle| z_t,\theta\right) p\left(z_t,\theta \middle| D_{1:k-1}\right) \tag{7}$$

where $y_k$ denotes the new monitoring, inspection or testing data assimilated at update *k*, $D_{1:k}$ denotes the accumulated data from updates 1 through *k*, $z_t$ denotes the latent structural or degradation state at time *t*, and *θ* denotes model, load, resistance or bias parameters. Equation (7) uses only the likelihood contribution of the newly acquired data $y_k$ and therefore avoids double-counting observations already contained in $D_{1:k-1}$. The conditional failure probability given the accumulated monitoring data $D_{1:k}$ is then expressed as

$$P_f(t \mid D_{1:k}) = P[g(X_t, z_t, \theta, t) \le 0 \mid D_{1:k}] \tag{8}$$

For a prediction horizon, posterior predictive failure probability further requires future loads, degradation evolution and parameter uncertainty:

$$P_f(t+\tau \mid D_{1:k}) = \int I[g(x, z_{t+\tau}, \theta, t+\tau) \le 0] p(x, z_{t+\tau}, \theta \mid D_{1:k}) dx dz d\theta \tag{9}$$

When the assessment object expands from a component to a system or network, multiple failure modes and component correlations must be considered. Under a series-system assumption, the probability that all m components remain in the safe domain is written in Equation (10). Thus, Equation (10) represents conditional series-system reliability, not system failure probability. The corresponding series-system failure probability is its complement, namely one minus the right-hand side of Equation (10). Parallel, mixed, and network systems require their own system event definitions and dependence models; Equation (10) is not a general formula for those configurations.

$$R_{\text{sys}}(t \mid D_{1:k}) = P\left(\bigcap_{j=1}^{m} \{g_j(X_t, z_t, \theta, t) > 0\} \mid D_{1:k}\right) \tag{10}$$

For long-term bridge operation, interval failure probability is more relevant than point-in-time failure probability:

$$P_f([t, t+T] \mid D_{1:k}) = P\left[\min_{\tau \in [t, t+T]} g(X_\tau, z_\tau, \theta, \tau) \le 0 \mid D_{1:k}\right] \tag{11}$$

At the maintenance-decision level, reliability indicators are transformed into expected loss or utility. The unified decision expression is

$$a^*(D_{1:k}) = \arg\min_{a\in A} E[L(a, X_{t:t+T}, z_{t:t+T}, \theta) \mid D_{1:k}] \qquad (12)$$

where a denotes candidate actions such as inspection, repair, strengthening, load restriction, closure or additional monitoring, and $L(.)$ is a loss function combining failure consequences, maintenance costs and operational impacts. The final goal of monitoring-data-driven reliability assessment is therefore not a single reliability index, but the selection of an engineering action with minimum expected loss under uncertainty.

## 4. Bridge monitoring data and structural state characterization

### 4.1. Data types and structural state descriptors

Table 1 maps typical monitoring or inspection data to extractable structural state descriptors, model updating targets and reliability inputs. It emphasizes that monitoring data are not reliability indicators by themselves; only after state characterization and model updating can they become probabilistic inputs to limit-state functions or risk assessment.

**Table 1.** Mapping from bridge monitoring data to structural state descriptors, model updating targets and reliability inputs.

| Monitoring/inspection data | Structural state descriptors extracted | Updating target and reliability input | Representative references |
|---|---|---|---|
| Vibration, acceleration and modal data | Natural frequencies, damping ratios, mode shapes, dynamic responses and input-state estimates | Updating targets: stiffness, damping, boundary conditions, damage indices and state-space models. Reliability inputs: dynamic performance limit states, posterior parameter distributions and damage-state probabilities. | [10,11,15,20–26] |
| Strain, stress and fiber-optic measurements | Strain fields, stress ranges, local load effects and fatigue-sensitive locations | Updating targets: sectional stiffness, material parameters, stress-response models and local damage variables. Reliability inputs: fatigue reliability, strength limit states and component reliability indices. | [6,23,27–32] |
| Displacement, deflection and deformation data | Deflection histories, deformation shapes, settlements and global deformation states | Updating targets: stiffness degradation, boundary conditions, serviceability states and deformation prediction models. Reliability inputs: serviceability reliability, probability of deformation exceedance and safety margin. | [33–37] |
| Crack, corrosion, visual and NDT data | Crack width, corrosion degree, damage location, damage grade and local degradation state | Updating targets: resistance models, degradation variables, damage states and spatial corrosion fields. Reliability inputs: resistance distributions, time-dependent reliability, fragility curves and remaining life. | [32,38–43] |
| Traffic loads, bridge weigh-in-motion, temperature and environmental data | Vehicle-load spectra, traffic actions, thermal effects, environmental conditions and load-response relationships | Updating targets: stochastic load models, vehicle-bridge interaction models and thermal-load models. Reliability inputs: load-effect distributions, fatigue reliability, time-varying risk and network serviceability. | [1,13,31,37,44–48] |
| Load testing and proof-load testing data | Load-response curves, lower bounds on capacity and pass/fail proof-load events | Updating targets: resistance distributions, model bias, boundary conditions and calibration factors. Reliability inputs: updated failure probability, reliability index, safety reserve and value of testing. | [5,6,49] |
| Long-term monitoring, digital twins and multi-source data fusion | State trajectories, anomaly patterns, degradation trends and physical-virtual mapping errors | Updating targets: digital-twin states, model posteriors, parameter evolution and data-assimilation models. Reliability inputs: dynamic risk, system reliability, maintenance decisions and life prediction. | [2,3,26,34,50–55] |
| Inspection ratings, maintenance records and expert information | Discrete condition states, defect records, maintenance history and inspection evidence | Updating targets: discrete state transitions, Bayesian-network nodes and degradation-state probabilities. Reliability inputs: inspection planning, value of information, maintenance priority and risk-management variables. | [16,30,43,56–62] |

The core issue in Section 4 is not to list all measurable bridge data, but to explain how monitoring data enter limit-state functions through state representation. In reliability terms, raw data are useful only when they constrain load effects, resistance degradation, model parameters, boundary conditions, model bias or failure consequences. The mapping in Table 1 can be interpreted as a data-state-variable-limit-state chain: data provide observational evidence; structural state descriptors provide physical interpretation; random variables and model parameters carry uncertainty; and the limit-state function converts them into failure probability or reliability index.

Demand-side data mainly update load effects and their time-varying distributions. Traffic-load observations and bridge weigh-in-motion (BWIM/WIM) data can provide information on vehicle weights, axle configurations, lane occupancy, vehicle classes, traffic sequences, and load effects. These data may therefore be used to update vehicle-load spectra, lane-specific load distributions, dynamic load effects, influence lines, dynamic load amplification, and load-rating-related variables. Zhou *et al.* [63] showed that uneven multilane truck loading can produce substantially different load effects and reliability levels among the girders of a multigirder bridge. Hou *et al.* [64] further demonstrated that the combination of measured structural responses and WIM truck data can support data-driven analytical load-rating assessment. At the network level, Zhou *et al.* [65] used site-specific electronic toll collection (ETC) data, traffic microsimulation, and bridge responses to evaluate the load effects of ongoing traffic and the feasibility of traffic-based load testing. These studies confirm that traffic and WIM data are important sources of demand-side evidence in bridge reliability updating. However, such data do not by themselves demonstrate an increase in structural resistance or a reduction in deterioration.

Resistance-side data provide relatively direct evidence for bridge reliability. Crack width [39], corrosion observations [41], non-destructive testing [32], ground-penetrating radar for bridge decks [43], fiber-optic strain [28], load testing and proof-load testing [4,6] can constrain sectional stiffness, reinforcement corrosion, fatigue damage, lower bounds on capacity and resistance-model bias. Field load testing is particularly valuable because it provides measured structural responses under known or partially known vehicle loads. Depending on the test design, the resulting information may include influence lines, influence surfaces, live-load distribution factors, load-rating factors, deflection-based serviceability indicators, modal properties, dynamic amplification factors and lower-bound evidence on load-carrying capacity. Limited-vehicle testing can reduce the logistical burden of conventional testing while still providing useful influence-line and capacity information [66]. Network-level and conventional static or dynamic load tests can also support the joint assessment of load distribution, serviceability, dynamic response and finite-element model calibration [67,68]. Therefore, load-testing data provide a relatively direct connection between measured bridge responses and capacity-related reliability variables, although measurement uncertainty, model-form error and extrapolation to untested load levels must still be considered. These data are more directly informative about the resistance model $R(\mathbf{X}_R, t)$ and associated degradation processes, and they often provide stronger evidence for remaining capacity and time-dependent reliability than global vibration features alone.

Model-side data mainly correct the discrepancy between structural models and real bridges [10,15]. Acceleration, modal parameters, dynamic strain, displacement responses, foundation responses, railway-bridge operating responses and long-term structural health monitoring (SHM) data can identify global stiffness, boundary conditions, connection states, foundation constraints, substructure parameters or state-space models [24–26,69,70]. Their advantage is that they provide continuous evidence for finite-element updating and state estimation, but their ability to identify local fatigue cracks, hidden corrosion or sectional resistance degradation is often indirect.

