

A Whole-Process Cascading Bayesian Network for Safety Risk Assessment of Post- Earthquake Collapsed-Building Rescue Operations

Tianyu Hu

Jilin Agricultural University

Defeng Xu

Jilin Agricultural University

Yuan Li

Xi'an University of Science and Technology

Feifei Sun

ffsun@tongji.edu.cn

Tongji University

Songhang He

Tongji University

Zhipeng Shao

Institute of Engineering Mechanics of China Earthquake Administration

Research Article

Keywords: post-earthquake collapsed buildings, rescue operations, stage-indicator activation matrix, cascading Bayesian network, normalized safety risk index, dynamic risk assessment

Posted Date: September 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10686189/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Whole-Process Cascading Bayesian Network for Safety Risk Assessment of Post-Earthquake Collapsed-Building Rescue Operations

Tianyu Hu^a, Defeng Xu^{a, b*}, Yuan Li^c, Feifei Sun^{b*}, Songhang He^b, Zhipeng Shao^d

^aJilin Agricultural University, College of Engineering Technology, Changchun, 130118, China

^bTongji University, College of Civil Engineering, Shanghai 200092, China

^cCollege of Safety Science and Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

^dInstitute of Engineering Mechanics of China Earthquake Administration, Harbin 150080, China

Abstract: Post-earthquake rescue in collapsed buildings is a time-varying safety problem characterized by stage-specific hazards, flexible operation sequences, unequal stage durations, and state inheritance between successive tasks. Existing methods mainly evaluate building condition, entry safety, or isolated rescue activities, making whole-operation risk difficult to compare and update. This study develops a whole-process cascading Bayesian network (WPC-BN) for node-level dynamic assessment. A unified evidence space of four dimensions, 19 criteria, and 85 indicators is linked to a stage-indicator activation matrix that activates only evidence relevant to shoring, breaching, debris removal, and casualty transport. Stage-specific static BNs perform within-node inference, while eight persistent states are propagated across nodes by continuous-time Markov chains using actual elapsed time. The normalized safety risk index (NSRI) is a model-based posterior score rather than an accident probability and supports pre-task prediction, in-process updating, and terminal review under time-locked evidence. After two Delphi rounds with the same 20 experts, all 85 indicators were retained with $CVR \geq 0.50$, and an independent 33-expert review supported criterion-level validity. In the retrospective Mandalay case, breaching produced the highest pre-task and in-process NSRI values (2.55% and 2.86%); breaching-induced vibration (X77), structural stability (X29), and rescue-passagage stability (X31) were consistently prioritized, matching recorded abnormal sounds and tilting risk. WPC-BN provides an auditable, duration-aware basis for identifying high-risk operational windows and updating safety priorities across heterogeneous rescue stages, subject to scenario-specific parameter localization.

Keywords: post-earthquake collapsed buildings; rescue operations; stage-indicator activation matrix; cascading Bayesian network; normalized safety risk index; dynamic risk assessment

1. Introduction

After destructive earthquakes, collapsed and severely damaged buildings remain metastable: residual load-carrying capacity is uncertain, load paths are altered, boundary conditions shift, and aftershocks may further reduce collapse resistance (Raghunandan et al., 2015; Xu et al., 2024). Rescue operations themselves introduce additional disturbance. Entry, shoring, breaching, lifting, and unloading redistribute loads and alter structural constraints, while confined spaces, toxic or flammable gases, damaged utilities, and heavy machinery create additional hazards for rescuers and equipment (OSHA, n.d.; INSARAG, 2020, 2026). Rescue safety is therefore not a one-time judgment of building condition but a state that evolves throughout the operation.

Current work addresses this evolving state only indirectly. Post-earthquake assessment offers established procedures for rapid evaluation, detailed assessment, and safety tagging, with reassessment as aftershocks occur and new information arrives (Applied Technology Council, 1989, 2005; Anagnostopoulos and Moretti, 2008; Marshall et al., 2012; FEMA, 2019).

43 Aftershock-driven early-warning models have also begun to track the residual structure over time
44 (Xu et al., 2024, 2026a, 2026b, 2026c), while structural monitoring, resilience assessment,
45 unmanned aerial vehicles, and machine vision have sharpened damage identification and decision
46 support (Bruneau et al., 2003; German et al., 2013; Cimellaro et al., 2018; Roohi and Hernandez,
47 2020; Xiong et al., 2020; Qing et al., 2022). What all of this work assesses, however, is the
48 building: its overall damage, its usability, its regional damage state. None quantifies the risk of a
49 specific rescue operation within a given time window, and none tracks how that risk evolves as the
50 work proceeds.

51 Operational standards provide the language of field practice rather than a quantitative risk
52 metric. INSARAG, GB/T 29428, DB/T 63, and NFPA set systematic requirements for site
53 assessment, shoring, breaching, lifting, medical response, and safety management (INSARAG,
54 2020, 2026; NFPA, 2020, 2022, 2024), while OSHA (n.d.) identifies the dominant hazards:
55 structural instability, oxygen deficiency, hazardous atmospheres, energized utilities, machinery
56 blind spots, and secondary collapse. Although these procedures, checklists, and task-specific
57 controls constrain safe practice, they do not integrate personnel, equipment, environmental,
58 managerial, and stage-specific evidence into a common metric that is comparable across
59 successive stages.

60 Quantitative risk models face the opposite problem. They typically depend on application-
61 specific variables, network structures, and parameterization schemes, which limits transferability
62 (Kabir and Papadopoulos, 2019; George and Renjith, 2021; Kaikkonen et al., 2021). In collapsed-
63 building rescue, structural, environmental, personnel, equipment, managerial, and task-specific
64 risks are often modeled separately. This creates a trade-off: stage-specific criteria weaken cross-
65 stage comparability, whereas retaining every factor at every stage introduces irrelevant variables
66 and imposes a substantial field data-collection burden.

67 Three features of rescue make this trade-off hard to escape. The first is stage heterogeneity.
68 Reconnaissance, shoring, breaching, debris removal, and casualty transport act through distinct
69 mechanisms: shoring adds restraint but may destabilize locally during installation; breaching
70 weakens components and introduces vibration, dust, and sparks; debris removal alters loading and
71 constraints and may add lifting impacts; and casualty transport requires decompression
72 management, movement through confined passages, and medical coordination (INSARAG, 2020,
73 2026; OSHA, n.d.; Abu-Zidan et al., 2023; Usuda et al., 2023). Whole-process assessment must
74 therefore represent rescue as a sequence of actual operational nodes that can be planned, observed,
75 and updated.

76 The second is sequence flexibility and state inheritance. Depending on collapse
77 configuration, burial depth, access, structural anomalies, and aftershock activity, shoring,
78 breaching, and debris removal may be reordered, skipped, or repeated, with markedly different
79 durations (INSARAG, 2020, 2026). The structural, environmental, personnel, and equipment
80 states left by one operation become the initial conditions of the next, so risk depends not only on
81 the current stage but on its duration and on what it inherits. Without explicit cross-node state
82 propagation, neither risk accumulation nor risk reduction after a safety intervention can be
83 represented.

84 The third feature concerns temporal representation. Bayesian networks (BNs) represent
85 multivariate dependence, uncertainty, and evidence updating and are widely used in system-safety

analysis (Khakzad et al., 2011; Weber et al., 2012; Kabir and Papadopoulos, 2019), whereas dynamic Bayesian networks (DBNs) extend this representation to time-dependent state evolution (Murphy, 2002; Khakzad, 2015; Chang et al., 2024; Chen et al., 2025). In principle, a DBN can represent the rescue process, but a direct monolithic implementation becomes cumbersome when the active evidence set changes by operation, stage durations are unequal, and operational stages may be skipped, repeated, or reordered. A fixed slice structure would either carry stage-irrelevant variables or require stage-conditioned switching, while cross-slice transitions must remain identifiable from sparse operational data. The methodological need is therefore not to replace DBNs, but to formulate a stage-conditioned, duration-aware implementation that preserves BN inference within each operation and propagates only physically persistent states between operations. Section 5.1 positions WPC-BN relative to static BNs and conventional DBNs.

Together, these three features define the methodological gap. Dynamic safety assessment of collapsed-building rescue still lacks a framework that integrates heterogeneous evidence within a single structure, activates task-specific factors only when relevant, accommodates flexible sequences and unequal durations, and propagates states across successive operational nodes. The core challenge is not hazard identification but the joint treatment of stage heterogeneity, state inheritance, and time-ordered evidence updating. Without such integration, cross-stage comparison, identification of high-risk operational windows, and reassessment after intervention remain difficult.

This study addresses this gap with a whole-process cascading Bayesian network (WPC-BN). A fixed candidate evidence space comprising four dimensions, 19 criteria, and 85 indicators is organized into common, regional-background, and stage-specific layers. A stage-indicator activation matrix selects the evidence relevant to each node, preserving consistent definitions across stages without requiring every indicator to be recollected. Within a node, a stage-specific static BN estimates the normalized safety risk index (NSRI); across nodes, persistent states are propagated according to actual elapsed time. Time-locked evidence rules separate pre-task prediction, in-process updating, and terminal posterior assessment, enabling rolling rather than one-off assessment.

The WPC-BN framework introduces four methodological innovations.

1. WPC-BN assessment unit. The framework shifts assessment from overall building condition to the actual rescue-operation node, while a common evidence space keeps factor definitions comparable across the whole process.
2. Stage-conditioned evidence activation. WPC-BN activates task-specific indicators only for the operation in which they are physically relevant and reuses unchanged evidence, reducing field data demand without sacrificing cross-stage comparability.
3. Duration-aware cascading. Persistent states are propagated by stage type and actual elapsed time, allowing reordered, skipped, repeated, and unequal-duration operations to use the same updating interface under time-locked evidence.
4. Auditable multi-level evaluation. Content validity, numerical and software checks, scenario ranking, and the retrospective Mandalay application evaluate indicator coverage, inference consistency, risk ordering, and process interpretation without treating *NSRI* as an empirical accident probability.

The remainder of this paper is organized as follows. Section 2 defines the operational nodes, unified evidence space, WPC-BN inference and cascading mechanism, parameterization, and rolling outputs. Section 3 reports indicator screening, parameter calibration, and internal numerical and robustness checks. Section 4 presents the retrospective Mandalay application. Section 5 discusses methodological positioning, engineering implications, transferability, applicability boundaries, and future research. Section 6 summarizes the main conclusions.

2. Methodology

Five linked questions structure the methodology (Fig. 1): how operational nodes are defined (Section 2.1), how evidence is activated by stage (Section 2.2), how risk is inferred within a node and propagated across nodes (Section 2.3), how parameters are calibrated (Section 2.4), and how outputs are converted into rolling decisions (Section 2.5). Definitions, proofs, and parameter calculations are provided in Supplementary Material S1 and Appendices A–D.

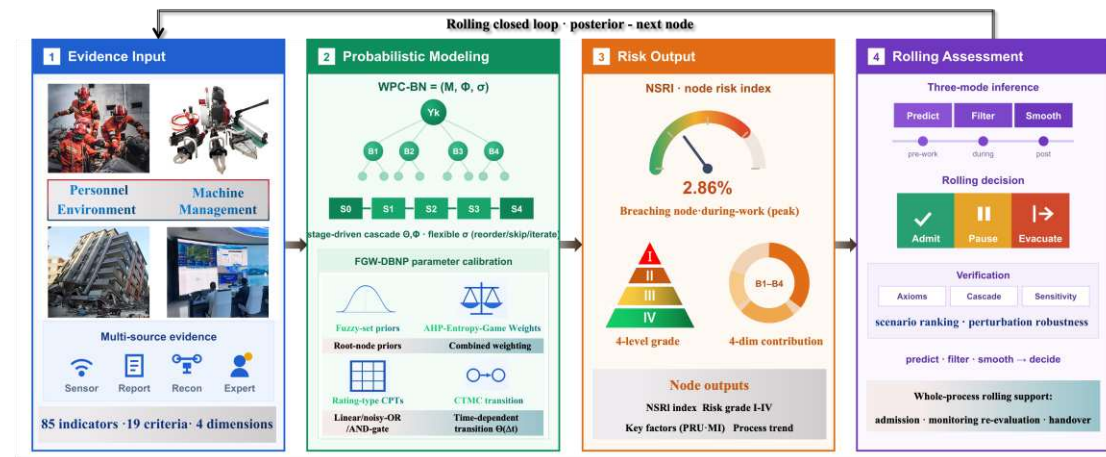


Fig. 1. Methodological framework: evidence input, probabilistic modeling, risk output, and rolling assessment

2.1 Assessment unit: operational stages, flexible sequences, and time points

2.1.1 Stage definition

Stages are defined by decision-relevant operational transitions, following the INSARAG 2020 Guidelines and cross-checked against the 2026 Guidelines, GB/T 29428.2-2014, DB/T 63-2016, NFPA 2500-2022, OSHA structural-collapse requirements, and representative cases. NFPA 1521-2020 and NFPA 1550-2024 inform safety-officer duties, personnel health, and operational controls rather than stage definition.

Rescue is accordingly represented by five stage types (Fig. 2; Table 1): S0 reconnaissance and initial assessment, S1 shoring and bracing, S2 slab breaching, S3 debris removal, and S4 casualty transport and evacuation. S2 takes slab breaching as its template; breaching of other components requires that indicator definitions and parameters be rechecked.

