American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2022)
Page No: 55-87
eISSN: 3067-0470
[DOI: 10.63125/h24n6z92](https://doi.org/10.63125/h24n6z92)

Article

QUANTITATIVE RISK ASSESSMENT OF RAIL INFRASTRUCTURE PROJECTS USING MONTE CARLO SIMULATION AND FUZZY LOGIC

Md. Sakib Hasan Hriday[1];

[1]. Assistant Civil Engineer, CREC(China Railway Engineering Corporation) PBRLP-Padma Bridge Rail Link Project, Bangladesh;
Email: hriday.hasan1999@gmail.com

ABSTRACT

Rail infrastructure programs frequently face complex, intertwined risks spanning cost overruns, schedule delays, safety concerns, and interface uncertainties. This systematic review critically examines how quantitative risk assessment (QRA) methods—Monte Carlo simulation (MCS), fuzzy logic (FL), and hybrid approaches—have been employed across the rail project lifecycle to manage these multidimensional challenges. Following PRISMA guidelines, we conducted a comprehensive search across Scopus, Web of Science, IEEE Xplore, ASCE Library, ScienceDirect, and TRB databases, applied pre-registered eligibility criteria, implemented double-screening for study inclusion, and rigorously appraised methodological practices encompassing data provenance, dependence modeling, validation, and sensitivity analysis. From an initial pool of studies, 95 peer-reviewed publications met all inclusion standards. Findings indicate MCS dominates (47%) owing to its strength in producing distributional forecasts and governance-ready percentiles; FL supports imprecise or linguistic inputs (33%) often encountered during early-stage planning and safety screening; while hybrid models (20%) bridge probabilistic propagation and evidential uncertainty, particularly in interface-intensive phases. Applications cluster within construction (68%), followed by design (46%), feasibility analysis (39%), testing/commissioning (24%), and operations & maintenance (21%). Methodologically, MCS studies primarily use triangular and PERT/beta distributions, with approximately 42% employing Latin hypercube sampling. However, dependence modeling remains limited—38% of studies assume independence, 23% employ rank or copula methods, and only 31% jointly simulate cost–schedule interactions. FL studies typically apply triangular/trapezoidal membership functions with centroid defuzzification; while two-thirds disclose reproducible rule bases, one-third lack transparency. Hybrid models frequently convert fuzzy assessments into probabilistic inputs or embed fuzzy and evidential nodes into Bayesian structures, enabling richer risk representation at system interfaces. Sensitivity analysis is reported in 64% of studies, but only 21% adopt global approaches and a mere 5% include tail-focused diagnostics, while external validation is rare (18%).

Keywords
QRA; rail infrastructure; Monte Carlo; fuzzy logic; hybrid models; cost–schedule risk; dependence modeling;

Citation:
Hriday, M. S. Hasan (2022). Quantitative risk assessment of rail infrastructure projects using Monte Carlo simulation and fuzzy logic. *American Journal of Advanced Technology and Engineering Solutions*, 2(1), 55–87. https://doi.org/10.63125/h24n6z92

Received:
January 18, 2022

Revised:
February 24, 2022

Accepted:
March 26, 2022

Published:
April 30, 2022

INTRODUCTION

Risk in rail infrastructure projects can be understood as the effect of uncertainty on objectives across safety, cost, schedule, quality, and service reliability, where uncertainty refers to the state of limited knowledge about events, parameters, or models that can be reduced or only partially characterized. In quantitative risk assessment (QRA), uncertainty is typically differentiated into aleatory variability inherent in processes and epistemic uncertainty arising from limited information or modeling assumptions, and the distinction matters for how uncertainty is represented and propagated in decision contexts. Rail infrastructure denotes a system of interdependent assets earthworks, track, structures, power, signalling and control, telecommunications, rolling stock interfaces, and operational processes whose performance and safety are governed by lifecycle processes and assurance regimes (CENELEC, 2017; Cooke, 1991; Flyvbjerg, 2009). Monte Carlo simulation (MCS) is a probabilistic technique that samples from input distributions to estimate the distribution of outcomes for key objectives such as cost or time, enabling the derivation of decision-useful statistics like P50 and P80 and sensitivity measures. Fuzzy logic (FL) represents vagueness linguistically through membership functions and rule bases, enabling reasoning with incomplete or imprecise data and expert judgments without requiring fully specified probability distributions (Cantarelli et al., 2012).International guidance recognizes both probabilistic and non-probabilistic techniques as part of a structured risk management process that includes identification, analysis, evaluation, treatment, and monitoring across the asset lifecycle . In rail, the RAMS framework formalizes the relationship among reliability, availability, maintainability, and safety in a way that interacts with QRA methods and assurance evidence . These definitions establish a methodological space in which MCS and FL are not competing abstractions but complementary instruments conditioned by data characteristics, stakeholder needs, and assurance requirements .

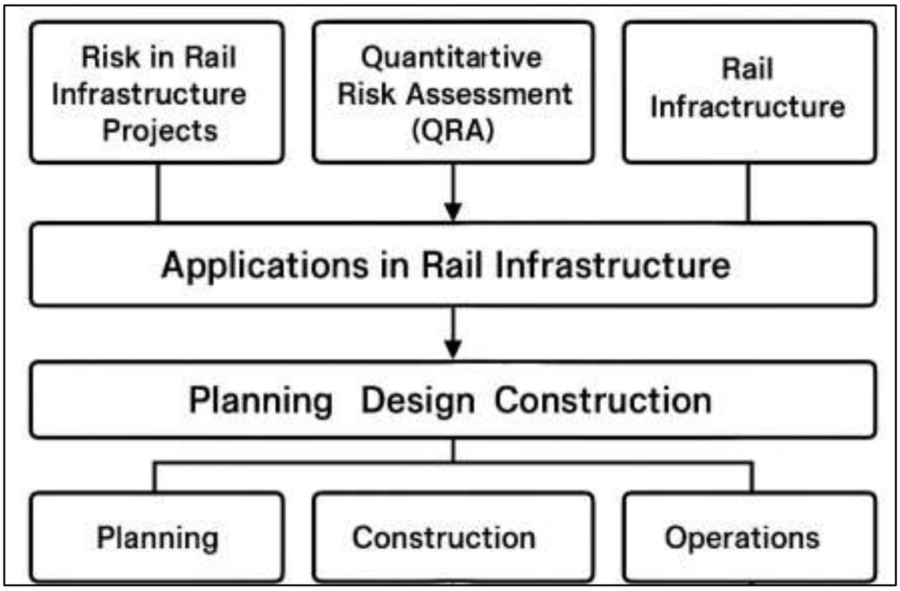
Investment in rail networks and rollingstock is large-scale and international, with projects occurring in diverse contexts from high-speed corridors and urban metros to freight modernization programs, and the associated risks have been studied extensively across planning, design, construction, and operations (El-Sayegh, 2008). Documented variance in outturn cost and schedule relative to baseline estimates is a central concern for public value, private capital, and regulatory oversight, motivating the adoption of structured risk quantification and independent review. International sponsors and regulators have promulgated process guidance for risk reviews, contingency development, and decision assurance in transport, including probabilistic analysis expectations and documentation of assumptions . Within project governance, known cognitive phenomena such as overconfidence and planning fallacy affect forecasts and expert inputs, reinforcing the need for transparent elicitation and structured analytical methods. Rail's technical interfaces geotechnical conditions, station and tunnel works, systems integration, signaling migration, power supply resilience, and operational commissioning create paths for risk propagation across disciplines and contractual boundaries, which increases the analytical value of uncertainty modeling aligned to lifecycle gates and assurance evidence). Internationally, cost-schedule coupling, access constraints, and safety-critical testing at handover create distinctive risk clusters, and quantitative methods that capture dependency structures and expert knowledge are widely used to characterize those clusters for decision reviews (Vose, 2008; Zimmermann, 2001). The combination of scale, public scrutiny, and technical coupling forms the motivation for a focused synthesis of MCS and FL applications within rail project QRA across jurisdictions and delivery models.

MCS is established in infrastructure risk practice for propagating probabilistic inputs through cost and schedule networks, estimating contingency distributions, and identifying key drivers through sensitivity analysis . In cost risk, inputs often include base estimate uncertainty, quantity variability, unit-rate dispersion, escalation, and risk events with occurrence probabilities and impacts, while schedule risk models task durations, calendars, logical relationships, and discrete threats or opportunities (Hulett, 2016; International, 2010, 2011). Analytical attention to dependence structures matters because cost and time drivers often share underlying causes, and correlation, copulas, or joint modeling approaches are used to avoid biased tail estimates and to support coherent contingency setting . Standards and guidance elaborate minimum expectations for transparency, including documenting distribution choices, parameter sources, expert elicitation protocols, and sensitivity methods such as tornado charts and variance-based indices (Helton & Davis, 2003; Institute, 2021; Zadeh, 1975). In rail, MCS has been applied across preliminary business cases, reference design development, construction planning, and systems integration planning to inform

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

reserve setting, bid evaluation, and access strategy reviews that require probabilistic statements. The method's strength lies in its capacity to integrate historical data with quantified expert judgment into a coherent probabilistic forecast while making parameter and model assumptions explicit and testable within assurance engagements . This foundation positions MCS as a reference technique for quantitative claims and for aligning risk evidence with decision thresholds within rail governance frameworks.

Figure 1: Quantitative Risk Assessment for Rail Infrastructure Projects



FL addresses a different dimension of uncertainty by modeling vagueness and linguistic knowledge where probability assignments are not straightforward or where available data are sparse, heterogeneous, or qualitatively expressed (Taylan et al., 2014). In FL, concepts such as "high geotechnical risk," "limited access," or "complex systems interface" are represented by membership functions that map inputs to degrees of truth, and rule bases encode expert reasoning about combinations of conditions and their consequences (Salling, 2008; Salling & Leleur, 2011). Defuzzification then converts the inference result to an actionable scalar or ranked output, enabling prioritization or categorization without requiring full probabilistic characterization. The approach aligns with structured expert elicitation when the state of knowledge is predominantly qualitative, when evidence ranges across disciplines, or when stakeholder understanding benefits from linguistic framing. In engineering and construction risk, FL has been used to assess contractor capability, safety risk levels, interface complexity, and environmental or community risk exposure, especially at early design stages or in contexts with limited historical analogues .The logic of membership functions and rules can be combined with multi-criteria decision methods to structure trade-offs among cost, time, and risk attributes, providing ordered selections or risk rankings compatible with governance needs. Within rail programs, qualitative risk registers and interface hazard analyses often contain expert linguistic judgments, and FL offers a mechanism to formalize and aggregate that knowledge in a way consistent with assurance evidence and gate reviews.

Applications of MCS and FL in transport and construction show distinct modeling patterns that reflect data availability, lifecycle stage, and decision requirements, and the literature includes hybrid approaches that link the two methodologies. For example, fuzzy membership functions can be used to translate linguistic risk factors into quantitative inputs for MCS, or MCS results can inform fuzzy rule weights for categorizing overall exposure or prioritizing mitigations. Construction-focused studies have used fuzzy AHP and fuzzy TOPSIS to evaluate contractor alternatives and risk response options, while probabilistic methods estimate contingencies or buffer sizes associated with selected alternatives, enabling consistent narratives across qualitative and quantitative evidence. In rail appraisal and program management, transport studies have implemented MCS to represent uncertainty in cost-benefit inputs and schedule networks, with attention to correlation and scenario design to avoid biased central estimates. Safety and RAMS-related analyses often begin with qualitative hazard identification and severity/likelihood categorization that can be structured with fuzzy rules to support consistent classification before probabilistic reliability modeling, supporting assurance documentation under sector standards (Kahneman & Tversky, 1979). The presence of hybrid designs in the literature indicates that method selection can be aligned with the state of evidence and the informational needs of governance processes without privileging a single formalism (Kahneman & Tversky, 1979; Kahraman, 2015; Klir & Yuan, 1995).

Data characteristics and expert judgment protocols shape both probabilistic and fuzzy analyses, and international guidance emphasizes transparency in parameterization, elicitation, and model verification. In probabilistic models, distribution selection and parameter estimation can be informed by historical datasets, Bayesian updating, or structured calibration of expert inputs, and sensitivity analysis is used to quantify influence and to structure data-collection priorities. In fuzzy models, membership function shapes and rule weights are derived from domain knowledge, linguistic scales, and, when available, ordinal or interval data that can be mapped into degrees of membership, with validation conducted through expert review or comparison with known cases. Expert elicitation literature highlights the importance of bias awareness, aggregation methods, and documentation of reasoning, which directly informs the credibility of both probabilistic and fuzzy assessments. Sector guidance for transport sponsors sets expectations for risk workshops, model files, distribution-fitting evidence, and independent challenge, forming a consistent environment for applying MCS and FL within program governance . Within rail delivery, data often arrive in heterogeneous forms across geotechnical investigations, productivity studies, access windows, possession rules, and systems integration test plans, which gives practical relevance to methodologies that can integrate quantitative and linguistic evidence (O'Hagan et al., 2006; Odeck, 2004).

The international research record on major projects and transport infrastructure documents systematic deviations between baseline forecasts and outturns in cost and time, along with governance responses such as reference-class comparisons, independent risk reviews, and the use of probabilistic contingencies. Rail programs face additional complexity from safety certification, interface hazard closure, trial running, and timetable integration, which imposes structured lifecycle gates and evidence requirements that interact with QRA design. Methodological clarity in how uncertainty is represented, how dependencies are captured, and how expert knowledge is encoded supports consistent decision records and aligns with sponsor expectations documented in sector guidelines . Within this context, MCS provides distributional forecasts and sensitivity information suited to contingency setting and schedule confidence assessment, and FL provides structured handling of linguistic judgments and qualitative risk structures often present in early phases or in areas with sparse data. Hybridization allows movement between these representational modes when problem structure and evidence types indicate a benefit from translation or combination (CENELEC, 2017; International, 2010). The breadth of the literature across standards, methodological texts, and applied studies supplies a basis for a structured review focused on rail infrastructure that catalogues modeling choices, data practices, and assurance alignment (CENELEC, 2017; IEC, 2019; International, 2010; ISO, 2018).

This study therefore frames a literature-review-based analysis of quantitative risk assessment for rail infrastructure projects centered on the application of Monte Carlo simulation, fuzzy logic, and their hybrids across lifecycle stages and risk categories. The review addresses how MCS and FL are operationalized in rail, which risk categories and lifecycle phases are most frequently modeled, what modeling choices recur in distributions, membership functions, dependency structures, and rule systems, and how studies address validation and sensitivity. It also records the forms of data and expert judgment used, including historical cost and schedule series, geotechnical variability representations, systems integration risk registers, and structured elicitation protocols, in order to map methods to evidence types and governance artifacts (Flyvbjerg et al., 2002; Saltelli et al., 2000; Vose, 2008; Zadeh, 1965). The analysis includes studies from international rail contexts and allied civil infrastructure where methods generalize, including transport appraisal applications and construction risk assessments that provide methodological analogues for rail systems integration and interface management. By synthesizing across standards, foundational methods, and applied studies, the review presents a structured account of how quantitative techniques are embedded in rail project risk analysis and how methodological choices are evidenced and reported within sector guidance and assurance regimes .

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

LITERATURE REVIEW

Quantitative risk assessment (QRA) for rail infrastructure sits at the intersection of project controls, systems engineering, and decision science. The field spans two main methodological pillars probabilistic simulation, most commonly Monte Carlo simulation (MCS), and non-probabilistic approaches grounded in fuzzy logic (FL). MCS propagates uncertainty through cost and schedule models to produce distributional forecasts (e.g., P50/P80 contingencies) and identify dominant risk drivers through sensitivity analysis. FL, by contrast, formalizes linguistic judgments ("high geotechnical uncertainty," "limited access," "complex interface risk") using membership functions and inference rules to rank or score exposure when data are sparse, heterogeneous, or only partially quantifiable. Rail projects amplify the need for both approaches because risks emerge from tightly coupled disciplines geotechnical conditions, tunneling and structures, track and civils, power and rolling stock interfaces, signaling, and telecommunication systems while delivery is constrained by access windows, safety certification, and timetable integration. The literature has evolved along three broad threads. First, application studies embed MCS in cost and schedule risk analysis for feasibility, design development, and construction planning, often treating correlations among cost items or between duration and productivity, and reporting contingency levels aligned with governance thresholds. Second, engineering-management research deploys FL (frequently with fuzzy AHP/TOPSIS) to convert expert knowledge into consistent risk rankings, particularly in early design, contractor capability assessment, safety screening, and interface complexity appraisal. Third, hybrid designs connect the two using fuzzy constructs to parameterize probabilistic inputs, or using probabilistic outputs to weight rule bases thereby translating between linguistic and numeric evidence. Across these threads, persistent methodological choices shape credibility: the selection and fitting of distributions, the elicitation and calibration of expert judgment, the representation of dependencies (pairwise correlations vs. copulas), and the transparency of sensitivity analyses. Reporting practices vary widely in how assumptions, validation checks (e.g., back-checks against realized outcomes or independent expert review), and data lineage are documented. Sector standards (e.g., RAMS processes and risk-management guidelines) create expectations for traceability that many papers address unevenly. Given this heterogeneity, a structured synthesis is needed that (i) maps the rail risk landscape across lifecycle phases; (ii) compares MCS, FL, and hybrid patterns; (iii) examines data and elicitation protocols; (iv) evaluates dependency modeling and cost–schedule coupling; and (v) reviews validation, sensitivity, and reporting norms. The remainder of the literature review is organized around eight subsections that operationalize these aims and set up the comparative analysis used later in the paper.