The value of bridge monitoring data should therefore be assessed by the strength of their link to the limit state, rather than by data volume. High-frequency acceleration data may significantly improve modal identification but may not constrain fatigue crack growth. Conversely, a small amount of load-testing or crack-inspection information may substantially alter a resistance distribution or failure probability. Reliability assessment must identify whether each data type updates demand-side, resistance-side or model-side variables, and must avoid equating condition monitoring results with improved safety reliability.

### *4.2. Transformation from monitoring data to reliability inputs*

Mathematically, monitoring data enter reliability analysis through at least four layers. The first layer is observational data *D*, including sensor responses, inspection images, load-testing records, vehicle loads, environmental actions or condition ratings. The second layer consists of structural state descriptors $z_t$, such as frequency, strain range, deflection, crack width, corrosion degree, scour depth, bearing condition or stiffness degradation. The third layer consists of reliability variables, including load effect $\mathbf{X}_S$, resistance variable $\mathbf{X}_R$, degradation parameters, model bias and measurement error. The fourth layer is the limit-state function $g(\mathbf{X}, z_t, \theta, t)$. Monitoring information enters failure probability calculation only when *D* affects $\mathbf{X}_S$, $\mathbf{X}_R$, $\theta$ or $g(\bullet)$ through $z_t$.

Different data types have different strengths of association with limit states. Strongly associated data include load testing, proof-load testing [6], local strain, cracks [41], corrosion, weigh-in-motion [1,31,37,47] and bridge-deck NDT [43], which can directly constrain resistance, load effect or degradation variables. Moderately associated data include deflection, displacement, temperature, long-term response trends and influence lines; their reliability contribution depends on whether a stable mechanical mapping can be established [25,33,36]. Weakly or indirectly associated data include frequency and acceleration features used only for modal identification; their reliability contribution must be transferred through structural and error models [10,20,26].

This association strength determines the boundary of data use. For fatigue reliability, vehicle-load spectra and stress-range data are more important than global modal frequencies. For corrosion-degradation reliability, crack width, cover deterioration and corrosion evidence are more explanatory than a single dynamic test. For scour risk, water level, flow velocity, foundation response and scour depth are closer to the limit state than superstructure responses. For bridge-network reliability, single-bridge condition must be further mapped to traffic function, detour cost and recovery time.

Reliability-oriented structural state characterization should answer three questions: whether the extracted state corresponds to a limit-state variable, how state-estimation error enters failure probability, and whether state changes alter maintenance decisions. If a monitoring feature only improves response-fitting accuracy but does not change the load-effect distribution, resistance distribution, degradation model or decision threshold, its contribution to reliability assessment is limited. Conversely, even low-frequency inspection data may have high value if they alter a lower bound on resistance or the posterior failure probability.

The transformation logic can be summarized as follows: data *D* constrain states and model parameters through an observation model $p(D \mid z_t, \theta)$, modify the variable distributions in the limit-state function through the posterior distribution $p(z_t, \theta \mid D)$, and finally produce conditional failure probabilities, posterior predictive failure probabilities or maintenance decision losses. This logic separates data-processing accuracy from reliability contribution and forms the basis for subsequent model updating, uncertainty quantification and maintenance decision analysis.

Recent studies have further extended load testing toward rapid and network-level bridge assessment. Zheng *et al.* used a limited number of testing vehicles to identify bridge influence lines and evaluate the load-carrying capacity of a long-span suspension bridge [66]. Zhou *et al.* developed an open-traffic load-testing approach using moving artificial truck fleets, computer vision, traffic-load reconstruction, and hybrid virtual-real traffic simulation, and demonstrated its use across seven long-span bridges [67]. These studies indicate that load testing can evolve from a bridge-closure-based activity for a single bridge into a flexible source of structural-response evidence for network-level condition assessment. However, the reliability meaning of the resulting information still depends on whether the identified influence lines, load effects, or response indicators are explicitly linked to the relevant limit-state function.

The reliability contribution of traffic-load and BWIM/WIM data depends on the variable that is updated. At the demand level, these data can modify the distribution of vehicle weights, axle loads, lane-specific traffic effects, dynamic amplification, influence-line responses, and fatigue stress ranges [63–65]. At the model level, the combination of WIM data and measured structural responses can update influence lines, load-distribution factors, or load-rating-related parameters. However, the updated demand model must still be linked to the relevant limit-state function. Traffic-load observations collected during normal operation may not adequately characterize extreme traffic events or the tail of the load-effect distribution. Extreme-value modelling, extrapolation uncertainty, traffic growth, dependence between vehicles, and the representativeness of the monitored traffic records should therefore be explicitly considered.

## 5. Model updating methods based on bridge monitoring data

Model updating methods are best read as transformation operators between monitoring evidence and reliability variables. Finite-element updating, Bayesian updating, filtering, surrogate modelling, machine learning and digital twins differ in what they transform: physical model parameters, posterior uncertainty, online states, computational cost, data-driven mappings or closed-loop decision processes. Their value should therefore be judged by whether they change limit-state variables, uncertainty or decisions, not merely by response-fitting accuracy.

This section uses three criteria: applicable conditions, failure scenarios and reliability contribution. Applicable conditions define when data, models and objectives support a method; failure scenarios identify violated assumptions that may mislead conclusions; reliability contribution asks whether the method changes load effects, resistance, degradation, model bias, failure probability or maintenance action.

Tables 2 and 3 summarize these criteria. Table 2 clarifies method boundaries and complementarities, whereas Table 3 compares advantages, limitations and application scenarios. Load testing also illustrates the difference between response fitting and reliability-oriented model updating. Measured responses from controlled vehicle passages can be used to update influence lines, stiffness parameters, boundary conditions, load-distribution factors or finite-element model parameters. Zheng *et al.* [66] demonstrated the use of limited testing vehicles and prior information for influence-line correction and load-carrying-capacity evaluation, while Dong *et al.* [68] showed that static and dynamic load tests can jointly support load-rating assessment, serviceability evaluation, dynamic-response identification and finite-element model calibration. These studies provide physically interpretable updating targets. Nevertheless, a calibrated response model should not be regarded as a complete reliability model unless the updated parameters are propagated through the relevant limit-state function and their measurement, model-form and extrapolation uncertainties are quantified. Zhou *et al.* [65] further showed that moving artificial truck fleets can provide network-level response evidence under open-traffic conditions, but the applicability of such data depends on the achieved load level and on the representativeness of the tested traffic scenarios.

**Table 2.** Boundaries, applicable conditions and complementarities of model updating methods.

| Method | Applicable conditions and role | Failure scenarios and main limitations | Reliability contribution and representative studies |
| --- | --- | --- | --- |
| Bayesian model updating | Transforms monitoring or inspection data into probabilistic updates of parameters or states through priors, likelihood functions and posteriors. It is suitable for limited data, uncertainty representation and direct coupling with reliability updating. | It depends strongly on the prior, likelihood and error model. Sampling becomes difficult for high-dimensional or computationally expensive models. | It can be accelerated by surrogate models, implemented online by filtering and embedded in digital twins to form a closed loop. Representative studies: [6–9,12,48] |
| Filtering and sequential estimation | Recursively estimates structural states, inputs, parameters and noise statistics through state-space models. It is suitable for long-term monitoring, online warning, tracking of time-varying parameters and unknown-input estimation. | Results are sensitive to the state equation, noise statistics, linearization and filter stability. | It can be viewed as an online form of Bayesian updating and can provide real-time data assimilation for digital twins. Representative studies: [10,11,40,71,72] |
| Surrogate modelling and machine learning | Uses Kriging, neural networks, generative models or multi-task learning to approximate expensive models or learn data-state mappings. It is useful for high-dimensional, high-cost and multi-source-data problems. | Physical interpretability and extrapolation ability are limited; uncertainty calibration and surrogate-error control are critical. | It accelerates Bayesian updating and reliability analysis and supports fast prediction and missing-data treatment in digital twins. Representative studies: [23,50,73–79] |
| Digital twins | Organizes physical models, monitoring data, data assimilation, condition assessment and maintenance decisions into a continuously evolving cyber-physical representation. | It is not a single algorithm. Implementation depends on data quality, model updating, reliability analysis and decision modules. | It provides a system-level integration platform for Bayesian updating, filtering, surrogate modelling and risk-based decision-making. Representative studies: [3,34,51,54,55] |
| Finite-element/ physics-based model updating | Physically interpretable and compatible with code checking and engineering judgement. It is suitable when a credible physical model exists and measured responses can be linked to model outputs. | It depends on the initial model. Model-form error, parameter correlation and sparse sensors may lead to non-unique updating. Use with caution when structural details or damage mechanisms are unclear. | [5,15,24,26,32,70,80,81] |
| Bayesian model updating | Unifies prior knowledge, monitoring evidence and posterior uncertainty. It can pass updated parameter distributions directly to reliability analysis. | It depends on the prior, likelihood and error model; high-dimensional sampling is expensive and model-form error may be underestimated. | [6–9,12,48] |
| Filtering and sequential estimation | Suitable for online monitoring and recursive state/parameter tracking, including unknown inputs and noise statistics. | Sensitive to state equations, noise covariance and linearization; abrupt damage or strong nonlinearity may cause filter drift. | [10,11,22,40,71] |
| Surrogate models and active-learning reliability | Greatly reduces calls to expensive models and is suitable for small failure probabilities, high-dimensional random variables and complex limit states. | Surrogate errors near the failure boundary may dominate the reliability result; extrapolation and uncertainty calibration are difficult. | [73,76,77,79,82–84] |
| Machine-learning-assisted updating | Handles high-dimensional monitoring, video/images, multi-source sensing and simulation-to-measurement gaps with high computational efficiency. | Weak physical interpretability and dependence on training data; credibility near extreme states and failure boundaries is limited without physics constraints. | [23,50,74,75,85] |
| Digital twins and multi-source fusion | Integrates monitoring, physical models, data assimilation, reliability assessment and maintenance decisions for life-cycle management. | Implementation is complex; without uncertainty propagation and decision modules, a twin may remain a visualization or dashboard. | [3,34,51,54,55,86] |