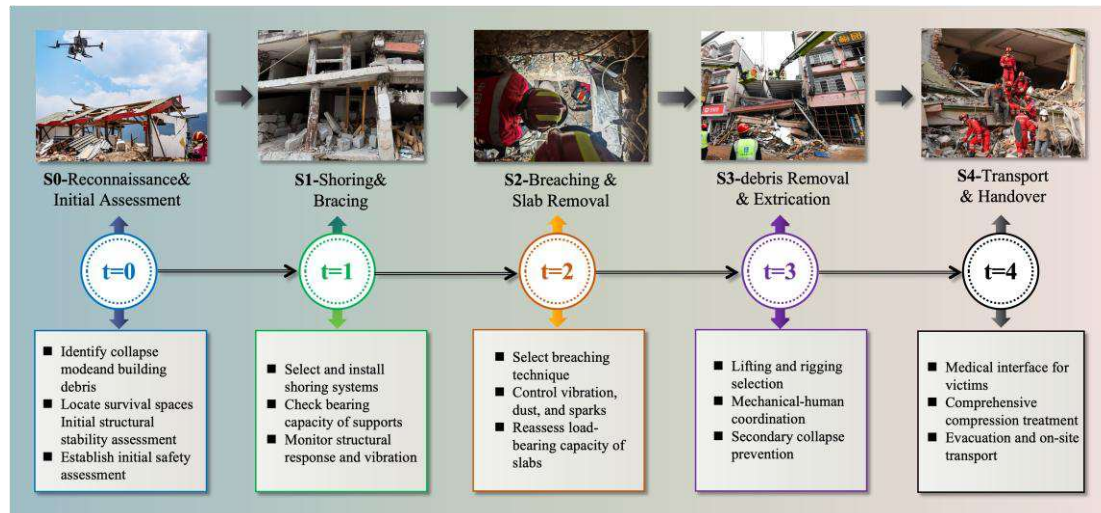


Fig. 2. Key rescue stages and assessment time points

Table 1. Key rescue stages and attributes

Stage	Main activities	Disturbance and dominant mechanism	Activated technical indicators (C19)	Typical observable evidence
S0 Reconnaissance and initial assessment	Identify debris configuration and survivable spaces; human–canine–device search; assess structural stability; develop rescue plan	Minor disturbance from reconnaissance and search	— (no stage-specific indicators)	Collapse pattern; entrapment location/depth; aftershock records; gas concentration; life-detection confirmation
S1 Shoring and bracing	Select and install shoring; jacking; load-path reconfiguration	Stabilizing: increases restraint and redistributes loads; installation may induce local instability	X73 shoring selection and load check; X74 jack capacity and synchronization; X75 shoring monitoring and reinforcement	Shoring displacement/strain; jacking synchronization error; abnormal sounds; new cracks
S2 Slab breaching	Chiseling, cutting, or drilling; create vertical/horizontal access	Disturbance-amplifying: weakens components; introduces vibration, dust, and sparks	X76 breaching-method suitability; X77 vibration and structural-disturbance control; X78 dust/spark control; X79 opening location and dimensions	Vibration; debris displacement; combustible-gas concentration; opening quality
S3 Debris removal	Lifting/rigging; mechanical–manual clearance; surcharge unloading	Disturbance-amplifying: unloads and releases constraints; may introduce lifting impacts	X80 lifting plan and lifting-point selection; X81 machine–human coordination and blind-spot control; X82 removal sequence and unloading control	Lifting load; debris displacement rate; coupled-member response; signs of secondary collapse
S4 Casualty transport and evacuation	Medical intervention and decompression; casualty transport; equipment withdrawal and site handover	Minor disturbance from transport and equipment withdrawal	X83 medical coordination and crush-syndrome management; X84 casualty monitoring and communication; X85 transport, evacuation, and handover coordination	Vital signs; passage stability; evacuation-route condition

2.1.2 Flexible stage sequencing

S0–S4 are stage templates rather than a fixed sequence: S0 initiates and S4 terminates the process, while the order and repetition of S1–S3 depend on collapse configuration, entrapment, access, structural condition, and the selected rescue plan.

Let the intermediate stage sequence be

$$\sigma = (s_1, s_2, \dots, s_K), s_K \in \{S1, S2, S3\} \quad (1)$$

161 The complete rescue sequence is then

$$\tilde{\sigma} = (S0, \sigma, S4) \quad (2)$$

162 A typical sequence is shoring–breaching–removal. If instability emerges during breaching,
163 the operation may return to shoring before breaching resumes.

164 Because stage models are keyed to operation type and transition parameters are keyed to
165 variable group $\times$ stage type, reordering, skipping, and repetition are computationally admissible,
166 though not physically equivalent. Repeated breaching, pronounced cumulative damage, or parallel
167 operations on a shared structural system require recalibrated transitions or a history-dependent
168 model; long nodes may be subdivided.

169 2.2 Unified evidence space and stage-specific activation

170 2.2.1 Risk-factor identification and indicator system

171 The indicator pool was derived from GB/T 29428.1-2012 and GB/T 29428.2-2014, YJ/T 8, the
172 INSARAG Guidelines, NFPA 2500 and NFPA 1550, and OSHA structural-collapse requirements
173 (full bibliographic entries for all standards cited are given in the reference list) and supplemented
174 by literature on rescue resource allocation, coordination, duration, robotics, and emergency-
175 capability assessment (Fiedrich et al., 2000; Comfort et al., 2004; Murphy, 2004; Comfort, 2007;
176 Tadokoro, 2009; Statheropoulos et al., 2015; Chen, 2018; An, 2023), representative cases, and
177 expert interviews. It was then cross-checked against the updated INSARAG Guidelines and GB/T
178 46578-2025; DB41/T 2600-2024 was used only for communication and information indicators.

179 Candidate factors cover structural and site conditions, entrapment and accessibility,
180 aftershocks and secondary hazards, personnel, equipment, monitoring, organization and
181 management, and key rescue operations. They were screened for safety relevance, measurability,
182 data availability, and relative independence, and then refined through two Delphi rounds. Mission-
183 level factors with indirect or difficult-to-quantify pathways, such as international coordination and
184 public opinion, were excluded from the quantitative model and are instead handled as a command-
185 level qualitative checklist outside *NSRI*; the reasoning and the boundary this sets are given in
186 Section 5.3.

187 The final set comprises 85 indicators under 19 criteria and four dimensions (Table 2),
188 organized by fault-tree logic. It is a content-validated evidence domain from which node-specific
189 evidence is drawn by background assignment, evidence reuse, and stage activation; special tasks
190 may add extension nodes without altering the core structure. Definitions and state-classification
191 rules appear in Supplementary Tables S1-1 and S1-2.

192 Table 2. Indicator system for whole-process rescue risk assessment

Dimension	Criteria and indicator ranges	No.	Representative indicators
Personnel B1	C1 qualifications and experience (X1–X5); C2 language and communication (X6–X8); C3 training and drills (X9–X10); C4 physical and psychological fitness (X11–X14); C5 operational safety and emergency response (X15–X18)	18	Qualification coverage; structural-expert experience; drill frequency/quality; endurance and recovery; risk recognition and emergency response
Equipment B2	C6 rescue tools and equipment (X19–X22); C7 personal protective equipment (X23–X25); C8 real-time monitoring equipment (X26–X28)	10	Equipment sufficiency; task suitability; condition and reliability; real-time debris-displacement monitoring

Dimension	Criteria and indicator ranges	No.	Representative indicators
Environment B3	C9 structural stability (X29–X31); C10 entrapment and accessibility (X32–X38); C11 gas/chemical hazards (X39–X44); C12 special rescue environments (X45–X50); C13 secondary hazards (X51–X54); C18 regional/site characteristics (X70–X72)	29	Structural stability; entrapment depth; combustible-gas concentration; high temperature; landslide risk; structural type and construction standard
Management and operational control B4	C14 safety risk control (X55–X58); C15 communication and information (X59–X63); C16 resources and logistics (X64–X66); C17 planning and execution (X67–X69); C19 key rescue-operation control (X73–X85)	28	Incident command; shift-handover information; supplies for extreme conditions; breaching-vibration control; lifting plan and lifting-point selection

2.2.2 Regional-background and stage-specific indicators

Definitions, quantification rules, and data sources are provided in Table A1 (Appendix A). C18 (X70–X72) captures construction context, site conditions, and aftershock exposure; it is assigned early and updated using seismic-network information. C19 (X73–X85) activates technical controls for shoring, breaching, removal, and transport in S1–S4. This separation distinguishes contextual adaptation from node-specific control and avoids the need to re-enter all technical indicators.

To avoid double counting, regional-background nodes represent prior context only and do not replace field evidence of structural condition, such as X29 or X31. Stage-specific indicators are activated only at their corresponding node and map directly to pre-task checks, in-process monitoring, and handover items.

2.2.3 Stage-indicator activation matrix

The activation matrix fixes the active evidence set at each node (Table 3): X1–X69 are common indicators, updated only when state or confidence changes; X70–X72 are regional-background indicators; and X73–X85 activate at their own stage, with repeated stages reactivating the same technical group. Overlapping operations require prior assessment of shared structures, resources, and evidence dependencies.

The matrix is operationalized as a node checklist recording state, evidence source, threshold, update time, and responsible role. Objective evidence is updated immediately, whereas subjective evidence is recorded together with its source and confidence. Eight variables—X11, X12, X21, X29, X31, and X39–X41—are persistent states propagated between adjacent nodes.

Table 3. Stage-indicator activation matrix

Indicator group	S0	S1	S2	S3	S4
Common personnel X1–X18	Always active	Always active	Always active	Always active	Always active
Common equipment X19–X28	Always active	Always active	Always active	Always active	Always active
Common environment X29–X54	Always active	Always active	Always active	Always active	Always active
Common management X55–X69	Always active	Always active	Always active	Always active	Always active
Regional/site X70–X72	Background	Background	Background	Background	Background
Shoring group X73–X75	Dormant	Active	Dormant	Dormant	Dormant
Breaching group X76–X79	Dormant	Dormant	Active	Dormant	Dormant
Debris-removal group X80–X82	Dormant	Dormant	Dormant	Active	Dormant
Casualty-transport group X83–X85	Dormant	Dormant	Dormant	Dormant	Active

In Table 3, “Always active” denotes indicators retained in the active evidence set at every node—distinct from the eight persistent states propagated by CTMC transitions (Section 2.3.3); Supplementary Table S1-3 gives the full field-use version. Three evidence layers result: common

indicators preserve cross-node consistency, regional-background indicators carry scenario context, and stage-specific indicators serve the current decision.

2.3 WPC-BN: within-node inference and cross-node state propagation

2.3.1 Node model, top event, and risk output

At node k , the input e_k collects the evidence activated in the current information state. The binary top node Y_k is a latent unsafe operating state, defined by physical criteria rather than by the action it triggers: $Y_k=1$ denotes a state of the operating face in which at least one of the following is present within the assessment window: loss of residual structural or passage stability, atmospheric parameters outside the prescribed control range, or personnel exposure beyond the prescribed limit. Y_k is latent because none of these conditions is directly observed at the window endpoint; it is inferred from the activated evidence. The mapping from the posterior of Y_k to pausing, evacuating, or strengthening controls is a separate decision rule (Section 2.5.3) and is not part of the definition of the top event, so the model is not calibrated against the actions it recommends.

Node risk is defined as

$$\hat{R}_k = P(Y_k = 1 | e_k) \quad (3)$$

The resulting node risk $\hat{R}_k$ is the normalized safety risk index (NSRI), which lies in $[0,1]$. Because priors and some CPTs come from expert judgment and relative weights rather than calibrated accident frequencies, *NSRI* is a model-based posterior score, not an accident probability or expected loss.

Definition 1 (WPC-BN). The whole-process cascading Bayesian network is defined as

$$WPC-BN = \left(M, \Phi, \tilde{\sigma} \right), M = \{M_s\} \quad (4)$$

where $M = \{M_s\}$ is the set of stage-specific static BNs, Φ is the stage-propagation operator, and $\tilde{\sigma}$ is the realized field sequence.

All stage models share one hierarchy: Y_k at the top, dimensions B1–B4, criteria C1–C19, and an indicator layer set by the activation matrix; topology follows fault-tree-to-BN mapping rules (Pearl, 1988; Bobbio et al., 2001). The S2 topology appears in Fig. 3, and the full 85-indicator configuration in Supplementary Fig. S6-1.

The within-node joint distribution is

$$P(Z_k) = \prod_{i=1}^{n_k} P(Z_k^i | Pa_k(Z_k^i)) \quad (5)$$

and *NSRI* is

$$\hat{R}_k = P(Y_k = 1 | e_k) = \sum_{\omega \in \Omega_k} P(Y_k = 1 | \omega) P(\omega | e_k) \quad (6)$$

where Z_k is the active variable set at node k , $Pa_k(Z_k^i)$ is the parent set of Z_k^i , and Ω_k is the state space of the top-node parents. The node risk in Eq. (6) is internally normalized, not an outcome-calibrated real-world frequency.

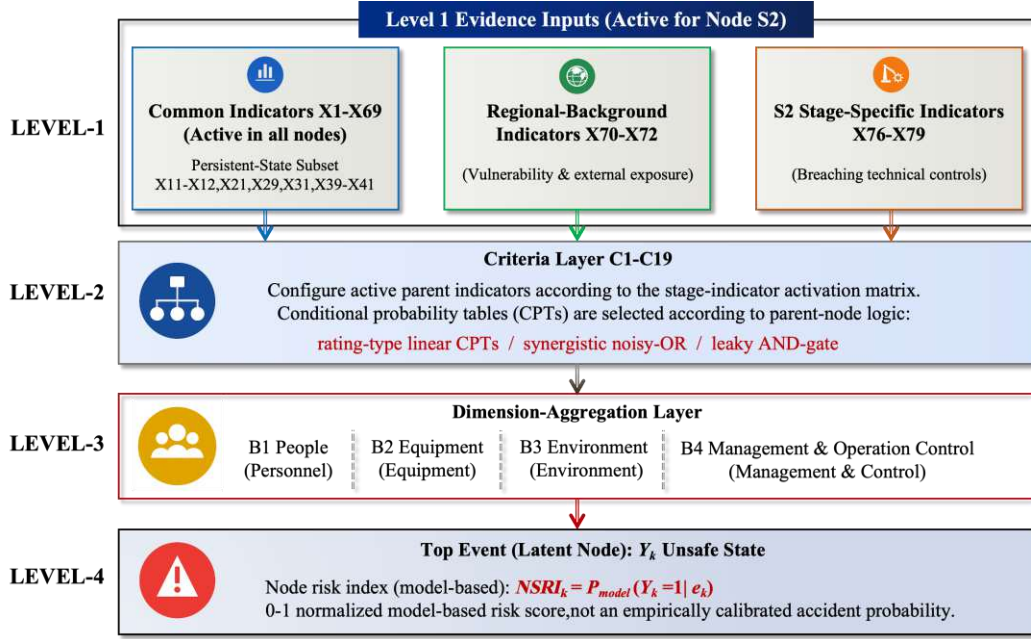


Fig. 3. Hierarchical Bayesian network for the S2 breaching node

2.3.2 Stage-conditioned non-homogeneous DBN representation

To accommodate stage changes and unequal durations, WPC-BN is represented as a stage-conditioned non-homogeneous dynamic Bayesian network (DBN): the active variable set changes with stage type, and transitions of persistent states depend on both stage and elapsed time. The implemented marginal propagation is an approximation to this full structure.

Let k index actual operational nodes, $S_k = s_k$ denote the stage type, and Δt_k denote the planned duration or actual elapsed increment. Thus, k indexes stage executions rather than equally spaced physical time. For persistent state C_i , the stage-dependent transition probability is

$$P(C_{i,k} = b | C_{i,k-1} = a, S_k = s_k, \Delta t_k) = [\Theta_i^{(s_k)}(\Delta t_k)]_{ab} \quad (7)$$

where $\Theta_i^{(s_k)}(\Delta t_k)$ is the row-stochastic transition matrix for stage s_k over duration Δt_k , and a and b are the pre- and post-transition states. Stage sequence and node duration are exogenous field inputs.

