**Structural Engineering and Mechanics,** *Vol. 97, No. 6 (2026) 829-850*
https://doi.org/10.12989/sem.2026.97.6.829

Techno-Press

# Network screening of derailment consequence severity using exposure proxies: Gyeongbu High-Speed Line

Jaehoon Lim[1a], Seong Keun Kang[1], Kyumin Na[2],
Hyunkyu Jun[1], Chanwoo Park[1], Jung Sik Kong[*3]

[1]*System Safety Research Department, Korea Railroad Research Institute, Uiwang-si, Gyenggi-do, 16105, Korea*
[2]*Department of Mechanical Design Engineering, Pukyong National University, Busan, 48547, Korea*
[3]*School of Civil, Environmental, and Architectural Engineering, Korea University, Seoul, 02841, Korea*

*(Received February 20, 2026, Revised March 5, 2026, Accepted March 11, 2026)*

**Abstract.** This study presents a deterministic analytic protocol for consequence-severity screening that focuses on evaluating the behaviour of screening methods rather than on probabilistic risk estimation. The protocol represents escalation potential using interpretable structural context descriptors and geospatial exposure proxies, and evaluates screening outputs using rank-stability and screening-efficiency metrics. The analytic metrics include Spearman rank correlation, top-k overlap (Jaccard similarity), grade reclassification rates, and candidate inflation to quantify how modelling choices and baseline definitions expand or contract prioritised candidate sets. Corridor-scale decision utility is reported using contiguous high-grade clusters rather than isolated section scores. The protocol is demonstrated on the Gyeongbu High-Speed Line (Republic of Korea) under alternative segmentation resolutions (500 m and 1 km) and grading variants. Results show that combined structural-exposure screening yields more compact candidate sets than structural-only or exposure-only baselines while remaining robust to key analytical choices. It is intended for prioritisation and does not estimate derailment probability, calibrate losses, or constitute a full risk assessment.

**Keywords:** derailment consequences; geospatial proxies; high-speed rail; infrastructure vulnerability; network screening; railway safety; risk management

## 1. Introduction

Railway derailments are relatively infrequent, yet their consequences can be severe and are strongly conditioned by local infrastructure configuration and surrounding exposure [1-3]. Beyond direct damage, major events may involve prolonged disruption and impacts to passengers and third parties, keeping derailment risk management a continuing concern in railway safety engineering and regulation [2-4].

Conventional safety assessment largely emphasises frequency- or probability-based indicators [5], which are indispensable for understanding initiation and monitoring performance but provide limited insight into consequence severity conditional on derailment. At network scale, this can

∗Corresponding author, Professor, E-mail: jskong@korea.ac.kr
[a] Ph.D., E-mail: jhlim24@krri.re.kr


http://www.techno-press.org/sem/5
ISSN: 1225-4568 (Print), 1598-6217 (Online)

obscure locations where likelihood appears low while escalation potential is elevated due to confinement, structural interaction, or nearby exposure, consistent with risk perspectives that treat likelihood and consequence as distinct decision-support dimensions [6-8].

A key practical requirement is that corridor-scale consequence assessment must remain transparent, reproducible, and feasible under heterogeneous data availability. Detailed dynamic simulation or calibrated loss estimation is rarely viable for routine network screening, and externally imposed thresholds are difficult to justify consistently across jurisdictions and data regimes [9-11]. A structured screening approach that preserves interpretable links to escalation mechanisms while operating on routinely available geospatial information is therefore valuable [12, 13].

This paper proposes a deterministic, consequence-oriented framework for network-level screening of derailment severity potential. Using interpretable structural descriptors (e.g., tunnels and bridges) and geospatial exposure proxies (e.g., built-up context, sensitive receptors, and surface water proximity), the framework grades sections in relative terms to support prioritisation as a complement to frequency-oriented assessment rather than a replacement for likelihood modelling [2, 8]. The contribution is threefold: (i) a transparent deterministic grading structure for corridor-scale escalation potential, (ii) an analytic protocol to evaluate screening behaviour under key analytical choices (segmentation resolution, grading variant, and baseline definition) using rank agreement, top-k overlap, grade reclassification, and candidate inflation metrics, and (iii) decision-ready corridor outputs that summarise contiguous high-grade clusters for directing follow-on investigations and interventions.

The remainder of the paper is organised as follows: Section 2 reviews related work; Section 3 presents the framework and grading logic; Section 4 describes the case study and data preparation; Section 5 reports results; Section 6 discusses implications and limitations; and Section 7 concludes.

## 2. Background and related work

### 2.1 Likelihood-consequence separation in risk and reliability assessment

Risk and reliability science commonly treats likelihood and consequence as distinct analytical dimensions because they support different decisions and rest on different evidence bases [6-9, 14, 15]. In early-stage screening, limited data and network heterogeneity often preclude consistent probabilistic modelling, so consequence-focused characterisation can guide where detailed follow-up should be concentrated without implying a full probabilistic risk estimate [10, 11]. At network scale, probability-oriented models typically require strong and non-uniform assumptions about exposure and transferability, introducing uncertainties that vary with purpose and evidence availability [8, 11, 16]. For this reason, transparent and reproducible screening layers remain a common preparatory step when scope and limitations are stated explicitly [8, 11].

### 2.2 Railway derailment assessment and frequency-oriented approaches

Derailment assessment has primarily focused on occurrence modelling and frequency estimation using historical records, track-condition indicators, operating characteristics, and traffic composition [1, 5, 17]. These methods are valuable for understanding initiation contributors and monitoring safety performance, and they are embedded in regulatory and industry practice [3, 18], although broader reviews note persistent challenges due to heterogeneous data and reporting [19].

However, frequency-oriented indicators do not directly address consequence severity conditional on derailment. For corridor prioritisation, this is limiting because sections with low apparent frequency may still exhibit high escalation potential due to confinement, rigid-structure interaction, or proximity to exposed receptors. Industry guidance recognises the need to consider consequence-relevant context alongside likelihood, yet operationalising this consistently at corridor scale remains challenging [2, 18]. Meanwhile, probabilistic derailment studies continue to evolve for specific contexts (e.g., yards or hazardous materials transport), reinforcing the value of keeping probability-oriented modelling separate from consequence-focused screening when comparability and interpretability across sections are required [20-23].

### 2.3 Derailment consequences and escalation mechanisms

Evidence from derailment investigations and studies indicates that severity is often governed by contextual escalation mechanisms rather than the initiating failure mode alone. Reported contributors include confinement in tunnels and constrained alignments, interactions with rigid structural elements (e.g., bridge structures and parapets), and proximity to population or environmentally sensitive receptors [2, 12, 23]. These mechanisms imply that consequence potential can vary systematically along a network and can be represented using infrastructure context and surrounding exposure descriptors; this is also consistent with broader dynamics literature on train-track-bridge interaction under abnormal conditions [24]. Quantifying consequences in absolute terms (e.g., monetary loss, casualties, monetised environmental damage) remains difficult due to transferability and valuation assumptions, sensitivity to local context, and data completeness, which can undermine corridor-wide comparability without extensive calibration [9, 10, 14]. Given additional disruption and indirect-cost components, relative grading can therefore be a pragmatic option for early-stage network screening when the grading logic is explicit and interpretable [25].

### 2.4 Network-level screening and positioning of the present study

Screening and classification are widely used in infrastructure risk management to prioritise attention rather than to predict outcomes precisely, translating heterogeneous information into ordered categories that support comparison and staged decision-making [8, 11]. In railways, consequence-oriented screening has been developed less systematically than frequency-based assessment, and practice often embeds consequence implicitly within composite indices or relies on site-specific studies, limiting transparency and corridor-wide comparability [2, 18].

This study proposes a deterministic, consequence-oriented screening framework that makes the consequence dimension explicit and operational at network level. Interpretable structural descriptors (e.g., tunnels and bridges) and geospatial exposure proxies (e.g., built-up context, sensitive receptors, surface water proximity) are used to support reproducible corridor-wide grading. The framework is positioned as an early-stage prioritisation input: it does not estimate derailment probability, simulate accident dynamics, or provide calibrated loss predictions; instead, it supports transparent comparison of relative escalation potential and produces corridor-level outputs suitable for directing subsequent detailed analyses and interventions [6, 8, 11].