Across method families, reliability contribution depends on the data-state-limit-state association. Finite-element and physics-based updating are useful when monitored responses can identify interpretable stiffness, boundary [15,69], damping [80], damage [87,88] or resistance parameters; Bayesian updating is valuable when limited monitoring, inspection or load-testing evidence can be expressed through defensible priors, likelihoods and error models [26,32,70,89]. Filtering supports online recursion when state equations and noise models are credible [25,40,71,72,74], while machine learning [75,78,79,90], surrogate models [76,91–95], and active learning [77,82,96–98] mainly contribute by reducing computational cost, accelerating failure-domain search or handling high-dimensional inverse problems. Digital twins add reliability value only when they integrate data assimilation, uncertainty propagation, failure-probability prediction and maintenance feedback; otherwise they remain condition-display platforms [3,34,51,54,55].

The representative papers in Table 4 illustrate the same point: substructure Bayesian updating [8], proof-load reliability updating [6], Kriging error control [76] and digital-twin [3,51,54] studies are important only when they occupy a clear position in the reliability chain and change posterior failure probability, reliability index, risk loss or maintenance action. Without this verification layer, model updating remains response fitting rather than reliability updating.

**Table 4.** In-depth analysis of representative studies.

| Representative study | Core problem and methodological contribution | Main limitation | Implication for this review |
|---|---|---|---|
| Bayesian framework [7] | Addresses how observation bias enters structural model updating; replaces point estimates with posterior distributions. | Depends on prior assumptions and the error model. | Provides the probabilistic basis for reliability updating. |
| Bridge substructure updating [8] | Reduces the dimensionality of full-bridge updating by combining substructuring and adaptive sampling. | Affected by substructure partitioning and sensor layout. | Shows that localized identification can be more informative than blind global calibration. |
| Time-dependent reliability updating [9] | Updates reliability trajectories of degrading structures using inspection evidence. | Sensitive to degradation and inspection-error models. | Directly couples model updating with time-dependent reliability analysis. |
| Proof-load testing plus monitoring [6] | Uses proof-load events and monitoring data to verify capacity of existing structures. | Information is limited by the applied proof-load level. | Suitable for evaluating safety reserves of existing bridges. |
| Kalman tracking [10] | Tracks parameters online as monitoring data evolve through recursive state-parameter estimation. | Sensitive to noise assumptions and linearization. | Useful for long-term monitoring and warning. |
| FE-aided Kalman filtering [11] | Estimates responses and parameters under unknown loads by combining FE models with filtering. | Computationally demanding and dependent on model quality. | Illustrates the complementarity between physics-based models and online assimilation. |
| Kriging error control [76] | Quantifies and controls surrogate errors in reliability analysis with equality information. | Mainly addresses computational error. | Reminds that tail reliability requires control of boundary errors. |
| Multi-task learning updating [23] | Jointly identifies parameters using strain and acceleration through multi-task learning. | Affected by task weights and noise levels. | Represents a route for combining multi-source data fusion with probabilistic updating. |
| Digital-twin framework [3] | Connects monitoring, models and decisions in a digital-twin framework for civil structures. | Framework-oriented and requires implementation details. | Clarifies that a digital twin is a system integration platform. |
| Physics-informed twin [34] | Embeds physical constraints into deformation monitoring of dams. | Generalization across operating conditions still needs verification. | Represents a move from data display toward safety-assessment closure. |

## 6. Uncertainty quantification and reliability updating

The goal of uncertainty quantification in bridge reliability is not simply to reduce uncertainty, but to establish a traceable error budget. For traffic-load updating, the error budget should include vehicle-weight estimation error, axle-spacing and lane-identification error, incomplete traffic-event capture, spatial transfer error, temporal non-stationarity, dependence among successive vehicles, and uncertainty in extrapolating observed traffic to design or extreme traffic conditions. The use of site-specific WIM or ETC data can reduce epistemic uncertainty in the demand model, but it does not eliminate aleatory variability or tail-model uncertainty. Monitoring data may reduce epistemic uncertainty in material parameters, boundary conditions or load models; they may also reveal model biases, environmental effects and degradation mechanisms not represented in the original model, thereby increasing model-level uncertainty. Reliability updating should therefore distinguish aleatory uncertainty, epistemic uncertainty, model-form error, measurement error, degradation-process uncertainty and computational error.

Bayesian reliability updating is credible when observations constrain reliability variables through likelihood functions and when priors, error models and limit-state functions are explicit [6,9,12]. The posterior should not be interpreted as safer merely because it is more concentrated: model-form error, measurement error, environmental effects, bearing constraints, connection details and traffic-load bias may dominate the updated result [27,72,99].

Small failure-probability estimation is difficult because bridge monitoring data mostly describe normal operation, whereas failure probability is controlled by extreme loads, fatigue tails, advanced corrosion, scour, or seismic intensity. Active-learning Kriging, subset simulation, importance sampling, Langevin sampling, Hamiltonian neural networks, and graph-neural-network surrogates reduce computational burden, but credibility still requires evidence that the data constrain variables near the failure domain and that surrogate or sampling error is controlled there. Load-testing evidence can provide a useful intermediate link between ordinary monitoring responses and capacity-related variables, but it does not remove the uncertainty associated with extrapolation to untested load levels.

Time-dependent, system and network reliability require uncertainty propagation across time and scale. Fatigue [13], corrosion [41], crack growth [42], thermal loading [46] and ageing-bridge fragility [100,101] studies must connect observed degradation states to future load and resistance processes rather than simply add time to equations. At system or network scale, posterior component reliabilities must further pass through dependence, common hazards [102], traffic demand [17,58,62], alternative paths and recovery strategies before becoming service-risk indicators [103–106].

Overall, a credible bridge reliability update should report the updated object, error sources, posterior distributions, sensitive variables, computational error and decision implications, rather than only the numerical change in failure probability or reliability index.

## 7. Reliability, risk and fragility assessment methods for bridges

This section concerns how updated states and uncertainties become bridge safety, risk and fragility indicators. The central issue is not which algorithm computes failure probability, but whether the chosen limit state, random variables, failure mode and consequence model match the engineering decision.

Component and time-dependent reliability are appropriate when failure modes and degradation mechanisms are sufficiently identifiable. Component analysis converts local monitoring, inspection, load-testing [66,67] or proof-load [6] evidence into failure probability, safety reserve, load-rating information or capacity-related model parameters for steel members, cracked RC members, bridge decks and existing bridge systems [4,16,43,49,107]. In particular, load-testing studies provide an important intermediate link between measured structural responses and resistance-related reliability variables. Influence lines, load-distribution factors, load-rating factors, serviceability responses and calibrated finite-element parameters can be used to update the demand or resistance model. However, these outputs should not be treated as direct substitutes for failure probability, because the transition from test-level responses to ultimate-limit-state reliability requires explicit consideration of uncertainty and load-level extrapolation. Time-dependent reliability extends this logic to remaining life, inspection intervals and maintenance timing when degradation state or rate can be updated from observations [9,13,42,100,101]. Their shared risk is over-extrapolating sparse local data or short-term trends without spatial representativeness or physical degradation support.

Fragility, system reliability, network risk and maintenance decision analysis extend reliability assessment from components to hazards, systems, service functions and actions. Fragility is useful for seismic [108–110], flood [111,112], scour [113,114], wind, ship-impact or extreme-traffic scenarios when hazard intensity, damage state and functional loss are defined. System [59,79,115,116] and network [17,58,62,103,106] methods are needed when dependence, redundant load paths, traffic connectivity, detour cost, recovery time and maintenance resources control risk. Decision-oriented assessment then translates failure probability or reliability index into expected loss, value of information, load restrictions, repair priorities or budget allocation [56,57,59–61].

Thus, reliability, risk, fragility and resilience are not independent method labels; they express the consequences of the same monitoring evidence at different spatial scales and for different decision objectives. Accordingly, this review emphasizes when these indicators can be transformed or connected, rather than presenting the calculation methods as unrelated lists.

## 8. Bridge infrastructure application scenarios

### *8.1. Scenario classification*

Bridge infrastructure itself contains distinct application scenarios. Table 5 summarizes the bridge application matrix, indicating that the review covers not only single bridge components but also single-bridge systems, existing bridges, railway bridges, highway bridges and bridge networks.