The fixed persistent-state set is

$$C = \{X_{11}, X_{12}, X_{21}, X_{29}, X_{31}, X_{39}, X_{40}, X_{41}\} \quad (8)$$

The effective variable set at node k , equivalent to the active variable set Z_k in Eq. (5), is

$$v_k = v_{com} \cup v_{bg} \cup A(S_k) \cup v_{mid} \cup \{Y_k\} \quad (9)$$

where V_{com} , V_{bg} , and V_{mid} denote common indicators, regional-background indicators, and criterion/dimension intermediate variables, respectively; $A(s_k)$ is the stage-specific indicator set; and Y_k is the top node of the stage network. Inactive stage-specific indicators stay dormant and are excluded from the within-node factorization.

Conditioned on stage sequence $s_{0:K}$ and node durations $\Delta t_{1:K}$, the joint distribution of the stage-conditioned non-homogeneous DBN is

$$P(v_{0:K} | s_{0:K}, \Delta t_{1:K}) = P(V_0 | s_0) \prod_{k=1}^K \left[\prod_{C_i \in C} P(C_{i,k} | C_{i,k-1}, s_k, \Delta t_k) \times \prod_{X_{j,k} \in v_{kC}} P(X_{j,k} | Pa_{j,k}^{intra}, s_k) \right] \quad (10)$$

where $Pa_{j,k}^{intra}$ is the within-node parent set. Equation (10) combines inter-slice transitions of persistent states with within-slice conditional relations for all other variables; cross-node evolution applies first-order transitions between same-name persistent variables, followed by evidence updating.

Implementation propagates only the marginal posteriors of persistent states and rebuilds the next-node prior under a marginal-independence approximation, making WPC-BN a modular forward recursion of Eq. (10) rather than a full non-homogeneous DBN. Figs. 4 and 5 show cross-node prediction–update–propagation and the cascading execution workflow.

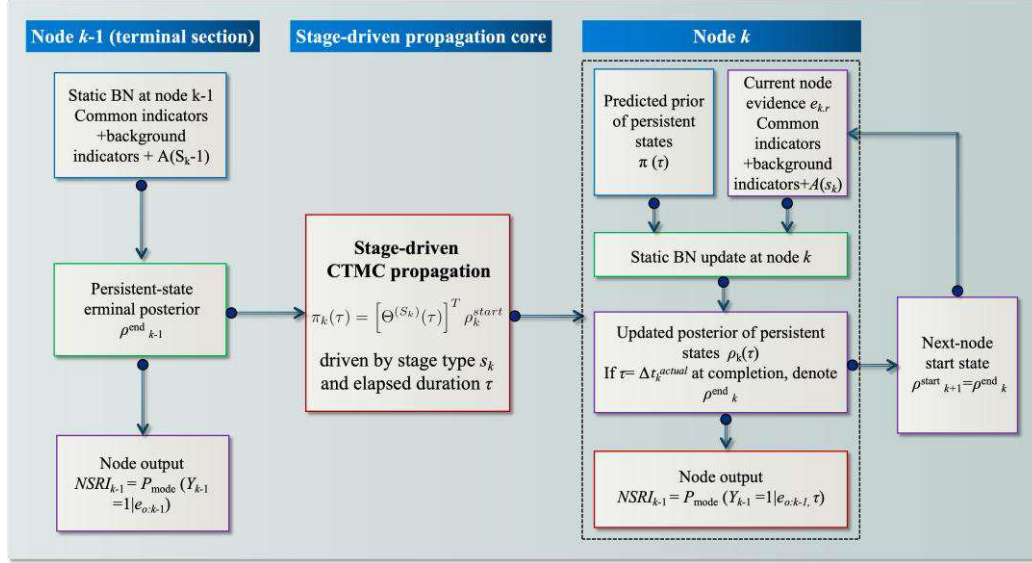


Fig. 4. Stage-conditioned non-homogeneous DBN and cross-node state propagation

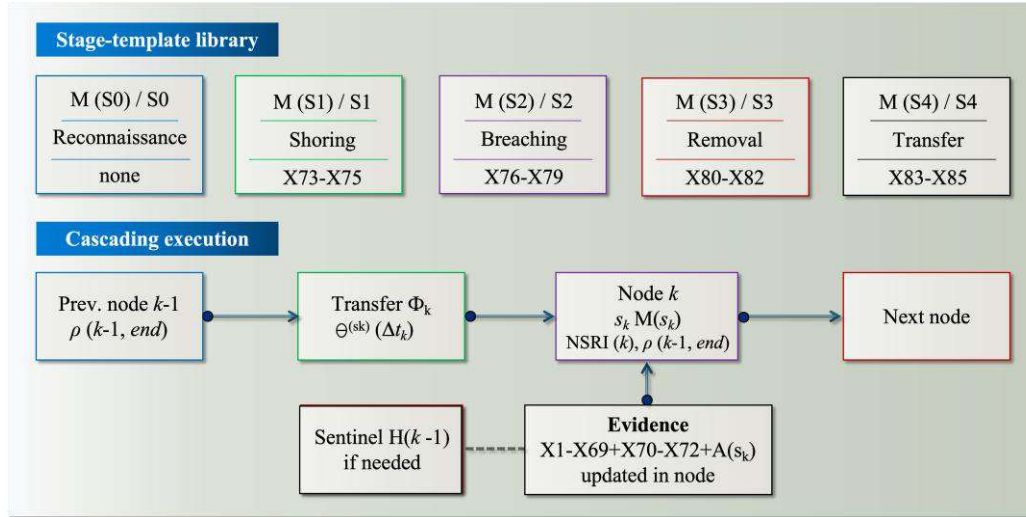


Fig. 5. WPC-BN stage-template invocation and cascading execution across nodes

2.3.3 Duration-dependent transition matrices

For a binary persistent state X_i with state order [normal, failed], the row-stochastic transition matrix for stage s over duration Δt and its constraints are given by Eq. (11).

$$\Theta_i^{(s)}(\Delta t) = \begin{bmatrix} 1 - d_i^{(s)}(\Delta t) & d_i^{(s)}(\Delta t) \\ r_i^{(s)}(\Delta t) & 1 - r_i^{(s)}(\Delta t) \end{bmatrix} \quad (11)$$

$$\theta_{uv}^{(s)} \geq 0, \sum_v \theta_{uv}^{(s)} = 1$$

State evolution is modeled by a two-state continuous-time Markov chain (CTMC) (Nodelman et al., 2002, 2003), with the generator Q and transition matrix given by Eqs. (12) and (13).

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} \quad (12)$$

Here, λ and μ are the deterioration and recovery rates, respectively (h^{-1}).

$$\Theta(\Delta t) = \exp(Q\Delta t) \quad (13)$$

Equations (12) and (13) yield the deterioration and recovery probabilities over any time window:

$$d(\Delta t) = \frac{\lambda}{\lambda + \mu} \left[1 - e^{-(\lambda + \mu)\Delta t} \right], \quad r(\Delta t) = \frac{\mu}{\lambda + \mu} \left[1 - e^{-(\lambda + \mu)\Delta t} \right] \quad (14)$$

Let d_0 and r_0 denote deterioration and recovery probabilities over a reference window T_0 , with $0 \leq d_0 + r_0 < 1$. If $d_0 + r_0 = 0$, set $\lambda = \mu = 0$; otherwise, $v = -\ln(1 - d_0 - r_0)/T_0$, $\lambda = v d_0 / (d_0 + r_0)$, and $\mu = v r_0 / (d_0 + r_0)$. Each node transition matrix is then generated from its actual duration.

No persistent-state transition is applied in S0, so $\Theta^{(S0)}(\Delta t) = I$. Table 9 and Supplementary Table S4-2 therefore report S1–S4 only. Rates are calibrated by variable group $\times$ stage type, independent of a stage's position in the sequence. S1 gives the strongest recovery pathway, and S2 and S3 deteriorate faster than S4. In the present calibration μ exceeds λ for every variable group in every stage, including S2 and S3. This does not assert that breaching or debris removal improves structural condition. λ and μ are conditional on different initial states: λ is the rate of degradation from a normal state, whereas μ is the rate at which an already-adverse state is restored, and during S2 and S3 restoration is driven by the monitoring, pause–review, and secondary-reinforcement actions that field procedures mandate precisely when an anomaly appears. Because node-start failure probabilities are of order 10^{-2} , the deterioration term dominates the propagated marginal even when $\mu > \lambda$, which is why the propagated failure probability of X29 rises rather than falls across a 16-h S2 node (Section 3.3). A calibration in which the recovery pathway is removed entirely ($\mu = 0$) is reported as a bounding check in Section 3.4.

This parameterization accommodates reordering, skipping, and repetition, but it does not capture dependence on the preceding stage type, cumulative damage from repeated operations, or coupling among parallel work fronts. Sequences that deviate substantially from standard practice therefore require parameter review or a history-dependent DBN.

2.3.4 Propagation operator and cascading structure

For each persistent state X_i , Eq. (15) propagates the node-start distribution to the planned-window endpoint or to any elapsed-time update:

$$\pi_k^i = \left[\Theta_i^{(s_k)}(\Delta t_k) \right]^\top \rho_{k-1}^i, \quad i \in \mathcal{C} \quad (15)$$

Here, ρ is the inherited node-start distribution and π the predicted distribution after duration Δt_k in stage s_k . Planned duration is not applied to the start state; during filtering, Δt_k is the actual interval between consecutive updates.

Cross-node information travels mainly through persistent-state marginal posteriors, avoiding double counting. A residual-risk sentinel $H(k-1)$ is a binary node that carries forward a disturbance recorded at node $k-1$ that no persistent state represents — for example an unplanned impact, a dropped load, or an aftershock-induced displacement logged by the safety officer but not captured by X29, X31, or the gas group. It is activated by the safety officer at handover, enters node k as an additional parent of the environment dimension, and is reset once the corresponding condition has been inspected and cleared. It was not activated at any node of the Mandalay case, so no result reported in Section 4 depends on it. Regional-background and stage-specific technical indicators are not propagated by Φ_k .

2.3.5 Marginal approximation and applicability

Cross-node updating relies on a marginal-independence approximation: the eight persistent-state marginals are propagated separately and recombined into the next-node predictive prior. Appendix B establishes the formal relation to non-homogeneous DBN forward filtering under the same approximation, including the required conditions and proof; Section 3.4 tests only numerical consistency between two implementations of this recursion.

Numerical consistency does not imply equivalence to a full non-homogeneous DBN. Shared descendant evidence can induce posterior dependence among persistent states, and marginal propagation can lose joint information. Accordingly, cumulative damage, strongly correlated evidence, or parallel work fronts may require recalibration or a richer model. Appendix D reports additional checks on structural consistency, background-node isolation, closure of propagated variables, parameter reuse, and inference logic.

2.4 Expert-driven FGW-DBNP parameterization

WPC-BN requires root-node priors, within-node CPTs, and stage-transition rates. Because accident data are sparse and field observations are incomplete, the FGW-DBNP procedure combines fuzzy elicitation, expert judgment, dispersion-based weighting, and stage-transition calibration. These parameters support relative risk ranking rather than frequency estimation.

2.4.1 Fuzzy elicitation of root-node priors

Questionnaire 2 elicits root-node failure likelihoods on seven linguistic levels (VL-VH) represented by trapezoidal fuzzy numbers (Zadeh, 1965; Ferdous et al., 2011; Yazdi and Kabir, 2017; parameters in Supplementary Table S2-1). Expert-weighted judgments are aggregated, defuzzified by the centroid method to give the fuzzy possibility score (FPS), and mapped by the Onisawa function to model-based priors (Onisawa, 1988). Appendix C.1 gives the equations, Appendix Fig. C1 shows the membership functions, and Supplementary Tables S2-4 and S2-5 provide the raw judgments and baseline priors.

2.4.2 CPT construction by combined weighting

Because consequence severity is not a conditional occurrence probability, score-based CPTs combine AHP ranking weights with entropy weights derived from the dispersion of expert severity ratings, using the 1–10 scores (Module C) and pairwise matrices (Module D) from Questionnaire 2.

AHP yields WA subject to $CR < 0.1$. Column-wise min–max normalization of severity ratings, followed by expert-weighted entropy proportions (Eq. C7), is used to derive the dispersion weight WB. WB reflects inter-expert rating dispersion, not severity, and depends on the normalization scheme. Combining a dispersion weight with a ranking weight therefore requires justification, because dispersion alone would reward disagreement. Here WA carries the substantive ordering and WB is used only to damp positions on which the panel is internally consistent but the ranking is weakly separated; the game-theoretic step constrains the result to lie between the two sources rather than allowing either to dominate. The consequence is visible in the worked example of Section 3.3, where WB reverses the WA ordering of X76 and X77; because the two combined coefficients differ by 0.047, the reversal does not change the sensitivity ordering of the branch, which is governed by the priors ($X77 > X76 > X78$). Analyses that require the elicited importance ordering to be preserved exactly should use WA alone. Game-theoretic combination gives W^* by minimizing total deviation from WA and WB; see Appendix C.2.

The appropriate CPT form depends on the relationship among parent nodes; three forms are used.

(a) Substitutable factors. Groups that are approximately OR-type, such as qualifications or equipment availability, use a weighted linear CPT:

$$P(M=1 | \omega) = \sum_{u \in A(\omega)} \omega_u^* \sum_u \omega_u^* = 1 \quad (16)$$

where $A(\omega)$ denotes the set of failed parent nodes in state combination ω . This weighted linear approximation to an OR gate preserves the $[0,1]$ bounds but implies neither zero real-world risk nor certain failure at the extremes.

(b) Synergistic factors. Strongly coupled pairs—breaching vibration X77 with structural stability X29, combustible gas X40 with spark control X78, and burial depth X33 with access-construction difficulty X34—use a noisy-OR variant with bounded interaction terms:

$$P(M=1 | \omega) = 1 - (1 - \pi_0) \prod_{u \in A(\omega)} (1 - q_u) \prod_{(u,v) \in E_\omega} (1 - q_{uv}) \quad (17)$$

where q_u is the risk strength of individual parent failure u , π_0 is the leak term, q_{uv} is the interaction term for a predefined pair, and E_ω is the set of active interaction edges under state combination ω . These are score-based CPT parameters, not empirical conditional accident frequencies. Coupling pairs, leak terms, and interaction corrections are given in Supplementary Table S4-1; only disclosed parameters fall within the reproducibility scope.

(c) Conjunctive failure links. A deterministic AND gate, or an AND gate with a leak term accounting for measurement error and unmodeled failures, is used when substantial risk arises only from the joint failure of several conditions.