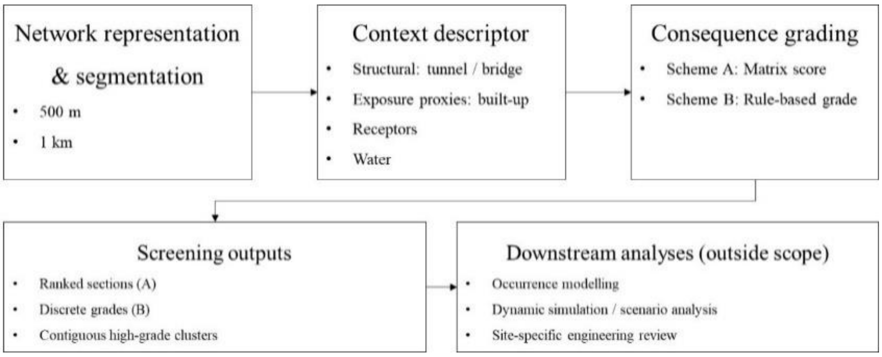



Figure 1. Position of the consequence-oriented screening framework within a reliability and risk assessment workflow

## 3. Consequence-oriented screening framework

### 3.1 Framework overview and role within a reliability workflow

The proposed framework is a deterministic screening tool to characterise relative derailment consequence potential at network scale. Its position within a reliability and risk assessment workflow is shown in Fig. 1. The framework supports early-stage prioritisation by organising corridor sections according to contextual escalation potential before occurrence modelling, probabilistic reliability analysis, or dynamic simulation [6, 8, 11]. It does not estimate derailment probability, reconstruct post-derailment dynamics, or provide calibrated loss estimates; instead, it focuses on infrastructure and exposure contexts that may condition post-derailment severity, enabling transparent comparison where data availability is uneven and calibration is typically infeasible [9, 10].

### 3.2 Analytical units and network representation

For network-level screening, the corridor is represented using a single representative alignment as the spatial reference for processing and aggregation. The alignment is segmented into fixed-length analytical sections that serve as comparable units; two resolutions (500 m and 1 km) are used to examine sensitivity to analytical scale within an otherwise consistent pipeline. Fixed-length segmentation is adopted for reproducibility and corridor-wide consistency and is not intended to represent the physical extent of derailment consequences. Where discrete structural contexts are present (notably tunnels and bridges), segmentation is handled using minimum-split processing to avoid diluting context across section boundaries and to preserve localised conditions relevant to consequence escalation.

### 3.3 Infrastructure and environmental context descriptors

Consequence escalation potential is represented using a limited set of interpretable descriptors capturing (i) structural context and (ii) surrounding exposure context. Structural context is represented using binary indicators for tunnels and bridges: tunnel presence reflects a confined setting in which post-derailment motion may be constrained and interaction with rigid boundaries

Table 1. Context descriptors and proxy indicators used for consequence grading

| Domain | Variable | Symbol | Type | Section-level definition | Role in grading |
|---|---|---|---|---|---|
| Structural | Tunnel presence | $I_{tu}(s)$ | Binary | $I_{tu}(s) = 1$ if section $s$ intersects a tunnel feature; else 0 | Input to Scheme A/B |
| | Bridge presence | $I_{br}(s)$ | Binary | $I_{br}(s) = 1$ if section $s$ intersects a bridge feature; else 0 | Input to Scheme A/B |
| Exposure | Built-up ratio (within buffer) | $R_b(s)$ | Continuous | $R_b(s) = A_{bldg}(s)/A_{buf}(s)$, where $A_{bldg}(s)$ is total building footprint area within the 500 m buffer of section $s$, and $A_{buf}(s)$ is the buffer area | Defines built-up high flag $U(s)$ |
| | Receptor count (within buffer) | $N_r(s)$ | Count | $N_r(s)$ = number of selected receptor point features within the 500 m buffer of section $s$ | Defines receptor-high flag $R(s)$ |
| | Surface water proximity | $I_w(s)$ | Binary | $I_w(s) = 1$ if mapped water features intersect the 500 m buffer of section $s$; else 0 | Defines water-high flag $W(s)$ |
| Derived | Built-up high flag | $U(s)$ | Binary | $U(s) = 1[R_b(s) \geq \tau_b]$, where $\tau_b$ is the corridor-internal percentile threshold (base setting: 70th percentile), computed separately for each segmentation resolution | Scheme B input (via $H(s)$) |
| | Receptor-high flag | $R(s)$ | Binary | $R(s) = 1[N_r(s) \geq 1]$ (presence criterion) | Scheme B input (via $H(s)$) |
| | Human/built exposure | $H(s)$ | Binary | $H(s) = U(s) \vee R(s)$ | Scheme B input |
| | Water exposure | $Q(s)$ | Binary | $Q(s) = W(s) = I_w(s)$ | Scheme B input |

The table summarises the structural descriptors (tunnel and bridge presence) and exposure proxies (built-up context, sensitive receptors, and surface water proximity) used as inputs to the grading schemes (A and B). Section-level proxy computation is described in Appendix A, while the derivation of binary flags and the deterministic grading rules are detailed in Appendix B.

may be more likely, while bridge presence reflects elevated potential for structural interaction and restricted lateral dispersion. These indicators describe context only and do not imply derailment initiation mechanisms or causality [2]. Environmental context is represented using proxy indicators derived from routinely available geospatial layers beyond the railway boundary, characterising built-up context, sensitive receptors, and proximity to surface water; they support relative comparison at screening scale and are not calibrated measures of damage magnitude or impact [12]. Operational definitions, thresholds, and aggregation rules are specified in Section 4 and Appendix A to support reproducibility and to bound interpretation.

### 3.4 Deterministic grading logic and scoring variants

At corridor scale, screening requires a grading structure that is transparent and reproducible under heterogeneous data availability. The present study therefore adopts deterministic grading logic rather than calibrated loss models or scenario-based dynamic simulation, avoiding numerical weights that would require external calibration and may not transfer consistently across the

Table 2. Deterministic grading rules for Scheme B based on structural and exposure context indicators

| Rule ID | Structural condition | Exposure condition | Assigned ($G_B$) grade | Interpretation |
|---|---|---|---|---|
| 1 | $I_{tu} = 1$ AND $I_{br} = 1$ | Any exposure state $(H, Q) \in \{0,1\}^2$ | Grade 4 | Compound structural context (tunnel + bridge) |
| 2 | $I_{tu} = 1$ AND $I_{br} = 0$ | $H = 1$ AND $Q = 1$ | Grade 4 | Tunnel with both exposure dimensions present |
| 3 | $I_{tu} = 1$ AND $I_{br} = 0$ | $(H, Q) \in \{(1,0), (0,1)\}$ | Grade 3 | Tunnel with a single exposure dimension present |
| 4 | $I_{tu} = 0$ AND $I_{br} = 1$ | $H = 1$ AND $Q = 1$ | Grade 3 | Bridge with both exposure dimensions present |
| 5 | $I_{tu} = 0$ AND $I_{br} = 1$ | $(H, Q) \in \{(1,0), (0,1)\}$ | Grade 2 | Bridge with a single exposure dimension present |
| 6 | $I_{tu} = 0$ AND $I_{br} = 0$ | $H = 1$ AND $Q = 1$ | Grade 2 | Open track with both exposure dimensions present |
| 7 | Otherwise | Otherwise | Grade 1 | Low consequence context under screening assumptions |

$I_{tu}$ and $I_{br}$ denote tunnel and bridge presence indicators, respectively. $H$ is the human/built exposure flag and $Q$ is the water exposure flag (definitions in Table 1/Appendix A-B).

corridor [11, 14].