**Table 5.** Application matrix for bridge infrastructure reliability assessment.

| Bridge application scenario | Data and methods | Assessment target and representative problem | Representative references |
|---|---|---|---|
| Highway bridges and bridge networks | Traffic loads, weigh-in-motion, acceleration, strain, deflection, cracks and long-term monitoring; FE updating, Bayesian updating, filtering, Kriging reliability and network resilience analysis. | Component reliability, fatigue life, load-rating decisions, network resilience and maintenance priority; examples include foundation identification, fatigue, proof-load updating and network resilience. | [1,5,6,8,13,15,17,30,31, 42,47,59,62,77,79,106, 113] |
| Railway bridges and dynamically sensitive bridges | Train loads, acceleration, dynamic strain, displacement and track-bridge interaction responses; ambient-vibration calibration, filtering, vehicle-bridge interaction models and state-space updating. | Operational safety, dynamic performance, parameter updating and load-effect identification; examples include railway-bridge monitoring and moving-load identification. | [11,15,24–26,44,45,81] |
| Existing bridges and degrading components | Crack width, corrosion information, NDT, proof-load data, condition ratings and material degradation data; Bayesian inversion, reliability updating, degradation models and inspection-data fusion. | Remaining life, resistance updating, service safety and repair decisions; examples include historic steel bridges, cracked RC beams, corrosion fields and existing-structure reliability. | [4,6,30,32,39,41,43,49, 61,62] |
| Long-span and multi-source monitored bridges | Long-term monitoring, temperature, wind, strain, displacement, vision and multi-source sensing; digital twins, Bayesian updating, machine learning and data fusion. | Dynamic risk, condition assessment, model updating and predictive maintenance; examples include digital-twin damage assessment and separation of environmental effects. | [2,3,26,34,36,50–55] |
| Bridge networks and system risk | Single-bridge condition, traffic flow, network topology, hazards and simulation data; Bayesian networks, system reliability, data-driven seismic analysis and network resilience assessment. | System performance, network reliability, long-term resilience and risk management; examples include large-scale transport risk, deteriorating bridge networks and maintenance-resource allocation. | [14,17,58,59,62,79, 102–104,106] |

Bridges are the core application object of this review. The literature covers bridge weigh-in-motion [1], boundary-condition updating, fatigue life [13], foundation identification [15], load testing [5], railway-bridge dynamic responses [81], bridge condition ratings [8], failure scenarios of steel truss bridges [107], thermal loads at bridge expansion joints [16,46] and bridge-network resilience [17]. Bridge-specific journals further add studies on scour monitoring [113], monitoring before and after railway-bridge retrofitting [24], strain-threshold condition assessment [30], hybrid strain-acceleration weigh-in-motion [31], Kriging-based bridge weigh-in-motion [47] and human-machine collaboration in bridge health monitoring [53], strengthening the bridge-oriented scope of the review.

Compared with bridges, buildings, dams, offshore platforms and transportation networks appear mainly as transfer-boundary and comparison scenarios. Building structures emphasize seismic response [23], modal identification [52], component or system fragility [108] and regional risk [111,117–119]. Dams [34], wharves [38] and offshore platforms [56,57] focus more on deformation safety, corrosion fatigue, inspection planning and high-consequence risk control. Transportation networks emphasize functional loss, alternative paths, recovery time and system resilience [14,17,58,103,104].

These non-bridge scenarios show that bridge reliability updating methods have transferable components, but their applicability must be re-evaluated. Limit states, data frequency, monitoring accessibility, failure consequences and maintenance constraints vary across infrastructure types. Priors, likelihood functions, degradation models and maintenance thresholds derived from bridges should not be directly applied to other infrastructure without reassessment.

Within bridge applications, existing bridges and special components constitute an important direction. Studies of historic steel bridges, multi-source NDT, proof-load testing, cracked RC beams, spatial corrosion distributions and reliability assessment of RC structures under environmental actions show that existing-bridge assessment requires integrating inspection, monitoring, material degradation and model updating evidence [39,41,49,120]. Studies on ground penetrating radar (GPR) inspection of bridge decks, strengthening assessment of continuous girder bridges and network-level deterioration prediction further demonstrate that inspection data and maintenance history can be used not only for condition rating but also as inputs for reliability updating and maintenance prioritization [32,43,62].

### *8.2. Scenario differences and method-transfer boundaries*

Scenario differences within bridge infrastructure are reflected not only in bridge type and structural configuration, but also in data availability, dominant failure modes, association strength between monitoring data and limit states, and maintenance-decision constraints. Highway bridges often provide traffic loads, strain, acceleration and long-term monitoring data, making them suitable for model updating, fatigue reliability and load-restriction studies. Railway bridges emphasize train loads, dynamic response, track-bridge interaction and operational safety. Long-span bridges emphasize wind, temperature, fatigue and multi-source monitoring-data fusion. Existing medium and small bridges are often limited by low inspection frequency, sparse data and incomplete structural information. Bridge networks further concern service function, alternative routes and post-disaster recovery.

Method transfer within the bridge domain also requires assumptions. Bayesian updating is relatively general, but priors, likelihood functions and error models must be rebuilt for different bridge types, materials and degradation mechanisms. Filtering and data assimilation are appropriate for continuous monitoring and recursive state estimation, but they are not always suitable for low-frequency manual inspection, discrete ratings or irregular detection data. Surrogate models and machine learning improve computational efficiency for high-dimensional reliability analysis, but their training samples, failure-boundary coverage and extrapolation capability must be validated for specific bridge types and load cases. Digital twins are suitable for large bridges with long-term monitoring, but they impose higher requirements on data governance, model maintenance, uncertainty propagation and decision modules.
Even within bridge engineering, mature case experience cannot be extrapolated unconditionally. Highway-bridge traffic-load models cannot replace railway train-load models. Wind-induced responses and temperature effects in long-span bridges cannot be simply transferred to ordinary girder bridges. Single-bridge component reliability cannot be equated with bridge-network risk. If differences in bridge type, load spectrum, monitoring frequency, failure consequence and maintenance constraint are ignored, method transfer may preserve only the algorithmic form while losing the physical meaning required for reliability assessment.

Transferable elements in bridge reliability updating are mainly probabilistic inference logic, posterior uncertainty propagation and risk-decision thinking. Non-transferable elements include specific priors, likelihood functions, limit-state functions, degradation models and maintenance thresholds. Fatigue-sensitive bridges require reconstruction of vehicle-load spectra and stress-range models. Corrosion-degrading bridges require crack, corrosion and resistance-degradation evidence. Railway bridges require train loads, track-bridge interaction and operating-safety constraints. Bridge networks require propagation of single-bridge or component failure probabilities to traffic capacity, detour cost and recovery time.

Current gaps remain substantial. The mapping from long-term monitoring data to local fatigue cracks, reinforcement corrosion and other key failure mechanisms is still insufficient. Bridge weigh-in-motion and traffic-load data mainly update demand-side variables and provide limited explanation of resistance degradation. The transfer from single-bridge reliability to network risk, traffic capacity and maintenance priority needs stronger mechanisms. Uncertainty propagation and reliability-decision modules in bridge digital twins remain underdeveloped. Future research should move from individual bridge cases toward cross-bridge-type, cross-scale and network comparisons, clarifying which parts of the data-model-reliability-decision chain are transferable and which are not.

## 9. Maintenance decision-making and risk management

The engineering value of bridge reliability assessment ultimately lies in decision support. For bridge managers, the purpose of monitoring data and model updating is not merely to obtain more accurate model parameters, but to support inspection planning, repair and strengthening, load restriction, risk transfer and asset management. Reliability results must therefore be translated into executable maintenance and risk-control strategies.

Inspection planning and value-of-information analysis link monitoring evidence to maintenance decisions. Studies on jacket platforms and marine concrete structures [56,57] are treated here only as supplementary method-transfer examples for reliability-based inspection planning; they are not counted as bridge-specific evidence. The bridge-focused evidence includes proof-load reliability updating [6], scour monitoring [113], bridge condition thresholds [30], GPR bridge-deck inspection [43] and bridge maintenance frameworks [61]. Together, these studies show how to determine when to inspect, what information to collect, whether to perform load testing and whether additional monitoring is worth its cost.

For maintenance and life-cycle management, monitoring-informed reliability assessment should not stop at describing deterioration; it should translate fatigue, corrosion and degradation evidence into decisions on inspection intervals, load restrictions, repair timing and strengthening schemes. Studies on post-tensioned concrete box-girder bridges show that reliability updating can revise life estimates under coupled traffic loading and corrosion [100], while Markov-chain degradation models provide a probabilistic basis for remaining-life prediction [101]. Work on historical steel bridges further demonstrates how non-destructive characterization and reliability assessment can prioritize intervention when structural capacity, documentation and inspection information are uncertain [4]. Recent bridge-management studies in Structure and Infrastructure Engineering and Journal of Infrastructure Systems extend this decision logic from component-level maintenance to life-cycle performance, network connectivity under corrosion and deterioration prediction after maintenance [59–62,106].

