Advances in Differential Equations and Control Processes 2026, 33(1), 3774.
doi: https://doi.org/10.59400/adecp3774

Article

Data-driven hierarchical decision support for civil aviation maintenance safety risk: A fusion of Bayesian network and system dynamics

Jirong Duan, **Ming Cheng***

School of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China
* **Corresponding author:** Ming Cheng, mcheng@cauc.edu.cn

CITATION

Duan J, Cheng M. Data-driven hierarchical decision support for civil aviation maintenance safety risk: A fusion of Bayesian network and system dynamics. Advances in Differential Equations and Control Processes. 2026; 33(1): 3774. https://doi.org/10.59400/adecp3774

ARTICLE INFO

Received: 5 December 2025
Revised: 12 January 2026
Accepted: 20 January 2026
Available online: 5 February 2026

Abstract: To enhance the scientificity and precision of risk analysis and management decision-making in aircraft maintenance operations, this study proposes a risk analysis-decision model tailored to maintenance events. Based on actual civil aviation maintenance scenarios, the model employs real data to conduct data-driven analysis and precisely calculates the occurrence probabilities of various risk factors by constructing a Bayesian risk probability network. Meanwhile, it selects three categories of key risk factors: personnel (A), management (B), and organization (C), to build a system dynamics scenario, thereby simulating the long-term implementation effects of different management strategies. The research findings indicate that the existing maintenance management system demonstrates a certain level of risk buffering efficacy under normal operating conditions, effectively preventing risks from evolving into higher severity levels. The combinations of key risk factors at different severity levels exhibit a hierarchical characteristic, specifically manifesting as three tiers dominated by organization and safety barriers, personnel capabilities and behaviors, and daily operations and slow-variable risks, respectively. It is proposed that maintenance safety risk governance should adopt a graded and differentiated management strategy. At the decision-making level, the model is capable of simulating the long-term impacts of different management strategies. The study reveals that increasing management investment can significantly reduce process risks, whereas systemic risks and frontline operational errors require sustained, long-term resource allocation for improvement.

Keywords: continuous endurance risk management; Bayesian network; system dynamics; risk analysis; data-driven

1. Introduction

According to statistics from the International Air Transport Association (IATA), it is projected that the global air passenger traffic volume will reach 5.2 billion person-times by 2025, representing a year-on-year increase of 6.7% and surpassing the 5 billion person-times milestone for the first time [1]. Given such a massive transportation scale, there is an extremely high demand for aircraft reliability. Efficiently and accurately identifying aircraft faults and enhancing the effectiveness and efficiency of aircraft maintenance are crucial for ensuring the stable operation of aircraft. It has become a trend to conduct big data analysis and modeling on operational data generated during the daily operation, maintenance, and repair of aircraft. The maintenance approach has shifted from the original diagnosis-based planned maintenance to the current "statistics-based predictive maintenance [2,3].