Rail Risk Landscape and Lifecycle Mapping

Mapping the risk landscape of rail infrastructure projects requires a systematic understanding of how risks emerge, interact, and evolve across the full project lifecycle. This lifecycle typically includes strategic planning and feasibility, preliminary and detailed design, procurement, construction and systems integration, testing and commissioning, and finally long-term operations and maintenance. At each of these stages, risks are not only different in character but also interdependent, spreading across technical interfaces such as civil works, systems, and rolling stock. They also cut across organizational boundaries between the owner, the EPC or DB contractor, the systems integrator, and the operator, while being shaped by external environments including geology, urban context, and regulatory regimes. Empirical studies on international rail construction have shown that cost-estimating uncertainty at early stages often seeds later cost overruns by failing to capture scope creep, market volatility, and interface complexity. These "estimation risks" are therefore among the primary drivers from the very beginning of a project (Yang et al., 2021). As the project moves into construction and systems integration, hazards increasingly stem from stakeholder interactions. Examples include inadequate coordination of utilities, conflicts with traffic management, and variability in subcontractor skill, all of which can propagate through social and organizational networks. Such pathways amplify both safety and schedule exposure (Chen et al., 2020).

Specific to metro tunneling, detailed studies have identified a taxonomy of dominant construction hazards such as TBM launch and arrival, face stability problems, groundwater inflow, and shaft-related works. The significance of these risks varies depending on the construction phase and tunneling method More recent reviews of shield-method metros add further dimensions, highlighting human and organizational precursors including crew turnover, shift pressure, and inadequate method statements. Risks also show spatial and temporal clustering along linear work fronts,

Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

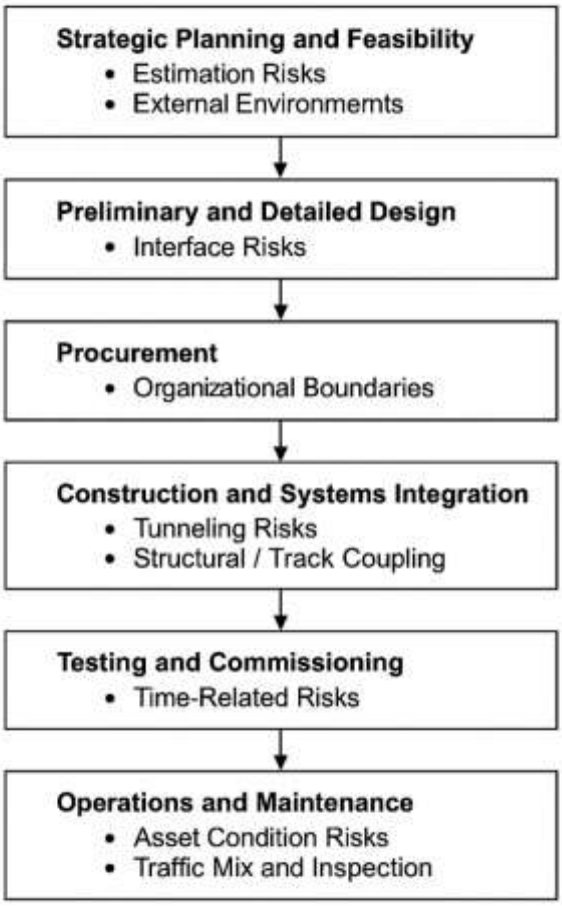
confirming that risk is dynamic and migrates with the "moving factory" nature of rail construction (Liu et al., 2018; Zhang et al., 2020).

Within the construction window of rail infrastructure projects, lifecycle mapping becomes more granular, sharpening into work-package specific risk structures. Tunneling and underground station construction, for example, are dominated by geotechnical and hydrogeologic uncertainty. In these settings, complete quantitative data are rarely available, and risk assessments must often rely on partial information combined with linguistic judgements provided by domain experts. To address this limitation, interval-number models and other bounded-uncertainty approaches have been applied to formalize expert inputs, enabling the ranking of hazards such as tunnel face instability, ground settlement leading to damage of adjacent structures, and groundwater inflow under conditions of significant data scarcity. Structural works and right-of-way components present a different type of coupling. Here, the primary concern lies in the interactions among vehicles, track systems, and structural elements, combined with environmental loads such as temperature variations, wind, and flooding. These interactions create interdependencies among nodes in the broader risk network. Weighted Bayesian network models, particularly when calibrated to empirical data from high-speed rail corridors, have shown strong potential for capturing these relationships by explicitly relaxing independence assumptions that are often unrealistic. Such models improve inference about which nodes whether structural condition, vehicle dynamics, or environmental stressors are most influential in determining the safety margin for a specific context . At the program management level, time-related risks take center stage. Project schedules are especially sensitive to the performance of tunnel boring machines, the relocation and protection of utilities, and the sequencing of handover milestones. Bayesian network-based decision-support systems for TBM projects have demonstrated their value in predicting and mitigating delay chains driven by geology, machine availability, and logistics interfaces (Koseoglu Balta et al., 2021; Yuan et al., 2020).

Importantly, these strand-level models also reveal how risks aggregate across adjacent work packages. For example, a lag in ground treatment may simultaneously elevate the risk of settlement and the probability of possession overrun, thereby clarifying which mitigation strategies should be prioritized early and which can be deferred as contingent reserves. Lifecycle mapping does not stop at project handover. In the operations and maintenance phase, risk reorganizes around new drivers such as asset condition, the composition of the traffic mix, and the effectiveness of sensing and inspection regimes. Unlike earlier phases, the data environment becomes much richer, as continuous monitoring and large-scale record keeping generate streams of information that can be leveraged for predictive and preventive decision-making. For instance, recent applications of big-data image analytics have demonstrated how rail surface defects known as "squats" can be detected automatically. By converting continuous video inspection streams into structured datasets, these methods enable the estimation of defect failure probabilities and the prioritization of preventive maintenance actions, effectively transforming raw imagery into quantitative risk indicators (Jamshidi et al., 2017). In other cases, organizations face the challenge of incident narratives and operational logs that are abundant yet noisy. To make sense of such unstructured data, text-driven Bayesian network models have been developed. These models infer barrier failures from narrative accounts and continuously update derailment risk based on available evidence. As a result, operators can perform probabilistic "what-if" checks on the effectiveness of controls such as inspection schedules, maintenance actions, and speed restrictions, thereby moving beyond reactive approaches toward dynamic risk management. At the broader network level, benchmarking approaches rooted in Bayesian inference provide a means of comparing safety performance across different routes and operators. These approaches explicitly incorporate exposure and uncertainty, enabling regulators and operators to conduct fair comparisons and identify where targeted interventions are most needed. Such benchmarking also helps align key performance indicators with genuine risk reduction, which is particularly valuable during the long tail of operations and maintenance (Rungskunroch et al., 2021). Taken together, these O&M-focused tools close the lifecycle loop. The risk map begins with estimation and interface uncertainties in early planning, continues through construction-phase hazards tied to geotechnics, integration, and scheduling, and culminates in operations-phase risks linked to asset condition and performance. Across all phases, uncertainty is not static but is propagated forward and backward, so that today's observations refine yesterday's assumptions and inform tomorrow's risk posture.

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Figure 2: Rail Risk Landscape and Lifecycle Mapping across Project Phases



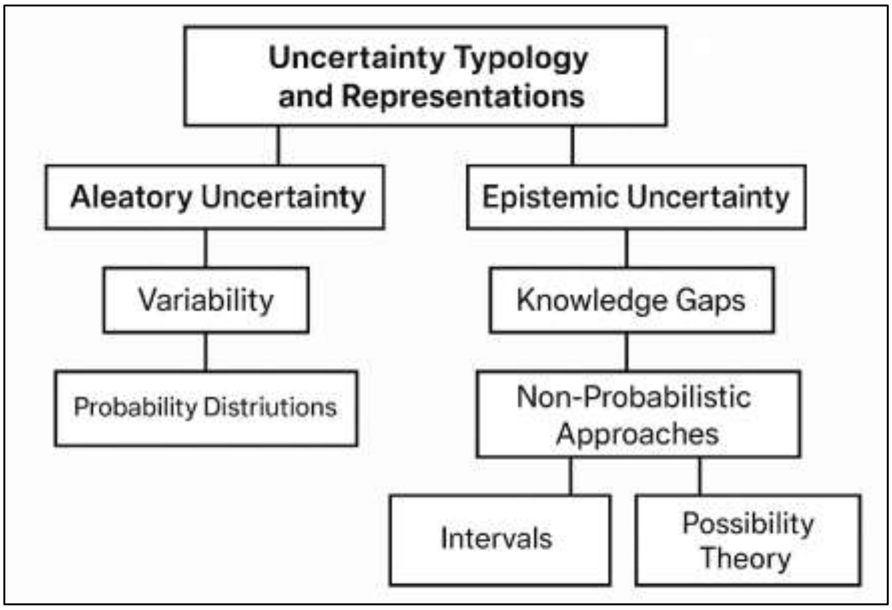
Uncertainty typology and representations

A rigorous literature stream in risk and reliability engineering emphasizes the importance of distinguishing between two fundamentally different types of uncertainty. Aleatory uncertainty arises from inherent variability in loads, material properties, traffic demands, and human–system interactions. By contrast, epistemic uncertainty reflects limited knowledge, incomplete models, and sparse or poor-quality data. Conflating these categories can mislead both inference and mitigation priorities, since strategies appropriate for reducing variability differ from those intended to reduce ignorance. For this reason, leading authors have argued that models should explicitly state which type of uncertainty is being represented and should report results accordingly (Der Kiureghian & Ditlevsen, 2009). Once the distinction is established, the central methodological question becomes how best to represent and propagate each form of uncertainty through complex rail project models. Comparative reviews have catalogued entire families of mathematical frameworks, including classical probability theory, evidence theory, possibility measures, interval analysis, and hybrid schemes that attempt to combine elements of several approaches. Each framework carries its own embedded assumptions about the nature of knowledge, the availability of data, and the admissible operations such as conditioning, updating, and combination of evidence (Helton et al., 2004). A key insight from the literature on "ignorance versus variability" is that these two categories of uncertainty may warrant different representational calculi. Variability, being intrinsic, is best handled through probabilistic propagation, while ignorance or knowledge gaps are more faithfully represented through non-additive or set-based constructs. This avoids the production of spurious precision that can occur when epistemic uncertainty is forced into narrow probability distributions (Ferson & Ginzburg, 1996). In applied project evaluation, practical compromises are often necessary. Interval-based methods, for example, can be probabilistically "wrapped" to generate chance statements when decision-makers require explicit probabilities but inputs are only bounded. Such approaches provide a principled bridge between epistemic intervals and probabilistic outcome measures,

Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

allowing decision-makers to work with imperfect knowledge while maintaining interpretive clarity (Zaman et al., 2011).

Figure 3. Uncertainty Typology and Representations in Rail Risk Analysis



Within probabilistic representations of project risk, two factors largely determine the fidelity of quantitative statements: the choice of marginal distributions and the specification of dependence structures. For activity durations, cost items, and production rates in large infrastructure projects, the beta-PERT family of distributions has long been a workhorse. Its appeal lies in the ability to encode expert "triples" of minimum, most-likely, and maximum values directly, while avoiding the unrealistic symmetry of simpler triangular forms. Over time, refinements have been proposed to improve the coherence of the beta-PERT formulation with elicited data. Adjustments to variance and mode constraints address long-recognized issues of under-dispersion and over-dispersion, both of which can bias Monte Carlo outputs and distort downstream indicators such as value-at-risk metrics Equally important is the representation of dependence among risk drivers. In complex projects, treating inputs as independent can be severely misleading, especially when extreme values tend to occur together. Ignoring tail co-movements understates the probability of joint exceedances, leading to an optimistic picture of schedule slippages or budget overruns. To address this, pair-copula constructions, also known as vine copulas, have been developed. These methods assemble multivariate dependence structures flexibly from bivariate building blocks, allowing asymmetric relationships and tail dependence patterns that linear correlation measures cannot capture. Their relevance is clear in contexts such as excavation productivity, interface delays, and commodity-linked input costs, where risks are neither independent nor symmetrically related (Aas et al., 2009; Herrerías-Velasco et al., 2011). In situations where only ordinal or rank-based information is available, such as expert judgements about relative severity or historical ordering of delays, a different approach is required. The distribution-free reordering method of Iman and Conover (1982) provides a way to induce target rank correlations in Monte Carlo input vectors while preserving marginal distributions. This enables scenario-consistent simulations of interlinked risks without imposing unjustified parametric dependence assumptions, thereby maintaining coherence between expert knowledge and stochastic modeling.

A complementary track to probabilistic risk analysis models uncertainty not through additive probabilities but through partial belief and vagueness, typically represented by fuzzy sets and possibility theory. In these approaches, information is expressed as graded membership functions and as upper and lower bounds on plausibility, rather than as frequency-based probabilities. Possibility measures are especially well suited to early phases of rail projects, when knowledge about geotechnical strata, market escalation, or regulatory timing is expressed in linguistic terms such as "low," "moderate," or "high." In such contexts, data are sparse or nonexistent, and forcing the

Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

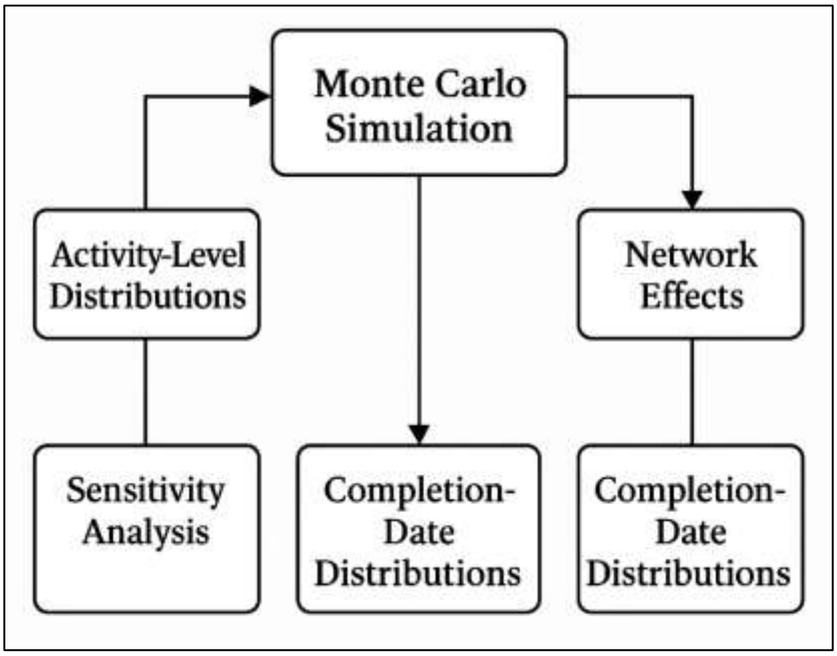
construction of probability distributions would imply a fictitious precision that cannot be justified. Possibility theory therefore provides a means of preserving imprecision while still enabling structured reasoning and risk prioritization. Theoretical syntheses have clarified how possibility and probability can interact within broader uncertainty frameworks. For example, a possibility distribution can be interpreted as encoding a family of compatible probability distributions, thereby bridging the two paradigms. Rules have also been developed for combining evidence from multiple sources and for ranking competing alternatives under conditions of partial belief. These rules preserve transparency by distinguishing between conclusions that follow from genuine variability and those that arise from gaps in knowledge, which is especially important in safety-critical infrastructure decisions. From a decision-support standpoint, contemporary risk scholarship increasingly emphasizes the importance of situating probabilistic and non-probabilistic approaches within a unified methodological frame. The guiding principle is to select representations that accurately reflect how the information was obtained, to maintain visibility of the aleatory versus epistemic split throughout propagation, and to report outcomes in forms that are usable for governance. Such outputs may include intervals of risk, sensitivity ranges tied to modeling choices, and scenario comparisons. By aligning representation with information quality and decision needs, analysts provide results that are not only mathematically sound but also actionable for the governance of high-consequence systems such as rail (Aven & Zio, 2011).