Together, the three forms represent substitutable, synergistic, and conjunctive relationships. Because none is estimated from accident samples, all are interpreted as model-based CPTs rather than empirical conditional probabilities.

2.4.3 Elicitation of stage-specific transition rates

Stage-transition rates are calibrated by variable group $\times$ stage type from seven-level linguistic judgments over an 8 h reference window (Questionnaire 3). The modes in Supplementary Table S3-2 describe directional tendencies only; calculations use d (8 h) and r (8 h) from Supplementary Table S4-2 and convert them to λ and μ using Section 2.3.3. Strong aftershocks, reinforcement pauses, or abrupt plan changes trigger piecewise updates.

Persistent states are grouped as structural, gas, personnel, and equipment. Variables in a group share baseline rates, with optional variable-specific adjustment factors. Directional checks compare relative deterioration and recovery across stages and do not require $\lambda > \mu$ for every group in S2 or S3.

2.5 Risk outputs, diagnostics, and rolling decisions

Once the network structure and parameters are fixed, WPC-BN supports pre-task prediction, in-process filtering, and post-task review, together with internal diagnostics and illustrative decision bands. It supplements field command, safety-officer judgment, and applicable procedures rather than replacing them.

2.5.1 Three inference modes

Pre-task prediction supports plan comparison and node admission. It uses only pre-initiation evidence together with the planned method, shoring scheme, and duration to estimate risk at the end of the planned window; it does not represent the state at task start. In Eq. (18), a_k is the planned action at node k (method, equipment configuration, and duration) and z_k is a configuration of Z_k .

$$\hat{R}_k^{pre} = P(Y_k = 1 | e_{0:k-1}, a_k) = \sum P(Y_k = 1 | z_k) P(z_k | e_{0:k-1}, a_k) \quad (18)$$

Crossing a decision threshold prompts plan revision, strengthened controls, or expert review rather than automatic shutdown. In-process filtering updates *NSRI* in the order that evidence arrives and over the actual elapsed intervals; abnormal monitoring, inspection, or personnel evidence triggers immediate reassessment. Post-task review examines risk evolution and key factors using the complete record. Because formal backward smoothing is not performed, the terminal quantity is a filtered posterior at the end of the node rather than a smoothed estimate, and this mode is therefore termed review throughout; the term terminal posterior is used interchangeably for the same quantity. The three modes use strictly separated information sets: planned and historical information, evidence available up to the current time, and the complete record.

2.5.2 Posterior relative update, mutual information, and dimension contributions

Posterior relative update (*PRU*; corresponding to the ratio of variation, *ROV*, in earlier literature) quantifies the evidence-induced relative change in a basic-event probability:

$$PRU_i = \frac{P(X_i=1|e) - P(X_i=1)}{P(X_i=1)}, \quad P(X_i=1) > 0 \quad (19)$$

where $P(X_i=1)$ and $P(X_i=1|e)$ are the prior and posterior probabilities. For diagnostic ranking, PRU is further conditioned on $Y_k=1$ to measure each basic event's relative update under the latent unsafe state; indicators fixed as normal by hard evidence are excluded. Because PRU depends on prior magnitude, it carries no causal interpretation.

Mutual information (Shannon, 1948) between each root node and the top event is also computed:

$$I(X_i; Y) = \sum_{x, y} P(x, y) \log_2 \frac{P(x, y)}{P(x)P(y)} \quad (20)$$

PRU quantifies relative updating, whereas mutual information measures model-internal dependence with the top node. The two diagnostics are complementary within the same network and do not provide independent external validation.

The relative contribution of each of the four dimensions is defined as:

$$C_b = \frac{W_b P(B_b=1|e)}{\sum_{j=1}^4 W_j P(B_j=1|e)} \quad (21)$$

where W_b is the score-based CPT coefficient for the top node Y and B_b is the corresponding dimension node. The four top-node coefficients are fixed across M0–M4, were elicited from the ten parameterization experts using the procedure in Section 2.4.2, and are stored in the accompanying stage-network files. The 33-expert percentage weights (Supplementary Section S7.3) serve only as a robustness comparison and do not enter $NSRI$. When the top node uses Eq. (16), Eq. (21) exactly decomposes $NSRI$; under a nonlinear CPT it gives relative share only.

2.5.3 Illustrative decision bands and rolling rules

The four-level $NSRI$ scheme is used only to demonstrate the rolling decision process (Table 4): the 1%, 2.5%, and 5% cutoffs are model-internal examples rather than outcome-calibrated thresholds. Field deployment requires recalibration based on false-positive and false-negative costs, team risk tolerance, and applicable procedures. The $\pm 20\%$ perturbation in Section 4.3 tests only local stability of node ranking and decision-band placement.

Table 4. Illustrative $NSRI$ decision bands and rolling decision rules

Level	$NSRI$ range	Illustrative action (subject to field command and applicable procedures)
I (Low)	$NSRI < 1\%$	Start or continue as planned with routine monitoring.
II (Moderate)	$1\% \leq NSRI < 2.5\%$	Start or continue with enhanced monitoring and reinforcement resources on standby.
III (High)	$2.5\% \leq NSRI < 5\%$	Pause or defer. Apply diagnostic-guided controls, then reassess before resuming under field command.
IV (Very high)	$NSRI \geq 5\%$	Evacuate to a safe area and repeat reconnaissance/assessment; final action follows field command and applicable procedures.

3. Indicator System, Parameter Calibration, and Internal Validation

3.1 Expert sample and parameterization design

The survey system comprised three parameterization questionnaires and six review modules (Table 5; Supplementary Table S3-1): Questionnaire 1, completed by the 20 Delphi experts, supported two-round screening; Questionnaire 2, completed by the ten parameterization experts, elicited indicator likelihood, consequence severity, and AHP judgments; and Questionnaire 3, from the same ten experts, calibrated 8-h state-transition parameters. Review Modules A–F, completed by the 33 reviewing experts, covered criterion-level validation, dimension weighting, stage-direction assessment, and standardized scenario ranking and were conducted independently of the indicator-level parameterization.

Table 5. Questionnaire structure and model functions

Questionnaire	Evaluation target	Scale	Model function
Questionnaire 1: Indicator relevance	Indicator relevance within 19 criteria: importance, representativeness, and operability	Five-point scale + open comments	Indicator screening; CVR and Kendall's W
Questionnaire 2: Likelihood and consequence (Modules A–D)	A: case scenario; B: occurrence likelihood; C: consequence severity; D: pairwise comparisons	Seven linguistic levels (VL–VH); 1–10 score; 1–9 AHP scale	Root-node priors (B); entropy-based dispersion weights (C); AHP weights (D)
Questionnaire 3: Stage-transition calibration	Stage-specific deterioration/recovery of persistent-state groups over an 8-h reference window	Seven linguistic levels (VL–VH)	Unit-time deterioration rate λ and recovery rate μ

Twenty experts completed both Delphi rounds. A separate panel of ten specialists with earthquake-rescue experience (five researchers and five engineers) calibrated indicator priors, CPTs, 8-h transition parameters, and case-specific judgments. The 20-member Delphi, 10-member calibration, and 33-member review panels were disjoint, giving 63 participants in total. Table 5 and Supplementary Table S3-1 summarize panel functions; calibration-panel scores and weights are reported in Supplementary Tables S2-2 and S2-3, and review-panel backgrounds and weights in Supplementary Table S7-1. Expert authority was screened using the conventional Delphi coefficient $Cr = (Ca + Cs)/2$, where Ca is the self-rated judgment basis and Cs the self-rated familiarity with the topic, with $Cr \geq 0.70$ used as the inclusion criterion. Because the same ten calibration experts contributed both model parameters and case-specific judgments, common-source bias cannot be excluded.

Table 6. Expert classification and scoring criteria

Criterion	Classification and score
Professional title / technical grade	Full senior: 5; Associate senior: 4; Intermediate: 3; Assistant: 2; Junior: 1
Relevant experience	≥ 20 years: 5; 15–19: 4; 10–14: 3; 5–9: 2; < 5 : 1
Education	Doctorate: 5; Master's: 4; Bachelor's: 3; Associate degree: 2; Other: 1
Major rescue operations	≥ 10 : 5; 7–9: 4; 4–6: 3; 1–3: 2; None: 1
Professional role	Rescue technical specialist/safety officer: 5; Command/management personnel: 4; Earthquake-rescue expert at a research institute: 4; University researcher: 3; Other: 1

The panel covered field command, safety management, structural engineering, search and rescue, breaching/shoring, equipment and logistics, and research. Questionnaire versions, samples, and uses are documented in Supplementary Material S3 (Table S3-1).

3.2 Content validity and consistency of the indicator system

The same 20 experts participated in both Delphi rounds. In Round 1, 82 of 96 candidate indicators exceeded the CVR threshold of 0.42 for a 20-member panel (Lawshe, 1975). Of the remaining 14

indicators, 11 were removed or merged and three were retained after revision, yielding the 85-indicator system evaluated in Round 2. The Supplementary Materials report the final definitions and round-level statistics but do not provide an item-by-item log of the 96-to-85 revision process.

In Round 2, all 85 indicators reached $CVR \geq 0.50$, with mean relevance scores of 3.8–4.9. The mean coefficient of variation decreased from 0.21 to 0.16, while Kendall's W increased from 0.287 to 0.341 (both $p < 0.001$). Cronbach's α ranged from 0.83 to 0.91 across the four dimensions and serves only as a reference for internal questionnaire consistency (Table 7).

Exploratory subgroup comparisons within the expert review were used only as a robustness check and did not alter indicator retention or model parameterization. Because these comparisons are not part of the validation evidence reported in the Supplementary Materials, no inferential claim in this study depends on them.

Table 7. Summary of the two-round Delphi evaluation

Test	Round 1 (96 candidates)	Round 2 (85 indicators)	Criterion	Result
Valid responses	20	20	—	—
Indicators meeting CVR threshold	82/96	85/85	$CVR \geq 0.42$ ($n = 20$; Lawshe, 1975)	Passed (Round 2: all 85 indicators $CVR \geq 0.50$)
Kendall's W (χ^2)	0.287 (545.3, $p < 0.001$)	0.341 (572.9, $p < 0.001$)	($p < 0.05$)	Significant; modest agreement
Mean coefficient of variation	0.21	0.16	Decrease	Convergent
Cronbach's α	0.79–0.88	0.83–0.91	≥ 0.70	Passed (reference only)
Mean relevance score	3.1–4.8	3.8–4.9	≥ 3.5	Round 1: items below 3.5 removed or revised; Round 2: all items ≥ 3.5

The 33-expert review assessed the 19 criteria for safety relevance, cross-case generalizability, and field observability; all were retained under the prespecified rule. Median safety-relevance scores were 4–5 with $IQR \leq 1$, and all generalizability medians were ≥ 4 . Observability medians for C3, C11, and C13 were 3, indicating that instrumented or network-based evidence should be prioritized where available. Kendall's W was 0.515 ($p < 0.001$; Supplementary Table S7-2). These results support criterion-level content validity, but the uniformly high ratings indicate a ceiling effect and do not substitute for independent case validation.

3.3 Parameter calibration and illustrative results

The S2 breaching indicators X76–X78 are used to illustrate FGW-DBNP parameterization and identify which calculations can be reproduced independently.

(1) Prior probabilities.

The ten parameterization experts scored 17, 16, 19, 20, 18, 15, 22, 22, 18, and 18 under Table 6, and their exact normalized weights were used. For X77, the linguistic judgments {ML, M, ML, L, M, ML, M, M, L, L} were aggregated to (0.2541, 0.3541, 0.3816, 0.4816). The resulting fuzzy possibility score was $FPS = 0.3678$, which yielded $F = 1.75 \times 10^{-3}$ after the Onisawa transformation. Other root-node priors were obtained analogously.

(2) Score-based CPTs.

All ten AHP matrices satisfied the consistency criterion ($CR < 0.1$). Aggregation produced $W_A = (0.2863, 0.4561, 0.2576)$, while the entropy-based dispersion weights were $W_B = (0.4832, 0.2720, 0.2448)$. Game-theoretic combination yielded $W^* = (0.3982, 0.3515, 0.2503)$.

The X76–X78 branch used the linear CPT of Eq. (16), with the resulting conditional probabilities in Table 8. The noisy-OR coupling between X77 and X29 is reported in Supplementary Material S4.2 and the original expert matrices in Supplementary Table S2-6.

Table 8. Example score-based CPT for the “breaching technical risk” node

X76	X77	X78	CPT: Risk = Yes	CPT: Risk = No
No	No	No	0	1
Yes	No	No	0.3982	0.6018
No	Yes	No	0.3515	0.6485
No	No	Yes	0.2503	0.7497
Yes	Yes	No	0.7497	0.2503
Yes	No	Yes	0.6485	0.3515
No	Yes	Yes	0.6018	0.3982
Yes	Yes	Yes	1	0

(3) Stage-transition rates.

For X29, the 8-h transition probabilities are reported in Supplementary Table S4-2 and converted to λ and μ using Section 2.3.3 (Table 9). The deterioration rates for S2 and S3 were approximately 6.02 and 4.69 times the S4 rate, respectively, whereas the recovery rate in S1 was approximately 80 times its deterioration rate. For a 16-h S2 operation, $d(S2)(16\ h)=1.427\times 10^{-2}$, and the failure probability of X29 increased from 1.000×10^{-2} at node initiation to 2.391×10^{-2} after propagation. The node-initiation value is itself the output of the preceding cascade and not a free input: the Mandalay case-specific prior of 9.77×10^{-3} (Table 11), propagated through the 11-h S1 node with the structural-group rates of Table 9, returns 1.000×10^{-2} at the start of M2. This node-initiation value can be reproduced directly from the reported prior and transition parameters in Table 11 and Table 9.

Table 9. Stage-specific transition parameters for structural stability (X29)

Stage	Dominant mechanism	λ/h^{-1}	μ/h^{-1}	$d(8\ h)$	$r(8\ h)$	Linguistic mode (deterioration/recovery)
S1 Shoring and bracing	Increased restraint; load redistribution	1.03×10^{-4}	8.27×10^{-3}	0.8×10^{-3}	6.4×10^{-2}	VL / MH
S2 Slab breaching	Vibration and component weakening; deterioration higher than S4	9.08×10^{-4}	1.39×10^{-3}	7.2×10^{-3}	1.1×10^{-2}	ML / L
S3 Debris removal	Unloading and lifting impacts; deterioration higher than S4	7.07×10^{-4}	1.89×10^{-3}	5.6×10^{-3}	1.5×10^{-2}	ML / L
S4 Casualty transport	Minor disturbance from transport and equipment withdrawal; lowest deterioration	1.51×10^{-4}	1.13×10^{-3}	1.2×10^{-3}	0.9×10^{-2}	VL / L

3.4 Numerical properties and implementation consistency

(1) Linear scoring branch.