At the system level, bridge risk management must account for network function and resilience. Large-scale transportation network risk, long-term resilience of deteriorating bridge networks, performance of complex infrastructure systems and seismic risk transfer for high-speed railway bridges indicate that bridge reliability assessment is expanding from single-bridge safety to network serviceability and risk-control strategies [14,17,58,121]. Network-level maintenance strategies, network deterioration prediction and surrogate seismic reliability for highway bridge systems further show that network decisions must jointly consider component correlation, maintenance intervention, traffic function and computational efficiency [62,79,106].

From a decision perspective, monitoring-data-driven bridge reliability assessment can support at least five engineering actions: optimizing inspection intervals according to reliability deterioration, information value and inspection cost; ranking repair or strengthening priorities under budget constraints; implementing load restriction or operational control when updated failure probability or risk exceeds thresholds; optimizing monitoring systems by evaluating whether new sensors, load tests or NDT have value; and optimizing life-cycle cost under safety, cost, serviceability and social-impact constraints.

To convert bridge reliability results into actions, updated failure probability, reliability index and consequence loss can be included in a decision loss function. Let $D_k$ be the information set after the *k*th monitoring, inspection or load-testing event, and let $a$ in A denote candidate actions such as continued monitoring, intensified inspection, repair, strengthening, load restriction or temporary closure. The expected total cost of action a can be written as

$$J(a \mid D_k) = C_I(a) + C_M(a) + C_U(a) + C_F(a) P_f(t + \Delta t \mid D_k, a) \tag{13}$$

where $C_I$, $C_M$, $C_U$ and $C_F$ denote inspection/monitoring cost, maintenance/strengthening cost, operational-impact cost and failure-consequence cost, respectively. $P_f(t + \Delta t \mid D_k, a)$ is the predicted failure probability in the decision horizon after action $a$. This expression places reliability updating, consequence assessment and operation cost on the same decision scale.

When costs are quantifiable and safety constraints exist, maintenance decision-making can be written as the constrained optimization problem

$$a^*(D_k) = \arg\min_{a \in A} J(a \mid D_k), \quad \text{s.t.}\ \beta(t + \Delta t \mid D_k, a) \ge \beta_{\text{req}} \tag{14}$$

In Equation (14), the reliability constraint should be defined using an assessment-specific target reliability level. This target may depend on the bridge type, structural redundancy, deterioration condition, traffic function, failure consequence, network importance, reference period, and available maintenance resources. Therefore, $\beta_{\text{t,assess}}$ should be determined within a risk-informed assessment framework rather than directly adopted from reliability-based design calibration. The basis for this assessment-specific target should be documented using risk, consequence, network, and cost considerations [18,19].

When consequences, traffic impacts, or social losses cannot be fully monetized, the monitoring-informed maintenance decision can be expressed as a threshold-based mapping from the updated information set $D_k$ to an engineering action:

$$a(D_k) = \begin{cases} \text{continue monitoring}, & \beta_k \ge \beta_{\text{req}}, \\ \text{intensified inspection or load restriction}, & \beta_{\text{warn}} \le \beta_k < \beta_{\text{req}}, \\ \text{repair, strengthening, or closure}, & \beta_k < \beta_{\text{warn}}. \end{cases} \tag{15}$$

Here, $\beta_k$ denotes the reliability index updated using the monitoring data up to decision epoch $k$, $\beta_{\text{req}}$ is the assessment-specific required reliability level for continued operation of an existing bridge, and $\beta_{\text{warn}}$ is an assessment-specific early-warning threshold that triggers more conservative intervention. Both $\beta_{\text{req}}$ and $\beta_{\text{warn}}$ should be calibrated according to the assessed bridge, limit state, reference period, failure consequence, redundancy, traffic function, and maintenance decision context. They should not be directly equated with the design target reliability index for new structures.

Target reliability selection is itself a decision problem in existing-bridge assessment. Kim *et al.* showed that bridge-network target reliability levels can be differentiated according to structural importance, maintenance cost, user cost, network reliability, and redundancy [18]. Su *et al.* further proposed a practical risk-informed framework for determining target reliability indices for existing structures by considering context definition, failure statistics, individual and societal risk criteria, as low as reasonably practicable (ALARP) principles, cost optimization, and marginal lifesaving cost [19]. These studies indicate that an assessment target should be linked to risk acceptance and management consequences. The design target reliability index may be used as a background benchmark, but it should not be directly imposed as the acceptance criterion for existing-bridge condition assessment.

For deciding whether to install additional sensors, perform NDT or conduct load testing, value of information can be used. Let $e$ be a candidate inspection or monitoring scheme, $y_e$ its possible observation and $C_e$ its implementation cost. The value of information can be written as

$$VOI(e \mid D_k) = \min_a J(a \mid D_k) - E_{y_e \mid D_k}\left[\min_a J(a \mid D_k, y_e)\right] - C_e \tag{16}$$

When $VOI(e \mid D_k)$ is positive, the expected improvement in the decision exceeds the cost of collecting the additional information. This makes the additional monitoring or inspection rational.

Based on these expressions, monitoring-data-driven maintenance can form a closed data-reliability-risk-action-feedback loop. First, monitoring data are quality-controlled and anomalies caused by sensor drift, environmental interference or obvious data errors are removed. Second, state variables related to limit states, such as load effects, degradation rates, cracks, corrosion, stiffness or boundary conditions, are extracted. Third, model parameters, state variables and model bias are updated, and posterior failure probabilities, reliability indices or risk indicators are calculated. Fourth, candidate actions are generated, including continued operation, intensified monitoring, inspection, repair, strengthening, load restriction, closure and risk transfer. Fifth, cost-risk optimization, threshold triggering or value-of-information analysis is performed using Equations (13)–(16). Sixth, new data after implementing the action are incorporated into $D_{k+1}$, initiating the next reliability-update and maintenance-decision cycle.

A limitation of many existing studies is that they compute updated reliability indices but do not explain how these indices change actual decisions. Does an increased reliability index justify a longer inspection interval? Does a decreased reliability index correspond to a load-restriction or repair threshold? Is the uncertainty reduction brought by monitoring sufficient to offset inspection and maintenance costs? Future research should move from updating reliability numbers to comparing expected risk and information value among alternative decisions.

For bridge networks, maintenance decisions should also extend from single-bridge or single-component safety to system serviceability. A bridge or component with lower reliability is not necessarily the highest network risk contributor. Its impact also depends on traffic flow, alternative routes, recovery time and failure correlation. System-level bridge maintenance therefore requires joint modelling of component posterior reliability, network topology, demand assignment and recovery strategies, rather than equating component-level reliability ranking with asset-management priority.

## 10. Major challenges and future research directions

The challenges in this field can be elevated to several scientific questions. First is the identifiability of reliability variables from monitoring data. Bridge measurements are often acceleration, strain, displacement, images or inspection ratings, whereas reliability variables are resistance, load effect, degradation rate and model bias. Future work must answer which monitoring quantities identify which reliability variables, how identification strength should be quantified, and whether weakly associated data may create a false sense of safety.

Second is the separation between model-form error and parameter uncertainty. Existing model updating often attributes finite-element response residuals to parameter deviations, but existing bridges may contain structural-model errors due to boundary conditions, connection details, construction deviations, local damage mechanisms and environmental effects. Future frameworks should jointly represent parameter uncertainty, model-class uncertainty and bias functions to avoid overly concentrated posteriors under incorrect models.

Third is the information mismatch between normal-operation monitoring and tail failure probability. Bridge SHM data mainly cover daily operating conditions and relatively low loading levels, whereas reliability may be controlled by extreme vehicle loads, fatigue tails, scour extremes, advanced corrosion, seismic intensity, nonlinear load redistribution, or ultimate load-carrying capacity. Good agreement between measured and simulated responses under ordinary service conditions therefore does not necessarily imply accurate estimation of small failure probabilities or ultimate-limit-state reliability. Future studies should connect normal-operation data to the failure domain through physical bridge-response models, degradation models, load testing, extreme-value statistics, active-learning sampling, and explicit model-discrepancy terms.

Fourth is the transfer of methods across bridge types and scales. Highway bridges, railway bridges, long-span bridges, existing medium and small bridges, and bridge networks differ in load spectra, failure consequences, monitoring frequency, maintenance constraints and function definitions. Future work should distinguish transferable probabilistic inference logic from non-transferable priors, likelihood functions, degradation models and maintenance thresholds, avoiding direct generalization from single-bridge cases to network-level risk management.

Fifth is the reliability meaning of digital twins. A bridge digital twin should not stop at three-dimensional visualization, data dashboards or deterministic condition assessment. It should answer how current failure probability changes, how future risk evolves, whether additional monitoring has information value, and which maintenance action minimizes expected loss. Evaluation criteria for bridge digital twins should therefore include data assimilation, uncertainty propagation, reliability prediction and decision feedback.

Sixth is the translation of reliability results into maintenance actions. Many studies provide updated $P_f$ or $\beta$ but do not specify how these results modify inspection intervals, load-restriction strategies, repair priorities or monitoring layouts. Future research should combine reliability updating with value of information, life-cycle cost, traffic impacts and risk-acceptance criteria to form a computable loop from monitoring evidence to action selection.