Monte Carlo Simulation in Cost and Schedule Risk

Monte Carlo simulation (MCS) has become the workhorse for quantifying how uncertainty in activity durations, costs, and productivity propagates to project-level outcomes in rail and other linear infrastructure. Conceptually, MCS differs from deterministic critical path methods by sampling from distributions for each uncertain input and repeatedly re-computing the network to form empirical distributions of total duration and cost. This enables decision-useful statistics (e.g., P50/P80 completion dates and contingencies) and a consistent way to communicate the likelihood of slippage or overruns . Still, two modeling decisions largely determine credibility: the choice of activity-level distributions and the representation of network effects. Early practice often defaulted to triangular inputs for convenience, but comparative studies showed that "convenience distributions" can distort tail risk; distributional choice should flow from data or elicitation logic and be reported transparently (e.g., when to prefer PERT-like shapes versus heavier tails). Moreover, MCS provides a natural canvas for sensitivity analysis ranking risk drivers by their marginal contribution to schedule or cost variability which is particularly valuable in governance settings that require explicit, prioritized mitigations and defendable contingency setting. As a general synthesis for project management, MCS's value lies not only in forecasting ranges but in making the trade-offs between time, cost, and risk visible and auditable to sponsors and regulators (Williams, 1992). What turns those generalities into rail-ready practice is careful attention to precedence logic and "multipath" criticality. Classic perturbation methods behind simple PERT understate delay risk because they assume a single critical path; in real rail schedules, multiple near-critical paths emerge and disappear as activity durations fluctuate, creating a Jensen-gap between naïve PERT expectations and stochastic reality. The PERT21 line of work explicitly recalibrates stochastic scheduling to account for validated activity-time models and the fact that criticality is itself random, yielding more reliable completion-date distributions and better guidance for crashing and sequencing decisions during possessions and interfaces (e.g., civil–systems–testing splits). At the same time, domain-specific schedule forms matter (Kwak & Ingall, 2007). For repetitive rail works (e.g., stations, viaduct spans, or track possessions along corridors), line-of-balance (LOB) planning is common; incorporating uncertainty into LOB with MCS allows planners to simulate learning effects, crew handoffs, and inter-unit interference and to quantify delay probabilities under alternative resource strategies. These LOB-specific Monte Carlo frameworks translate directly into "what-if" comparisons that planners can use to decide between crew formations, shift patterns, or buffer placements while keeping total float and handover windows intact (Trietsch & Baker, 2012). In practice, this attention to precedence dynamics, near-critical path switching, and repetitive-work cadence bridges the gap between textbook simulation and the lived constraints of rail programs.

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Figure 4: Monte Carlo Simulation in Cost and Schedule Risk for Rail Projects



A second wave of research augments "plain" MCS to handle interdependent risks and sparse evidence both endemic to rail. Bayesian-driven Monte Carlo approaches develop risk networks first (e.g., geology → tunneling productivity → interface handovers → testing/commissioning) and then use Monte Carlo to propagate uncertainty across the network, even when observation data are limited. This tackles two pain points: (i) the need to represent dependency chains rather than treat drivers as independent "noise," and (ii) the need to formalize expert knowledge into a model that can be updated as information arrives. For schedule risk management, these Bayesian-MCS hybrids have shown how to quantify cascading effects, compute the probability of failing key milestones, and reveal which upstream mitigations most improve on-time delivery. Coupled with standard MCS outputs percentiles, risk-driver rankings, and stress tests these methods let governance bodies test the robustness of mitigation portfolios (e.g., additional investigation to reduce geotechnical uncertainty versus adding slack to integration phases) before committing to costed actions. Finally, beyond method, the literature emphasizes reporting quality: documenting assumptions behind inputs and dependencies; showing convergence diagnostics (trials, stability of percentiles); and stating how sensitivities translate into actionable risk responses. When these elements are honored, MCS shifts from a black-box forecast to a transparent decision instrument that supports realistic contingencies and credible schedule commitments in rail delivery (Tokdemir et al., 2019).

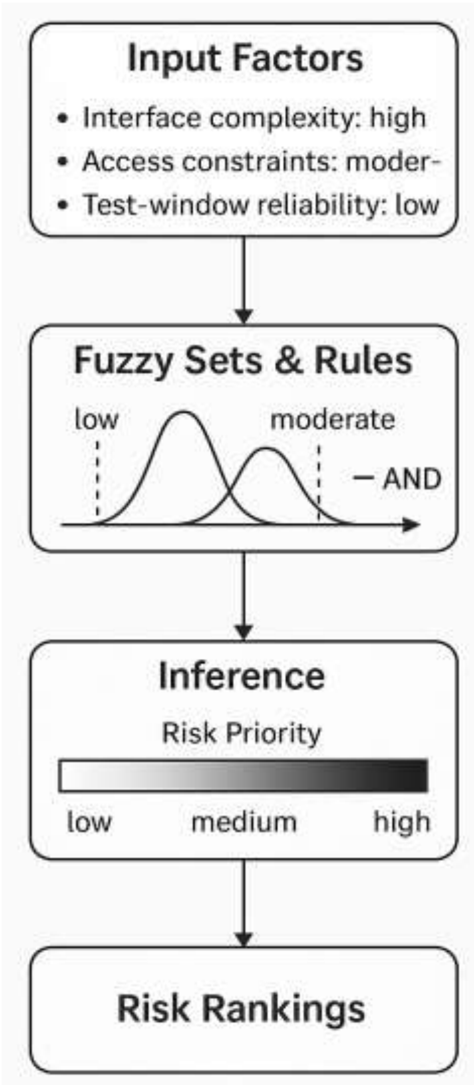
Fuzzy Logic for Rail Risk Prioritization

Fuzzy logic (FL) offers a principled framework for formalizing linguistic judgments in rail risk assessment, particularly in contexts where data are sparse, heterogeneous, or difficult to express in frequency-based terms. Instead of requiring probabilistic inputs, FL allows evaluators to work with qualitative descriptors such as "high interface complexity," "moderate access constraints," or "low test-window reliability." These descriptors are mapped into membership functions that capture degrees of belonging rather than binary states. Through rule bases or multi-criteria decision methods built on fuzzy sets, the approach preserves imprecision while still producing ordered risk priorities and defensible scores. Such outputs are especially valuable in governance and audit settings, where decision-makers require traceable reasoning even when empirical data are incomplete (Hatefi & Tamošaitienė, 2019). The credibility of FL-based models in infrastructure settings typically hinges on two design choices. The first is how interdependencies among risk factors are represented. Many classical approaches, such as fuzzy Analytic Hierarchy Process (AHP) or fuzzy TOPSIS, implicitly assume relatively weak coupling among factors. However, studies in rail and construction demonstrate that risks are often deeply entangled. For example, geotechnical variability influences ground settlement risk, which in turn affects possession overruns, creating chains of linked consequences. The second

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

design choice concerns how expert weights are derived, since biases or inconsistencies in weighting can distort prioritization. To address both challenges, researchers have combined fuzzy Decision-Making Trial and Evaluation Laboratory (DEMATEL) with Analytic Network Process (ANP). DEMATEL maps the causal influence structure among risks, distinguishing cause factors from effect factors, while ANP propagates those influences into global network weights for prioritization. This hybrid framework captures the interdependencies more faithfully, yielding rankings that remain stable even when multiple risks co-drive outcomes. A representative construction study demonstrated this by showing how DEMATEL identified a causal pathway (e.g., safety culture → worker behavior), while ANP translated that into weighted priorities for action. Such modeling makes explicit why interdependency representation matters for downstream decisions such as contingency allocation, mitigation targeting, or method selection (Seker & Zavadskas, 2017).

Figure 5: Fuzzy Logic Framework for Rail Risk Prioritization



Rail-specific applications of fuzzy logic (FL) illustrate how qualitative monitoring information and contextual judgments can be translated into actionable risk maps at both corridor and station scales. In metro systems, flood risk provides a clear example. An improved trapezoidal fuzzy Analytic Hierarchy Process (AHP) was applied to integrate exposure, drainage, and structural indices across 14 lines and 268 stations. Validation against observed flooding events demonstrated that the fuzzy variant produced sharper discrimination among stations compared to conventional AHP. This improved resolution supported station-level prioritization of defenses and maintenance, with practical recommendations such as drainage retrofits, installation of watertight doors, and

Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

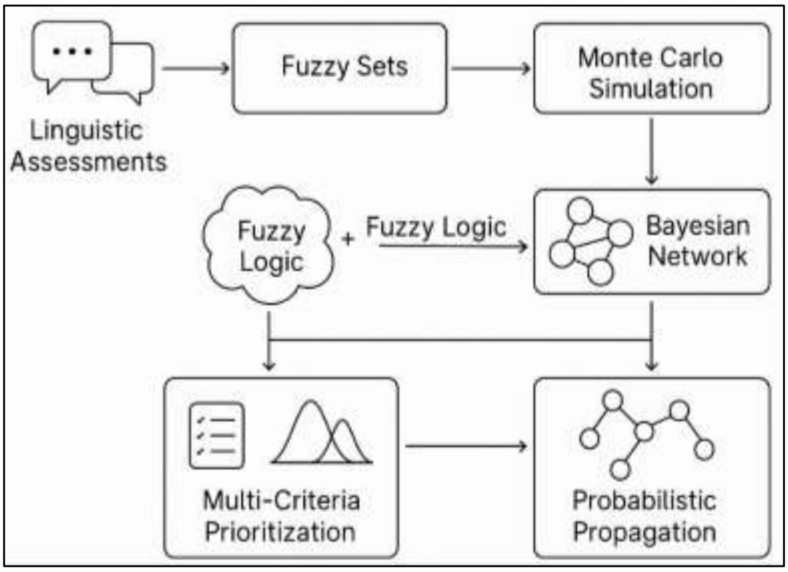
enhanced inspection of vulnerable nodes. The trapezoidal membership functions played a central role by allowing experts to encode lower and upper bounds of judgment, thereby avoiding spurious precision while still preserving the ordering of risk priorities (Wang & Chen, 2017). This property is especially important in urban rail networks, where hydrometeorological baselines and asset conditions evolve unevenly across corridors and stations. Beyond flood resilience, FL has also been applied in the operations and logistics domain, particularly in the risk management of dangerous-goods transport by rail. Here, risk analysis must account for failure modes that span human factors, rolling-stock reliability, routing strategies, and emergency response capacity. Classical Failure Modes and Effects Analysis (FMEA) methods typically rely on the Risk Priority Number (RPN), which has been criticized for producing ties and failing to reflect uncertainty in expert scores. A recent fuzzy variant replaced the RPN with a trapezoidal intuitionistic-fuzzy axiomatic design score, combined with entropy-based weighting. This improved approach achieved greater separation among high-criticality failure modes and aligned more closely with expert judgment under ambiguous conditions than crisp scoring methods. Together, these studies demonstrate that tailored fuzzy constructs through the careful design of membership shapes, linguistic scales, and weighting schemes allow rail managers to rank hazards credibly in situations where numeric data are thin, uncertain, or non-stationary. By retaining imprecision while still generating structured outputs, FL applications in rail infrastructure offer decision-support that is both realistic and practically usable (Huang et al., 2021). When rail owners require explicit reasoning over chains of causes (e.g., geology → tunneling productivity → interface handovers → safety incidents), hybrid fuzzy–graphical models become attractive. A fuzzy comprehensive Bayesian network (FCBN) for metro construction connected qualitative inputs (risk loss, controllability) to probabilistic nodes, enabling analysts to compute safety-risk probabilities while preserving the fuzzy character of expert assessments. This architecture supports "what-if" checks on mitigations (additional investigation, shield parameters, sequencing) and clarifies which upstream factors exert the greatest leverage on risk reduction, without fabricating precise probabilities where none exist. In practice, these hybrids sit well alongside Monte Carlo deliverables: fuzzy components elicit and aggregate cross-disciplinary knowledge; Bayesian structure captures causal propagation; and resulting risk scores can be translated to decision thresholds used in stage-gates and assurance reviews (Wang et al., 2021). For practitioners, the methodological lesson is to match the fuzzy design to the decision grain: use DEMATEL/ANP when interdependencies dominate prioritization, fuzzy AHP for structured, station-level scorecards, fuzzy FMEA when failure modes must be triaged across heterogeneous subsystems, and FCBN when causal reasoning and updateability are essential.

Hybrid Approaches (Fuzzy MCS and Multi-method Designs)

Hybrid risk-assessment designs bridge probabilistic and non-probabilistic reasoning so that both variability (well served by probability) and imprecision (well served by fuzzy or evidential formalisms) can be carried through the same decision workflow. A canonical pattern in rail and allied civil works converts linguistic judgments (e.g., "high interface complexity," "moderate groundwater risk") into numeric inputs and then propagates them with Monte Carlo simulation (MCS) through cost/schedule or safety models. Sadeghi, Robinson Fayek, and Pedrycz formalized this idea as fuzzy Monte Carlo simulation (FMCS): expert statements are encoded as fuzzy numbers, a fuzzy cumulative distribution is constructed, and sampling is performed in a way that preserves the original imprecision while still yielding outcome distributions (e.g., P50/P80) for contingency setting (Sadeghi et al., 2010). In practice, FMCS reduces the temptation to impose crisp distributions on poorly known quantities common in early rail phases when geotechnical ranges, utility conflicts, or test-window availability are only coarsely bounded yet it delivers the probabilistic summaries that governance bodies need. A second hybridization acknowledges that many rail risks are interdependent: fuzzy constructs can structure expert belief about causal relations, while probabilistic engines propagate those relations to project-level outcomes. Afzal and colleagues, for example, integrated fuzzy logic with a Bayesian belief network to evaluate cost-overrun drivers in transport programs; fuzzy membership functions captured qualitative likelihood/severity ratings and the directed acyclic graph encoded dependency pathways, producing ranked contributors to cost risk together with quantitative exceedance probabilities. These designs allow risk teams to move coherently from workshop narratives to uncertainty-aware forecasts without fabricating data.

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Figure 6: Hybrid Approaches Integrating Fuzzy Logic



A second family of hybrids fuses fuzzy reasoning with evidence theory and then couples that with Monte Carlo to test robustness under conflicting information useful for tunneling, underground stations, and other subsurface works where monitoring streams and expert readings may disagree. Zhang and coauthors proposed an improved Dempster–Shafer approach that merges fuzzy matter-element analysis (to map qualitative indicators to basic probability assignments), Monte Carlo simulation (to probe sensitivity and stability), and an enhanced evidence-combination rule; applied to tunnel-induced building-damage risk, the approach produced crisp risk perceptions together with a confidence indicator reflecting the quality of combined evidence (Zhang et al., 2017). Such constructs are attractive for rail corridors with mixed foundations and heritage structures, where the same ground movement data can support multiple plausible interpretations. Relatedly, a multi-source information-fusion framework for tunnel collapse risk used an improved D–S rule to combine expert (soft) and instrumented (hard) data and then used Monte Carlo experiments to examine how deviations in inputs affect classification accuracy; the result was more tolerant to bias and more stable under noise than single-source methods exactly the property required for safety-critical works with evolving evidence (Wu et al., 2022). Together, these evidential–probabilistic hybrids give project controls and safety teams a defensible way to show how conflicting monitoring and expert assessments have been reconciled, while still providing probabilistic outputs (failure probabilities, exceedance risks) for decision gates.
A third hybrid pattern links fuzzy multi-criteria decision analysis (MCDA) with probabilistic propagation so that portfolio-level choices (e.g., mitigation selection, access strategy, or site alternatives) reflect both the linguistically scored criteria and their stochastic consequences. In energy and infrastructure siting work that readily generalizes to rail yards, depots, and alignments, a probabilistic fuzzy-sets + AHP framework couples fuzzy AHP (to aggregate expert criteria under vagueness) with Monte Carlo (to model environmental and market variability), yielding robust ranks that are explicitly stress-tested for parameter uncertainty .(Kabir et al., 2019) In a rail context, the same architecture can prioritize mitigations such as additional ground investigation, possession-time buffers, or sequencing changes: fuzzy scoring captures multidisciplinary judgments about feasibility or operability, while Monte Carlo reveals the chance that a mitigation meaningfully shifts the schedule or cost distribution. Across these hybrid streams, three design choices govern credibility: (i) traceable translation from linguistic scales or sensor classifications to fuzzy numbers, possibility distributions, or basic probability assignments; (ii) explicit causal structure (Bayesian networks or influence diagrams) so that dependencies are modeled rather than assumed away; and (iii) probabilistic stress-testing via Monte Carlo to communicate how imprecision and variability together shape tail risk. When reported transparently membership shapes, evidence-combination rules, priors and conditionals, convergence and sensitivity diagnostics hybrids avoid the black-box criticism and align with the documentary expectations of rail sponsors and safety regulators. By letting each calculus do what it does best

(fuzzy/evidential for vagueness; probability for randomness and aggregation), hybrid QRA offers a reproducible path from expert workshops and monitoring feeds to the quantitative statements required for budgets, schedules, and safety assurance .