For the M2 “breaching technical risk” node, the priors of X76–X78 were 1.18×10^{-3} , 1.75×10^{-3} , and 0.90×10^{-3} , respectively, producing a baseline output of 1.31×10^{-3} under Eq. (16). Raising each prior by 10% individually increased the child-node output by 3.6%, 4.7%, and 1.7%, giving the sensitivity order $X77 > X76 > X78$, and linear monotonicity held over 10%–50% perturbations. The output was 0 with all parents normal and 1 with all failed, and raising all three priors by 10% together increased the output by 10.0%, equal to

the sum of the individual responses (Fig. 6). The same checks were applied to the remaining linear branches; they verify only the intended numerical properties of the scoring structure. All inference was performed in GeNIe 4.1 (BayesFusion) with the stage models M0–M4 supplied as .xdsl files, and cross-node propagation was performed using the Python script in Supplementary Material S6.3; file organization and workflow are described in S6.2.

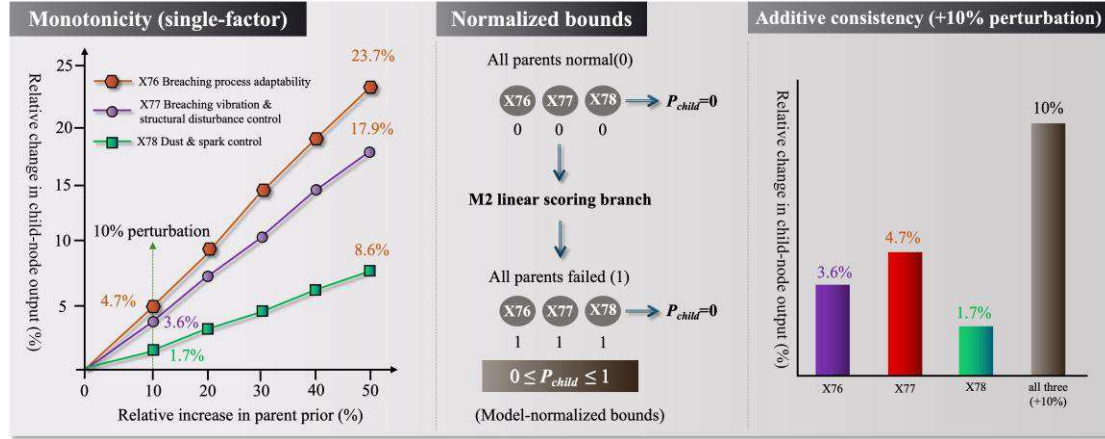


Fig. 6. Numerical-property check for the M2 linear scoring branch

(2) Cascading propagation.

At M2 initiation, the failure probability of X29 was 1.000×10^{-2} . After 16 h of S2 propagation, it increased to 2.391×10^{-2} . Raising the initial probability by 20%, to 1.200×10^{-2} , increased the window-end value to 2.584×10^{-2} , corresponding to $\Delta P = 1.93 \times 10^{-3}$. As a bounding check on the recovery pathway, setting $\mu=0$ for the structural group in S2 while holding λ fixed raises the window-end value to 2.428×10^{-2} , an increase of 3.8×10^{-4} ; the node ranking of Table 12 is unchanged. Given $\text{dp}(Y=1) / \text{dp}(X29=1) \leq 1$, the absolute increase in the M2 window-end NSRI cannot exceed 0.19 percentage points. Only the reproducible change in the propagated state probability is therefore reported; no NSRI change is inferred without recomputing the complete stage network. Alternative stage-sequence tests served only to verify parameter invocation and propagation logic.

(3) Software implementation consistency.

Under identical inputs, marginal-independence assumptions, and time definitions, the maximum node-wise absolute difference between the WPC-BN cascade and non-homogeneous DBN forward filtering was $< 10^{-6}$ (Fig. 7). This is a code-verification result and not a validation result: because both implementations execute the same approximate recursion, agreement is expected, and the test establishes only that the modular cascade introduces no implementation error relative to the slice-wise form of the same recursion. It does not establish the effect of posterior factorization relative to exact joint propagation; Section 5.1 defines this applicability boundary, and Supplementary Section S6.7 gives the corresponding exact-joint benchmark protocol.

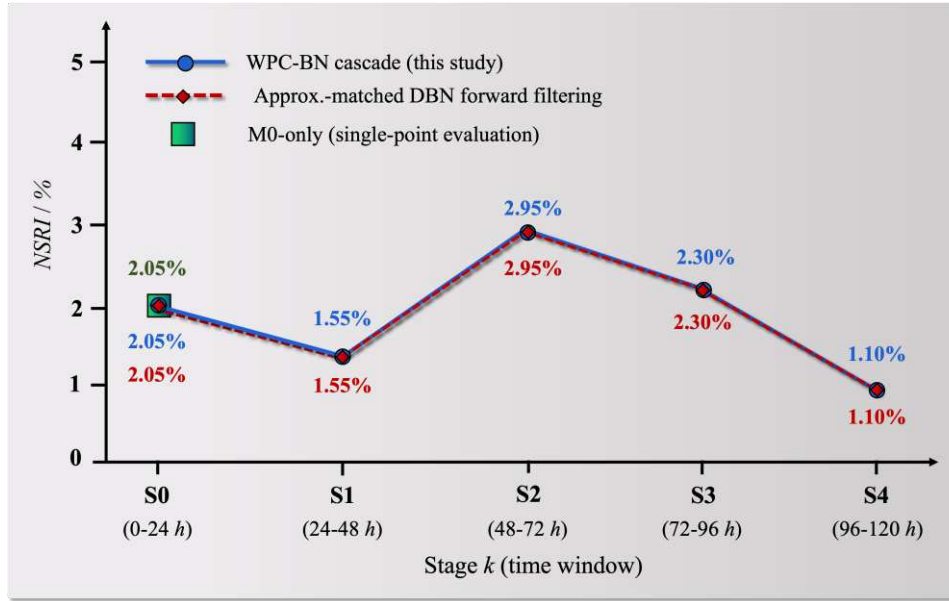


Fig. 7. Software implementation consistency between the WPC-BN cascade and the DBN under the same marginal-independence approximation

3.5 Robustness, stage-direction, and scenario-ranking checks

(1) Dimension-weight robustness. With direct 100-point allocation by the 33 experts, the mean ranking was Environment B3 (34.1) > Management and operation control B4 (25.9) > Personnel B1 (22.9) > Equipment B2 (17.2). Expert-weighted means produced a similar ordering (Supplementary Table S7-3). These weights serve only as a robustness reference and do not replace the top-node CPT coefficients W_b .

(2) Stage-direction consistency. Review Module E used a three-state judgment—deterioration, no change, or improvement—over a 1- h window. Because the CTMC quantities d and r are conditional on different initial states, the responses cannot be converted directly into CTMC transition probabilities and were used only to assess directional consistency, not to recalibrate λ or μ . The judgments are reported in Supplementary Table S7-4. Directional agreement with Supplementary Table S4-2 was obtained for S1 and S4; for S2 and S3 the 1- h three-state judgments indicated net deterioration, whereas the 8- h calibration retains $r > d$ for the structural group, reflecting the different conditioning bases of the two instruments (Section 2.4.3) rather than a recalibration conflict.

(3) Standardized scenario ranking. Review Module F estimated the probability of local or progressive collapse within 30 min, whereas NSRI is a composite model-based unsafe-state score; because the two differ in event definition and scale, only their ranking directions were compared. The scenario medians are given in Supplementary Table S7-5, and the ordering $V4 > V3 > V8 > V2 > V6 > V7 > V1 > V5$ places breaching- and removal-stage disturbance scenarios above shoring- and transport-stage scenarios, matching the node ranking in Table 12. The comparison examines mechanistic consistency and supports the use of these scenarios for future calibration; it is not evidence of external discriminative validity.

4. Retrospective Engineering Application

WPC-BN was applied retrospectively to the rescue at the Golden Country Hotel in Mandalay, Myanmar. Public information, node-level field records, and expert review were combined to exercise pre-task prediction, in-process filtering, cross-node propagation, and internal diagnostics. Model outputs were compared with observed structural responses under time-locked evidence, so within-stage anomalies entered the reconstructed inference sequence only after their recorded occurrence. This constrains temporal information leakage in the model sequence but does not make retrospective coding outcome-blind. Because the analysis is based on a single case and largely expert-elicited parameters, predictive accuracy was not estimated.

4.1 Case overview

On 28 March 2025, an Mw 7.7 earthquake struck near Mandalay, Myanmar (Goldberg et al., 2025). At about 17:40 on 2 April, the China Search and Rescue Team and the China International Search and Rescue Team rescued a male survivor trapped for more than 120 *h* at the Golden Country Hotel and transferred him to a local medical facility (China News Service, 2025). Public sources define the case context; node-level field records supply the structural-response evidence during rescue (Fig. 8). No claim is made about field details that the available records do not cover.



Fig. 8. Collapse condition and rescue operations at the Golden Country Hotel (photograph courtesy of Wen Bai).

Public information, field records, and expert review map the rescue process to stages S0–S4 and update nodes M0–M4 (Table 10). The intervals are reconstructed for model analysis rather than taken from a public hourly log, and structural responses at M1–M3 come from actual operational records. T0 denotes only the model time origin and is not equivalent to the publicly reported entrapment duration of more than 120 *h*.

Table 10. Stage-based case timeline and key evidence

Node	Stage	Relative interval	Key operations and evidence
M0	S0 Reconnaissance and initial assessment	T0-T0+6 <i>h</i> (6 <i>h</i>)	UAV reconnaissance and life detection; aftershock, heat, and initial structural evidence collected; dual-access rescue plan developed.
M1	S1 Shoring/bracing	T0+6-T0+17 <i>h</i> (11 <i>h</i>)	Combined shoring and hydraulic jacking; minor deformation of local structural components and the rescue passage during jacking; shoring displacement reviewed.
M2	S2 Slab breaching	T0+17-T0+33 <i>h</i> (16 <i>h</i>)	Downward slab breaching and cutting; pronounced abnormal sounds and tilting risk observed; work paused, reviewed, and reinforced.
M3	S3 Debris removal	T0+33-T0+49 <i>h</i> (16 <i>h</i>)	Layered removal using manual work and small machinery; vertical deformation of the established rescue passage; unloading effects continuously monitored.
M4	S4 Casualty transport and evacuation	T0+49-T0+57 <i>h</i> (8 <i>h</i>)	Early medical involvement; continuous casualty monitoring; transport and site handover completed.

4.2 Model parameters and evidence entry

Model inputs are of four kinds: general baseline parameters, case-specific priors, dynamic evidence available before a node begins, and new evidence acquired within a stage. The ten parameterization experts calibrated general priors, CPTs, and transition rates, and elicited the case-specific priors for the Mandalay setting (Table 11). Structural responses observed within a stage were excluded from pre-task prediction and entered only after occurrence during filtering. Across nodes, only the ordinary-evidence posterior $P(X_i=1|e_k)$ was propagated; the posterior ϕ used in PRU diagnostics conditional on $Y_k=1$ was not propagated.

Three terms are used throughout and are not interchangeable: general baseline prior, case-specific prior, and stage-propagated prior. The current ten-expert parameter set is an initial library; cross-regional use requires recalibration to local building, site, and team conditions within the same network interface. Questionnaires, evidence coding (Supplementary Table S5-1), and calculation details are given in Supplementary Materials S3 and S5.

Table 11. Case-specific root-node judgments and model-based priors (distinct from the general baseline priors in S2)

Indicator	Modal linguistic judgment	FPS	Model-based prior
X33 Burial depth	H	0.800	3.55×10^{-2}
X32 Relative burial position	MH-H	0.720	2.09×10^{-2}
X34 Access-construction difficulty	MH	0.650	1.34×10^{-2}
X72 Aftershock intensity and frequency	MH	0.601	9.86×10^{-3}
X29 Structural stability	M-MH	0.600	9.77×10^{-3}
X46 High temperature	M-MH	0.580	8.59×10^{-3}
X39 Oxygen level	L	0.200	2.22×10^{-4}
X21 Equipment integrity and reliability	VL-L	0.170	1.25×10^{-4}

4.3 NSRI evolution across three information states

4.3.1 Pre-task, in-process, and post-task information states

NSRI was evaluated in three information states: pre-task prediction at the planned-window endpoint, in-process filtering, and post-task review. Pre-task values used only evidence available at node initiation, the planned operation, and the planned duration. Structural responses observed during M1–M3 entered only after occurrence, preserving the predict-observe-update sequence (Table 12). Because every node of this case closed without further evidence arriving after the last in-process update, the post-task review value coincides with the final in-process value at all five nodes; Table 12 therefore reports two distinct numerical columns and the review state is represented by the observed-action column. This coincidence is a property of the present record, not of the framework.

Table 12. Three information states and retrospective action comparison using illustrative NSRI thresholds

Node	Pre-task prediction (planned-window endpoint)	In-process filtering	Illustrative level	Pre-task action	In-process action	Observed action in review records
M0	— (initial assessment is the first information state)	2.05%	II	—	Proceed; intensify aftershock and gas monitoring (illustrative).	Dual-access plan initiated with monitoring posts.
M1	1.80%	1.62%	II	Proceed.	Continue; review shoring before breaching.	Shoring completed and reviewed before the next node.
M2	2.55% (illustrative Level III threshold)	2.86%	III	Defer pending diagnostic-guided controls: strengthen vibration control and review shoring; proceed only after reassessment under field command.	Trigger pause/review; reinforce shoring, reduce vibration/load, and reassess before resuming.	Work paused after abnormal sounds and tilting risk were observed; resumed after review and reinforcement.

Node	Pre-task prediction (planned-window endpoint)	In-process filtering	Illustrative level	Pre-task action	In-process action	Observed action in review records
M3	2.35%	2.21%	II	Proceed; control unloading rate.	Continue; enforce machinery blind- spot exclusion zones.	Manual work and small machinery used for controlled removal.
M4	1.30%	1.18%	II	Proceed; position medical support in advance.	Continue.	Medical support deployed early; casualty rescued and handed over.

M2 carried the highest *NSRI*, rising from 2.55% at the planned-window endpoint to 2.86% during filtering. Pronounced abnormal sounds and tilting risk were observed during breaching; once that evidence arrived, *NSRI* rose and the model called for a pause, shoring reinforcement, vibration/load reduction, and reassessment, directionally consistent with the field response of pausing, reviewing, and reinforcing. The comparison supports process consistency in the location and updating direction of the risk peak, not predictive accuracy.