In recent years, Bayesian networks (BNs) have become one of the most widely adopted probabilistic modeling tools in civil aviation safety research, owing to their capability to represent causal dependencies, handle uncertainty, and integrate heterogeneous data sources. Existing studies span multiple aviation domains, including aircraft maintenance, flight operations, air traffic management, airport systems, and aviation supply chains, demonstrating the versatility and effectiveness of BN-based risk analysis frameworks. Bauranov and Rakas [4] developed a Bayesian network model to evaluate aviation safety under the introduction of new communication technologies, quantitatively assessing their influence on mid-air collision risk and highlighting the role of human–system interactions in collision prevention. Focusing on complex aviation logistics, Wang et al. [5] proposed a BN-based framework for risk and resilience analysis of humanitarian aviation supply chains. Their model integrates operational, environmental, and organizational factors to identify critical vulnerabilities and assess mission failure probabilities in highly uncertain operational environments. In air traffic management safety, Pan et al. [6] combined association rule mining with fault-tree Bayesian networks to analyze controller-caused aviation accidents, revealing dominant unsafe behavior paths and providing quantitative support for targeted safety interventions. Zhang and Mahadevan [7] automate the construction of probabilistic risk representations from aviation accident investigation reports, enabling the quantification of causal relationships and risk propagation based on historical data. This approach demonstrates how BN models can provide structured, data-driven insights into complex safety scenarios. In the commercial aviation domain, Zhou et al. [8] have employed Bayesian network modeling to systematically reduce safety risks by extracting causal events from large incident databases and optimizing resource allocation decisions. Using structure learning and mutual information, this model identifies key causal factors and supports decision-making for risk mitigation across the commercial air transportation system. Extending beyond flight risk, Bayesian networks have also been applied to ground operation safety in airport environments. Guo et al. [9] combined BN with accident trees to assess traffic safety risks in airport aircraft activity areas, identifying key human–vehicle interaction risk sources and formulating targeted safety measures. This work illustrates the versatility of BN methods across multi-factor operational domains. Additionally, complementary probabilistic modeling approaches, such as novel aviation risk level identification models based on enhanced text encoding and probabilistic inference [10], underscore the ongoing trend toward integrating data-driven and probabilistic methods for more accurate risk characterization in aviation safety research. From a broader methodological perspective, Dang et al. [11] proposed a BN-based aviation critical system risk analysis method that integrates empirical information fusion, providing methodological support for complex aviation system safety assessment. Chen and Lin [12] further extended BN-based modeling to multi-stakeholder unsafe event networks in airspace systems, emphasizing inter-organizational risk coupling and networked safety governance. Barry [13] leveraged large-scale flight data to learn a Bayesian Network (BN) and applied it to quantitatively evaluate airline runway excursion risk, thereby enhancing the understanding of the probabilistic impacts of meteorological conditions and flight operational factors on such risks. Qi et al. [14] combined interpretable learning with causal discovery to construct a BN, quantifying the triggering probabilities and joint risks of various contributing factors to runway excursions, and providing a structured approach for causal analysis of runway safety events. Zhao et al. [15], based on the Threat–Error–Undesired State (TEM) model, employed a BN to quantitatively analyze risks in air traffic control (ATC) operations, emphasizing the probabilistic influence of threats and errors on the occurrence of undesired states. Arnaldo Valdés et al. [16] adopted Bayesian inference and hierarchical modeling to predict aviation safety incidents, enabling the exploration of safety incident data, efficient identification of anomalies, and assessment of risk levels.

Beyond the aviation domain, Bayesian-network-based and system-oriented probabilistic modeling approaches have been widely applied to risk assessment and safety management in complex engineering and socio-technical systems, demonstrating their effectiveness in capturing uncertainty, dynamic evolution, and causal relationships across diverse application scenarios. Ifelebuegu et al. [17] proposed the use of Bayesian Layer of Protection Analysis (Bayesian-LOPA) to assess the risk of subsea gas compression systems, providing a more robust and credible approach for event scenario modeling. He et al. [18] proposed a method based on Bayesian networks and cloud models to address the deficiencies in coal mine gas explosion risk assessment and the handling of uncertainties in the risk evaluation process. Wang et al. [19] employed a hybrid approach based on Bayesian Networks (BN) to model the 2S2E (2 Subjects–2 Environments) framework for ship accidents, systematically integrating ship-related factors, human factors, internal environmental factors, and external environmental factors. Zhao et al. [20] constructed a Dynamic Bayesian Network (DBN) incorporating sliding-window and forgetting-factor mechanisms to enable adaptive updating of conditional probability tables, significantly enhancing the real-time responsiveness, interpretability, and predictive capability of bridge construction safety management. Li et al. [21] established a causal structure using fault tree analysis (FTA) and dynamically updated probabilities through a Bayesian network to evaluate the impact of blocked rescue routes on fire response at a fire station, and proposed targeted preventive measures. Huang et al. [22] developed a dynamic Bayesian network using an Interpretive Structural Model–Bayesian Network (ISM–BN) approach; through backward diagnostic reasoning, the major risk factors affecting navigation were identified, and corresponding recommendations were proposed, which is of significant importance for improving practical accident prevention and emergency response capabilities.