Two grading variants are used:

- Scheme A (matrix score): an ordinal/continuous score derived from combinations of structural and exposure descriptors, used for ranking and rank-based robustness checks.
- Scheme B (rule-based grade): a deterministic rule set that maps descriptor combinations to discrete consequence grades, used for decision-facing screening and prioritisation.

In Scheme B, grades are assigned from the structural context indicators $\{I_{tu}, I_{br}\}$ and exposure context indicators $\{H, Q\}$ defined in Sections 3.3 and 4.3. The fixed mapping used in this study is summarised in Table 2 and applied uniformly across the corridor.

Scheme B is intentionally conservative for screening under heterogeneous—and potentially incomplete—proxy layers; in particular, the tunnel-bridge combination is treated as a worst-case structural context (Rule 1) irrespective of exposure flags, because the absence of mapped exposure proxies does not necessarily imply the absence of exposure. The influence of conditioning this rule on exposure is examined via the limited what-if check reported in Section 6.2 (Table 8).

Both outputs are interpreted as relative indicators within the analysed network rather than absolute severity levels.

### 3.5 Interpretation of screening outputs

The screening outputs are a ranked ordering (Scheme A) and/or discrete grades (Scheme B) assigned to predefined analytical sections (Fig. 2). The outputs represent relative consequence escalation potential within the analysed corridor and should not be interpreted as expected loss estimates or as evidence that derailments will occur at the identified locations [8].

For corridor-level prioritisation, decision-making typically concerns continuous stretches rather than isolated points; accordingly, elevated sections should be interpreted with their spatial continuity, and downstream analyses—such as probabilistic occurrence modelling, dynamic

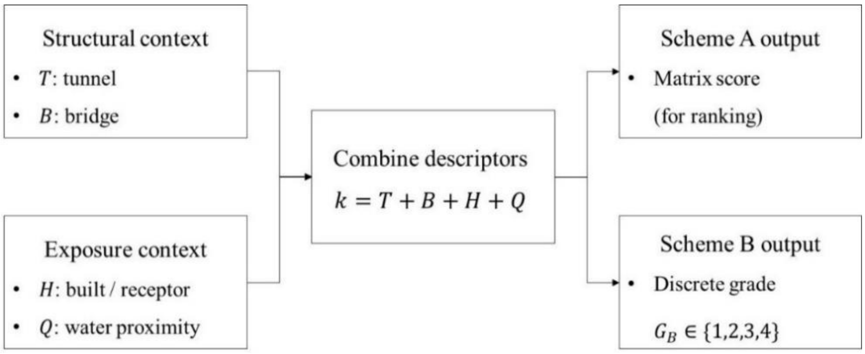



Figure 2. Schematic illustration of consequence grade assignment from combined structural and exposure descriptors

simulation, or site-specific engineering assessment—can be directed to prioritised segments.

## 4. Case study and data preparation

### 4.1 Case study overview

The case study considers the Gyeongbu High-Speed Line as a corridor-scale testbed for consequence-oriented screening. For network-level processing, the line is represented using a single representative alignment that serves as a consistent spatial reference for all spatial aggregation and comparison steps. The representation is intentionally simplified for screening purposes and does not seek to reproduce fine-scale track geometry or operational variation.

The representative alignment is segmented into fixed-length analytical sections that form the basic units of comparison. Two segmentation resolutions are considered: 500 m and 1 km. These resolutions are used to examine sensitivity to analytical scale while maintaining an otherwise consistent processing pipeline across the corridor. Section lengths are defined for analytical convenience and interpretability, rather than to represent derailment extents or accident-specific impact zones.

To preserve discrete structural contexts, segmentation is combined with a minimum-split handling. Where tunnel or bridge features intersect the representative alignment, section boundaries are adjusted to avoid artificial fragmentation of such features across multiple analytical sections. Detailed preprocessing procedures are provided in Appendix A.

### 4.2 Infrastructure data and structural classification

Infrastructure configuration is characterised using binary structural context indicators for tunnels and bridges. For each analytical section s, tunnel presence is represented by $I_{tu}(s) \in \{0,1\}$, and bridge presence is represented by $I_{br}(s) \in \{0,1\}$, assigned by spatial intersection between the section geometry and the corresponding infrastructure layers.

Tunnel presence is treated as a confinement context in which post-derailment motion may be constrained and interaction with rigid boundaries may be more likely. Bridge presence is treated as a context with elevated potential for structural interaction and restricted lateral dispersion [2].

These indicators are used solely to characterise contextual escalation potential at screening stage; they do not imply derailment initiation mechanisms, failure causes, or dynamic behaviours.

### 4.3 Environmental exposure proxies

Environmental context is characterised using spatial proxy indicators intended to represent exposure beyond the railway boundary. A fixed buffer distance of 500 m is applied to each analytical section, and environmental features within the buffer are aggregated to the section level using consistent rules applied uniformly across the corridor. The 500-m buffer is used as a pragmatic working scale for corridor-level screening, balancing spatial resolution and interpretability against the granularity of the available proxy layers. The selected distance is not intended to represent the physical extent of derailment consequences. It is used to define a consistent neighbourhood for comparative prioritisation along the corridor.

Three exposure categories are considered: surrounding built-up context derived from building footprints, sensitive receptors represented by selected points of interest, and proximity to surface water bodies represented by mapped water features. The proxy layers are derived from openly available geospatial sources. The proxies are used as relative descriptors for comparative screening; they are not calibrated measures of damage magnitude and do not imply causal impact.

Built-up context for section $s$ is represented by the building-area ratio within the buffer,

$$R_b(s) = \frac{A_{bidg}(s)}{A_{buf}(s)}, \tag{1}$$

where $A_{bldg}(s)$ is the total building footprint area within the section buffer and $A_{buf}(s)$ is the buffer area. A high built-up condition is defined using a network-internal percentile threshold to avoid externally imposed cut-offs. Specifically, $U(s) = 1$ if $R_b(s) \geq \tau_b$ is set to the 70th percentile of the corridor-wide $R_b$ distribution. The percentile is computed separately for each segmentation resolution to account for scale-dependent aggregation effects. Sensitivity to plausible threshold variation is examined in Appendix A.

Sensitive receptors are represented by the receptor count within the buffer, $N_r(s)$, defined as the number of receptor point features within the section buffer. Receptor points correspond to predefined point-of-interest categories associated with third-party presence; the category definition and tag mapping are reported in Appendix A. A receptor-high condition is defined using a presence criterion, $R(s) = 1$ if $N_r(s) \geq 1$. The effect of alternative receptor thresholds is assessed in Appendix A.

Surface water proximity is represented by a binary intersection indicator $I_w(s) \in \{0,1\}$, where $I_w(s) = 1$ indicates that mapped water features intersect the section buffer. For grading, the water exposure flag is defined as $W(s) = I_w(s)$. Following common practice in open geospatial data use, the proxy definitions are explicitly tested for threshold sensitivity to bound dependence on operational choices [26].

For downstream grading, the exposure context indicators are defined as

$$H(s) = U(s) \vee R(s), \tag{2}$$

$$Q(s) = W(s), \tag{3}$$

where $H(s)$ represents third-party exposure context and $Q(s)$ represents water-related escalation context. This separation preserves interpretability of distinct exposure mechanisms while keeping

Table 3. Operational definitions of environmental exposure proxies used for section-level screening (buffer=500 m)

| Proxy/indicator | Symbol | Operational definition (buffer=500 m) | Unit | Type | Threshold used (500 m seg.) | Threshold used (1 km seg.) |
|---|---|---|---|---|---|---|
| Built-up ratio | $R_b(s)$ | $R_b(s) = A_{bldg}(s)/A_{buf}(s)$ (total building footprint area within buffer / buffer area) | - | Continuous | - | - |
| Built-up high indicator | $U(s)$ | $U(s) = 1[R_b(s) \geq \tau_b]$ , ($\tau_b = q_{0.70}\{R_b\}$ (corridor-internal 70th percentile; computed separately by segmentation resolution) | - | Binary | 0.002560 | 0.003835 |
| Sensitive receptors count | $N_r(s)$ | $N_r(s)$ = count of selected receptor POI points within buffer | count | Count | - | - |
| Receptor presence indicator | $R(s)$ | $R(s) = 1[N_r(s) \geq 1]$ | - | Binary | 1 | 1 |
| Surface water intersection flag | $I_w(s)$ | $I_w(s) = 1$ if mapped water features intersect buffer; else 0 | - | Binary | - | - |
| Water exposure indicator | $W(s)$ | $W(s) = I_w(s)$ | - | Binary | - | - |
| Composite human/built exposure | $H(s)$ | $H(s) = U(s) \vee R(s)$ | - | Binary | - | - |
| Water exposure (grading input) | $Q(s)$ | $Q(s) = W(s)$ | - | Binary | - | - |

The table defines $R_b(s)$, $N_r(s)$, and $I_w(s)$, together with the corresponding binary flags $U(s)$, $R(s)$, and $W(s)$, and shows how they are combined into the exposure context indicators $H(s)$ and $Q(s)$.

the screening logic minimal. Operational definitions and threshold values for the exposure proxies are summarised in Table 3.