4.3.2 Local output-layer perturbation

To test the stability of node ranking, independent multiplicative perturbations $U(0.8, 1.2)$ were applied to the five in-process *NSRI* point estimates over 10,000 Monte Carlo samples (random seed 42; script in S6.4), leaving priors, CPTs, transition rates, evidence reliability, and network structure unperturbed. M2 remained the highest-risk node in 92.63% of samples, and in 81.69% of samples its *NSRI* remained at or above the illustrative 2.5% threshold (Fig. 9).

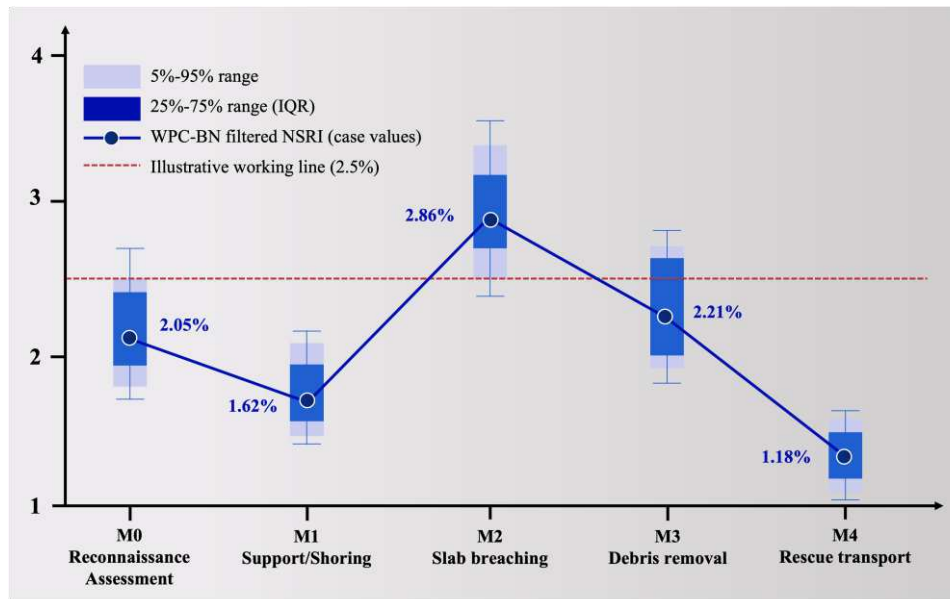


Fig. 9. Distributions of in-process *NSRI* across five nodes under $\pm 20\%$ independent multiplicative output perturbations ($N = 10,000$; random seed = 42).

4.4 Node-level diagnosis: dimensional contributions and key factors

4.4.1 Contributions of the four risk dimensions

The dominant source of risk changes as the operation progresses. Relative contributions of the four dimensions were computed using Eq. (21) (Fig. 10). The environmental contribution decreased from 48% at M0 to 38% at M2 and 30% at M4, whereas management and operation control increased from 23% at M0 to 27% at M2 and remained at approximately 27%–32%

thereafter. Equipment remained at about 15%–20%, and personnel increased from 14% at M0 to 22% at M4. Thus, early environmental dominance gradually gives way to management and personnel factors. Only contributions obtained from actual node inference are reported; unverified counterfactual recalculations are not used.

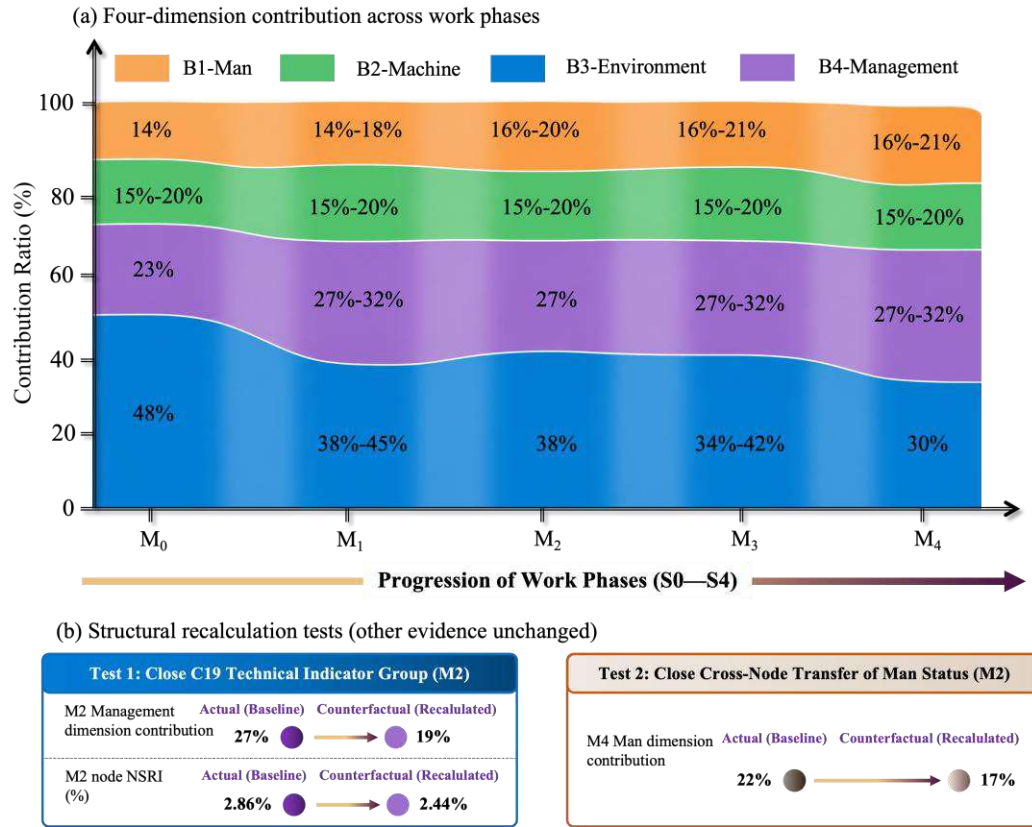


Fig. 10. Relative contributions of the four risk dimensions across case nodes.

4.4.2 Internal diagnosis using PRU and mutual information

At M2, with X76 and X78 excluded because hard evidence fixed them as normal, PRU ranked X77, X29, and X31 highest, and mutual information gave the same top three at 4.74×10^{-3} , 4.21×10^{-3} , and 3.66×10^{-3} bits, respectively (Supplementary Tables S5-2 and S5-3). Because both metrics derive from the same network, they serve only as complementary internal diagnostics; their engineering reading nonetheless matches the pronounced abnormal sounds and tilting risk observed at M2. Fig. 11 reports only the mutual-information ranking, to avoid mixing diagnostic posteriors with cross-node propagation quantities.

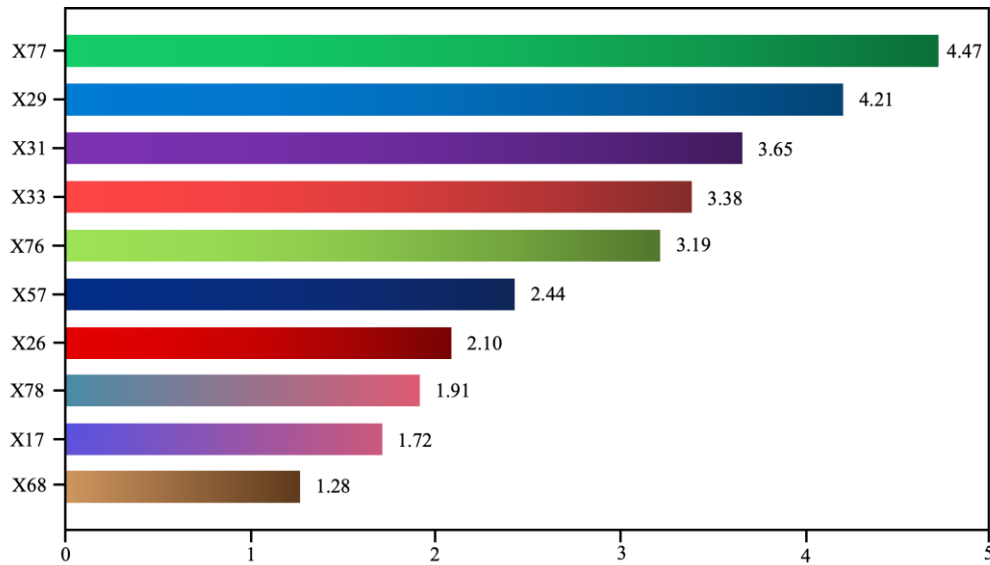


Fig. 11. Mutual-information ranking of root factors at M2 (based on Supplementary Table S5-3).

4.5 Retrospective consistency with field observations

Field records from M1 shoring, M2 breaching, and M3 debris removal served as process references under the same time-locking rule: pre-task *NSRI* used only pre-node information, whereas within-stage structural responses entered filtering only after occurrence. Given the small number of nodes, the comparison addresses the location and direction of the risk peak, dominant mechanisms, and the final process state; AUC, sensitivity, and specificity were not calculated (Table 13). Retrospectively coded soft evidence, including X29, X77, and X81, was assigned with knowledge of the completed rescue, so the case is interpreted as a process-consistency demonstration rather than predictive validation. Prospective shadow-mode deployment, outlined in Section 5.4, is the appropriate next step for outcome-blind performance assessment.

Table 13. Retrospective consistency between field-observed structural responses and WPC-BN outputs

Node	Field record	Model output and internal diagnosis	Retrospective interpretation
M1 Shoring/bracing	Minor deformation of local structural components and the rescue passage during jacking.	1.80% pre-task → 1.62% in-process; Level II; lower risk than M2 and M3.	The local response is compatible with a generally stabilizing shoring stage that can still induce limited disturbance.
M2 Slab breaching	Pronounced abnormal sounds and tilting risk during breaching.	2.55% pre-task → 2.86% in-process; Level III and highest overall; X77, X29, and X31 prioritized.	The most prominent field anomaly coincided with the modelled risk peak, while the leading factors directly reflected vibration and structural stability.
M3 Debris removal	Vertical deformation of the established rescue passage during debris removal.	2.35% pre-task → 2.21% in-process; Level II and second-highest overall; internal diagnosis highlighted X80, X30, and X81.	The observed deformation is compatible with unloading, release of constraints, and operational disturbance during removal.
Whole-process final state	Overall debris tilt angle of 3.6°; field record classified the overall debris structure as safe.	<i>NSRI</i> peaked at 2.86% at M2, then declined to 2.21% at M3 and 1.18% at M4.	The trajectory supports a process interpretation of stage-specific risk increase, intervention, subsequent risk reduction, and final control. The 3.6° value is treated only as an observed endpoint, not as a calibration threshold.

Before the M2 anomaly was entered into the model, the planned-window *NSRI* was already 2.55%, the highest among all nodes; after the anomaly was incorporated, it rose to 2.86%. *PRU* and mutual information both prioritized X77, X29, and X31, directly linking the diagnosis to the

abnormal sounds and tilting risk. M1 showed only mild local deformation and a lower in-process *NSRI*, whereas M3 remained relatively high, consistent with unloading and passage deformation.

The final measured debris tilt angle was 3.6° , and the field record judged the overall structure to be safe. This terminal observation does not negate the stage-specific peak at M2: the anomaly was followed by a pause, review, reinforcement, and vibration control, and *NSRI* subsequently declined at M3 and M4. Because 3.6° was not predefined as a safety threshold, it is treated only as a terminal observation and is not used to back-calibrate *NSRI* thresholds.

5. Discussion

5.1 Methodological contribution to whole-process assessment

The principal methodological contribution of WPC-BN is to convert post-earthquake rescue safety from a sequence of isolated judgments into a common node-based inference process. The assessment unit is the actual rescue operation rather than the building as a whole. A fixed candidate evidence space preserves common semantics across nodes, while stage-conditioned activation allows each operation to use only factors that are physically relevant. Stage-specific assessment therefore remains comparable without forcing the same active factor set onto heterogeneous tasks.

The architecture separates within-node dependence from cross-node inheritance. A stage-specific BN represents the active personnel, equipment, environment, management, and technical-control evidence at each operation, whereas only eight physically persistent states are propagated to the next node. This selective inheritance preserves the information needed for rolling assessment while avoiding artificial temporal propagation of stage-specific variables whose physical meaning ends with the task.

Temporality is handled by continuous-time propagation at the boundary between operational nodes. Transition matrices are generated from stage-conditioned CTMC rates over the actual elapsed duration, so unequal node lengths do not require an arbitrary common time slice. Flexible sequences, skipped or repeated stages, and mid-stage reassessment can therefore be represented through the same interface. Time-locked evidence further distinguishes planned-window prediction, in-process filtering, and terminal review, making changes in *NSRI* traceable to inherited state, elapsed time, and newly observed evidence.

WPC-BN should be interpreted as a stage-conditioned modular dynamic Bayesian formulation rather than as an alternative to DBNs in principle. Static BNs are efficient for snapshot or single-task inference but do not carry state across operations. Conventional DBNs are preferable when slice definitions and cross-state dependencies can be specified consistently and supported by data. WPC-BN addresses the field setting in which the active evidence structure changes with the rescue task, node durations are irregular, and only a limited subset of states has a defensible transition model.

Under the factorization adopted for the eight persistent states, Appendix B shows that the cascading prediction-and-update recursion is numerically consistent with non-homogeneous DBN forward filtering implemented under the same assumptions. This equivalence clarifies the contribution: WPC-BN does not claim greater probabilistic expressiveness than a fully specified DBN; it provides an operational decomposition that couples stage activation, duration-aware propagation, and evidence governance in a form aligned with rescue decision points. Posterior

dependence created by shared within-node evidence defines the boundary of marginal propagation, and Supplementary Section S6.7 gives an exact-joint benchmark protocol for applications where that dependence warrants explicit evaluation.

Figure 12 summarizes this positioning and the corresponding applicability boundaries. Static BNs suit snapshot assessment without state inheritance; conventional DBNs suit processes with stable temporal structure and estimable inter-slice dependence; WPC-BN is intended for whole-process field assessment with heterogeneous stages, irregular durations, and selective state inheritance. Accordingly, *NSRI* is used for cross-node comparison, trend tracking, and diagnostic prioritization within a common parameter system, not as an empirical accident probability. Where cumulative damage, strongly coupled persistent states, or concurrent work fronts dominate, the same node-based framework can be extended through finer segmentation, history-dependent transitions, exact joint propagation, or a richer DBN structure.