While most existing BN-based aviation studies focus on static probabilistic inference, recent research has highlighted the limitations of static models in representing long-term risk evolution and management feedback mechanisms. System dynamics (SD) has been increasingly recognized as a complementary tool for capturing cumulative risk effects and policy-driven changes. Zhang et al. [23] developed a comprehensive hybrid risk analysis framework that couples a three-layer Bayesian network with a System Dynamics model to enable both static probabilistic inference and dynamic risk feedback across multi-factor risk factors; this hybrid model was applied to complex drilling construction risk evaluation, demonstrating how the static BN risk structure can be extended with causal feedback loops for dynamic risk evolution analysis. In the context of interconnected critical infrastructure systems, Bakhtiari et al. [24] introduced a Dynamic Bayesian Network (DBN) framework for capturing multi-hazard risk progression and resilience. Their approach explicitly models disruption propagation and recovery processes over time, offering a dynamic perspective on risk and cascading effects unavailable in traditional static BN models. Likewise, You et al. [25] proposed a time-dependent reliability assessment method based on DBN for short-term multi-round situational awareness, representing dependencies across adjacent time slices and enabling reliability evaluation as time evolves. Li et al. [26] employed a system dynamics (SD) model to reveal the complex interdependencies and critical pathways among resilience dimensions, providing new conceptual and methodological insights into the measurement and evolution of urban flood resilience. Yuan et al. [27] developed a PSR–SD model for dynamic flood resilience assessment and identified a two-year policy lag effect in flood control engineering measures.

Overall, existing domestic and international studies demonstrate that Bayesian networks have become a mature and widely accepted tool for aviation risk analysis, covering flight operations, air traffic management, airport systems, maintenance activities, and aviation supply chains. Nevertheless, most studies either focus on isolated operational scenarios or rely on static probabilistic inference, with limited consideration of long-term risk dynamics and management strategy impacts. This gap motivates the present study, which integrates Bayesian network–based probabilistic risk analysis with system dynamics modeling to support graded, differentiated, and long-term risk governance in civil aviation maintenance systems operating under continuous and routine conditions. By jointly introducing Bayesian networks and system dynamics theory, this paper proposes an integrated risk analysis and decision-support framework that enables the effective fusion and utilization of large-scale multi-source operational data and fragmented unstructured textual information throughout the civil aviation maintenance process. The proposed framework facilitates both quantitative risk inference and dynamic risk evolution analysis, thereby providing a practical and scalable application tool for intelligent maintenance risk assessment and decision-making.

2. Materials and methods

Based on Bayesian network and system dynamics theories, this model utilizes unstructured textual experiential knowledge from Civil Aviation Maintenance processes to develop a maintenance risk analysis and decision-making model. It calculates risk probabilities, extracts risk factors, simulates the causal relationships and time-lag effects among risk factors, and evaluates the impact of different management strategies on the maintenance management environment. The model consists of three layers: data layer, Bayesian risk analysis layer, and system dynamics decision-making layer.

2.1. Data layer

The data layer primarily utilizes actual maintenance records as its core data source, integrating multi-source data such as incident reports and operational difficulty reports. This forms a two-tier tree structure composed of incident outcomes and risk factors, which is fed into the risk analysis layer. The risk factors analyzed and output from this layer are then introduced into the decision-making layer. Risk factors are classified using the Reason's [28] theoretical framework, while incident severity is ultimately categorized into three levels based on statistical analysis and the classification guidelines outlined in Doc9859, as shown in the **Tables 1** and **2**.

Table 1. Risk factor classification.