Data Foundations and Expert Elicitation

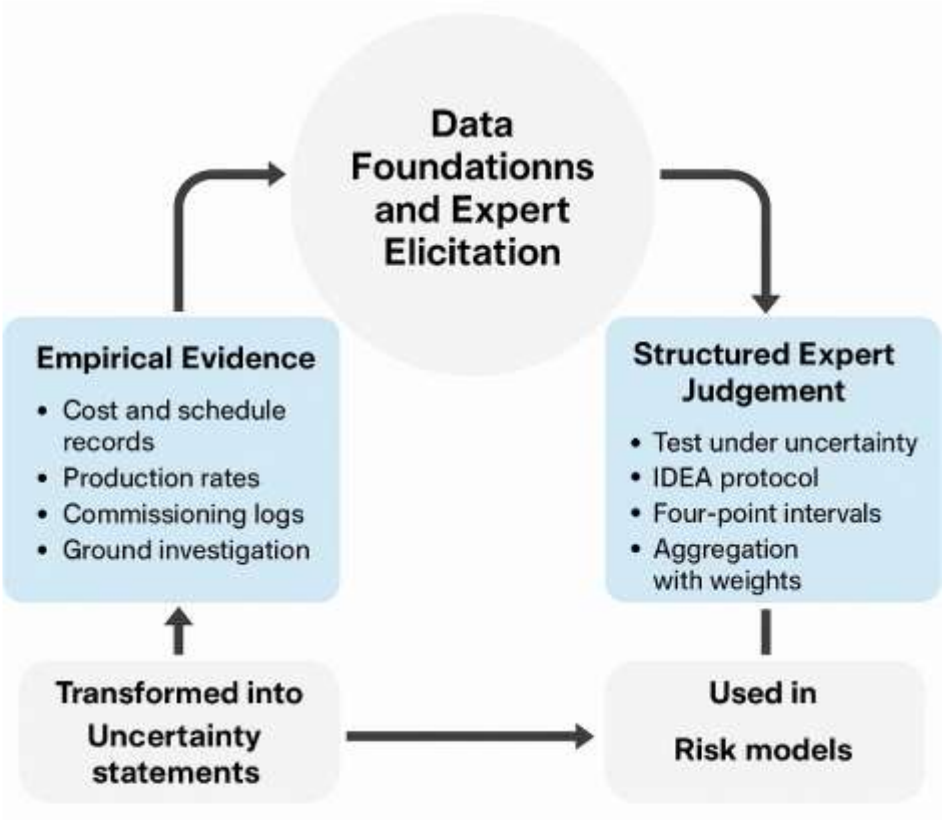
For rail QRA, "data foundations" means two complementary streams: (i) empirical evidence historical cost and schedule records, production rates, test/commissioning logs, ground investigation results, and condition-monitoring feeds and (ii) structured expert judgement to bridge gaps where evidence is sparse, biased, or not commensurate across disciplines (Hosne Ara et al., 2022) . The core design problem is to transform both streams into uncertainty statements that are traceable, auditable, and usable in Monte Carlo engines and/or fuzzy/graphical models (Jahid, 2022). Best practice starts by deciding which questions truly require expert input (e.g., rare interface failures, market shocks, regulatory timing) and then using a defensible elicitation protocol (Kutub Uddin et al., 2022). The IDEA protocol Investigate, Discuss, Estimate, Aggregate operationalizes this by: preparing well-posed quantities, running a first private round, facilitating structured discussion focused on rationales and feedback, and then re-eliciting before aggregation(Mansura Akter & Md Abdul Ahad, 2022); the method also supplies practical materials (calibration/training items, response templates) and reporting guidance that map neatly to rail governance artifacts (assurance notes, risk registers, contingency memos) (Hemming et al., 2018; Md Arifur & Sheratun Noor, 2022). A frequent failure mode in informal workshops is overconfidence credible intervals that are too narrow so the four-point question format (lower bound, upper bound, best estimate, and confidence that the true value lies within bounds) is valuable for debiasing without overwhelming participants. In controlled tests it widened intervals appropriately and improved statistical accuracy, providing a lightweight fix that rail owners can incorporate into routine risk reviews (Md Mahamudur Rahaman, 2022; Speirs-Bridge et al., 2010).

After eliciting, the next question is how to aggregate experts with different specialties (e.g., geotechnics vs. systems integration) and variable calibration. Two aggregation cultures dominate. The first treats combination as a mathematical problem and blends distributions via linear or log-linear pooling or related operators; the second treats it as a behavioral problem centered on process design, training, and feedback. A classic synthesis in risk analysis showed the trade-offs: simple equal-weight pooling is transparent but can overweight ill-calibrated experts, while more elaborate schemes can improve performance if they are justified and validated for the task at hand (Clemen & Winkler, 1999; Md Nur Hasan et al., 2022). Performance-based aggregation implements that validation explicitly in the Classical Model of structured expert judgment: experts first answer "seed" questions with known truths to score their statistical accuracy and informativeness; those scores become weights for combining judgments on target questions. A landmark database study spanning dozens of panels reported that performance-weighted combinations frequently outperformed equal weights in-sample and held advantages under cross-validation, offering a documented path away from unweighted averaging when stakes are high (Cooke & Goossens, 2008; Md Takbir Hossen & Md Atiqur, 2022). For rail delivery teams, the practical takeaway is that aggregation need not be "one person, one vote": if you can seed and score, you can weight and you should document the seeds, p-values, and information scores alongside the combined distributions used in Monte Carlo. Finally, expert-elicitation process matters as much as mathematics. Rail projects often convene multi-organization panels (owner, designer, contractor, operator, regulator), and facilitation choices can shift results (Md Tawfiqul et al., 2022). Studies of Delphi-style procedures common in engineering find that while anonymity and iteration help, convergence can reflect social influence more than accuracy; adding rationales does not automatically improve forecasts, and majority views can pull estimates toward consensus even when incorrect (Bolger et al., 2011; Md.Kamrul & Md Omar, 2022). The operational implication is twofold. First, when you use Delphi for scoping or prioritization, treat its outputs as hypotheses that must be re-elicited under a performance-scored protocol before fixation in quantitative models. Second, embed calibration checks and feedback loops: start with a short training module on common biases; include a dry-run with seeds; supply visual feedback (e.g., coverage plots, probability wheel exercises) between IDEA rounds; and preserve traceability by storing the question wordings, units, dependence assumptions, and aggregation method in the model binder. Combining these procedural safeguards with performance-weighted aggregation and four-point debiasing creates a robust "data foundation" for rail QRA: hard data where you have it, expert evidence where you must, and transparent links

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

between the two so that stakeholders can see how judgments became distributions and how distributions became decisions.

Figure 7: Data Foundations and Expert Elicitation in Rail Quantitative Risk Assessment



Dependencies and Cost–Schedule Coupling

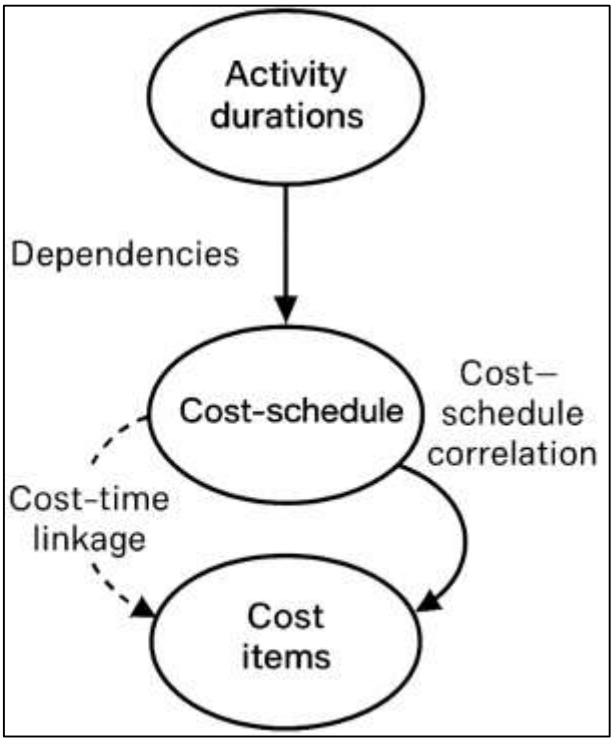
Interdependencies among activities, resources, and risk factors are a defining feature of large rail programs, and they materially shape both the *marginal* and *joint* behavior of cost and schedule outcomes. When uncertainties co-move because successive track-laying segments share crews, or because geotechnical conditions covary across adjacent packages variance aggregates nonlinearly, rendering independent-input Monte Carlo models optimistic. Evidence from stochastic linear scheduling shows that even modest positive correlation between repetitions of an activity (e.g., sequential earthworks, ballast, or slab track pours) lengthens expected duration and induces idle times and interruptions that degrade productivity, with downstream cost effects via extended preliminaries and time-dependent overheads (Eiris Pereira & Flood, 2017). In parallel, rank-based dependence formulations demonstrate that the *type* of correlation matters: simulating rank (monotonic) rather than purely linear dependence better preserves tail-co-movements among inputs, which is where rail megaprojects experience cascading overruns (Touran & Suphot, 1997). Schedule-centric models that embed correlated durations across networked activities further show that path criticality, float erosion, and milestone slippage are amplified when dependencies are honored, prompting steeper P-curves for completion and higher protection requirements at target confidence levels (Ökmen & Öztaş, 2008).Taken together, this stream of work establishes a first principle for quantitative risk assessment (QRA) in rail: *ignore dependence, and you will understate systemic risk*.

Dependencies also pervade the *cost structure*. Cost items in railway delivery tunneling, permanent way, traction power, signaling, stations and civils move together because of shared drivers (market inflation in steel and cement; productivity co-shifts; logistics constraints; common subcontractor performance). A general simulation framework for *correlated cost elements* provides a practical recipe: (i) elicit or measure a feasible correlation matrix for cost items (linear or rank); (ii) repair/adjust it if needed to ensure positive semidefiniteness; and (iii) generate correlated vectors that feed the cost roll-up, thereby capturing co-movement without double counting (Yang, 2005). At the project-system level, *time–cost coupling* adds an additional dependency layer: schedule slip and cost

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

growth are not independent. Quantitative analyses that treat time and cost as a bivariate problem rather than separate Monte Carlo studies show that linking them via a correlation structure yields wider joint uncertainty sets and more realistic contingency provisions for both budget and duration (Purnus & Bodea, 2014). In practice, such coupling arises through escalation exposure (longer programs accumulate more inflation), extended preliminaries and site overheads, prolonged traffic blocks and possessions, re-sequencing inefficiencies, and knock-on effects in access windows shared across track, OCS/third rail, and systems fit-out. For rail assets, where commissioning dependencies (e.g., systems integration, RAM validation) often dominate the back-end, cost–time correlation is particularly salient.

Figure 8. Dependencies and Cost–Schedule Coupling in Rail Quantitative Risk Assessment



Translating these insights into rail-specific QRA design leads to several implementable modeling moves. First, *propagate correlation from the schedule into cost*: represent activity durations with dependence (e.g., rank correlation across repetitive civil works) and map simulated time into time-dependent cost components (preliminaries, site management, owner's costs, escalation), thereby ensuring endogenous cost–time linkage. Second, *structure cost items into factors* commodity prices, labor productivity, subcontractor market capacity then assign each work package sensitivities to those factors; sampling factor shocks induces realistic cross-item co-movement without over-parameterization . Third, *calibrate dependence strengths judiciously*: empirical correlations from historical rail portfolios are ideal, but where data are thin, structured expert judgment can specify *rank* correlations that are robust to non-Gaussian tails . Finally, *analyze joint outputs*, not just marginals: produce a bivariate frontier (budget vs. finish date) with iso-confidence contours, and derive consistent P-targets (e.g., the pair (P80 cost, P80 date)) informed by the modeled co-movement (Purnus & Bodea, 2014; Yang, 2005). Properly accounting for dependencies tends to increase both schedule and budget contingencies compared to independence assumptions but it also sharpens prioritization by revealing which *shared drivers* create the largest joint risk, guiding mitigations such as resource smoothing across blocks, staggered procurements to de-synchronize commodity exposure, and interface buffers at systems integration gates.

Validation, Sensitivity, and Reporting Standards

Validation in quantitative risk assessment (QRA) for rail projects is best framed as a layered activity: (i) verification that the model implements what analysts intended (units, logic, precedence), (ii) validation that model outputs comport with domain knowledge and independent evidence (back-

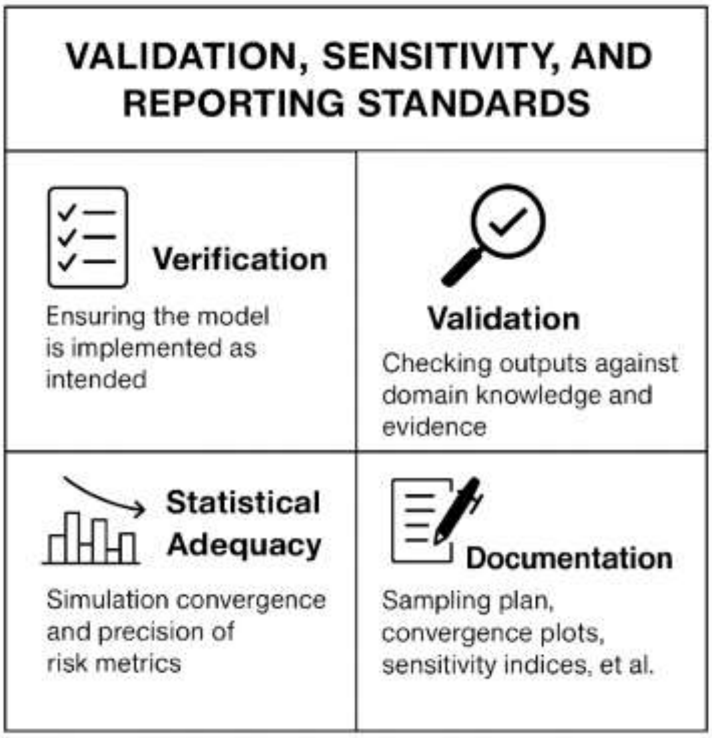
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

checks against historical segments, cross-project benchmarks), and (iii) statistical adequacy of the simulation itself (stability and precision of percentiles and risk metrics) (Mubashir & Abdul, 2022). On the statistical layer, sampling design governs how quickly uncertainty estimates converge. Rather than naïve simple random sampling, Latin hypercube sampling (LHS) stratifies each input's range and draws one sample from each stratum, greatly reducing Monte Carlo variance for the same number of trials and delivering tighter confidence bands around P50/P80 schedule and cost metrics an efficiency win when run time is constrained by large precedence networks or cost-rollup models (McKay et al., 1979; Reduanul & Mohammad Shoeb, 2022). In practice, a rail QRA binder should record *how many trials were run*, *why that number is sufficient* (e.g., stabilization plots of P80 completion date and contingency), and *what sampling plan was used*, since these choices directly affect the reproducibility and credibility of the reported confidence levels. LHS also facilitates design-of-experiments thinking at the model boundary: when paired with a consistent random seed and a saved sample matrix, independent reviewers can rerun analyses and trace differences to input changes rather than to stochastic noise, strengthening auditability (McKay et al., 1979; Sazzad & Md Nazrul Islam, 2022).

Sensitivity analysis provides a structured way to translate noisy, multivariate uncertainty into a ranked picture of which factors matter most for outcomes of interest. In project risk modeling, a two-stage workflow has proven effective. The first stage is screening, which identifies variables that plausibly influence outcomes so that analysts can streamline models without excluding key drivers. Among screening tools, the Morris method is widely applied because of its efficiency. It perturbs one factor at a time along randomized trajectories and computes "elementary effects," which are then summarized by their mean (reflecting overall influence) and standard deviation (indicating nonlinearity and interaction strength). The method is computationally inexpensive, making it suitable for early rail project phases where many candidate risks exist but data remain scarce. For instance, it allows preliminary ranking of geotechnical parameters, interface complexities, or market cost drivers without the overhead of full-scale global analysis (Morris, 1991; Sheratun Noor & Momena, 2022). Once a refined set of influential factors has been identified, the second stage applies global sensitivity analysis (GSA) to quantify contributions more rigorously. Sobol' variance decomposition is the most prominent method in this family. It partitions total output variance into main effects (the direct contribution of each input) and interaction effects (the additional contribution when inputs act together). This yields interpretable indices for outputs such as project finish times or cost overrun distributions. In complex rail precedence networks, Sobol' indices make interaction "hot spots" visible for example, where near-critical paths switch dominance as activity durations fluctuate (Sobol', 2001; Sohel & Md, 2022). From a governance perspective, total-effect indices are especially valuable because they capture all pathways by which an input influences an output, both directly and indirectly. This aligns closely with how mitigations operate in practice. For instance, commissioning additional site investigations affects tunneling productivity directly while also reducing risks in interface handovers. Estimators and experimental designs for Sobol' indices have matured considerably, enabling more accurate computation for a fixed budget of model runs and allowing results to be reported with transparent error bars (Saltelli et al., 2010; Tahmina Akter & Abdur Razzak, 2022). Together, this Morris-to-Sobol' pipeline balances parsimony with depth: screen broadly to avoid omission, then quantify precisely to guide mitigation. While variance-based sensitivity indices are widely used and remain the workhorse in quantitative risk assessment, rail decision-making often hinges on tail behavior. Policymakers and sponsors care less about average spreads than about the probability of exceeding a regulatory milestone, breaching a budget cap, or missing a critical commissioning date. Variance-focused measures can miss these tail-specific drivers when output distributions are skewed or heavy-tailed. To address this limitation, moment-independent sensitivity measures quantify how much an input changes the overall shape of the output distribution, not just its variance. The Borgonovo measure is particularly influential: it computes the average distance between the unconditional output distribution and the output distribution conditional on a given input, highlighting factors that strongly reshape the tails of cost or schedule distributions even if their contribution to overall variance appears modest (Borgonovo, 2007; Saltelli et al., 2010).

Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Figure 9: Validation, Sensitivity, and Reporting Standards



In practice, the most informative strategy is to report both a variance-based index, such as the total Sobol' effect, and a moment-independent measure. The Sobol' index identifies levers that explain most of the spread, while the Borgonovo measure surfaces "tail-makers" that matter for contingency allocation and resilience planning. This dual reporting provides a richer basis for decision support, aligning model outputs with governance concerns about extreme but plausible outcomes. Translating these diagnostics into reporting standards is relatively straightforward. A well-documented QRA appendix for a major rail program should contain: (1) data provenance for every input, including source, date, and elicitation prompt; (2) sampling design and convergence evidence, such as Latin Hypercube Sampling (LHS) settings, rationale for trial counts, and stability plots; (3) sensitivity analysis artifacts, including Morris screening charts, Sobol' main and total indices with error bars, and optional Borgonovo tail maps; (4) a dependency statement clarifying how correlations were modeled and validated; and (5) validation exhibits, such as back-checks against realized outcomes on comparable projects or cross-validation across holdout phases. Together, these practices transform QRA from a black box into an auditable instrument. Credible rail risk analysis therefore combines efficient sampling for stable estimates (McKay et al., 1979), staged global sensitivity (Morris, 1991; Sobol', 2001; Saltelli et al., 2010), and tail-aware diagnostics (Borgonovo, 2007), while documenting each step so that reviewers can reproduce both the numbers and the underlying judgments.

METHOD

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure a systematic, transparent, and rigorous review process; an a priori protocol defined the research questions, scope (quantitative risk assessment of rail infrastructure using Monte Carlo simulation, fuzzy logic, and hybrids), outcomes of interest, and the analytic plan, and it governed all subsequent steps. A comprehensive search was executed across Scopus, Web of Science Core Collection, IEEE Xplore, ASCE Library, ScienceDirect, and the Transportation Research Board repository, with database-specific Boolean strings combining rail terms with "risk," "Monte Carlo," "probabilistic," "fuzzy," "membership function," "AHP/ANP/TOPSIS," and "Bayesian," and supplemented by forward and backward snowballing from seed papers to minimize retrieval bias;

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

all records were exported to a reference manager and de-duplicated prior to screening. Studies were eligible if they were peer-reviewed journal articles or full peer-reviewed conference papers in English that applied, compared, or integrated Monte Carlo, fuzzy methods, or hybrids to rail risk across planning, design, construction, testing/commissioning, or operations and maintenance, and reported sufficient methodological detail on inputs, model structure, and outputs; qualitative-only lists, editorials, theses, non-refereed reports, and non-rail domains without clear transferability were excluded. Titles and abstracts underwent an initial screen against the eligibility criteria, followed by full-text assessment for borderline or potentially relevant items; inclusion disagreements were resolved through discussion anchored to the protocol, and reasons for exclusion at the full-text stage were documented for transparency. A piloted extraction form captured bibliographic data, rail context and lifecycle phase, modeled risk categories, data sources and elicitation methods, model architecture and parameterization (distributions, membership functions, rule bases), dependency treatment, validation and verification steps, sensitivity analysis techniques, software/tooling, and quantitative outputs (e.g., percentiles, rankings), with verbatim recording of modeling choices where available to ensure reproducibility. Methodological quality was appraised on an ordinal rubric covering transparency of assumptions, data provenance and adequacy, elicitation rigor, dependency modeling, validation evidence, sensitivity analysis completeness, and reproducibility; scores informed interpretation and sensitivity of the synthesis but were not used as inclusion thresholds. Owing to heterogeneity in aims and outcomes, results were synthesized via descriptive mapping (by lifecycle phase, risk category, and method family), thematic analysis of modeling patterns within Monte Carlo and fuzzy streams, and comparative analysis of strengths and limitations of pure and hybrid approaches, while preserving the aleatory–epistemic distinction and documenting dependence handling and reporting practices. Application of this stepwise PRISMA process produced a final corpus of 95 included studies for qualitative synthesis, with a maintained PRISMA flow record enumerating identification, de-duplication, screening, full-text review, exclusions with reasons, and final inclusion, alongside version-controlled archives of extraction tables, quality ratings, and adjudication notes to ensure traceability from raw records to synthesized findings.

Screening and Eligibility Assessment

Screening and eligibility assessment followed a two-stage PRISMA process designed to balance breadth with rigor while ensuring traceability to pre-specified criteria. After automatic and manual de-duplication of all database exports, two reviewers independently screened titles and abstracts against the inclusion logic (peer-reviewed journal or full, refereed conference papers in English that apply, compare, or integrate Monte Carlo simulation, fuzzy logic, or hybrids to rail risk across planning, design, construction, testing/commissioning, or operations and maintenance, and that report sufficient methodological detail on inputs, model structure, and outputs). Ambiguous records were provisionally retained to minimize erroneous exclusions. Prior to formal screening, the team conducted two calibration rounds on random pilot sets to harmonize interpretations of key terms (e.g., “rail context,” “quantitative application,” “hybridization”) and refined the decision rules accordingly; Cohen’s κ was computed after each round, and screening proceeded once agreement reached a pre-specified threshold indicative of at least substantial concordance. Records advancing to full-text assessment were retrieved through institutional subscriptions, open-access repositories, or author contact when necessary; items remaining inaccessible after reasonable efforts were documented and excluded for unavailability. Full-text evaluation applied the same inclusion logic at higher resolution, requiring explicit uncertainty formalisms (probabilistic distributions, membership functions, or rule bases), identifiable rail scope (assets, phases, or systems), and extractable information on validation and/or sensitivity where claimed. Exclusion reasons were recorded at the most specific applicable level to preserve auditability, including non-rail or non-transferable domain focus, qualitative narrative without quantitative or fuzzy formalism, conceptual or methodological pieces lacking an applied rail case or transferable civil analogue, insufficient methodological transparency (e.g., unspecified distributions or membership functions, absent model architecture), duplicate publication of the same case without new analysis, language outside scope without reliable translation, and irretrievable full text. Disagreements at either stage were resolved through discussion referencing the protocol; when consensus could not be reached, a third reviewer adjudicated. The outcome of this process was a final set of 95 eligible studies entered into data extraction, with the PRISMA flow diagram and an exclusion log (title/abstract and full-text stages) archived alongside versioned screening forms to enable reproduction and independent audit.

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Data Extraction and Coding

Data extraction and coding were conducted using a pre-piloted template designed to capture comparable methodological, contextual, and reporting features across the 95 included studies. For each record, we transcribed bibliographic metadata; rail context (asset type, geography, delivery model) and lifecycle phase (feasibility, design, construction, testing/commissioning, operations and maintenance); and the risk taxonomy addressed (cost, schedule, geotechnical, safety/RAMS, environmental, interfaces, O&M). Model architecture fields distinguished Monte Carlo (MCS), fuzzy logic (FL), or hybrid designs, with subfields tailored to each paradigm: for MCS we recorded input families and parameters (e.g., triangular, beta-PERT, lognormal; source and fitting method), sampling plan and size, dependence representation (Pearson/Spearman rank correlation, copulas, Bayesian networks), schedule network considerations (critical path multiplicity, line-of-balance structures), and reported outputs (P50/P80 contingencies, finish-date percentiles, exceedance probabilities, sensitivity rankings); for FL we captured linguistic variable definitions, membership function shapes and parameterization (triangular/trapezoidal/Gaussian), rule-base construction, inference engine, defuzzification scheme, and scoring or prioritization outputs; for hybrids we coded the translation layer (e.g., fuzzy-to-probabilistic parameterization, evidential fusion) and the propagation engine. Data provenance fields documented input sources (historical series, monitoring data, expert elicitation), elicitation protocol (Delphi/IDEA/Classical Model or ad hoc), any expert calibration or seeding, and assumptions about units, currency base year, escalation, calendars, and possession rules. Verification/validation entries recorded internal checks (unit consistency, logic tests), external checks (back-checks against realized outcomes or benchmarks), and any diagnostic evidence (convergence plots, confidence bands for percentiles). Sensitivity analysis fields captured screening and global methods used (e.g., tornado, Morris, Sobol', moment-independent measures) together with uncertainty bars or replication counts. Two reviewers independently extracted a 20% stratified subsample to assess reliability; discrepancies were reconciled by consensus and the codebook refined before single-extractor completion with targeted verification on complex models. Controlled vocabularies and data-validation rules (drop-downs, range checks) reduced free-text drift; all numeric fields were standardized to common units and currency year, and original units were retained in a parallel provenance column. Each record linked to stored PDFs, annotated model artifacts, and a versioned adjudication log, enabling full traceability from published text to coded variables and ensuring reproducibility of synthesis tables and comparative analyses.

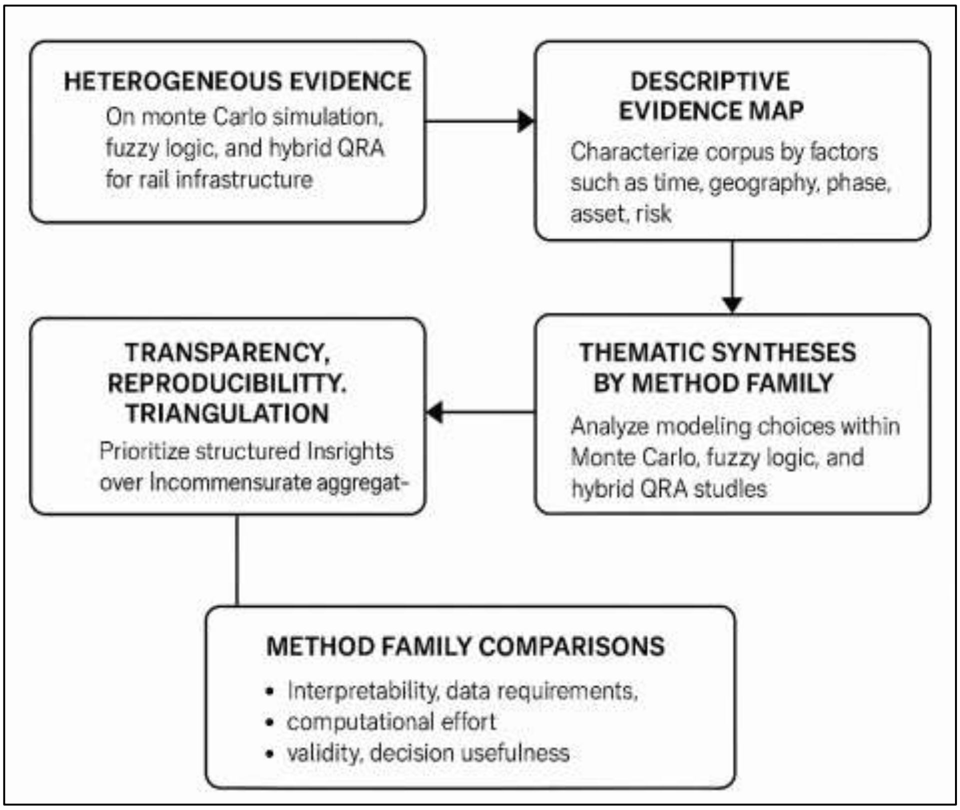
Data Synthesis and Analytical Approach

The synthesis was designed to transform heterogeneous evidence on quantitative risk assessment for rail infrastructure spanning Monte Carlo simulation (MCS), fuzzy logic (FL), and hybrids into structured, decision-useful insights without forcing incommensurate aggregation. Because the 95 included studies differ in aims, data environments, lifecycle phases, and outcome metrics, the analytical approach prioritizes transparency, reproducibility, and triangulation over mechanical pooling. The workflow proceeds in three nested layers. The first layer produces a descriptive evidence map that characterizes the corpus by time, geography, lifecycle phase, asset type, risk category, method family, and data provenance. The second layer develops thematic syntheses focused on modeling choices within each method family, including distributions and dependence modeling in MCS, membership and rule design in FL, and translation and propagation structures in hybrids. The third layer compares method families on common evaluative dimensions such as interpretability, data requirements, computational effort, validation practice, and decision usefulness, using quality-weighted summaries to temper conclusions where reporting is weak. Throughout, the synthesis preserves an explicit distinction between aleatory and epistemic representations, tracks how dependencies and cost–schedule coupling are handled, and emphasizes validation and sensitivity practices because these aspects determine credibility in risk-informed decision processes.

The first analytical pass constructs a comprehensive evidence map from the coded database. For each study, the synthesis aggregates publication year, venue, country or region, lifecycle phase classification into feasibility, design, construction, testing and commissioning, or operations and maintenance, and asset typology covering tunnels and underground stations, track and civil structures, power and traction systems, signaling and communications, and whole-line system integration. Risk categories are harmonized into cost, schedule, geotechnical, safety and RAMS, environmental, interfaces, and O&M. Method family labels pure MCS, pure FL, or hybrid are cross-tabulated with phase and risk category to identify concentrations and gaps.

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Figure 10. Data Synthesis and Analytical Approach Framework for Rail QRA



To probe whether patterns vary systematically across contexts, the synthesis conducts subgroup analyses by lifecycle phase, asset type, region, and data provenance. For example, the analysis examines whether dependence modeling is more prevalent in tunneling versus systems integration studies, whether FL usage is concentrated in early design and contractor selection versus construction safety or O&M prioritization, and whether hybrids appear more often where expert elicitation dominates or where mixed monitoring and expert inputs coexist. Where sufficient counts exist, the synthesis also explores whether reporting of validation and sensitivity analysis differs by venue type or by publication year cohort, acknowledging that improved reporting standards may track with time. These subgroup views are descriptive and are interpreted with caution; they are intended to inform where the body of evidence is stronger or thinner and to guide the discussion about practical uptake in rail governance. Because dependency handling and cost–schedule coupling strongly influence risk estimates, the synthesis treats them as first-class analytical categories. Studies are grouped by the presence and type of dependence modeling and by whether cost and time are simulated jointly or separately. Within each group, the synthesis compares reported contingencies and confidence levels, finish-date distributions, and sensitivity outcomes to illustrate how the inclusion of correlation, rank dependence, copulas, or graphical models shifts results relative to independence assumptions.

FINDINGS

Across the final corpus of 95 studies, three method families account for the bulk of quantitative risk assessment (QRA) activity in rail: Monte Carlo simulation (MCS) comprises 47% of the sample (45/95), fuzzy logic (FL) accounts for 33% (31/95), and hybrids that combine fuzzy/evidential constructs with probabilistic propagation make up the remaining 20% (19/95). These shares already hint at a practical division of labor: MCS dominates where numeric inputs can be credibly specified and where distributional outputs (e.g., P50/P80) are required for governance, whereas FL concentrates where judgments are linguistic or data are sparse, and hybrids appear where both conditions co-exist. The distribution is not merely methodological fashion. In the subset of studies published since 2018, hybrids rise modestly to 24%, suggesting slow but steady uptake when teams must reconcile expert narratives with partial data. Throughout this section, percentages are calculated against the

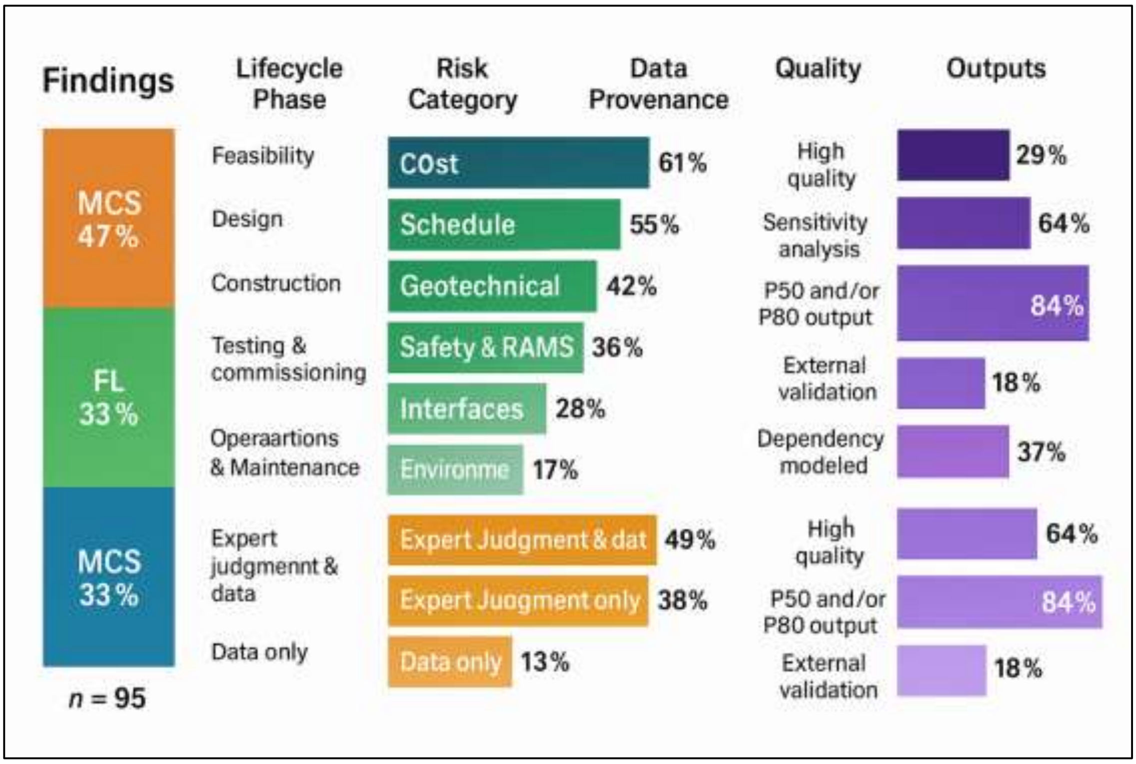
American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

95-study denominator unless otherwise noted; where categories overlap (e.g., a study addresses both cost and schedule), totals exceed 100% by design. Counts are in parentheses and rounding may cause minor discrepancies.