Seventh is the lack of reproducible benchmarks and engineering validation. Studies differ in data types, bridge types, model-error assumptions and reliability targets, making cross-method comparison difficult. Open benchmark cases for bridge reliability updating are needed, including raw monitoring data, structural models, inspection records, load-testing information, failure-mode definitions and maintenance-decision scenarios. Such benchmarks would allow different methods to be tested on the same reliability task and would clarify their real contributions.

## 11. Conclusions

This review critically synthesized how monitoring evidence supports reliability assessment and maintenance decisions for bridges. The 56-record core bridge dataset supports bridge-specific conclusions, while the remaining 58 records provide clearly labelled method-transfer or comparison evidence. The descriptive counts characterize this selected corpus and are not a census of the field.

The synthesis shows that monitoring data become decision-relevant only when they identify limit-state variables and their uncertainty is propagated into failure probability, risk and action selection. Finite-element updating, Bayesian inference, sequential filtering and surrogate modelling offer different combinations of physical interpretability, probabilistic updating, online recursion and computational efficiency. Their transfer across bridge types and scales requires rebuilding problem-specific priors, likelihoods, degradation models and decision thresholds.

Future work should strengthen traceable mappings from data to state variables and limit states, represent model-form and measurement uncertainty explicitly, and connect normal-operation observations to rare-event reliability through physical models and defensible extrapolation. Reproducible bridge benchmarks should couple monitoring data, structural models, failure-mode definitions and maintenance scenarios. Reliability-oriented digital twins should be judged by their ability to assimilate data, propagate uncertainty, predict reliability and feed results back into decisions, as defined in Sections 5 and 10.

## Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used generative AI tools only to improve language and readability. Specifically, the authors used ChatGPT for language polishing only. The authors take full responsibility for the content of the manuscript.

## Acknowledgments

The authors would like to acknowledge the support from the National Key Research and Development Program of China (No. 2023YFB2604400; 2023YFB2604401) and the research grant from the Natural Science Foundation of Beijing (No. JQ23036).

## Authors' contribution

Conceptualization, Y.B. and P.N.; investigation, P.N.; data curation, B.D.; writing—original draft preparation, B.D. and P.N.; writing—review and editing, P.N. and W.Y.; visualization, Q.L.; supervision, Y.B., J.L. and X.Z.; project administration, Q.H. All authors have read and agreed to the published version of the manuscript.

## Conflicts of interest

Jun Li holds the position of Associate Editor for *Smart Construction* and has not peer reviewed or made any editorial decisions for this paper. The authors declare no other conflicts of interest.

## References

[1] Yoshida I, Sekiya H, Mustafa S. Bayesian bridge weigh-in-motion and uncertainty estimation. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(1):04021001.

[2] Jiang W, Kim CW, Goi Y, Zhang F. Data normalization and anomaly detection in a steel plate-girder bridge using LSTM. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2022, 8(1):04021082.

[3] Torzoni M, Tezzele M, Mariani S, Manzoni A, Willcox KE. A digital twin framework for civil engineering structures. *Comput. Methods Appl. Mech. Eng.* 2024, 418:116584.

[4] Bouzas O, Conde B, Cabaleiro M, Riveiro B. A holistic methodology for the non-destructive experimental characterization and reliability-based structural assessment of historical steel bridges. *Eng. Struct.* 2022, 270:114867.

[5] Abedin M, De Caso y Basalo FJ, Kiani N, Mehrabi AB, Nanni A. Bridge load testing and damage evaluation using model updating method. *Eng. Struct.* 2022, 252:113648.

[6] de Vries R, Lantsoght EOL, Steenbergen RDJM, Hendriks MAN, Naaktgeboren M. Structural reliability updating on the basis of proof load testing and monitoring data. *Eng. Struct.* 2025, 330:119863.

[7] Yin T. A practical Bayesian framework for structural model updating and prediction. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2022, 8(1):04021073.

[8] Yang S, Yi T, Qu C, Zhang S, Li C. Adaptive sampling-based Bayesian model updating for bridges considering substructure approach. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2023, 9(3):04023024.

[9] Li P, Lu Z, Zhao Y. Bayesian updating of time-dependent structural reliability using the method of moment. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(4):04021066.

[10] Diaz M, Charbonnel PÉ, Chamoin L. A new Kalman filter approach for structural parameter tracking: application to the monitoring of damaging structures tested on shaking-tables. *Mech. Syst. Signal Process.* 2023, 182:109529.

[11] Caglio L, Stang H, Katsanos E. Bayesian model updating and nonlinear response estimation of structures subjected to unknown loads via FE-aided Kalman Filtering. *Mech. Syst. Signal Process.* 2025, 241:113505.

[12] Wang Z, Shafieezadeh A. Bayesian updating with adaptive, uncertainty-informed subset simulations: high-fidelity updating with multiple observations. *Reliab. Eng. Syst. Saf.* 2023, 230:108901.

[13] Guan Z, Yi T, Yang D, Zheng X, Li W, *et al.* Fatigue life uncertainty evaluation of steel box girder considering time-varying model of vehicle load and bridge. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2025, 11(3):04025043.

[14] Deriba AZ, Yang DY. Performance-based risk assessment for large-scale transportation networks using the transitional Markov chain Monte Carlo method. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2025, 11(1):04024090.

[15] Davis NT, Sanayei M. Foundation identification using dynamic strain and acceleration measurements. *Eng. Struct.* 2020, 208:109811.

[16] Saady AA, Ghasemi SH, Braley J, Balaguru P, Maher A. Reliability assessment of steel bridge elements in terms of the condition ratings under varying environmental conditions: a case study of New Jersey bridges. *Eng. Struct.* 2026, 353:122355.

[17] Liu Z, Zhang C, Ran X, Zhao W, Guo A. Long-term resilience assessment of bridge networks using data-driven seismic analysis method for deteriorating bridges. *Reliab. Eng. Syst. Saf.* 2026, 268:112048.

[18] Kim S, Ge B, Frangopol DM. Optimum target reliability determination for efficient service life management of bridge networks. *J. Bridge Eng.* 2020, 25(10):04020087.

[19] Su J, Zhang J, Caprani CC, Zhou J. A practical framework for determining target reliability indices for the assessment of existing structures based on risk-informed decision-making. *Struct. Saf.* 2025, 114:102583.

[20] O'Connell BJ, Champneys MD, Rogers TJ. A new perspective on Bayesian operational modal analysis. *Mech. Syst. Signal Process.* 2025, 236:112949.

[21] He X, Li D, Unjoh S. Joint state–parameter estimation for structures with seismic isolation and damping systems: theoretical observability and experimental verification. *Mech. Syst. Signal Process.* 2023, 203:110733.

[22] Taher SA, Li J, Fang H. Simultaneous seismic input and state estimation with optimal sensor placement for building structures using incomplete acceleration measurements. *Mech. Syst. Signal Process.* 2023, 188:110047.

[23] Yaoyama T, Itoi T, Iyama J. Probabilistic model updating of steel frame structures using strain and acceleration measurements: a multitask learning framework. *Struct. Saf.* 2024, 108:102442.

[24] Maes K, Lombaert G. Monitoring railway bridge KW51 before, during, and after retrofitting. *J. Bridge Eng.* 2021, 26(3):04721001.

[25] Buckley T, Pakrashi V, Ghosh B. A dynamic harmonic regression approach for bridge structural health monitoring. *Struct. Health Monit.* 2021, 20(6):3150–3181.

[26] Enshaeian A, Ghahremani B, Rizzo P. Structural health monitoring of a lenticular truss bridge: a comprehensive study. *Struct. Health Monit.* 2024, 23(6):3615–3639.

[27] Rema A, Krishna Swamy A. Use of Bayesian model averaging to estimate model uncertainty for predicting strain in a four-layered flexible pavement. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(1):04021002.

[28] Schranz B, Wagner PR, Czaderski C, Shahverdi M. Fibre optic measurements and model uncertainty quantification for Fe-SMA strengthened concrete structures. *Eng. Struct.* 2022, 256:114005.

[29] Tarpø M, Friis T, Georgakis C, Brincker R. Full-field strain estimation of subsystems within time-varying and nonlinear systems using modal expansion. *Mech. Syst. Signal Process.* 2021, 153:107505.

[30] Chen J, Chajes MJ, Shenton III HW, Aloupis C. Determining strain threshold values for bridge condition assessment using normal mixture models. *J. Bridge Eng.* 2022, 27(11):04022104.

[31] MacLeod E, Arjomandi K. Enhanced bridge weigh-in-motion system using hybrid strain–acceleration sensor data. *J. Bridge Eng.* 2022, 27(9):04022077.

[32] Zhang K, Qi T, Li D, Liu X, Xue X, *et al.* Health monitoring–based assessment of continuous girder bridge reinforcement. *J. Perform. Constr. Facil.* 2022, 36(2):04022001.

[33] Sun L, Li Y, Zhu W, Zhang W. Structural response reconstruction in physical coordinate from deficient measurements. *Eng. Struct.* 2020, 212:110484.

[34] Huang Z, Ma G, Ai Z, Wang J, Chang X, *et al.* Physics-constrained digital twin framework for deformation analysis and safety assessment of high earth-rock dams. *Reliab. Eng. Syst. Saf.* 2026, 273:112348.