## 4.4 Data integration and preparation for screening

Structural context indicators $\{I_{tu}(s), I_{br}(s)\}$ and environmental exposure proxies $\{R_b(s), N_r(s), I_w(s)\}$ are integrated at the analytical section level along the representative alignment. For each section, the resulting attribute set combines segmentation identifiers, structural indicators, and exposure proxies to form the inputs to the deterministic grading described in Section 3.

All preprocessing and integration steps are applied uniformly across the corridor without section-specific calibration, weighting, or manual adjustment. This ensures that spatial variation in screening outputs reflects contextual differences in the underlying descriptors rather than methodological artefacts.

For comparative evaluation in Section 5, two simplified baselines are defined from the same section-level attributes. A structural-only baseline uses $\{I_{tu}(s), I_{br}(s)\}$ only, and an exposure-only baseline uses $\{H(s), Q(s)\}$ only. These baselines are used to quantify the incremental discrimination provided by the combined grading relative to single-domain screening.

## 5. Results

### 5.1 Overview of screening outcomes

Screening outputs are reported for both segmentation resolutions (500 m and 1 km) and for both grading variants (Scheme A and Scheme B). Unless otherwise stated, corridor-level prioritisation outputs are presented using the 500 m segmentation because it provides finer localisation of contiguous high-grade segments without changing the underlying proxy construction rules. Scheme B is used for decision-facing grade summaries, while Scheme A is used to support rank-based comparisons.

### 5.2 Distribution of consequence grades

For corridor prioritisation, sections with elevated consequence escalation potential are interpreted in terms of spatial continuity. Contiguous clusters are formed by grouping adjacent sections that meet a chosen grade criterion. In this study, a high-grade segment is defined as $G_B(s) \geq 3$ under Scheme B at 500 m resolution, where $G_B(s)$ is the discrete consequence grade assigned to section $s$. Cluster length is computed as the number of consecutive high-grade sections multiplied by the section length (500 m). Fig. 3 shows the empirical distribution of contiguous high-grade cluster lengths, and Table 4 summarises the corresponding corridor-level statistics for prioritisation.

- For $G_B(s) \geq 3$, 111 clusters are identified with a combined length of 137.0 km, corresponding to 34.4% of the analysed corridor length (398.0 km). Cluster lengths are short-to-moderate (median 1.0 km, 90th percentile 2.5 km), with a maximum of 8.0 km.

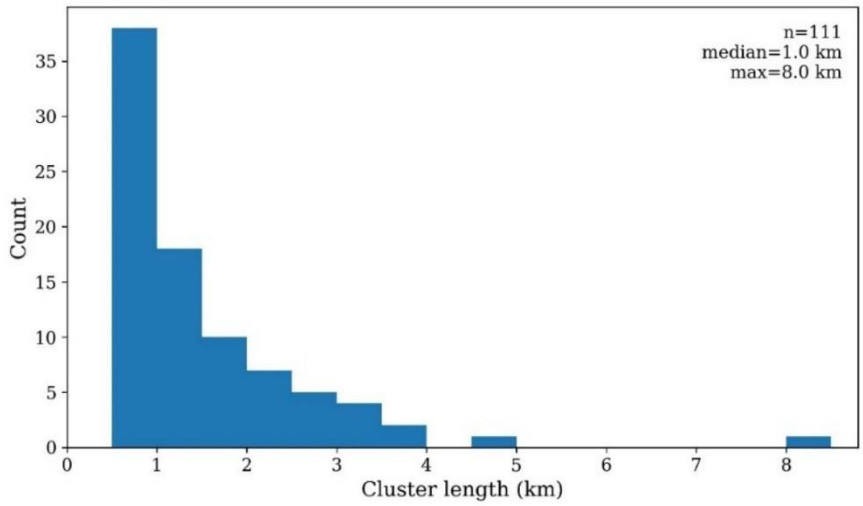



Figure 3. Distribution of contiguous high-grade cluster lengths along the corridor (Scheme B, 500 m)

Table 4. Summary of contiguous high-grade clusters (Scheme B, 500 m; criterion $G_B \geq 3$)

| Metric | Value |
|---|---|
| High-grade length (km) | 137.004 |
| High-grade share of corridor length (%) | 34.423 |
| Number of contiguous high-grade clusters | 111 |
| Median cluster length (km) | 1.0 |
| P90 cluster length (km) | 2.5 |
| Max cluster length (km) | 8.0 |

Clusters are contiguous runs of adjacent sections satisfying $G_B(s) \geq 3$(Grades 3-4). Cluster length is the sum of constituent section lengths (500 m per section).

Table 5. Baseline screening comparison against combined grading (Scheme B, 500 m; criterion $G_B(s) \geq 3$)

| Method | High set $N$(%) | Inflation | Spearman ($\rho$) | Top-80 overlap ($n$) | Top-80 Jaccard |
|---|---|---|---|---|---|
| Combined (Scheme B) | 275 (34.504) | 1.000 | 1.000 | 80 | 1.000 |
| Structural-only | 536 (67.252) | 1.949 | 0.436 | 25 | 0.185 |
| Exposure-only | 631 (79.172) | 2.295 | 0.780 | 26 | 0.194 |

Top-80 denotes the top 10% of sections ( $k = \lceil 0.10\,\mathrm{N} \rceil$ ; $k = 80$ for $N = 797$ ); ties are broken deterministically by chainage/section index. *Jaccard*=$|A \cap B|/|A \cup B|$. for the baseline vs. combined Top-80 sets. *Candidate inflation*= $|H_{baseline}|/|H_{combined}|$, where $H_{combined} = \{s: G_B(s) \geq 3\}$ and $H_{baseline}$ is defined by the corresponding baseline rule (structural-only: $I_{tu}$ or $I_{br}$; exposure-only: $H$ or $Q$).

- For $G_B(s) = 4$, 30 clusters are identified with a combined length of 25.0 km (6.3% of the corridor). These clusters are typically shorter (median 0.5 km), with a maximum of 3.0 km.

These outputs are directly usable as corridor-level decision inputs because they provide the extent and continuity of high-grade segments rather than isolated section-level scores.

### 5.3 Baseline comparisons: structural-only and exposure-only screening

To quantify what the combined grading adds beyond single-domain screening, two simplified baselines are constructed from the same section-level attributes (Section 4):

(i) a structural-only baseline using $\{I_{tu}(s), I_{br}(s)\}$ only, and

(ii) an exposure-only baseline using $\{H(s), Q(s)\}$ only.

The combined output is represented by Scheme B high-grade sections $G_B(s) \geq 3$ (500 m). Candidate inflation is defined as the ratio of baseline-selected candidates to combined high-grade sections.

Table 5 summarises the baseline comparison results against the combined grading, including candidate inflation, rank agreement, and Top-k set overlap.