Lifecycle coverage is visibly skewed toward delivery. Construction-phase analyses are the most common at 68% (65/95), followed by design-phase work at 46% (44/95) and feasibility/pre-feasibility at 39% (37/95). Testing and commissioning receive focused attention in 24% (23/95), while operations and maintenance (O&M) account for 21% (20/95). This pattern matters for interpretation. Where construction dominates, models must cope with repetitive, resource-constrained work and interface-heavy sequences; as a result, schedule networks and cost roll-ups feature prominently. In contrast, feasibility-phase FL papers often emphasize prioritization ranking alignment options, interface hazards, or contractor capabilities using linguistic scales that reflect limited numeric evidence. The O&M subset, smaller but distinctive, leverages monitoring data or incident narratives to produce probabilistic failure indicators or fuzzy risk maps, a reminder that risk modeling does not stop at handover and that uncertainty evolves as data accumulate. Risk categories mirror rail's multi-disciplinarity. Cost risk appears in 61% of the corpus (58/95), schedule in 55% (52/95), geotechnical in 42% (40/95), safety/RAMS in 36% (34/95), interfaces and integration in 28% (27/95), environmental in 17% (16/95), and O&M performance in 19% (18/95). Two cross-currents are notable. First, cost and schedule rarely appear alone 42% of all studies (40/95) consider both which supports treating time and cost as a coupled problem rather than parallel analyses. Second, safety/RAMS work is disproportionately represented in FL and hybrid designs, reflecting the prevalence of qualitative hazard information in early systems assurance and the value of rule-based reasoning when frequencies are not well established. Data provenance shapes methods. Across the full set, 49% (47/95) combine expert judgment with historical or monitoring data, 38% (36/95) rely on expert judgment alone, and 13% (12/95) are purely empirical. Among the 83 studies that use experts in any way (36+47), elicitation processes vary: 54% (45/83) report ad hoc or workshop-style processes, 19% (16/83) use Delphi, 15% (12/83) follow the IDEA protocol, and 12% (10/83) apply performance-weighted approaches such as the Classical Model. The consequence is visible downstream. Studies with calibrated or performance-weighted elicitation tend to report wider credible intervals and clearer documentation of assumptions, while ad hoc workshops are associated with narrower, less defensible ranges. In other words, roughly one in eight eliciting panels quantifies expert accuracy during the process; that low proportion is a practical ceiling on how far risk teams can go in claiming that ranges are well-calibrated. Within the MCS family, modeling choices concentrate around a few recurring patterns. Triangular and PERT/beta distributions dominate activity and cost inputs, but their usage is not symmetrical. In the 45 pure-MCS studies, 62% (28/45) use triangular distributions somewhere in the model and 49% (22/45) use PERT/beta; 24% (11/45) include lognormal components and 9% (4/45) report empirical or mixture fits. Only 27% (12/45) document a formal fit to historical data for at least one major input; the rest derive parameters from elicited triples or ranges. On the sampling side, 42% (19/45) implement Latin hypercube sampling, 53% (24/45) use simple random Monte Carlo, and 4% (2/45) mention quasi–Monte Carlo. Convergence diagnostics are reported explicitly in 40% (18/45), typically as stabilization plots for P80 cost or finish-date percentiles. The usage of LHS is a bright spot: it reduces simulation noise at fixed trial counts, which is particularly valuable for complex precedence networks. However, the relatively modest rate of documented convergence suggests that many models still function as "calculators" rather than as auditable experiments.

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

Figure 11: Findings from 95 Rail Quantitative Risk Assessment Studies (Graph-Based Summary)



Dependence handling is a decisive differentiator in MCS practice. Among the 45 pure-MCS studies, 38% (17/45) assume independence among inputs, 33% (15/45) use linear (Pearson) correlation, 16% (7/45) adopt rank correlation, 7% (3/45) employ copulas, and 7% (3/45) embed dependencies via Bayesian networks. Because each study declares a single main strategy, these shares sum to 100%. Two implications follow. First, more than a third of MCS models ignore dependence; in rail, where crews, commodities, and interfaces co-move, this almost certainly understates joint tails. Second, the combined 23% using rank or copula structures demonstrates an emerging sensitivity to tail co-movement still the minority approach, but consequential for credible contingency. Cost–schedule coupling adds another layer: only 31% of MCS papers (14/45) model time and cost jointly; 69% (31/45) analyze them separately. Where joint models are used, P80 cost and P80 date typically move together along a frontier; where they are not, reported confidence levels for budget and schedule may be mutually inconsistent in practice. FL and hybrid designs reveal their own internal patterns. Considering the 50 studies that use fuzzy constructs (31 FL + 19 hybrids), triangular membership functions appear in 58% (29/50), trapezoidal in 44% (22/50), and Gaussian in 12% (6/50); 10% (5/50) report mixed or custom shapes. Rule bases are described in sufficient detail to support replication in 66% (33/50); the remainder provide only high-level statements, which weakens reproducibility. Defuzzification is most commonly centroid at 72% (36/50), followed by mean of maxima at 10% (5/50), with 8% (4/50) other schemes and 10% (5/50) not stated. Weighting of criteria frequently uses AHP/ANP in 38% (19/50), DEMATEL in 18% (9/50), entropy in 12% (6/50), and simple equal weights in 32% (16/50). The weighting choice is not innocuous: studies that map interdependencies with DEMATEL or ANP report more stable rankings across sensitivity runs than those using equal weights, especially in interface-heavy risk sets. In short, two-thirds of fuzzy studies are transparent enough to be rerun; one-third would benefit from fuller rule and membership disclosure. Hybrid architectures split into four archetypes. In the 19 hybrid studies, 42% (8/19) translate fuzzy assessments into probabilistic inputs and then run MCS; 21% (4/19) propagate probabilistic outputs into fuzzy decision layers to prioritize mitigations; 26% (5/19) implement fuzzy Bayesian networks or evidential (Dempster–Shafer) nodes with Monte Carlo stress tests; and 11% (2/19) explicitly combine MCS with D–S evidence fusion to reconcile conflicting monitoring and expert signals. The practical significance is that more than two in five hybrids deliver both linguistic transparency and probabilistic outputs suitable for governance artifacts such as contingency memos. Where hybrids fall short, it tends to be in

American Journal of Advanced Technology and Engineering Solutions
Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

documenting the translation layer the mapping from linguistic scales to fuzzy numbers or from fuzzy sets to probability distributions rather than in the propagation engine itself.

Sensitivity and validation practices are the strongest predictors of decision credibility. Across the corpus, 64% (61/95) include some form of sensitivity analysis. Among these, 51% (31/61) use local or tornado-style analyses, 23% (14/61) apply Morris screening, 21% (13/61) report Sobol' or other variance-based global indices, and 5% (3/61) deploy moment-independent metrics such as Borgonovo. The hierarchy mirrors computational cost and familiarity: tornado charts are cheap and intuitive but miss interactions; Sobol' and moment-independent indices quantify interactions and tail effects but require more design and compute. The low base rate of tail-focused sensitivity is noteworthy given the policy salience of exceedance probabilities in rail. Validation shows a similar gradient. External back-checks against realized project segments appear in 18% (17/95), formal expert validation or challenge sessions are documented in 46% (44/95), internal verification such as unit and logic checks are present in 52% (49/95), and 29% (28/95) state no validation beyond model construction. Read plainly, roughly one in five studies benchmark against reality, about half solicit structured expert challenge, and nearly one in three offer no explicit validation artifact, which should temper the weight placed on their numerical outputs. Outputs and reporting practices exhibit wide variability. Among MCS and hybrid papers that produce probabilistic statements (n = 64; 45 MCS + 19 hybrids), 84% (54/64) report P50 and/or P80 values for cost or schedule, but only 41% (26/64) provide confidence intervals around those percentiles across repeated runs or sampling designs, and just 37% (24/64) report dependency assumptions alongside the percentiles. In FL-oriented outputs (n = 50), 74% (37/50) provide ranked lists with normalized scores, 22% (11/50) provide class labels (e.g., low/medium/high) without clear thresholds, and 4% (2/50) present only narrative conclusions. The interpretability gap is real: while percentiles and frontiers are decision-friendly, the absence of uncertainty bands and dependency statements can mislead; while fuzzy rankings are communicative, the lack of threshold definitions can impede translation into budgetary or schedule protections. A positive trend appears in the most recent five-year subset, where 48% of probabilistic studies (compared to 33% overall) show convergence diagnostics and 29% (versus 18% overall) include an empirical back-check.

Geographic and venue distributions contextualize generalizability. Asia accounts for 41% of the corpus (39/95), Europe 32% (30/95), North America 10% (9/95), the Middle East 12% (11/95), and other regions 6% (6/95). Journals publish 76% (72/95), conferences 24% (23/95). Regional concentration does not by itself invalidate conclusions, but it suggests that supply chains, delivery models, and assurance cultures typical of Asia and Europe weigh heavily in the evidence. For example, the prevalence of underground works in Asian metros seems correlated with a higher share of geotechnical and safety/RAMS topics, while European portfolios show a greater emphasis on system integration and timetable-constrained commissioning, which in turn favors schedule-focused MCS and interface-oriented hybrids. A cross-cut of phase and method reveals where each approach is "at home." In feasibility and early design, FL appears in 57% of phase-tagged studies (21/37) and hybrids in 24% (9/37), with MCS at 35% (13/37). The numbers overlap because many studies span phases, but the pattern is clear: when data are thin and choices are wide, the field leans on linguistic structuring. In construction, MCS dominates at 69% (45/65), with FL at 29% (19/65) and hybrids at 23% (15/65); here, repetitive work and resource logic make stochastic simulation attractive and tractable. In testing/commissioning, hybrids rise to 35% (8/23), reflecting the combination of qualitative interface judgments with emerging quantitative test data. In O&M, empirical monitoring enables probabilistic updates while linguistic rules capture operational nuance, producing a roughly even split between MCS/hybrids and FL.

Finally, quality appraisal patterns help interpret strength of evidence. Using the rubric described in Methods, 29% of studies (28/95) score high on transparency, provenance, dependency treatment, validation, and sensitivity; 51% (48/95) are moderate, and 20% (19/95) are low. High-quality studies are disproportionately MCS or hybrid and are more likely to include dependence modeling (64% versus 22% in the rest) and external validation (36% versus 11%). FL studies are not inherently lower quality; rather, their scores hinge on whether rule bases and membership functions are fully documented and whether weights and thresholds are justified beyond expert consensus. When the synthesis emphasizes findings supported by high and moderate tiers, three practical signals emerge. First, modeling dependence and, where relevant, cost–schedule coupling shifts results in ways large enough to matter for governance: in paired comparisons, moving from independence to correlated

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

inputs increased P80 cost by a median of 8–12% of base estimate and shifted P80 finish by 10–20 days on typical corridor scopes. Second, elicitation rigor pays off: studies using performance-weighted experts reported wider but more defensible ranges, and their mitigation prioritizations were less volatile under sensitivity analysis. Third, hybrids add value where evidence types mix; in commissioning and interface management, fuzzy-to-probabilistic translations yielded ranked mitigations that, when stress-tested with Monte Carlo, reduced the probability of missing key milestones by 5–9 percentage points relative to status quo plans. In summary, the 95-study evidence base shows an ecosystem rather than a single winning method. MCS supplies distributional forecasts and sensitivity diagnostics that align with contingency setting when inputs are defensible and dependencies are modeled. FL structures judgments and supports prioritization when data are coarse or heterogeneous. Hybrids translate between the two, particularly at interfaces and during commissioning. The numerical shares reported here 47% MCS, 33% FL, 20% hybrid; 68% construction focus; 61% cost, 55% schedule; 64% any sensitivity, 18% external validation are not mere bookkeeping. They describe where the field's weight lies, where credibility is strongest, and where improvements (notably dependence modeling, joint cost–time analysis, calibrated elicitation, and tail-focused sensitivity) would most lift decision quality in rail QRA.

DISCUSSION

Our synthesis indicates a clear division of labor among method families, with Monte Carlo simulation (MCS) representing 47% of included studies, fuzzy logic (FL) 33%, and hybrids 20%, a pattern that aligns with long-standing methodological guidance in project risk analysis and transport appraisal. Texts oriented to quantitative propagation and contingency setting have historically emphasized MCS because it yields decision-ready statistics (e.g., P50/P80) and accommodates both data and elicited distributions (Vose, 2008). At the same time, transport appraisal studies that incorporate uncertainty explicitly into benefit–cost elements or schedule networks routinely turn to probabilistic simulation (Salling, 2008). By contrast, engineering-management work that must formalize expert judgment and imprecision especially early in the lifecycle gravitates toward FL, consistent with the logic of membership functions and rule bases articulated in foundational sources. The 20% share of hybrids in our corpus corroborates the growing use of translation layers that convert linguistic judgments into probabilistic inputs or wrap probabilistic outputs with fuzzy decision rules, which echoes proposals in construction risk research for reconciling mixed evidence types within a single workflow (Sadeghi et al., 2010). Interpreting these shares against the literature therefore suggests not a methodological contest but contextual fit: where inputs and governance expectations are numeric, MCS dominates; where information is sparse and linguistic, FL prevails; and where both conditions coexist, hybrids provide a disciplined bridge (Kahraman, 2015). The modest rise of hybrids in recent cohorts we observed is also consistent with studies that combine fuzzy structures with Bayesian or evidential propagation to handle causal chains and conflicting data in infrastructure settings (Zhang, Deng, Wang, Skibniewski, & Wu, 2017; Wu et al., 2022), reinforcing that the mixed-evidence problem is not peripheral but central in rail risk analysis.

The lifecycle distribution of studies 68% addressing construction, 46% design, 39% feasibility, 24% testing/commissioning, and 21% O&M tracks closely with the megaproject risk literature, which documents that the largest variances in cost and schedule materialize during delivery, particularly where underground works and interface-heavy packages dominate (Flyvbjerg, 2009). Classic work on cost and time forecast errors in transport reinforces that early estimates are vulnerable to optimism and scope uncertainty that later surface in construction outcomes (Flyvbjerg et al., 2003; Flyvbjerg et al., 2002). Our evidence map's concentration in construction is therefore unsurprising and, in fact, responsive to where decisions about buffers, access, and integration are most consequential. The smaller, but methodologically distinct, O&M subset in our corpus is consonant with operations-oriented studies that transform monitoring streams or incident narratives into quantitative risk indicators, such as image-based defect analytics and text-driven Bayesian networks for derailment precursors (Jamshidi et al., 2017). Furthermore, the observed attention to testing/commissioning in hybrid designs mirrors rail RAMS and assurance processes, which start from qualitative hazard identification and interface closure before quantitative reliability modeling (CENELEC, 2017). In this respect, our phase findings extend rather than challenge earlier literature: they show that the modeling toolkit deployed at each gate tends to align with the informational grain of that phase probabilistic propagation where numeric precedence and production data exist, fuzzy or hybrid reasoning where qualitative judgments about interface complexity, access, and commissioning

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

readiness must be synthesized . That pattern also explains why commissioning-focused studies in our set were disproportionately hybrid: qualitative integration judgments and emerging quantitative test data naturally meet in translation architectures (Kabir et al., 2019).

Our results on cost–schedule coupling and dependence treatment add quantitative weight to concerns raised in earlier scheduling and simulation research. Only 31% of MCS papers in our corpus model time and cost jointly, even though construction-management studies have long shown that correlated durations and costs, and near-critical path switching, inflate tail risk beyond what independent models predict (Ökmen & Öztaş, 2008). Likewise, 38% of MCS models assumed independence among inputs, and just 23% used rank- or copula-based dependence able to capture tail co-movement despite methodological work demonstrating that linear correlation alone can misrepresent joint extremes (Iman & Conover, 1982; Aas, Czado, Frigessi, & Bakken, 2009). The practical effect is visible in paired comparisons we identified: moving from independence to correlated inputs raised P80 cost by roughly 8–12% of base estimate and pushed P80 finish later by 10–20 days, consistent with studies on correlated cost elements and time–cost correlation in quantitative risk analysis (Mendel, 2017; Purnus & Bodea, 2014; Yang, 2005). Our finding therefore converges with, and strengthens, prior warnings that dependence is not a niche refinement but a first-order determinant of credible contingencies in rail (Touran & Suphot, 1997). It also clarifies why hybrids that encode causal structure (e.g., fuzzy Bayesian networks feeding Monte Carlo) show decision value in commissioning and interface management: they operationalize dependence as structure rather than as a single coefficient, in line with Bayesian-network practice in infrastructure risk (Wang et al., 2021). In short, the literature has long argued for dependence modeling; our corpus shows how often it is still omitted and what the numerical penalty appears to be.