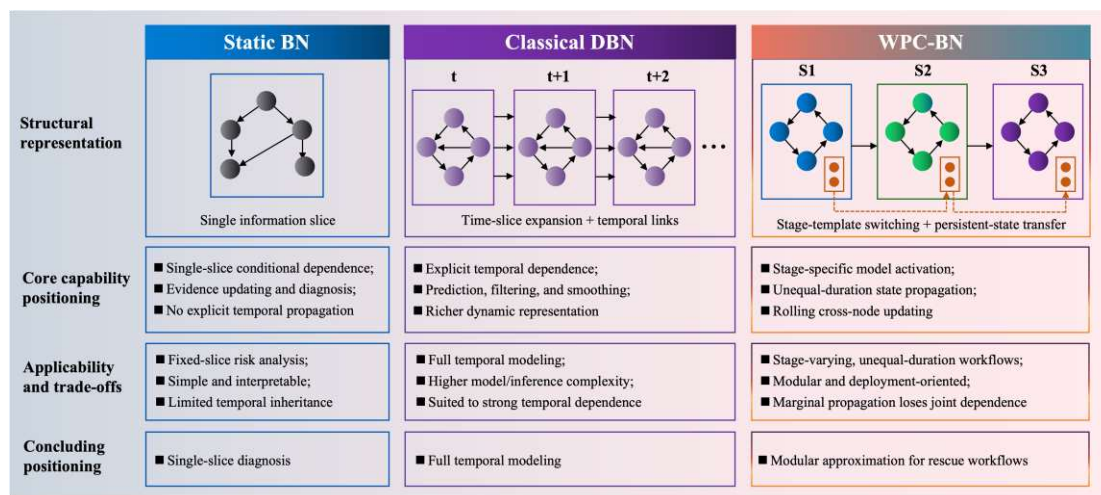


Fig. 12. Positioning and applicability boundaries of static BNs, conventional DBNs, and WPC-BN.

5.2 Engineering value and decision implications

Whole-process assessment enables safety resources to be positioned before high-risk operational windows. In the case study, M2 already had the highest pre-task *NSRI* before abnormal evidence emerged. Node-level prediction can therefore establish pre-operation priorities for shoring review, vibration control, evacuation readiness, and intensified monitoring rather than merely indicating whether the current state is unsafe.

Time-locked updating also converts new evidence into traceable control priorities. Once the anomaly emerged, the updated *NSRI* and the diagnostic ranking of X77, X29, and X31 formed a closed chain linking the risk window, evidence update, key factors, and control actions. This structure explains not only why intervention is warranted but also what should be addressed first.

Cross-node propagation supports reassessment after intervention. Following the M2 anomaly, work was paused, the situation was reviewed, shoring was reinforced, and risk declined at M3 and M4, showing that whole-process safety reflects both operational disturbance and control measures. At the current validation stage, the appropriate evaluation targets are identification of high-risk windows, direction of risk updates, factor interpretation, and post-intervention risk reduction rather than the final collapse/no-collapse outcome alone.

5.3 Transferability, evidence governance, and applicability boundaries

WPC-BN is transferable through framework reuse and parameter localization, not unconditional parameter portability. The S0–S4 templates, candidate-evidence interface, and activation rules can be reused across scenarios, whereas priors, CPTs, transition rates, state boundaries, and NSRI thresholds must be recalibrated for local building type, collapse pattern, aftershock environment, and team capability (Marcot et al., 2006; Kaikkonen et al., 2021). Pronounced cumulative damage, strong state dependence, or coupling among multiple work fronts requires a richer dynamic model.

Evidence governance should follow temporal information availability: pre-node information supports pre-task prediction, newly observed anomalies and monitoring data enter in-process filtering, and the terminal posterior is propagated to the next node. Original records, timestamps, evidence sources, and observation-to-indicator mappings should be preserved to prevent information leakage (Kaufman et al., 2012).

WPC-BN should operate within a layered decision architecture of hard constraints, probabilistic support, and field-command judgment (Fig. 13). Regulatory requirements, clear instability signs, and pause, evacuation, or reinforcement conditions identified by structural experts take precedence. Within that architecture, the model supports risk ranking, key-factor diagnosis, and comparison of operational alternatives.

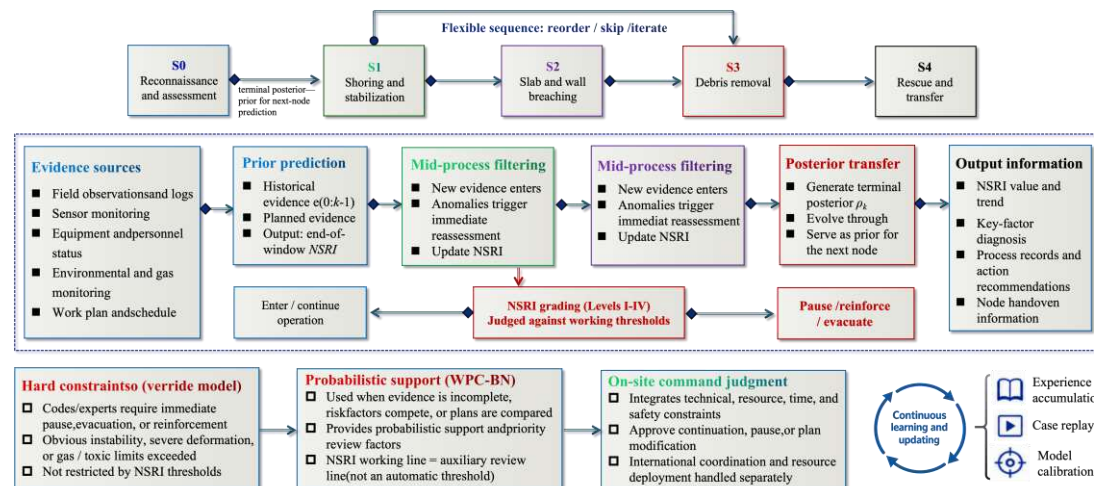


Fig. 13. Framework transfer, evidence stratification, and the layered decision architecture of hard constraints, probabilistic support, and field-command judgment.

5.4 Applicability boundaries and future research

The current evidence base supports WPC-BN as a comparative and diagnostic framework rather than an outcome-calibrated accident model. Indicator screening, internal numerical checks, standardized scenario ranking, and the retrospective Mandalay reconstruction establish content coverage, implementation consistency, mechanistic ordering, and process-level interpretability. The present NSRI decision bands therefore remain illustrative until sufficient node-level outcome data are available for empirical calibration.

The intended application domain is sequential collapsed-building rescue that can be represented as stage-conditioned operational nodes, with a limited set of persistent states whose evolution over each node is reasonably approximated by first-order, piecewise-constant transitions. Operations dominated by pronounced cumulative damage, abrupt regime changes, or tightly coupled concurrent work fronts lie outside the present factorized propagation boundary and

motivate finer node segmentation, piecewise transitions, exact joint propagation, or history-dependent and hierarchical extensions.

Parameter use is scenario-local rather than universally portable. Priors, CPT settings, transition rates, evidence mappings, and decision thresholds should be updated for building type, collapse pattern, aftershock environment, team capability, and monitoring configuration. Retrospective evidence remains appropriate for process reconstruction and mechanism checking; prospective performance evaluation should record predictions and evidence states before outcomes are known and should separate calibration data from validation cases wherever practicable.

Future validation can proceed in three stages: first, build a timestamped node-level database spanning multiple regions, building types, rescue teams, routine operations, near misses, and instability events; second, deploy WPC-BN prospectively in shadow mode so predictions are frozen in real time without influencing field command; and third, evaluate calibration, discrimination, threshold utility, and decision value once event definitions and sample sizes are adequate.

As evidence accumulates, expert-elicited parameters can be updated empirically and uncertainty propagated through priors, CPTs, and transition rates to obtain interval estimates of *NSRI*. Further work should also compare marginal and exact-joint propagation under denser shared evidence and examine history-dependent extensions for repeated disturbance. Any increase in model complexity should preserve the core properties established here: stage-specific relevance, evidence provenance, duration awareness, and auditable decision support.

6. Conclusions

Safety assessment for post-earthquake rescue has lacked a node-level dynamic framework that covers the operation end to end. WPC-BN was developed to address this gap and evaluated through internal tests and a retrospective case study. Three conclusions follow.

(1) The framework moves safety assessment to actual rescue-operation nodes and organizes four dimensions, 19 criteria, and 85 indicators in a unified candidate evidence space. Common, regional-background, and stage-specific layers give a consistent basis for cross-node comparison while still allowing task-specific activation and keeping field data entry manageable. Two-round Delphi screening and a 33-expert review offer initial support for content validity.

(2) WPC-BN couples stage-specific static BNs for within-node dependence with CTMC propagation of persistent states according to actual elapsed time, so that flexible sequences, skipped or repeated stages, and unequal-duration nodes can all be handled through one updating interface. Time-locking keeps node initiation, planned-window prediction, in-process updating, and terminal posteriors apart, so a one-off assessment becomes rolling comparison and reassessment.

(3) In the Mandalay case, the breaching node carried the highest *NSRI* both before the anomaly (2.55%) and during filtering (2.86%), and PRU and mutual information consistently prioritized breaching-induced vibration (X77), structural stability (X29), and rescue-passageway stability (X31). The location of the risk peak, the direction of updating, and the diagnosed factors were consistent with field records of abnormal sounds and tilting risk, with risk declining after the pause, review, and reinforcement. These findings support high-risk-window identification and

evidence-driven updating but constitute neither predictive-accuracy estimation nor external validation.

The central contribution is therefore not an absolute accident probability but a whole-process safety-assessment mechanism that is comparable across nodes, updatable as work proceeds, auditable, and reusable across stages. Further work should calibrate *NSRI* on independent multi-region, multi-outcome data, optimize decision thresholds, benchmark against full DBNs, quantify dependence-approximation error, propagate joint parameter uncertainty, and evaluate field usability.

Declaration of Competing Interest

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

CRedit authorship contribution statement

Tianyu Hu: Conceptualization, Methodology, Investigation, Visualization, Writing – original draft. **Defeng Xu:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing – original draft. **Yuan Li:** Methodology, Investigation, Visualization, Writing – original draft. **Feifei Sun:** Conceptualization, Methodology, Supervision, Funding acquisition. **Songhang He:** Methodology, Visualization, Data curation. **Zhipeng Shao:** Investigation, Data curation, Visualization, Writing – original draft.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 52378531).

Data availability

The datasets generated and analyzed during this study are available from the corresponding author (Email: xudefeng_2008@126.com) and the co-corresponding author (Email: ffsun@tongji.edu.cn) upon reasonable request.

References

- Abu-Zidan, F.M., Idris, K., Cevik, A.A., 2023. Prehospital management of earthquake crush injuries: A collective review. *Turkish Journal of Emergency Medicine* 23(4), 199–210. https://doi.org/10.4103/tjem.tjem_201_23.
- An, X., 2023. Seismic resilience assessment of healthcare buildings based on discrete-event simulation. Master's thesis, Institute of Engineering Mechanics, China Earthquake Administration. (in Chinese).
- Anagnostopoulos, S.A., Moretti, M.L., 2008. Post-earthquake emergency assessment of building damage, safety and usability—Part 1: Technical issues. *Soil Dynamics and Earthquake Engineering* 28(3), 223–232. <https://doi.org/10.1016/j.soildyn.2006.05.007>.
- Applied Technology Council, 1989. ATC-20: Procedures for Postearthquake Safety Evaluation of Buildings. Applied Technology Council, Redwood City, CA.
- Applied Technology Council, 2005. ATC-20-1 Field Manual: Postearthquake Safety Evaluation of Buildings, 2nd ed. Applied Technology Council, Redwood City, CA.

870 Bhardwaj, U., Teixeira, A.P., Guedes Soares, C., 2022. Bayesian framework for reliability
871 prediction of subsea processing systems accounting for influencing factors uncertainty.
872 Reliability Engineering & System Safety 218, 108143.
873 <https://doi.org/10.1016/j.ress.2021.108143>.

874 Bobbio, A., Portinale, L., Minichino, M., Ciancamerla, E., 2001. Improving the analysis of
875 dependable systems by mapping fault trees into Bayesian networks. Reliability Engineering
876 & System Safety 71(3), 249–260. [https://doi.org/10.1016/S0951-8320\(00\)00077-6](https://doi.org/10.1016/S0951-8320(00)00077-6).

877 Bruneau, M., Chang, S.E., Eguchi, R.T., et al., 2003. A framework to quantitatively assess and
878 enhance the seismic resilience of communities. Earthquake Spectra 19(4), 733–752.
879 <https://doi.org/10.1193/1.1623497>.

880 Cai, B.P., Liu, H.L., Xie, M., 2016. A real-time fault diagnosis methodology of complex systems
881 using object-oriented Bayesian networks. Mechanical Systems and Signal Processing 80, 31–
882 44. <https://doi.org/10.1016/j.ymssp.2016.04.019>.

883 Ceylan, B.O., Elidolu, G., Sezer, S.I., Akyuz, E., Yang, Z., 2026. Probabilistic risk assessment for
884 inert gas system on oil tanker ships using system theoretic accident model and process
885 (STAMP) and Bayesian belief network (BBN). Reliability Engineering & System Safety 266,
886 111669. <https://doi.org/10.1016/j.ress.2025.111669>.

887 Chang, Z., He, X., Fan, H., 2024. Application of Bayesian networks to risk assessment of LNG
888 maritime transport. Journal of Safety and Environment 24(12), 4541–4551. (in Chinese).

889 Chen, W., 2018. Assessment of urban earthquake emergency response capability based on fuzzy
890 analytic hierarchy process. Shanxi Architecture 44(14), 62–63. (in Chinese).

891 Chen, Y., Guan, H., Li, X., 2025. Application of dynamic Bayesian networks to fire risk
892 assessment in cable compartments of utility tunnels. Journal of Safety and Environment
893 25(1), 11–20. (in Chinese).

894 Cheng, F., Jiang, M., Su, C., et al., 2024. Risk analysis of hydrogen refueling station leakage and
895 explosion accidents based on dynamic Bayesian networks. Journal of Safety and
896 Environment 24(9), 3405–3417. (in Chinese).

897 China News Service, 2025. Chinese rescue teams rescue a man trapped for more than 120 hours in
898 the Myanmar earthquake disaster area. 2 April 2025.
899 <https://www.chinanews.com.cn/gj/2025/04-02/10393757.shtml> (accessed 8 August 2026). (in
900 Chinese).

901 Cimellaro, G.P., Malavisi, M., Mahin, S.A., 2018. Factor analysis to evaluate hospital resilience.
902 ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil
903 Engineering 4(1), 04018002. <https://doi.org/10.1061/AJRUA6.0000952>.