[35] Jia S, Akiyama M, Han B, Xie H, Frangopol DM. Structural identification via the inference of the stochastic volatility model conditioned on the time-dependent bridge deflection. *Struct. Saf.* 2023, 100:102279.

[36] Cai Q, Chen Z, Zhu S. A comprehensive review of bridge influence lines: from identifications to applications in bridge health monitoring. *Struct. Health Monit.* 2025.

[37] Žnidarič A, Kalin J. Using bridge weigh-in-motion systems to monitor single-span bridge influence lines. *J. Civ. Struct. Health Monit.* 2020, 10(5):743–756.

[38] Abdelmaksoud AM, Hassan A, Oudah F. Fuzzy logic-visual inspection based evaluation charts for lateral capacity prediction of deteriorating concrete piled wharves. *Eng. Struct.* 2025, 325:119312.

[39] Reta AK, Taube C, Morgenthal G. Structural condition assessment of cracked RC beams based on digital crack measurements. *Eng. Struct.* 2025, 323:119211.

[40] Abtahi S, Liu Z, Li Y. Corrosion-related parameter estimation for RC structures using UKF-based Bayesian nonlinear finite element model updating with seismic data. *Mech. Syst. Signal Process.* 2023, 191:110169.

[41] Jia S, Akiyama M, Frangopol DM, Xu Z. Bayesian inference of the spatial distribution of steel corrosion in reinforced concrete structures using corrosion-induced crack width. *Struct. Saf.* 2024, 111:102518.

[42] Molaioni F, Andriotis CP, Rinaldi Z. Life-cycle fragility analysis of aging reinforced concrete bridges: a dynamic Bayesian network approach. *Struct. Saf.* 2026, 119:102654.

[43] Goulias DG, Cafiso S, Di Graziano A, Saremi SG, Currao V. Condition assessment of bridge decks through ground-penetrating radar in bridge management systems. *J. Perform. Constr. Facil.* 2020, 34(5):04020100.

[44] Hou Z, Liang Y, Yu L. A novel semi-convex function for simultaneous identification of moving vehicle loads and bridge damage. *Mech. Syst. Signal Process.* 2025, 223:111900.

[45] Karyofyllas G, Giagopoulos D. Optimal reduced-order model for vehicle-bridge interaction and bridge damage detection using model based SHM-ML techniques. *Mech. Syst. Signal Process.* 2026, 250:114188.

[46] Wang Y, Yang D, Au FTK. A probabilistic framework with recurrent mixture density network for reliability analysis of bridge expansion joint under thermal loading. *Reliab. Eng. Syst. Saf.* 2025, 264:111341.

[47] Agüero-Barrantes P, Christenson RE, Lobo-Aguilar S. Utilizing Kriging metamodeling to provide practical and effective bridge weigh-in-motion. *J. Bridge Eng.* 2023, 28(2):04022138.

[48] Paul D, Roy K. Application of bridge weigh-in-motion system in bridge health monitoring: a state-of-the-art review. *Struct. Health Monit.* 2023, 22(6):4194–4232.

[49] Diamantidis D, Tanner P, Holicky M, Madsen HO, Sykora M. On reliability assessment of existing structures. *Struct. Saf.* 2025, 113:102452.

[50] Gong F, Xia Y, Ling Z, Lozano F, He T. Bayesian deep learning based bridge condition assessment considering uncertainty quantification of missing data. *Eng. Struct.* 2026, 348:121753.

[51] Chen D, Kim CW, Yoshida E, Sim C. Enhancing structural damage evaluation of PC girder bridges through digital twinning Bayesian model updating. *Eng. Struct.* 2024, 321:118974.

[52] Aloisio A, Pasca DP, Kurent B, Tomasi R. Long-term continuous dynamic monitoring of an eight-story CLT building. *Mech. Syst. Signal Process.* 2025, 224:112094.

[53] Muin S, Chern C, Mosalam KM. Human–machine collaboration framework for bridge health monitoring. *J. Bridge Eng.* 2024, 29(7):04024041.
[54] Gao Y, Li H, Xiong G, Song H. AIoT-informed digital twin communication for bridge maintenance. *Autom. Constr.* 2023, 150:104835.
[55] Costin A, Adibfar A, Bridge J. Digital twin framework for bridge structural health monitoring utilizing existing technologies: new paradigm for enhanced management, operation, and maintenance. *Transp. Res. Rec.* 2024, 2678(6):1095–1106.
[56] Gholami H, Asgarian B, Asil Gharebaghi S. Practical approach for reliability-based inspection planning of jacket platforms using Bayesian networks. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2020, 6(3):04020029.
[57] Wu L, Ni Q, Ni X. A Bayesian value-of-information framework for inspection planning of marine concrete structures under epistemic uncertainty. *Reliab. Eng. Syst. Saf.* 2026, 275:112817.
[58] Songhori MJ, Fecarotti C, van Houtum GJ. Simulation supported Bayesian network approach for performance assessment of complex infrastructure systems. *Reliab. Eng. Syst. Saf.* 2025, 261:111045.
[59] Yokota H, Frangopol DM, Akiyama M, Tsompanakis Y. Bridge maintenance, monitoring, management, risk, life-cycle performance and optimization. *Struct. Infrastruct. Eng.* 2022, 18(10–11):1377–1378.
[60] Casas JR, Frangopol DM, Turmo J, Tsompanakis Y. Bridge safety assessment, maintenance, monitoring, and life-cycle performance. *Struct. Infrastruct. Eng.* 2024, 20(7–8):957–959.
[61] Liljefors F, Köhler J. Framework for rational decision-making in bridge maintenance. *Struct. Infrastruct. Eng.* 2026, 22(7):1266–1279.
[62] Wang F, Lee CCB, Gharaibeh NG. Network-level bridge deterioration prediction models that consider the effect of maintenance and rehabilitation. *J. Infrastruct. Syst.* 2022, 28(1):05021009.
[63] Zhou J, Liu Y, Yi J. Effect of uneven multilane truck loading of multigirder bridges on component reliability. *Struct. Concr.* 2020, 21(4):1644–1661.
[64] Hou R, Jeong S, Lynch JP, Ettouney MM, Law KH. Data-driven analytical load rating method of bridges using integrated bridge structural response and weigh-in-motion truck data. *Mech. Syst. Signal Process.* 2022, 163:108128.
[65] Zhou J, Zheng Y, Tan Z, Wei B, Zhou X, *et al.* Feasibility study of bridge load testing using ongoing traffic through rapid estimation of traffic load effects for network-level highway bridges assisted by site-specific ETC data. *J. Civ. Struct. Health Monit.* 2025, 15(6):1793–1813.
[66] Zheng X, Yi T, Zhong J, Yang D. Rapid evaluation of load-carrying capacity of long-span bridges using limited testing vehicles. *J. Bridge Eng.* 2022, 27(4):04022008.
[67] Zhou J, Zheng Q, Tang T, Wei B, Zhou X, *et al.* A novel load testing method for condition assessment of network-level highway bridges using moving artificial truck fleets in an open traffic environment. *Eng. Struct.* 2025, 340:120730.
[68] Dong C, Bas S, Debees M, Alver N, Catbas FN. Bridge load testing for identifying live load distribution, load rating, serviceability and dynamic response. *Front. Built Environ.* 2020, 6:46.
[69] Yang S, Yi T, Qu C, Zhang S, Li H. Bayesian finite-element model updating of bridges considering boundary condition uncertainties. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2024, 10(4):04024059.

[70] Zhu H, Li J, Tian W, Weng S, Peng Y, *et al.* An enhanced substructure-based response sensitivity method for finite element model updating of large-scale structures. *Mech. Syst. Signal Process.* 2021, 154:107359.

[71] Asghari Varzaneh MA, Mohammadi M. A novel fuzzy adaptive unscented Kalman filter for model updating using biased measurements. *Mech. Syst. Signal Process.* 2025, 230:112519.

[72] Bilgin N, Olivier A. Joint Bayesian estimation of process and measurement noise statistics in nonlinear Kalman filtering. *Mech. Syst. Signal Process.* 2025, 223:111836.

[73] Hong L, Li H, Fu J. Novel Kriging-based variance reduction sampling method for hybrid reliability analysis with small failure probability. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2022, 8(2):04022017.

[74] Zeng J, Todd MD, Hu Z. A recursive inference method based on invertible neural network for multi-level model updating using video monitoring data. *Mech. Syst. Signal Process.* 2023, 203:110736.

[75] Mo J, Yan W. StocIPNet: a novel probabilistic interpretable network with affine-embedded reparameterization layer for high-dimensional stochastic inverse problems. *Mech. Syst. Signal Process.* 2024, 220:111623.

[76] Zhang C, Wang Z, Shafieezadeh A. Error quantification and control for adaptive Kriging-based reliability updating with equality information. *Reliab. Eng. Syst. Saf.* 2021, 207:107323.

[77] Godin-Hebert E, Khorramian K, Oudah F. Recommendations for active-learning Kriging reliability analysis of bridge structures. *J. Bridge Eng.* 2025, 30(1):04024099.

[78] Kaewnuratchadasorn C, Wang J, Kim CW. Neural operator for structural simulation and bridge health monitoring. *Comput.-Aided Civ. Infrastruct. Eng.* 2024, 39(6):872–890.