- The combined grading identifies 275 sections (34.5%) as high-grade ($G_B(s) \geq 3$).
- The structural-only baseline selects 536 sections (67.3%), inflating the candidate set by 1.95× relative to the combined high-grade set. Regardless of overlap structure, at least 48.7% of structural-only candidates are not high-grade under the combined grading (lower bound $(536 - 275)/536$), i.e., they fall to $G_B(s) \leq 2$.

- The exposure-only baseline selects 631 sections (79.2%), inflating the candidate set by 2.29×. Similarly, at least 56.4% of exposure-only candidates are not high-grade under the combined grading (lower bound $(631-275)/631$). In addition, because Scheme B high grades can be driven by structural context, exposure-only screening may omit a subset of combined high-grade sections when $H=0$ and $Q=0$.

To characterise re-ordering, rank agreement is evaluated between a baseline severity score (simple additive score from the baseline indicators) and the combined Scheme B ranking. Rank correspondence is limited for the structural-only baseline (Spearman $\rho=0.436$), while the exposure-only baseline shows higher rank association ($\rho=0.780$) but still produces substantial candidate inflation. Top-$k$ agreement at the decision-relevant tail is low for both baselines, where $k=\lceil 0.10\ \mathrm{N}\rceil$ (500 m: $k=80$). The baseline and combined top-$k$ sets show limited overlap (structural-only: $|A\cap B|=25$, Jaccard=0.185; exposure-only: $|A\cap B|=26$, Jaccard=0.194), indicating that combined grading materially reorders prioritisation beyond single-domain screening.

### 5.4 Plausibility and internal consistency checks using representative high-grade cases

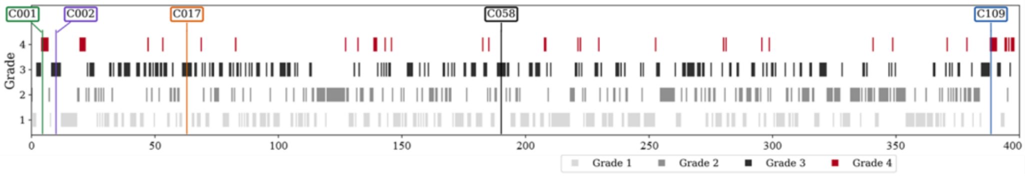
Because the study is positioned as consequence-oriented screening rather than empirical validation, plausibility is assessed through internal consistency checks: whether the highest grades concentrate where structural context and exposure context co-occur in a manner consistent with the intended interpretation of descriptors [2, 12]. Representative cases are selected from high-grade clusters and reported using chainage along the representative alignment. Fig. 4 visualises the selected representative high-grade clusters and their local context indicators, while Table 6 summarises the corresponding cluster attributes and proxy evidence used for plausibility checking.

Across the reported cases, high grades are consistently associated with the co-occurrence of at least one structural context indicator and elevated exposure context. Importantly, the cases include

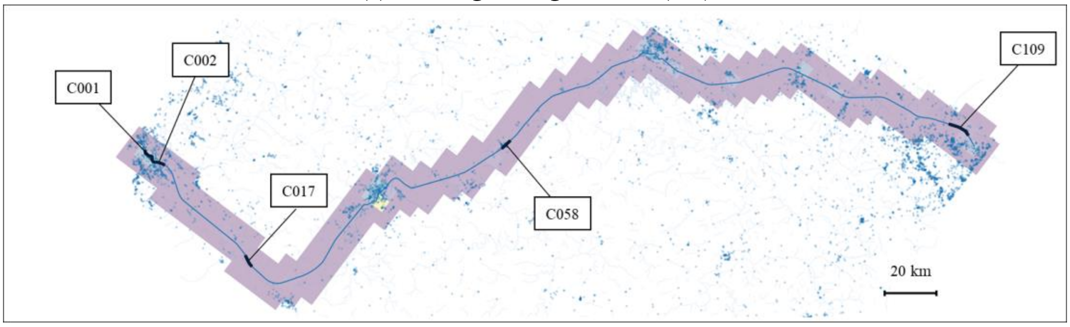
Table 6. Representative contiguous high-grade clusters for plausibility checking (Scheme B, 500 m; criterion $G_B(s)\geq 3$)

| Cluster ID | Chainage range (km) | Length (km) | Peak ($G_B$) | Composition (G4/G3) | Dominant context | Proxy evidence (representative values) |
|---|---|---|---|---|---|---|
| 109 | 384.5-392.5 | 8.0 | 4 | 6/10 | $T+H+Q$ | $R_b=0.0369$, $N_r=0$, water=1 |
| 1 | 2.0-7.0 | 5.0 | 4 | 6/4 | $T+H+Q$ | $R_b=0.2076$, $N_r=53$, water=1 |
| 2 | 8.0-12.0 | 4.0 | 3 | 0/8 | $T+H$ | $R_b=0.1000$, $N_r=2$, water=0 |
| 17 | 61.0-65.0 | 4.0 | 3 | 0/8 | $T+Q$ | $R_b=0.0000$, $N_r=0$, water=1 |
| 58 | 188.5-192.0 | 3.5 | 3 | 0/7 | $B+H+Q$ | $R_b=0.0188$, $N_r=8$, water=1 |

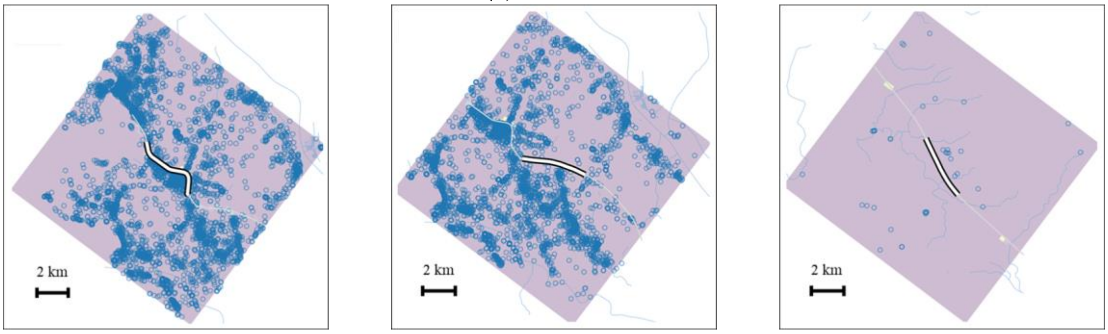
"Dominant context" summarises the prevailing $(T,B,H,Q)$ configuration within each cluster. Proxy evidence reports representative values of the built-up ratio $R_b$, receptor count $N_r$, and surface-water intersection flag (water) to support plausibility checks. All symbols and indicator definitions follow Tables 1-5 and Appendix A-B. High grades may still occur with low receptor-proxy counts because the rule set treats certain structural contexts conservatively under specific descriptor combinations.



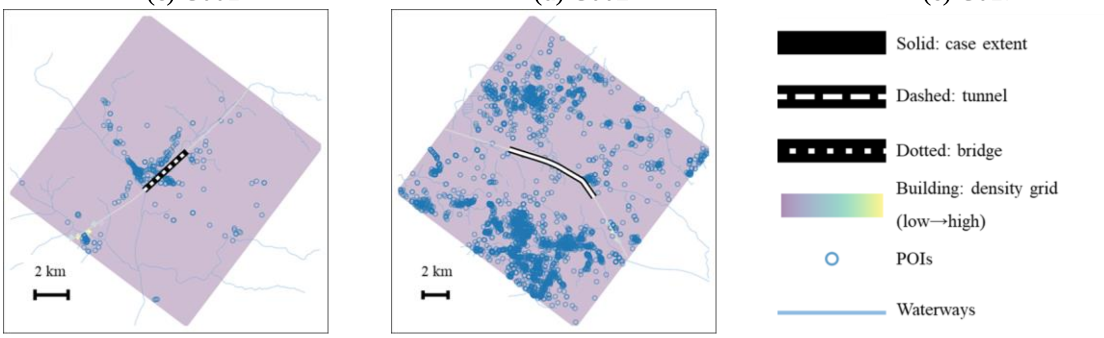

(a) Chainage along corridor (km)