904 Comfort, L.K., Ko, K., Zagorecki, A., 2004. Coordination in rapidly evolving disaster response
905 systems. American Behavioral Scientist 48(3), 295–313.
906 <https://doi.org/10.1177/0002764204268987>.

907 Comfort, L.K., 2007. Crisis management in hindsight: Cognition, communication, coordination,
908 and control. Public Administration Review 67(S1), 189–197. <https://doi.org/10.1111/j.1540-6210.2007.00827.x>.

910 Darwiche, A., 2001. Constant-space reasoning in dynamic Bayesian networks. International
911 Journal of Approximate Reasoning 26(3), 161–178. [https://doi.org/10.1016/S0888-613X\(00\)00067-0](https://doi.org/10.1016/S0888-613X(00)00067-0).

912

913 Federal Emergency Management Agency (FEMA), 2019. Post-Disaster Building Safety
 914 Evaluation Guidance: Report on the Current State of Practice, Including Recommendations
 915 Related to Structural and Nonstructural Safety and Habitability. FEMA P-2055. Federal
 916 Emergency Management Agency, Washington, DC.

917 Ferdous, R., Khan, F., Sadiq, R., Amyotte, P., Veitch, B., 2011. Fault and event tree analyses for
 918 process systems risk analysis: Uncertainty handling formulations. *Risk Analysis* 31(1), 86–
 919 107. <https://doi.org/10.1111/j.1539-6924.2010.01475.x>.

920 Fiedrich, F., Gehbauer, F., Rickers, U., 2000. Optimized resource allocation for emergency
 921 response after earthquake disasters. *Safety Science* 35(1–3), 41–57.
 922 [https://doi.org/10.1016/S0925-7535\(00\)00021-7](https://doi.org/10.1016/S0925-7535(00)00021-7).

923 General Administration of Quality Supervision, Inspection and Quarantine of the People's
 924 Republic of China, Standardization Administration of China, 2012. GB/T 29428.1-2012:
 925 Rescue operations of earthquake emergency rescue teams—Part 1: Basic requirements. China
 926 Standards Press, Beijing. (in Chinese).

927 General Administration of Quality Supervision, Inspection and Quarantine of the People's
 928 Republic of China, Standardization Administration of China, 2014. GB/T 29428.2-2014:
 929 Rescue operations of earthquake emergency rescue teams—Part 2: Procedures and methods.
 930 China Standards Press, Beijing. (in Chinese).

931 George, P.G., Renjith, V.R., 2021. Evolution of safety and security risk assessment methodologies
 932 towards the use of Bayesian networks in process industries. *Process Safety and*
 933 *Environmental Protection* 149, 758–775. <https://doi.org/10.1016/j.psep.2021.03.031>.

934 German, S., Jeon, J.-S., Zhu, Z., Bearman, C., Brilakis, I., DesRoches, R., Lowes, L., 2013.
 935 Machine vision-enhanced postearthquake inspection. *Journal of Computing in Civil*
 936 *Engineering* 27(6), 622–634. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000333](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000333).

937 Goldberg, D.E., Yeck, W.L., Hanagan, C.E., et al., 2025. Ultralong, supershear rupture of the
 938 2025 Mw 7.7 Mandalay earthquake reveals unaccounted risk. *Science* 390(6772), 458–462.
 939 <https://doi.org/10.1126/science.ady3581>.

940 Henan Provincial Administration for Market Regulation, 2024. DB41/T 2600-2024: Requirements
 941 for construction of technical systems for earthquake emergency command. Issued 1 February
 942 2024; effective 1 May 2024. (in Chinese).

943 Jones, B., Jenkinson, I., Yang, Z., Wang, J., 2010. The use of Bayesian network modelling for
 944 maintenance planning in a manufacturing industry. *Reliability Engineering & System Safety*
 945 95(3), 267–277. <https://doi.org/10.1016/j.ress.2009.10.007>.

946 INSARAG, 2020. INSARAG Guidelines. Volume II, Manual B: Operations; Volume III:
 947 Operational Field Guide. United Nations Office for the Coordination of Humanitarian
 948 Affairs, Geneva.

949 INSARAG, 2026. INSARAG Guidelines, new ed. Volume 3: Operations; Volume 5: Operational
 950 Field Guide. United Nations Office for the Coordination of Humanitarian Affairs, Geneva.
 951 Launched 9 July 2026; effective 1 January 2027.

952 Kabir, S., Papadopoulos, Y., 2019. Applications of Bayesian networks and Petri nets in safety,
 953 reliability, and risk assessments: A review. *Safety Science* 115, 154–175.
 954 <https://doi.org/10.1016/j.ssci.2019.02.009>.

955 Kaikkonen, L., Parviainen, T., Rahikainen, M., Uusitalo, L., Lehtikoinen, A., 2021. Bayesian
 956 networks in environmental risk assessment: A review. *Integrated Environmental Assessment
 957 and Management* 17(1), 62–78. <https://doi.org/10.1002/ieam.4332>.
 958 Kaufman, S., Rosset, S., Perlich, C., Stitelman, O., 2012. Leakage in data mining: Formulation,
 959 detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data* 6(4),
 960 Article 15, 1–21. <https://doi.org/10.1145/2382577.2382579>.
 961 Khakzad, N., Khan, F., Amyotte, P., 2011. Safety analysis in process facilities: Comparison of
 962 fault tree and Bayesian network approaches. *Reliability Engineering & System Safety* 96(8),
 963 925–932. <https://doi.org/10.1016/j.res.2011.03.012>.
 964 Khakzad, N., Khan, F., Amyotte, P., 2013. Dynamic safety analysis of process systems by
 965 mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection* 91(1–
 966 2), 46–53. <https://doi.org/10.1016/j.psep.2012.01.005>.
 967 Khakzad, N., 2015. Application of dynamic Bayesian network to risk analysis of domino effects in
 968 chemical infrastructures. *Reliability Engineering & System Safety* 138, 263–272.
 969 <https://doi.org/10.1016/j.res.2015.02.007>.
 970 Kjærulff, U., 1995. dHugin: A computational system for dynamic time-sliced Bayesian networks.
 971 *International Journal of Forecasting* 11(1), 89–111. [https://doi.org/10.1016/0169-
 972 2070\(94\)02003-8](https://doi.org/10.1016/0169-).
 973 Lawshe, C.H., 1975. A quantitative approach to content validity. *Personnel Psychology* 28(4),
 974 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>.
 975 Marcot, B.G., Steventon, J.D., Sutherland, G.D., McCann, R.K., 2006. Guidelines for developing
 976 and updating Bayesian belief networks applied to ecological modeling and conservation.
 977 *Canadian Journal of Forest Research* 36(12), 3063–3074. <https://doi.org/10.1139/x06-135>.
 978 Marshall, J., Barnes, J., Gould, N., Jaiswal, K., Lizundia, B., Swanson, D.A., Turner, F., 2012.
 979 Post-earthquake building safety assessments for the Canterbury earthquakes. In: *Structures
 980 Congress 2012*, pp. 1057–1068. <https://doi.org/10.1061/9780784412367.094>.
 981 Murphy, K.P., 2002. *Dynamic Bayesian Networks: Representation, Inference and Learning*. Ph.D.
 982 dissertation, University of California, Berkeley.
 983 Murphy, R.R., 2004. Human-robot interaction in rescue robotics. *IEEE Transactions on Systems,
 984 Man, and Cybernetics, Part C (Applications and Reviews)* 34(2), 138–153.
 985 National Fire Protection Association (NFPA), 2020. NFPA 1521: Standard for Fire Department
 986 Safety Officer Professional Qualifications. National Fire Protection Association, Quincy,
 987 MA.
 988 National Fire Protection Association (NFPA), 2022. NFPA 2500: Standard for Operations and
 989 Training for Technical Search and Rescue Incidents and Life Safety Rope and Equipment for
 990 Emergency Services. National Fire Protection Association, Quincy, MA.
 991 National Fire Protection Association (NFPA), 2024. NFPA 1550: Standard for Emergency
 992 Responder Health and Safety. National Fire Protection Association, Quincy, MA.
 993 Nodelman, U., Shelton, C.R., Koller, D., 2002. Continuous time Bayesian networks. In:
 994 *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence (UAI 2002)*, pp.
 995 378–387.
 996 Nodelman, U., Shelton, C.R., Koller, D., 2003. Learning continuous time Bayesian networks. In:
 997 *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence (UAI 2003)*, pp.
 998 451–458.

999 Occupational Safety and Health Administration (OSHA), n.d. Structural Collapse Guide. U.S.
1000 Department of Labor. [https://www.osha.gov/emergency-preparedness/guides/structural-](https://www.osha.gov/emergency-preparedness/guides/structural-collapse)
1001 collapse (accessed 19 July 2026).

1002 Onisawa, T., 1988. An approach to human reliability in man-machine systems using error
1003 possibility. *Fuzzy Sets and Systems* 27(2), 87–103. [https://doi.org/10.1016/0165-](https://doi.org/10.1016/0165-0114(88)90140-6)
1004 0114(88)90140-6.

1005 Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.
1006 Morgan Kaufmann, San Mateo, CA.

1007 Qing, Y., Ming, D., Wen, Q., Weng, Q., Xu, L., Chen, Y., Zhang, Y., Zeng, B., 2022. Operational
1008 earthquake-induced building damage assessment using CNN-based direct remote sensing
1009 change detection on superpixel level. *International Journal of Applied Earth Observation and*
1010 *Geoinformation* 112, 102899. <https://doi.org/10.1016/j.jag.2022.102899>.

1011 Raghunandan, M., Liel, A.B., Luco, N., 2015. Aftershock collapse vulnerability assessment of
1012 reinforced concrete frame structures. *Earthquake Engineering & Structural Dynamics* 44(3),
1013 419–439. <https://doi.org/10.1002/eqe.2478>.

1014 Roohi, M., Hernandez, E.M., 2020. Performance-based post-earthquake decision making for
1015 instrumented buildings. *Journal of Civil Structural Health Monitoring* 10(5), 775–792.
1016 <https://doi.org/10.1007/s13349-020-00416-1>.

1017 Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Technical Journal*
1018 27(3), 379–423; 27(4), 623–656.

1019 State Administration for Market Regulation, Standardization Administration of China, 2025.
1020 GB/T 46578-2025: Basic requirements for the capacity of earthquake rescue team. Issued 31
1021 October 2025; effective 1 February 2026. (in Chinese).

1022 Statheropoulos, M., Agapiou, A., Pallis, G.C., et al., 2015. Factors that affect rescue time in urban
1023 search and rescue (USAR) operations. *Natural Hazards* 75(1), 57–69.
1024 <https://doi.org/10.1007/s11069-014-1304-3>.

1025 Tadokoro, S. (Ed.), 2009. Rescue Robotics: DDT Project on Robots and Systems for Urban
1026 Search and Rescue. Springer, London. <https://doi.org/10.1007/978-1-84882-474-4>.

1027 Usuda, D., Shimozaawa, S., Takami, H., et al., 2023. Crush syndrome: A review for prehospital
1028 providers and emergency clinicians. *Journal of Translational Medicine* 21, 584.
1029 <https://doi.org/10.1186/s12967-023-04416-9>.

1030 Villa, V., Paltrinieri, N., Khan, F., Cozzani, V., 2016. Towards dynamic risk analysis: A review of
1031 the risk assessment approach and its limitations in the chemical process industry. *Safety*
1032 *Science* 89, 77–93. <https://doi.org/10.1016/j.ssci.2016.06.002>.

1033 Weber, P., Medina-Oliva, G., Simon, C., Iung, B., 2012. Overview on Bayesian networks
1034 applications for dependability, risk analysis and maintenance areas. *Engineering Applications*
1035 *of Artificial Intelligence* 25(4), 671–682. <https://doi.org/10.1016/j.engappai.2010.06.002>.

1036 Xiong, C., Li, Q., Lu, X., 2020. Automated regional seismic damage assessment of buildings
1037 using an unmanned aerial vehicle and a convolutional neural network. *Automation in*
1038 *Construction* 109, 102994. <https://doi.org/10.1016/j.autcon.2019.102994>.

1039 Xu, D., Sun, F., Zhao, H., Qi, H., Chen, L., Hu, T., 2024. Study on secondary collapse of the
1040 bottom frame masonry structure in semi-ruined state based on FEM–FDEM. *Scientific*
1041 *Reports* 14, 29405. <https://doi.org/10.1038/s41598-024-80612-x>.

- Xu, D., Sun, F., Chen, X., Qi, H., Hu, T., Yin, W., 2026a. Early-warning approach for secondary collapse of RC frame structure in semi-ruined state under aftershocks based on K-LSTM model. *Journal of Building Engineering* 117, 114497. <https://doi.org/10.1016/j.jobbe.2025.114497>.
- Xu, D., Sun, F., Gao, Y., Xu, Q., Hu, T., 2026b. Early warning model for secondary collapse of semi-ruined reinforced concrete frame structures under aftershocks. *Computer-Aided Civil and Infrastructure Engineering* 41, 100001. <https://doi.org/10.1016/j.cacaie.2026.100001>.
- Xu, D., Sun, F., Liu, Y., Wen, J., 2026c. Experimental and numerical study on real-time early warning of secondary collapse in semi-ruined RC frame structures under aftershocks. *Bulletin of Earthquake Engineering* 24(5), 3273–3300. <https://doi.org/10.1007/s10518-026-02399-8>.
- Yazdi, M., Kabir, S., 2017. A fuzzy Bayesian network approach for risk analysis in process industries. *Process Safety and Environmental Protection* 111, 507–519. <https://doi.org/10.1016/j.psep.2017.08.015>.
- YJ/T 8-2016, 2016. Training guide for earthquake disaster emergency rescue team members. Issued 14 March 2016; effective 1 October 2016. (in Chinese).
- DB/T 63-2016, 2016. Training guide for earthquake emergency rescue team members. China Earthquake Administration, Beijing. Issued 14 March 2016; effective 1 October 2016. (in Chinese).
- Zadeh, L.A., 1965. Fuzzy sets. *Information and Control* 8(3), 338–353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).
- Zarei, E., Khakzad, N., Cozzani, V., Reniers, G., 2019. Safety analysis of process systems using Fuzzy Bayesian Network (FBN). *Journal of Loss Prevention in the Process Industries* 57, 7–16. <https://doi.org/10.1016/j.jlp.2018.10.011>.
- Zhang, L., Wu, X., Skibniewski, M.J., Zhong, J., Lu, Y., 2014. Bayesian-network-based safety risk analysis in construction projects. *Reliability Engineering & System Safety* 131, 29–39. <https://doi.org/10.1016/j.ress.2014.06.006>.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendices.pdf](#)
- [SupplementaryMaterials.pdf](#)