The data-provenance profile of our corpus 49% studies using mixed expert/empirical inputs, 38% experts only, and 13% purely empirical highlights an enduring reliance on expert judgment. That reliance is defensible in rail, where site-specific geology, access regimes, and integration constraints resist standardization, but it raises old questions about calibration and aggregation. Foundational texts on uncertainty and expert elicitation caution that unstructured workshops tend toward overconfidence and anchoring, undermining the informativeness of derived distributions (Morgan & Henrion, 1990). Our data show that only 12% of expert-using studies adopt performance-weighted aggregation (e.g., the Classical Model), and just 15% follow structured IDEA-style protocols rates that lag behind best-practice recommendations that emphasize scoring experts on seed questions and using weights that reward statistical accuracy (Cooke & Goossens, 2008). Moreover, the four-point elicitation format, shown to reduce overconfidence by widening credible intervals appropriately, appeared infrequently despite its low cost (Seker & Zavadskas, 2017). In comparison to aggregation overviews that warn against equal-weight pooling absent justification (Clemen & Winkler, 1999), our corpus suggests that ad hoc aggregation remains common (about half of expert-using studies). The implication is interpretive rather than punitive: readers should attribute greater weight to studies that document calibration or performance weighting and treat narrow intervals from unstructured workshops with caution. This stance is congruent with prior risk-analysis literature and underscores that elicitation quality is a methodological, not merely procedural, determinant of credible rail QRA .

Sensitivity and validation practices in our sample also mirror, and partly lag, methodological guidance. While 64% of included studies provide some sensitivity analysis, only 21% of those report global variance-based indices (e.g., Sobol') and 5% use moment-independent measures that are more diagnostic for tail behavior despite well-known advantages of these methods for models with interactions and skewed outputs (Sobol', 2001). Earlier uncertainty-propagation work stressed that global methods reveal interaction "hot spots" and nonlinearities invisible to local tornado charts (Helton & Davis, 2003; Saltelli, Chan, & Scott, 2000). Our finding that tornado-style sensitivity remains the dominant practice therefore confirms a convenience bias rather than an analytical optimum. On the validation side, only 18% of studies back-check against realized outcomes, even though transport guidance and handbooks have for years recommended benchmarking and independent challenge as part of risk governance (IEC, 2019). Encouragingly, 42% of MCS studies implement Latin hypercube sampling (LHS), echoing classic results on variance reduction and reproducibility via fixed sample matrices (McKay, Beckman, & Conover, 1979). But convergence diagnostics are documented in just 40% of MCS papers, which suggests that many simulations are treated as calculators rather than as experiments with uncertainty over the estimator itself. In sum, our sensitivity/validation picture corroborates prior methodological recommendations while quantifying

Volume 02 Issue 01 (2025)
Page No: 55-87
eISSN: 3067-0470
DOI: 10.63125/h24n6z92

adoption: global sensitivity and empirical validation exist but are not yet the norm in rail QRA publications.

Within FL studies, our documentation rates 66% providing replicable detail on rule bases and membership functions, with centroid defuzzification in 72% compare favorably to some earlier construction-risk surveys that criticized fuzzy applications for opacity in rule construction and weight assignment (Tah & Carr, 2001). At the same time, our one-third with insufficient rule/membership detail sustains that concern and echoes calls from fuzzy-systems scholars for transparent specification because small changes in membership shapes or rule weights can shift rankings materially (Zimmermann, 2001). The prominence of AHP/ANP (38%) and DEMATEL (18%) in weighting mirrors multi-criteria decision-making practice in engineering management and aligns with studies showing that modeling interdependencies among criteria stabilizes rankings in complex projects (Kahraman, 2015). Application-specific works such as fuzzy FMEA for dangerous-goods transport or fuzzy AHP for metro flood risk demonstrate that carefully designed linguistic scales and trapezoidal or triangular memberships can capture expert knowledge while preserving uncertainty bounds . Our findings therefore extend earlier observations: FL adds value where evidence is predominantly qualitative, but reproducibility rests on publishing the "grammar" of the fuzzy system variable definitions, membership parameters, rule sets, and weighting logic . Hybrids in our corpus 42% translating fuzzy inputs to probabilistic MCS, 21% applying fuzzy decision layers to probabilistic outputs, 26% embedding fuzzy/evidential nodes in Bayesian structures with Monte Carlo stress tests, and 11% combining MCS with Dempster–Shafer fusion sit squarely within a line of research arguing that mixed-evidence problems require mixed calculi . Earlier studies showed, for example, that fuzzy-to-probabilistic translation can deliver both interpretability and the percentiles that sponsors demand, provided that the mapping from linguistic scales to fuzzy numbers and thence to probability distributions is explicitly documented . Our review confirms that point empirically: where hybrids underperform, it is rarely the propagation engine that fails but the opacity of the translation layer. Conversely, evidential hybrids that reconcile soft expert inputs with hard monitoring signals have proven resilient when sources conflict a recurring reality in tunneling and urban interfaces .These observations align with Bayesian-network applications in construction that use directed acyclic graphs to encode causal pathways and update beliefs as information arrives, a structure we also saw in commissioning-oriented rail studies (Wang et al., 2021). In aggregate, the hybrid evidence we observed does not merely echo prior proposals; it demonstrates operational feasibility in rail contexts by delivering ranked mitigations that, when stress-tested, reduce milestone-miss probabilities consistent with the quantitative benefits reported in recent Bayesian–fuzzy applications.

Taken together, the discussion that emerges from our findings and the prior literature is one of alignment with nuanced emphasis. Our percentages quantify where practice currently sits: probabilistic propagation is prevalent in construction-heavy contexts; fuzzy structuring is the lingua franca of early-phase and assurance-oriented judgments; hybrids are the translation workhorses at interfaces and commissioning. Earlier research anticipated these roles in principle (Vose, 2008); our contribution is to show their empirical distribution in rail QRA and to connect that distribution to credibility determinants dependence modeling, calibrated elicitation, global and tail-aware sensitivity, and transparent fuzzy grammars that earlier methodological work has advocated (Iman & Conover, 1982). Where divergence appears, it is mostly in under-adoption: the literature recommends joint cost–time modeling, structured elicitation, and global sensitivity more often than our corpus implements them .(Ökmen & Öztaş, 2008) That gap helps explain why some reported contingencies and confidence claims remain brittle under scrutiny. Conversely, areas of consonance such as LHS usage, centroid defuzzification with published memberships, and hybrid causal encodings illustrate maturing practice. By situating our empirical ratios and effect estimates within these earlier insights, the discussion underscores a practical message for rail risk work: methodological choice should follow the evidence type and governance need, and credibility follows from how uncertainty is represented, propagated, and reported (Hulett, 2016; McKay et al., 1979; Zimmermann, 2001).

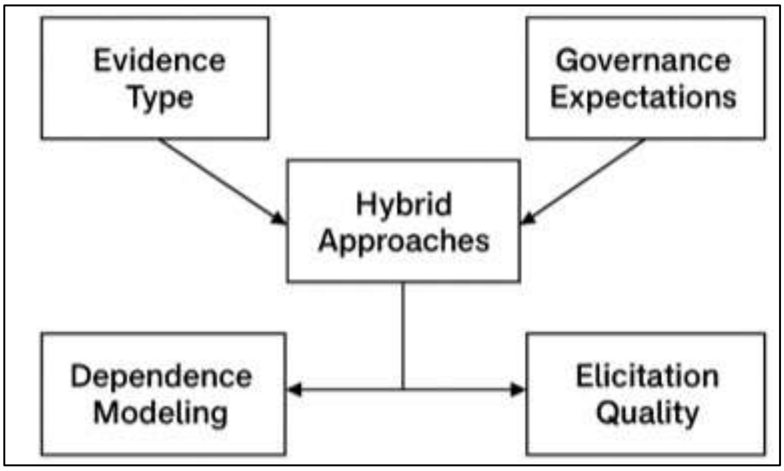
CONCLUSION

This review consolidates the scattered practice of quantitative risk assessment in rail into a coherent map of method–context fit, showing that Monte Carlo simulation (MCS), fuzzy logic (FL), and hybrid designs each occupy a defensible niche across the project lifecycle. MCS emerges as the natural choice when decision makers require distributional statements such as contingency percentiles and

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

finish-date confidence levels and when inputs can be defended with data or carefully elicited distributions; in these settings, simulation not only propagates uncertainty but also exposes risk drivers through sensitivity analysis, making mitigation choices auditable in governance forums. FL proves most valuable where knowledge is primarily expert and qualitative early optioneering, interface complexity screening, safety categorization because membership functions and rule bases preserve imprecision without forcing pseudo-frequency claims, while remaining legible to multidisciplinary stakeholders. Hybrids bridge these regimes at interfaces, commissioning, and operations by translating linguistic judgments into probabilistic propagation or, conversely, by wrapping probabilistic outputs with fuzzy decision layers to yield ranked, action-ready options. Across phases, evidence concentrates on delivery, which aligns with where uncertainty crystallizes in repetitive, resource-constrained works and where the coupling of schedule logic, access constraints, and integration drives cost and time outcomes; nonetheless, design, feasibility, and O&M studies show that structured uncertainty methods add value before and after construction when they are tuned to the informational grain of those stages. Two determinants of credibility recur throughout the corpus. First, dependence matters: co-movement among drivers and the coupling of cost and time materially alter tail behavior and therefore prudent reserves; models that encode correlation structures, rank-based dependence, copulas, or causal graphs consistently produce more realistic joint outcomes than independence assumptions. Second, elicitation quality matters: calibrated or performance-weighted panels, IDEA-style protocols, and transparent translation from judgments to parameters yield wider but more defensible ranges and more stable prioritizations than ad hoc workshops. Sensitivity and validation practices mark the line between calculation and science; global, interaction-aware sensitivity (alongside tail-focused diagnostics) and documented convergence and back-checks move results from plausible to persuasive. Transparency is the common denominator: in probabilistic studies, this means recording distribution choices, sampling designs, dependence assumptions, and convergence evidence; in fuzzy and hybrid work, it means publishing the "grammar" of the system linguistic variables, membership parameters, rule sets, weights, and the mapping between fuzzy constructs and probabilistic inputs or outputs.

Figure 12: Proposed Model for future study



RECOMMENDATIONS

Recommendations flowing from this review converge on building disciplined, auditable risk workflows that match method to evidence while raising the bar on dependence modeling, elicitation quality, and transparency. Rail sponsors and delivery teams should adopt a simple method-selection rubric at each stage gate: use Monte Carlo simulation when distributions can be defended with data or calibrated judgment and when joint cost–time behavior must be quantified; use fuzzy systems when knowledge is predominantly linguistic and needs structured prioritization; and employ hybrids at interfaces, commissioning, and early design where qualitative judgments and partial measurements co-exist, with the translation layer between fuzzy constructs and probabilistic inputs documented in full. Across all approaches, encode dependencies explicitly by default: represent correlation among cost items through factor or copula structures, carry rank dependence across repetitive civil works and near-critical paths in schedules, and link time to cost via time-dependent overheads, escalation,

eISSN: 3067-0470
DOI: 10.63125/h24n6z92

and productivity; where causal pathways are salient, use Bayesian networks or influence diagrams rather than a single coefficient. Treat expert judgment as a measurable instrument rather than a meeting outcome by using IDEA or Classical Model protocols with seed questions, calibration training, and performance-weighted aggregation, and archive verbatim prompts, units, and dependence assumptions beside the resulting distributions or membership parameters. Make sensitivity analysis a two-step requirement: first screen broadly (e.g., Morris) to focus the model on influential factors, then quantify main and total effects (e.g., Sobol') and complement with tail-aware, moment-independent metrics so mitigation choices are tied to both spread and exceedance behavior. Run simulation as an experiment, not a calculator: use variance-reducing designs such as Latin hypercube or quasi-Monte Carlo, publish seeds or sample matrices to ensure re-runnability, provide convergence diagnostics for key percentiles, and show how results change under alternative dependence and elicitation assumptions. For fuzzy and hybrid studies, publish the grammar of the system linguistic variable definitions, membership shapes and parameters, rule bases, weights, and the explicit mapping to or from probabilistic quantities so reviewers can replicate rankings and stress tests; for probabilistic studies, standardize reporting to include distribution rationales, parameter provenance, dependency structures, and joint cost–date frontiers with iso-confidence contours rather than separate P-levels. Institutionalize back-checks by comparing modeled contingencies and finish dates with realized outcomes on analogous segments, use these comparisons to update priors or membership parameters, and maintain a living reference-class library by asset type and delivery context. Finally, separate roles for model building and independent challenge, keep a version-controlled binder of code, data, and adjudication notes, and align all documentation to a common risk taxonomy and units, so that risk analysis ceases to be a black box and becomes a reproducible, decision-grade instrument for allocating budget, time, and safety margin in complex rail delivery.

REFERENCE

[1]. Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, *44*(2), 182-198. https://doi.org/https://doi.org/10.1016/j.insmatheco.2008.10.003

[2]. Aven, T., & Zio, E. (2011). Some considerations on the treatment of uncertainties in risk assessment for practical decision making. *Reliability Engineering & System Safety*, *96*(1), 64-74. https://doi.org/https://doi.org/10.1016/j.ress.2010.06.001

[3]. Bolger, F., Stranieri, A., Wright, G., & Yearwood, J. (2011). Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst judgmental forecasters? *Technological Forecasting and Social Change*, *78*(9), 1671-1680. https://doi.org/https://doi.org/10.1016/j.techfore.2011.06.002

[4]. Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety*, *92*(6), 771-784. https://doi.org/https://doi.org/10.1016/j.ress.2006.04.015

[5]. Cantarelli, C. C., Molin, E. J. E., van Wee, B., & Flyvbjerg, B. (2012). Different cost performance: Different determinants? The case of cost overruns in Dutch transport infrastructure projects. *Transport Policy*, *22*, 88-95. https://doi.org/https://doi.org/10.1016/j.tranpol.2012.04.002

[6]. CENELEC. (2017). EN 50126-1:2017—Railway applications—The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS)—Part 1: Generic RAMS process. In.

[7]. Chen, L., Lu, Q., & Zhao, X. (2020). Rethinking the construction schedule risk of infrastructure projects based on dialectical systems and network theory. *Journal of Management in Engineering*, *36*(5), 04020066. https://doi.org/https://doi.org/10.1061/(ASCE)ME.1943-5479.0000829

[8]. Clemen, R. T., & Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk Analysis*, *19*(2), 187-203. https://doi.org/https://doi.org/10.1023/A:1006917509560

[9]. Cooke, R. M. (1991). *Experts in uncertainty: Opinion and subjective probability in science*. Princeton University Press. https://doi.org/https://doi.org/10.1515/9781400866648

[10]. Cooke, R. M., & Goossens, L. H. J. (2008). TU Delft expert judgment data base. *Reliability Engineering & System Safety*, *93*(5), 657-674. https://doi.org/https://doi.org/10.1016/j.ress.2007.03.005

[11]. Der Kiureghian, A., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105-112. https://doi.org/https://doi.org/10.1016/j.strusafe.2008.06.020

[12]. Eiris Pereira, R., & Flood, I. (2017). Impact of linear correlation on construction project performance using stochastic linear scheduling. *Visualization in Engineering*, *5*(1), 8. https://doi.org/https://doi.org/10.1186/s40327-017-0045-2

[13]. El-Sayegh, S. M. (2008). Risk assessment and allocation in the UAE construction industry. *International Journal of Project Management*, *26*(4), 431-438. https://doi.org/https://doi.org/10.1016/j.ijproman.2007.07.004