(b) Overview




(c) C001 (d) C002 (e) C017




(f) C058 (g) C109 (h) Legend

Figure 4. Representative high-grade clusters and local context indicators (Scheme B, 500 m) (a) corridor chainage view (b) overview (c)-(g) selected cluster examples (h) legend. Basemap data © OpenStreetMap contributors (ODbL 1.0)

both configurations with $Q(s) = 1$ (water-related context present) and configurations where high grades arise primarily through $H(s)$ (third-party exposure context) in combination with tunnel/bridge presence. This pattern supports the intended interpretation that grades are not driven

Table 7. Robustness and sensitivity across segmentation resolution and grading schemes

| Test | Units | $N$ | Spearman $\rho$ | Top-$k$ Jaccard | Top-$k$ reclass. | Grade mismatch |
|---|---|---|---|---|---|---|
| Segmentation sensitivity (Scheme A) | 1 km | 399 | 0.786 | 0.739 | 0.030 | 0.180 |
| Segmentation sensitivity (Scheme B) | 1 km | 399 | 0.873 | 0.481 | 0.070 | 0.185 |
| Scheme sensitivity (A vs B) | 500 m | 797 | 0.527 | 0.212 | 0.130 | - |
| Scheme sensitivity (A vs B) | 1 km | 399 | 0.597 | 0.159 | 0.145 | - |

Top-k refers to the top 10% of units (500 m: $k = 80$; 1 km: $k = 40$); ties are broken deterministically by chainage/section index. For segmentation sensitivity (500 m→1 km), each 1 km unit is mapped to the maximum grade (Scheme A) or maximum Scheme-B rank key among its constituent 500 m sections. For Scheme B, $\rho$, Jaccard, and reclassification are computed on the rank key, whereas the grade mismatch is computed on discrete grades $G_B$ after the same mapping.

by a single proxy category alone but reflect combined structural-exposure context at screening scale.

### 5.5 Robustness to segmentation resolution and grading variants

Robustness is quantified in terms of (i) top-$k$ overlap, (ii) rank correlation, and (iii) grade reclassification. Top-$k$ is defined as the top 10% of sections under the relevant score/rank. Top-$k$ overlap is reported as Jaccard similarity between top-$k$ sets. Spearman $\rho$ is computed for rank agreement, and grade reclassification rate is computed as the proportion of sections whose discrete grade differs between the compared settings. For 500 m versus 1 km comparisons, 500 m sections are mapped to their parent 1 km segment by spatial containment to enable consistent comparison. Table 7 summarises robustness metrics across segmentation resolution (500 m vs 1 km) and grading variants (Scheme A vs Scheme B), including rank agreement, top-k overlap, reclassification, and grade mismatch.

Key observations are as follows:

- Segmentation sensitivity (500 m vs 1 km): rank agreement remains moderate to high (Spearman $\rho \approx 0.75$), while discrete grades show non-negligible reclassification (16.8% under Scheme A; 25.5% under Scheme B), reflecting the effect of aggregation and minimum-split handling on discrete boundary assignments.
- Variant sensitivity (Scheme A vs Scheme B): rank correlation is moderate ($\rho = 0.527$ at 500 m; $\rho = 0.597$ at 1 km), while top-$k$ overlap is low (Jaccard 0.212 at 500 m; 0.159 at 1 km) and grade reclassification is higher (≈37%), consistent with the fact that the two variants encode different aggregation and decision thresholds rather than representing minor perturbations of the same scale.

### 5.6 Sensitivity of proxy operational definitions

Dependence on operational thresholds is assessed through a limited sensitivity check around the built-up and receptor definitions. Results are reported in Appendix A using threshold variation ranges that preserve interpretability at screening stage. In brief, varying the built-up threshold within a plausible percentile range and tightening the receptor presence criterion produces limited reclassification and preserves substantial top-k overlap relative to the base settings. These results bound dependence on a single cut-off choice and support the use of percentile-based operationalisation under heterogeneous corridor conditions [26].

## 6. Discussion

### 6.1 Interpretation of corridor-scale screening outputs

This study introduces a consequence-oriented screening layer that organises corridor sections by relative escalation potential, rather than estimating derailment probability or calibrated loss. The grades should therefore be interpreted as network-internal prioritisation signals derived from the co-occurrence of structural context $\{I_{tu}, I_{br}\}$ and exposure context $\{H, Q\}$, not as predictions of event occurrence. Two outputs are retained for distinct roles: Scheme A supports rank-based comparisons, while Scheme B provides an interpretable discrete grading for operational screening [8, 11].

### 6.2 Decision utility at corridor scale

Prioritisation actions are typically applied over continuous stretches rather than isolated points. The contiguous high-grade clusters (Fig. 3; Table 4) provide corridor-level decision inputs by summarising the extent and continuity of elevated grades, while section-level grades retain within-cluster detail for targeted follow-up.

To assess whether corridor-scale prioritisation is unduly driven by Scheme B's tunnel-bridge hard rule, a limited what-if condition was applied in which tunnel-bridge segments receive Grade 4 only when $H$=1 and $Q$=1 (otherwise Grade 3). Under this stricter condition, Grade 4 500-m segments decreased from 51 to 47, while the high-grade candidate set (Grades 3-4) remained 275 and rankings were essentially unchanged (top-50 overlap=0.94; Spearman $\rho$=0.9994; Table 8); no tunnel-bridge cases occurred at 1-km aggregation.

### 6.3 Added discrimination beyond single-domain screening

Baseline comparisons (Table 5) indicate that structural-only and exposure-only screening inflate the candidate set substantially, weakening prioritisation by diluting attention. The combined grading reduces candidate inflation and materially reorders the prioritisation list beyond single-domain screening, consistent with the intended role of integrating structural and exposure contexts without calibrated weighting or external valuation assumptions [9, 10, 14].

### 6.4 Plausibility and internal consistency checks

As the study is positioned as screening rather than empirical validation, plausibility is assessed through internal consistency: whether the highest grades concentrate where structural and exposure contexts co-occur in a manner consistent with the grading logic. The representative cases (Fig. 4; Table 6) show repeated configurations in which high grades arise from combined structural context and elevated exposure context, including cases driven primarily by $H$ and cases involving $Q$. This supports interpretation consistency, while not constituting predictive validation of realised consequences [2, 12].

Table 8. Limited what-if check on Scheme B Rule 1 (tunnel-bridge) and ranking stability

| Resolution | Scenario | Grade 4 count | High-grade count (Grades 3-4) | Top-50 overlap | Spearman $\rho$ |
|---|---|---|---|---|---|
| 500 m | Baseline Scheme B | 51 | 275 | N/A | N/A |
| 500 m | What-if: tunnel-bridge requires $H = 1$ and $Q = 1$ for Grade 4 | 47 | 275 | 0.94 (47/50) | 0.9994 |
| 1 km | Baseline / What-if | unchanged | unchanged | N/A | N/A |

### 6.5 Robustness to analytical choices and proxy uncertainty

Robustness results (Table 7) show moderate-to-high rank agreement across segmentation resolutions, with non-negligible discrete reclassification reflecting aggregation and boundary handling effects. Differences between Scheme A and Scheme B should be interpreted as consequences of different aggregation and thresholding rather than instability under small perturbations. Proxy threshold sensitivity (Appendix Table A1) bounds dependence on operational cut-offs, but limitations of open geospatial data remain; prioritised clusters should therefore trigger local verification where decisions are high-stakes [26].

### 6.6 Scope, limitations, and recommended use

The framework does not model derailment probability, simulate post-derailment dynamics, or produce calibrated loss estimates. It is recommended as an upstream screening layer to direct follow-on analyses—such as occurrence modelling, dynamic simulation, or site-specific engineering review—towards contiguous high-grade clusters identified at corridor scale [8, 11, 18]. The proxy layers used here are derived from OpenStreetMap (OSM) and therefore inherit completeness and attribution uncertainties, including missing or uneven POI tagging and spatial variability in bridge/tunnel and water-feature mapping [27]. The outputs should accordingly be interpreted as screening priorities conditional on the available OSM layers, rather than as definitive exposure inventories.