[79] Liu T, Meidani H. Graph neural network surrogate for seismic reliability analysis of highway bridge systems. *J. Infrastruct. Syst.* 2024, 30(4):05024004.

[80] Barazzutti LF, Gomes HM, Drehmer LRC. Sensitivity-based model update for estimating generalized proportional damping parameters in a finite-element model. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2024, 10(2):04024011.

[81] Sánchez-Quesada JC, Moliner E, Romero A, Galvín P, Martínez-Rodrigo MD. Ballasted track interaction effects in railway bridges with simply-supported spans composed by adjacent twin single-track decks. *Eng. Struct.* 2021, 247:113062.

[82] Wang Z, Zhang W, Du F, Lu S, Geng Z. A novel active learning reliability analysis method based on ensemble of Kriging models and importance sampling for small failure probabilities. *Comput. Methods Appl. Mech. Eng.* 2025, 446:118322.

[83] Zhang X, Quek ST. Efficient subset simulation with active learning Kriging model for low failure probability prediction. *Probab. Eng. Mech.* 2022, 68:103256.

[84] Zhang X, Wang L, Sørensen JD. AKOIS: an adaptive Kriging oriented importance sampling method for structural system reliability analysis. *Struct. Saf.* 2020, 82:101876.

[85] Wang T, Bi S, Zhao Y, Dinh L, Mottershead J. Data-driven stochastic model updating and damage detection with deep generative model. *Mech. Syst. Signal Process.* 2025, 232:112743.

[86] Yu G, Xu W, Peng C, Chen C, Guo T. Exploration of multi-fidelity co-kriging and real-time hybrid simulation for time-dependent seismic performance evaluation. *Eng. Struct.* 2025, 345:121497.

[87] Das A, Debnath N. Gibbs sampling for damage detection using complex modal data from multiple setups. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(2):04021018.
[88] Teymouri D, Sedehi O, Katafygiotis LS, Papadimitriou C. Input-state-parameter-noise identification and virtual sensing in dynamical systems: a Bayesian expectation-maximization (BEM) perspective. *Mech. Syst. Signal Process.* 2023, 185:109758.
[89] Ma P, Zhang Y, Cai E, Luo M, Guo J, *et al.* A copula-based transitional Markov chain monte Carlo method for Bayesian model updating. *Reliab. Eng. Syst. Saf.* 2026, 265:111572.
[90] Zhou X, Yu W, Liu Y, Elchalakani M. Cross-domain structural damage identification based on updated FEM and WGAN-GP: Bridging the gap between FEM simulations and actual structural responses. *Eng. Struct.* 2026, 349:121888.
[91] Ahmadivala M, Mattrand C, Gayton N, Orcesi A, Yalamas T. AK-SYS-t: new time-dependent reliability method based on Kriging metamodeling. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(4):04021038.
[92] Dessena G, Ignatyev DI, Whidborne JF, Zanotti Fragonara L. A global-local meta-modelling technique for model updating. *Comput. Methods Appl. Mech. Eng.* 2024, 418:116511.
[93] Guo H, Dong Y, Gardoni P. Efficient subset simulation for rare-event integrating point-evolution kernel density and adaptive polynomial chaos Kriging. *Mech. Syst. Signal Process.* 2022, 169:108762.
[94] Coletti K, Davis RB, Schultz R. Structural response reconstruction using a system-equivalent singular vector basis. *Mech. Syst. Signal Process.* 2025, 227:112374.
[95] Tabandeh A, Jia G, Gardoni P. Langevin importance sampling for reliability analysis. *Reliab. Eng. Syst. Saf.* 2026, 266:111634.
[96] Nitzler J, Temür BZ, Koutsourelakis PS, Wall WA. Efficient Bayesian multi-fidelity inverse analysis for expensive and non-differentiable physics-based simulations in high stochastic dimensions. *Comput. Methods Appl. Mech. Eng.* 2026, 448(Part A):118442.
[97] Liu Y, Li L, Zhao S, Zhou C. A reliability analysis method based on adaptive Kriging and partial least squares. *Probab. Eng. Mech.* 2022, 70:103342.
[98] Prentzas I, Fragiadakis M. Quantified active learning Kriging model for structural reliability analysis. *Probab. Eng. Mech.* 2024, 78:103699.
[99] Ebrahimi M, Nobahar E, Mohammadi RK, Farsangi EN, Noori M, *et al.* The influence of model and measurement uncertainties on damage detection of experimental structures through recursive algorithms. *Reliab. Eng. Syst. Saf.* 2023, 239:109531.
[100] Kim J, Song J. Time-dependent reliability assessment and updating of post-tensioned concrete box girder bridges considering traffic environment and corrosion. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(4):04021062.
[101] Chiachío J, Jalón ML, Chiachío M, Kolios A. A Markov chains prognostics framework for complex degradation processes. *Reliab. Eng. Syst. Saf.* 2020, 195:106621.
[102] Politis SS, Zhang Z, Han Z, Hasenbein JJ, Arellano M. Stochastic analysis of network-level bridge maintenance needs using Latin hypercube sampling. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(1):04020049.

[103] Zwirglmaier K, Chan J, Papaioannou I, Song J, Straub D. Hybrid Bayesian networks for reliability assessment of infrastructure systems. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2024, 10(2):04024019.

[104] Davila-Frias A, Yodo N, Le T, Yadav OP. A deep neural network and Bayesian method based framework for all-terminal network reliability estimation considering degradation. *Reliab. Eng. Syst. Saf.* 2023, 229:108881.

[105] Li T, Wu S, Huang Y, Beck JL, Li H. Probabilistic graphical model using graph neural networks for Bayesian inversion of discrete structural component states. *Reliab. Eng. Syst. Saf.* 2026, 272:112478.

[106] Han X, Frangopol DM. Life-cycle connectivity-based maintenance strategy for bridge networks subjected to corrosion considering correlation of bridge resistances. *Struct. Infrastruct. Eng.* 2022, 18(12):1614–1637.

[107] López S, Barros B, Buitrago M, Morales-Nápoles O, Adam JM, *et al.* Reliability-based structural assessment of steel truss bridges subjected to failure scenarios. *Eng. Struct.* 2025, 341:120850.

[108] Moosapoor M, Bessason B, Hrafnkelsson B, Rupakhety R, Bjarnason JÖ, *et al.* Damage-informed seismic fragility of non-structural elements in South Iceland: a Bayesian hierarchical approach. *Eng. Struct.* 2025, 338:120504.

[109] Pang Y, Luo Q, Wei K, Deng K, Chen X. Seismic resilience assessment of reinforced concrete bridges constructed with tall piers in the deep-water environment. *Eng. Struct.* 2025, 325:119387.

[110] Franchini A, Sebastian W, D'Ayala D. Surrogate-based fragility analysis and probabilistic optimisation of cable-stayed bridges subject to seismic loads. *Eng. Struct.* 2022, 256:113949.

[111] Su P, Xiong F, Lu Y, Hu Q, Zhang B. Efficient Bayesian metamodeling for fine-grained and robust fragility analysis of buildings at a regional scale. *Struct. Saf.* 2023, 102:102324.

[112] Patsialis D, Taflanidis AA. Multi-fidelity Monte Carlo for seismic risk assessment applications. *Struct. Saf.* 2021, 93:102129.

[113] Azhari F, Loh KJ. Warning time–based framework for bridge scour monitoring. *J. Bridge Eng.* 2020, 25(7):04020040.

[114] Yousefpour N, Downie S, Walker S, Perkins N, Dikanski H. Machine learning solutions for bridge scour forecast based on monitoring data. *Transp. Res. Rec.* 2021, 2675(10):745–763.

[115] Song J, Kang WH, Lee YJ, Chun J. Structural system reliability: overview of theories and applications to optimization. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 2021, 7(2):03121001.

[116] Thaler D, Dhulipala SLN, Bamer F, Markert B, Shields MD. Reliability analysis of complex systems using subset simulations with Hamiltonian Neural Networks. *Struct. Saf.* 2024, 109:102475.

[117] Ierimonti L, Cavalagli N, Venanzi I, García-Macías E, Ubertini F. A transfer Bayesian learning methodology for structural health monitoring of monumental structures. *Eng. Struct.* 2021, 247:113089.

[118] Cao J, Xiong H, Ghahari SF, Taciroglu E. A validated lateral response model for mass timber frames with knee-braces. *Eng. Struct.* 2021, 239:112278.

[119] Tulebekova S, Malo KA, Rønnquist A, Nåvik P. Modeling stiffness of connections and non-structural elements for dynamic response of taller glulam timber frame buildings. *Eng. Struct.* 2022, 261:114209.

[120] Guo H, Dong Y, Bastidas-Arteaga E. Mixed Bayesian Network for reliability assessment of RC structures subjected to environmental actions. *Struct. Saf.* 2024, 106:102392.
[121] Xiong L, Zhou W, Jiang L, Wu L, Zhong B. Seismic risk transfer control strategies and reliability assessment of high-speed railway simply supported beam bridges. *Reliab. Eng. Syst. Saf.* 2026, 269:112094.