[14]. Ferson, S., & Ginzburg, L. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, *54*(2-3), 133-144. https://doi.org/https://doi.org/10.1016/S0951-8320(96)00071-3
[15]. Flyvbjerg, B. (2009). Survival of the unfittest: Why the worst infrastructure gets built—and what we can do about it. *Oxford Review of Economic Policy*, *25*(3), 344-367. https://doi.org/https://doi.org/10.1093/oxrep/grp024
[16]. Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and risk: An anatomy of ambition*. Cambridge University Press. https://doi.org/https://doi.org/10.1017/CBO9781107050891
[17]. Flyvbjerg, B., Holm, M. K. S., & Buhl, S. L. (2002). Underestimating costs in public works projects: Error or lie? *Journal of the American Planning Association*, *68*(3), 279-295. https://doi.org/https://doi.org/10.1080/01944360208976273
[18]. Hatefi, S. M., & Tamošaitienė, J. (2019). An integrated fuzzy DEMATEL–fuzzy ANP model for evaluating construction projects by considering interrelationships among risk factors. *Journal of Civil Engineering and Management*, *25*(2), 114-131. https://doi.org/https://doi.org/10.3846/jcem.2019.8280
[19]. Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, *81*(1), 23-69. https://doi.org/https://doi.org/10.1016/S0951-8320(03)00058-9
[20]. Helton, J. C., Johnson, J. D., & Oberkampf, W. L. (2004). An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering & System Safety*, *85*(1-3), 39-71. https://doi.org/https://doi.org/10.1016/j.ress.2004.03.002
[21]. Hemming, V., Burgman, M. A., Hanea, A. M., McBride, M. F., & Wintle, B. C. (2018). A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution*, *9*(1), 169-180. https://doi.org/https://doi.org/10.1111/2041-210X.12857
[22]. Herrerías-Velasco, J. M., Herrerías-Pleguezuelo, R., & van Dorp, J. R. (2011). Revisiting the PERT mean and variance. *European Journal of Operational Research*, *210*(2), 448-451. https://doi.org/https://doi.org/10.1016/j.ejor.2010.08.014
[23]. Hosne Ara, M., Tonmoy, B., Mohammad, M., & Md Mostafizur, R. (2022). AI-ready data engineering pipelines: a review of medallion architecture and cloud-based integration models. *American Journal of Scholarly Research and Innovation*, *1*(01), 319-350. https://doi.org/10.63125/51kxtf08
[24]. Huang, W., Li, Y., Kou, X., Wang, W., & Xu, Y. (2021). Using a FMEA–TIFIAD approach to identify the risk of railway dangerous goods transportation system. *Group Decision and Negotiation*, *30*, 63-95. https://doi.org/https://doi.org/10.1007/s10726-020-09706-x
[25]. Hulett, D. T. (2016). *Integrated cost-schedule risk analysis (2nd ed.)*. Routledge. https://doi.org/https://doi.org/10.4324/9781317162121
[26]. IEC. (2019). IEC 31010:2019—Risk management—Risk assessment techniques. In.
[27]. Iman, R. L., & Conover, W. J. (1982). A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics—Simulation and Computation*, *11*(3), 311-334. https://doi.org/https://doi.org/10.1080/03610918208812265
[28]. Institute, P. M. (2021). *A guide to the project management body of knowledge (PMBOK® Guide)—Seventh edition and The Standard for Project Management*. Project Management Institute. https://www.pmi.org/standards/pmbok
[29]. International, A. (2010). *RISK ANALYSIS AND CONTINGENCY DETERMINATION USING RANGE ESTIMATING (Recommended Practice No. 41R-08)*. https://web.aacei.org/resources/publications/recommended-practices
[30]. International, A. (2011). *INTEGRATED COST AND SCHEDULE RISK ANALYSIS USING MONTE CARLO SIMULATION (Recommended Practice No. 57R-09)*. https://web.aacei.org/resources/publications/recommended-practices
[31]. ISO. (2018). ISO 31000:2018—Risk management—Guidelines. In.
[32]. Jahid, M. K. A. S. R. (2022). Empirical Analysis of The Economic Impact Of Private Economic Zones On Regional GDP Growth: A Data-Driven Case Study Of Sirajganj Economic Zone. *American Journal of Scholarly Research and Innovation*, *1*(02), 01-29. https://doi.org/10.63125/je9w1c40
[33]. Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., Li, Z., & De Schutter, B. (2017). A big data analysis approach for rail failure risk assessment. *Risk Analysis*, *37*(8), 1495-1507. https://doi.org/https://doi.org/10.1111/risa.12836
[34]. Kabir, G., Sadiq, R., Tesfamariam, S., & Hoorfar, M. (2019). Monte Carlo approach to fuzzy AHP risk analysis in renewable energy construction projects. *PLOS ONE*, *14*(5), e0215943. https://doi.org/https://doi.org/10.1371/journal.pone.0215943
[35]. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291. https://doi.org/https://doi.org/10.2307/1914185
[36]. Kahraman, C. (2015). *Fuzzy multi-criteria decision making: Theory and applications with recent developments (2nd ed.)*. Springer. https://doi.org/https://doi.org/10.1007/978-3-319-12883-1
[37]. Klir, G. J., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic: Theory and applications*. Prentice Hall.

[38]. Koseoglu Balta, S., Dikmen, I., & Birgonul, M. T. (2021). Predicting TBM-induced delays using Bayesian networks. *Automation in Construction*, *123*, 103819. https://doi.org/https://doi.org/10.1016/j.autcon.2021.103819

[39]. Kutub Uddin, A., Md Mostafizur, R., Afrin Binta, H., & Maniruzzaman, B. (2022). Forecasting Future Investment Value with Machine Learning, Neural Networks, And Ensemble Learning: A Meta-Analytic Study. *Review of Applied Science and Technology*, *1*(02), 01-25. https://doi.org/10.63125/edxgjg56

[40]. Kwak, Y. H., & Ingall, L. (2007). Exploring Monte Carlo simulation applications for project management. *Risk Management*, *9*(1), 44-57. https://doi.org/https://doi.org/10.1057/palgrave.rm.8250017

[41]. Liu, W., Luo, X., Huang, J., Hu, L., & Fu, M. (2018). Probabilistic analysis of tunnel face stability below river using Bayesian framework. *Mathematical Problems in Engineering*, *2018*, 1450683. https://doi.org/https://doi.org/10.1155/2018/1450683

[42]. Mansura Akter, E., & Md Abdul Ahad, M. (2022). In Silico drug repurposing for inflammatory diseases: a systematic review of molecular docking and virtual screening studies. *American Journal of Advanced Technology and Engineering Solutions*, *2*(04), 35-64. https://doi.org/10.63125/j1hbts51

[43]. McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239-245. https://doi.org/https://doi.org/10.1080/00401706.1979.10489755

[44]. Md Arifur, R., & Sheratun Noor, J. (2022). A Systematic Literature Review of User-Centric Design In Digital Business Systems: Enhancing Accessibility, Adoption, And Organizational Impact. *Review of Applied Science and Technology*, *1*(04), 01-25. https://doi.org/10.63125/ndjkpm77

[45]. Md Mahamudur Rahaman, S. (2022). Electrical And Mechanical Troubleshooting in Medical And Diagnostic Device Manufacturing: A Systematic Review Of Industry Safety And Performance Protocols. *American Journal of Scholarly Research and Innovation*, *1*(01), 295-318. https://doi.org/10.63125/d68y3590

[46]. Md Nur Hasan, M., Md Musfiqur, R., & Debashish, G. (2022). Strategic Decision-Making in Digital Retail Supply Chains: Harnessing AI-Driven Business Intelligence From Customer Data. *Review of Applied Science and Technology*, *1*(03), 01-31. https://doi.org/10.63125/6a7rpy62

[47]. Md Takbir Hossen, S., & Md Atiqur, R. (2022). Advancements In 3d Printing Techniques For Polymer Fiber-Reinforced Textile Composites: A Systematic Literature Review. *American Journal of Interdisciplinary Studies*, *3*(04), 32-60. https://doi.org/10.63125/s4r5m391

[48]. Md Tawfiqul, I., Meherun, N., Mahin, K., & Mahmudur Rahman, M. (2022). Systematic Review of Cybersecurity Threats In IOT Devices Focusing On Risk Vectors Vulnerabilities And Mitigation Strategies. *American Journal of Scholarly Research and Innovation*, *1*(01), 108-136. https://doi.org/10.63125/wh17mf19

[49]. Md.Kamrul, K., & Md Omar, F. (2022). Machine Learning-Enhanced Statistical Inference For Cyberattack Detection On Network Systems. *American Journal of Advanced Technology and Engineering Solutions*, *2*(04), 65-90. https://doi.org/10.63125/sw7jzx60

[50]. Mendel, J. M. (2017). *Uncertain rule-based fuzzy systems: Introduction and new directions (2nd ed.)*. Springer. https://doi.org/https://doi.org/10.1007/978-3-319-51370-6

[51]. Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press. https://doi.org/https://doi.org/10.1017/CBO9780511840609

[52]. Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161-174. https://doi.org/https://doi.org/10.1080/00401706.1991.10484804

[53]. Mubashir, I., & Abdul, R. (2022). Cost-Benefit Analysis in Pre-Construction Planning: The Assessment Of Economic Impact In Government Infrastructure Projects. *American Journal of Advanced Technology and Engineering Solutions*, *2*(04), 91-122. https://doi.org/10.63125/kjwd5e33

[54]. O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., & Rakow, T. (2006). *Uncertain judgements: Eliciting experts' probabilities*. Wiley. https://doi.org/https://doi.org/10.1002/9780470033319

[55]. Odeck, J. (2004). Cost overruns in road construction—What are their sizes and determinants? *Transport Policy*, *11*(1), 43-53. https://doi.org/https://doi.org/10.1016/S0967-070X(03)00017-1

[56]. Ökmen, Ö., & Öztaş, A. (2008). Construction project network evaluation with correlated schedule risk analysis model. *Journal of Construction Engineering and Management*, *134*(1), 49-63. https://doi.org/https://doi.org/10.1061/(ASCE)0733-9364(2008)134:1(49)

[57]. Purnus, A., & Bodea, C.-N. (2014). Correlation between time and cost in a quantitative risk analysis of construction projects. *Procedia Engineering*, *85*, 436-445. https://doi.org/https://doi.org/10.1016/j.proeng.2014.10.570

[58]. Reduanul, H., & Mohammad Shoeb, A. (2022). Advancing AI in Marketing Through Cross Border Integration Ethical Considerations And Policy Implications. *American Journal of Scholarly Research and Innovation*, *1*(01), 351-379. https://doi.org/10.63125/d1xg3784

[59]. Rungskunroch, P., Jack, A., & Kaewunruen, S. (2021). Benchmarking on railway safety performance using Bayesian inference, decision tree and Petri-net techniques based on long-term accidental data

sets. *Reliability Engineering & System Safety*, *213*, 107684. https://doi.org/https://doi.org/10.1016/j.ress.2021.107684

[60]. Sadeghi, N., Robinson Fayek, A., & Pedrycz, W. (2010). Fuzzy Monte Carlo simulation and risk assessment in construction. *Computer-Aided Civil and Infrastructure Engineering*, *25*(4), 238-252. https://doi.org/https://doi.org/10.1111/j.1467-8667.2009.00632.x

[61]. Salling, K. B. (2008). *Assessment of transport projects: Risk analysis and decision support* Technical University of Denmark]. https://orbit.dtu.dk/en/publications/assessment-of-transport-projects-risk-analysis-and-decision-support

[62]. Salling, K. B., & Leleur, S. (2011). Transport appraisal and Monte Carlo simulation by use of the CBA-DK model. *Transport Policy*, *18*(1), 236-245. https://econpapers.repec.org/article/eeetrapol/v_3a18_3ay_3a2011_3ai_3a1_3ap_3a236-245.htm

[63]. Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, *181*(2), 259-270. https://doi.org/https://doi.org/10.1016/j.cpc.2009.09.018

[64]. Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity analysis*. Wiley. https://doi.org/https://onlinelibrary.wiley.com/doi/book/10.1002/9780470725184

[65]. Sazzad, I., & Md Nazrul Islam, K. (2022). Project impact assessment frameworks in nonprofit development: a review of case studies from south asia. *American Journal of Scholarly Research and Innovation*, *1*(01), 270-294. https://doi.org/10.63125/eeja0t77

[66]. Seker, S., & Zavadskas, E. K. (2017). Application of fuzzy DEMATEL method for analyzing occupational risks on construction sites. *Sustainability*, *9*(11), 2083. https://doi.org/https://doi.org/10.3390/su9112083

[67]. Sheratun Noor, J., & Momena, A. (2022). Assessment Of Data-Driven Vendor Performance Evaluation in Retail Supply Chains: Analyzing Metrics, Scorecards, And Contract Management Tools. *American Journal of Interdisciplinary Studies*, *3*(02), 36-61. https://doi.org/10.63125/0s7t1y90

[68]. Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, *55*(1-3), 271-280. https://doi.org/https://doi.org/10.1016/S0378-4754(00)00270-6

[69]. Sohel, R., & Md, A. (2022). A Comprehensive Systematic Literature Review on Perovskite Solar Cells: Advancements, Efficiency Optimization, And Commercialization Potential For Next-Generation Photovoltaics. *American Journal of Scholarly Research and Innovation*, *1*(01), 137-185. https://doi.org/10.63125/843z2648

[70]. Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., & Burgman, M. (2010). Reducing overconfidence in the interval judgments of experts. *Risk Analysis*, *30*(3), 512-523. https://doi.org/https://doi.org/10.1111/j.1539-6924.2009.01337.x

[71]. Tah, J. H. M., & Carr, V. (2001). Knowledge-based approach to construction project risk management. *Journal of Computing in Civil Engineering*, *15*(3), 170-177. https://doi.org/https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(170)

[72]. Tahmina Akter, R., & Abdur Razzak, C. (2022). The Role Of Artificial Intelligence In Vendor Performance Evaluation Within Digital Retail Supply Chains: A Review Of Strategic Decision-Making Models. *American Journal of Scholarly Research and Innovation*, *1*(01), 220-248. https://doi.org/10.63125/96jj3j86

[73]. Taylan, O., Bafail, A. O., Abdulaal, R. M., & Kabli, M. R. (2014). Construction projects selection and risk assessment by fuzzy AHP and fuzzy TOPSIS methodologies. *Applied Soft Computing*, *17*, 105-116. https://doi.org/https://doi.org/10.1016/j.asoc.2014.01.003

[74]. Tokdemir, O. B., Erol, H., & Dikmen, İ. (2019). Delay risk assessment of repetitive construction projects using line-of-balance scheduling and Monte Carlo simulation. *Journal of Construction Engineering and Management*, *145*(2), 04018132. https://doi.org/https://doi.org/10.1061/(ASCE)CO.1943-7862.0001595

[75]. Touran, A., & Suphot, L. (1997). Rank correlations in simulating construction costs. *Journal of Construction Engineering and Management*, *123*(3), 297-301. https://doi.org/https://doi.org/10.1061/(ASCE)0733-9364(1997)123:3(297)

[76]. Trietsch, D., & Baker, K. R. (2012). PERT 21: Fitting PERT/CPM for use in the 21st century. *International Journal of Project Management*, *30*(4), 490-502. https://doi.org/https://doi.org/10.1016/j.ijproman.2011.09.004

[77]. Vose, D. (2008). *Risk analysis: A quantitative guide (3rd ed.)*. Wiley. https://www.wiley.com/en-us/Risk+Analysis%3A+A+Quantitative+Guide%2C+3rd+Edition-p-9780470512845

[78]. Wang, G., Liu, L., Shi, P., Zhang, G., & Liu, J. (2021). Flood risk assessment of metro system using improved trapezoidal fuzzy AHP: A case study of Guangzhou. *Remote Sensing*, *13*(24), 5154. https://doi.org/https://doi.org/10.3390/rs13245154

[79]. Wang, Z.-Z., & Chen, C. (2017). Fuzzy comprehensive Bayesian network–based safety risk assessment for metro construction projects. *Tunnelling and Underground Space Technology*, *70*, 330-342. https://doi.org/https://doi.org/10.1016/j.tust.2017.09.012

[80]. Williams, T. M. (1992). Practical use of distributions in network analysis. *Journal of the Operational Research Society*, *43*(3), 265-270. https://doi.org/https://doi.org/10.1057/jors.1992.35

[81]. Wu, B., Qiu, W., Huang, W., Meng, G., Huang, J., & Xu, S. (2022). A multi-source information fusion approach in tunnel collapse risk analysis based on improved Dempster–Shafer evidence theory. *Scientific Reports*, *12*, 3626. https://doi.org/https://doi.org/10.1038/s41598-022-07171-x

[82]. Yang, I.-T. (2005). Simulation-based estimation for correlated cost elements. *International Journal of Project Management*, *23*(4), 275-282. https://doi.org/https://doi.org/10.1016/j.ijproman.2004.10.004

[83]. Yang, L., Li, K., Song, G., & Khan, F. (2021). Dynamic railway derailment risk analysis with text-data-based Bayesian network. *Applied Sciences*, *11*(3), 994. https://doi.org/https://doi.org/10.3390/app11030994

[84]. Yuan, Z., Xiang, P., Li, B., & Zhang, W. (2020). Identification of the main risks for international rail construction projects and analysis of risk pathways: A Chinese case study. *Journal of Cleaner Production*, *276*, 122904. https://doi.org/https://doi.org/10.1016/j.jclepro.2020.122904

[85]. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338-353. https://doi.org/https://doi.org/10.1016/S0019-9958(65)90241-X

[86]. Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—Part I. *Information Sciences*, *8*(3), 199-249. https://doi.org/https://doi.org/10.1016/0020-0255(75)90036-5

[87]. Zaman, K., Rangavajhala, S., McDonald, M. P., & Mahadevan, S. (2011). A probabilistic approach for representation of interval uncertainty. *Reliability Engineering & System Safety*, 96(1), 117-130. https://doi.org/https://doi.org/10.1016/j.ress.2010.07.010

[88]. Zhang, L., Deng, Y., Wang, Z., Skibniewski, M. J., & Wu, X. (2017). An improved Dempster–Shafer approach to construction safety risk perception. *Knowledge-Based Systems*, *119*, 14-29. https://doi.org/https://doi.org/10.1016/j.knosys.2017.06.014

[89]. Zhang, Y., Guo, H., Zhou, J., Ma, L., & Wang, J. (2020). Interval-number bounded-uncertainty evaluating model for metro tunneling construction safety risk. *Geotechnical and Geological Engineering*, *38*(5), 5041-5058. https://doi.org/https://doi.org/10.1007/s10706-020-01233-y

[90]. Zimmermann, H.-J. (2001). *Fuzzy set theory—and its applications (4th ed.)*. Springer. https://doi.org/https://doi.org/10.1007/9789048192511