## 7. Conclusions

This paper presents a deterministic framework for consequence-oriented screening of railway derailment severity at the network level. The framework was designed to organise contextual information related to derailment escalation potential and to support comparative prioritisation across large and heterogeneous networks. By focusing on infrastructure configuration and surrounding environmental context, the approach differentiates between network sections without relying on probabilistic modelling or scenario simulation.

Applying the framework to the Gyeongbu High-Speed Line illustrates that relative consequence potentials are unevenly distributed along the corridor. Sections associated with elevated grades are concentrated within limited spatial segments, rather than being uniformly dispersed. These patterns remain broadly consistent across alternative segmentation resolutions, while discrete grade assignments exhibit expected aggregation effects (Table 7).

The proposed framework is an early-stage screening tool within a layered reliability and risk assessment process. It is not intended to predict accident outcomes or provide absolute measures of consequence magnitude, and it does not estimate derailment probability or calibrate losses. The grades should therefore be interpreted as screening priorities conditional on the available proxy layers rather than as probabilistic risk or loss estimates. Instead, it supports transparent and reproducible prioritisation by identifying contiguous segments where contextual conditions may plausibly amplify derailment consequences and where more detailed analysis may therefore be most effectively directed (Fig. 3).

Within its defined scope, the framework complements frequency-oriented indicators and detailed analytical methods by explicitly addressing consequence-related considerations at the network level. Its deterministic structure and reliance on interpretable descriptors enable the use of routinely available geospatial sources, while sensitivity checks bound dependence on operational assumptions (Appendix Table A1).

Future work can extend the framework by applying it to other corridors and by exploring alternative contextual descriptors and data sources. Such extensions would clarify how consequence-oriented screening can be systematically combined with likelihood modelling within layered railway safety decision processes.

## Acknowledgements

This research was supported by the Ministry of Land, Infrastructure, and Transport of the Korean Government (grant numbers: RS-2025-02222955 and RS-2021-KA163289).

## References

1. Anderson, R.T., Barkan, C.P. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record, 1863(1), 88-98. https://doi.org/10.3141/1863-12.
2. Evans, A.W. (2013). The economics of railway safety. Research in Transportation Economics, 43(1), 137-147. https://doi.org/10.1016/j.retrec.2012.12.003.
3. European union agency for railways (2018). Report on railway safety and interoperability in the EU. Publications Office of the European Union, Luxembourg. https://doi.org/10.2821/205360.
4. Elvik, R. (2013). International transferability of accident modification functions for horizontal curves. Accident Analysis & Prevention, 59, 487-496. https://doi.org/10.1016/j.aap.2013.07.010.
5. Liu, X., Saat, M.R., Barkan, C.P. (2017). Freight-train derailment rates for railroad safety and risk analysis. Accident Analysis & Prevention, 98, 1-9. https://doi.org/10.1016/j.aap.2016.09.012.
6. Kaplan, S., Garrick, B.J. (1981). On the quantitative definition of risk. Risk Analysis, 1(1), 11-27. https://doi.org/10.1111/j.1539-6924.1981.tb01350.x.
7. Hansson, S.O. (1989). Dimensions of risk. Risk Analysis, 9(1), 107-112. https://doi.org/10.1111/j.1539-6924.1989.tb01225.x.
8. Aven, T., Zio, E. (2014). Foundational issues in risk assessment and risk management. Risk Analysis, 34(7), 1164-1172. https://doi.org/10.1111/risa.12132.
9. Haimes, Y.Y. (2009). Risk modeling, assessment, and management. John Wiley & Sons, Hoboken, New Jersey.
10. Anthony (Tony) Cox Jr, L. (2008). What's wrong with risk matrices?. Risk Analysis, 28(2), 497-512. https://doi.org/10.1111/j.1539-6924.2008.01030.x.
11. Zio, E. (2018). The future of risk assessment. Reliability Engineering & System Safety, 177, 176-190. https://doi.org/10.1016/j.ress.2018.04.020.

12. Forsberg, R., Björnstig, U. (2011). One hundred years of railway disasters and recent trends. Prehospital and Disaster Medicine, 26(5), 367-373. https://doi.org/10.1017/S1049023X1100639X.
13. Tang, Z., Hu, Y., Wang, S., Ling, L., Zhang, J., Wang, K. (2024). Train post-derailment behaviours and containment methods: a review. Railway Engineering Science, 32(1), 59-80. https://doi.org/10.1007/s40534-023-00313-5.
14. Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. European Journal of Operational Research, 253(1), 1-13. https://doi.org/10.1016/j.ejor.2015.12.023.
15. Zio, E. (2007). An introduction to the basics of reliability and risk analysis. Vol. 13, World Scientific.
16. Paté-Cornell, M.E. (1996). Uncertainties in risk analysis: Six levels of treatment. Reliability Engineering & System Safety, 54(2-3), 95-111. https://doi.org/10.1016/S0951-8320(96)00067-1.
17. Evans, A.W. (2000). Fatal train accidents on Britain's mainline railways. Journal of the Royal Statistical Society Series A: Statistics in Society, 163(1), 99-119. https://doi.org/10.1111/1467-985X.00159.
18. Rail Safety and Standards Board (2019). Taking safe decisions issue 3. London.
19. Hong, W.T., Clifton, G., Nelson, J.D. (2023). Railway accident causation analysis: current approaches, challenges and potential solutions. Accident Analysis & Prevention, 186, 107049. https://doi.org/10.1016/j.aap.2023.107049.
20. Saat, M.R., Barkan, C.P. (2005). Release risk and optimization of railroad tank car safety design. Transportation Research Record, 1916(1), 78-87. https://doi.org/10.1177/0361198105191600111.
21. Saat, M.R., Lin, C.Y. (2016). Shared rail corridor adjacent track accident risk analysis. National University Rail Center, Urbana-Champaign.
22. Lai, J., Wang, K., Xu, J., Wang, P., Chen, R., Wang, S., Beer, M. (2023). A failure probability assessment method for train derailments in railway yards based on IFFTA and NGBN. Engineering Failure Analysis, 154, 107675. https://doi.org/10.1016/j.engfailanal.2023.107675.
23. Liu, X., Turla, T., Zhang, Z. (2018). Accident-cause-specific risk analysis of rail transport of hazardous materials. Transportation Research Record, 2672(10), 176-187. https://doi.org/10.1177/0361198118794532.
24. Zhai, W., Han, Z., Chen, Z., Ling, L., Zhu, S. (2019). Train–track–bridge dynamic interaction: a state-of-the-art review. Vehicle System Dynamics, 57(7), 984-1027. https://doi.org/10.1080/00423114.2019.1605085.
25. Kyriakidis, M., Simanjuntak, S., Singh, S., Majumdar, A. (2019). The indirect costs assessment of railway incidents and their relationship to human error-the case of signals passed at danger. Journal of Rail Transport Planning & Management, 9, 34-45. https://doi.org/10.1016/j.jrtpm.2019.01.001.
26. Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B: Planning and Design, 37(4), 682-703. https://doi.org/10.1068/b35097.
27. OpenStreetMap contributors. OpenStreetMap data [dataset]. 2026-01-20 snapshot; accessed 2026-01-20. Available from: https://www.openstreetmap.org

## Appendix A. Data, network representation, and preprocessing

This appendix summarises the geospatial inputs and preprocessing used to derive section-level structural descriptors and exposure proxies for deterministic grading (Appendix B).

### A.1 Data sources and layers

All contextual layers are derived from OpenStreetMap (OSM) [27], including the corridor alignment, tunnels/bridges, buildings, sensitive receptor POIs, and surface-water features. Layers are processed from a single snapshot in a metric projected coordinate system.

Sensitive receptor POIs are defined using a fixed tag list applied uniformly across the corridor; the full tag list is provided in the Supplementary material/scripts.

### A.2 Network representation and segmentation

A single representative alignment is used as the spatial reference for all processing. The alignment is segmented into fixed-length analytical sections at two resolutions for sensitivity checks: 500 m ($N = 797$) and 1 km ($N = 399$), covering a total corridor length of 398.004 km. To minimise boundary artefacts, section boundaries are aligned to tunnel/bridge endpoints (minimum-split) so that each analytical section has internally consistent structural indicators $I_{tu}(s)$ and $I_{br}(s)$. Chainage (start/end position along the alignment) is stored for each section and is used for deterministic ordering and tie-breaking where required.

### A.3 Structural descriptors

Two binary structural indicators are computed by spatial intersection between each section and the corresponding OSM structural feature layer:

- $I_{tu}(s) = 1$[section $s$ intersects an OSM tunnel feature],
- $I_{br}(s) = 1$[section $s$ intersects an OSM bridge feature].

These indicators are used directly as deterministic inputs in Appendix B.

Table A1. OSM layers and tag filters used for structural and exposure proxy extraction

| Category | Feature type | Category-level OSM tags | Geometry | Used to derive |
|---|---|---|---|---|
| Corridor reference | Representative alignment | Corridor alignment extracted as a polyline reference | Line | Section segmentation, chainage |
| Structural | Tunnel features | tunnel=* | Line/Relation | $I_{tu}(s)$ |
| | Bridge features | bridge=* | Line/Relation | $I_{br}(s)$ |
| Exposure | Building footprints | building=* | Polygon | $A_{bldg}(s)$, $R_b(s)$, $U(s)$ |
| | Sensitive receptor POIs | Fixed tag list (Supplementary) | Point | $N_r(s)$, $R(s)$ |
| | Surface water features | natural=water, water=*, waterway=* | Line/Polygon | $I_w(s)$, $Q(s)$ |

OSM features were extracted from a single snapshot and processed consistently along the corridor (download date: 2026-01-20). Category-level tags are reported for concision; the full sensitive receptor POI tag list is provided in the Supplementary material (or processing scripts).

### A.4 Exposure proxies within a fixed buffer

A fixed-radius buffer of 500 m is generated around each section. Within this buffer, three exposure proxies are computed and then converted into binary flags for Scheme B (Table 3). Unless stated otherwise, polygon- and line-based features are evaluated using standard intersection operations against the buffer polygon.

Built-up context is represented by the ratio of building footprint area to buffer area. Let $A_{buf}(s)$ be the buffer area and $A_{bldg}(s)$ be the sum of intersection areas of OSM building footprint polygons with the buffer polygon. The built-up ratio is:

$$R_b(s) = \frac{A_{bldg}(s)}{A_{buf}(s)} \tag{A.1}$$

A built-up high flag is defined using a corridor-internal percentile threshold, computed separately for each segmentation resolution:

$$U(s) = 1[R_b(s) \geq \tau_b], \tau_b = q_{0.70}\{R_b\} \tag{A.2}$$

Sensitive receptors are represented by the count of selected OSM POIs within the buffer:

$$N_r(s) = \# \,\{\text{selected receptor POIs within } s - \text{buffer}\} \tag{A.3}$$

$$R(s) = 1[N_r(s) \geq 1] \tag{A.4}$$

Surface water proximity is represented by a buffer-intersection flag using OSM hydrographic features:

$$I_w(s) = 1[\text{OSM surface water features intersect } s - \text{buffer}] \tag{A.5}$$

$$Q(s) = I_w(s) \tag{A.6}$$

For Scheme B, the composite human/built exposure flag is defined as:

$$H(s) = U(s) \vee R(s) \tag{A.7}$$

### A.5 Integrated section-level inputs and scope

For each segmentation resolution, the section-level dataset integrates $\{I_{tu}(s), I_{br}(s)\}$ with $\{R_b(s), N_r(s), I_w(s), U(s), R(s), H(s), Q(s)\}$. These inputs are subsequently mapped to deterministic grades and rank keys in Appendix B and summarised for prioritisation in the main text.

The resulting indicators are proxies intended for corridor-scale screening. They are not calibrated to external loss metrics and are subject to completeness and attribution uncertainties in OSM tagging and feature coverage.

## Appendix B. Deterministic consequence grading

This appendix defines deterministic rules used to assign ordinal consequence screening grades to each section for corridor-internal prioritisation (conditional on derailment occurrence).

### B.1 Indicator set

Four binary indicators (Appendix A) are used: tunnel, bridge, human/built exposure, and surface-water exposure.

$$k(s) = I_{tu}(s) + I_{br}(s) + H(s) + Q(s) \in \{0,1,2,3,4\} \tag{B.1}$$

### B.2 Scheme A (matrix score) for rank-based comparisons

Scheme A produces an ordinal matrix score $G_A(s)$ by combining a structural grade and an exposure grade.

The structural grade is defined deterministically from $(T, B)$:

$$S(s) = \begin{cases} 4, & T = 1 \land B = 1 \\ 3, & T = 1 \land B = 0 \\ 2, & T = 0 \land B = 1 \\ 1, & T = 0 \land B = 0 \end{cases} \tag{B.2}$$

The exposure grade is defined from whether any exposure pathway is present:

$$E(s) = \begin{cases} 2, & H = 1 \lor Q = 1 \\ 1, & H = 0 \land Q = 0 \end{cases} \tag{B.3}$$

The Scheme A matrix score is then:

$$G_A(s) = 10S(s) + E(s) \tag{B.4}$$

so that higher values indicate higher screening severity in the matrix sense (e.g., 42>32>22>12>11).

For rank-based outputs, sections are ordered by decreasing $G_A(s)$. Where ties occur, ordering is made deterministic using chainage/section index.

### B.3 Scheme B (rule-based discrete grades) for decision-facing screening

Scheme B maps $(T, B, H, Q)$ to a four-level discrete grade $G_B(s) \in \{1,2,3,4\}$ using the fixed rule set summarised in Table 2 and reproduced here for completeness:

$$G_B(s) = \begin{cases} 4, & T = 1 \land B = 1 \\ 4, & T = 1 \land B = 0 \land H = 1 \land Q = 1 \\ 3, & T = 1 \land B = 0 \land (H, Q) \in \{(1,0), (0,1)\} \\ 3, & T = 0 \land B = 1 \land H = 1 \land Q = 1 \\ 2, & T = 0 \land B = 1 \land (H, Q) \in \{(1,0), (0,1)\} \\ 2, & T = 0 \land B = 0 \land H = 1 \land Q = 1 \\ 1. & \text{otherwise.} \end{cases} \tag{B.5}$$

The "high-grade" set used for corridor screening and prioritisation is defined as $G_B(s) \geq 3$.

### B.4 Scheme B rank key for total ordering

Some analyses require a strict total ordering rather than only discrete grades (e.g., Spearman rank correlation and Top-$k$ overlap metrics). A Scheme B rank key is therefore defined to provide deterministic ordering within and across grades.

Ordering precedence is defined as follows (higher first at each step):

1. $G_B(s)$ (primary),
2. Built-up ratio $R_b(s)$,
3. Receptor count $N_r(s)$,
4. Water intersection flag $I_w(s)$,
5. Chainage/section index (final deterministic tie-break)

In implementation, this precedence may be encoded into a single scalar key $G_{B,rank}(s)$ for convenience, but the ordering rule is fully determined by the precedence above (i.e., $G_{B,rank}(s)$ is used only as a consistent representation of the same lexicographic order).

### B.5 Cross-resolution mapping used in sensitivity checks

For segmentation sensitivity checks (500 m→1 km), each 1 km unit is compared against the most severe constituent 500 m section under a conservative screening interpretation. For a 1 km segment $S$ composed of 500 m sections $s \in S$, the mapped values are:

$$G_A^{Max}(s) = \max_{s \in S} G_A(s) \tag{B.6}$$

$$G_{B,rank}^{Max}(s) = \max_{s \in S} G_{B,rank}(s) \tag{B.7}$$

$$G_B^{Max}(s) = \max_{s \in S} G_B(s) \tag{B.8}$$

This mapping avoids dilution of localised high-grade contexts when aggregating to a coarser resolution and supports like-for-like comparisons between segmentation settings.