Layer	Type	Symbol	Classification
Risk of personnel behavior and competence	Adverse Weather	F1	Strong Wind/Storm/Thunderstorm/Hail/Heavy Fog (F11), No (F12)
	Night Work	F2	Night (F21), No (F22)
	Inherent Defects of Equipment and Components	F3	Manufacturing Defects (F31), Use of Expired Components (F32)
Human behavior and capability risks	Inadequate Skills and Knowledge	F4	Incompetence (F41), Lack of Training/Insufficient Training (F42)
	Unsafe Behavior and Attitude	F5	Insufficient Responsibility (F51), Insufficient Communication (F52), Failure to Strictly Implement Maintenance Procedures (F53)
	Cognitive and Judgment Bias	F6	Inadequate Instructions (F61), Failure to Detect Defects During Inspection (F62), Failure to Detect Faults (F63)
Technical and management process risks	Defects in Procedures and Inspection Systems	F7	Procedure Information Missing/Inconsistent/Incorrect (F71), Insufficient Inspection (F72), No Inspection Conducted (F73), Insufficient Acceptance (F74), No Acceptance Inspection Conducted (F75)
	Tool and Component Management Oversight	F8	Improper installation/adjustment (F81), Incorrect installation (F82), Misalignment error (F83), Component not reinstalled/missing installation (F84), Tools left inside equipment (F85), Unsecured components or latches (F86), Missing components (F87)
	Documentation and Traceability Defects	F9	Incomplete Maintenance records (F91), Misleading/missing documents (F92), Inadequate reporting (F93)
Organizational and regulatory risks	Safety Culture and Resource Deficiencies	F10	Time pressure (F101), Inadequate resource planning (F102)
	Regulatory and Rectification Mechanism Failures	F11	Poor organizational management (F111), Unrectified issues (F112), Insufficient maintenance supervision by operator (F113), Insufficient supervision by regulatory authority (F114)
	Unauthorized Release	F12	Release of unairworthy components (F121), No (F122)

Table 2. Factor severity classification.

Type	Classification
Severity(S)	Hazardous (S1) Major (S2) Minor (S3)

This study collected a total of 1616 unsafe incidents from airlines between 2013 and 2025, supplemented by 90 incident cases from the Civil Aviation Safety Information Database of China covering 2018 to 2025. Initial data cleaning involved using Python software to match and remove special symbols via regular expressions; employing jieba to segment Chinese text into words; filtering meaningless terms using a predefined stopword list; standardizing text formats; and eliminating duplicates and short texts by removing duplicate rows and excessively brief entries, yielding 1364 valid records. Then, rule-based matching was applied to compare case texts against a keyword list, enabling rapid subcategory classification based on match results. For the remaining ambiguous texts, 20% were manually annotated for risk category training. TF-IDF was used to convert texts into numerical features, while the remaining 80% data trained a random forest model for text classification. Finally, classification results underwent manual review to correct misclassifications.

2.2. Bayesian risk analysis layer

The specific steps for constructing a Bayesian network model are as follows:

(a) Construct the network structure. Build a two-layer tree-structured BN composed of event outcomes and risk factors.

(b) Determine prior probabilities. Parameter learning mainly involves determining the conditional probability tables for each node. Machine learning is performed using XX sample data from a certain airline as the training set to determine the conditional probabilities of each node.

(c) Conduct risk analysis. Perform diagnostic inference to analyze the risk of event occurrence.

2.2.1. Structure learning

For the cleaned data, a tree structure is introduced to construct the Bayesian Network (BN) model. Tree construction involves structural learning using the TAN algorithm and parameter learning using the Counting Learning algorithm. This step is implemented through Python programming.

Structure learning first traverses all risk variable pairs, calculating the CMI value for each pair under a given "event severity" state, as shown in Equation (1). This serves as a quantitative indicator of their potential dependency. Next, using all risk variables as nodes and the computed CMI values as edge weights, an undirected graph encompassing all possible edges is constructed. Based on this fully undirected graph, the maximum spanning tree algorithm is applied to identify a spanning tree with the highest total weight. Specifically, the maximum spanning tree is constructed using the Kruskal algorithm, in which all edges are sorted in descending order according to their CMI weights and iteratively added while avoiding cycle formation. In this study, the MST construction is implemented in Python using the networkx library, where the maximum_spanning_tree() function is applied to the weighted undirected graph, ensuring an efficient and reproducible tree construction process. A root node is selected from the spanning tree, and all edges are directed outward from this root, converting the undirected graph into a directed graph. The "event severity" is introduced into the network as a categorical variable, and directed edges are added from this categorical variable to each risk variable. Equation (2) represents the functional relationship of the TAN structure, forming the tree structure shown in **Figure 1**.

$$S_{CMI}\left(F_I, F_J|I\right) = \sum_{F_{ii},F_{ji},S_i} P\left(F_{ii}, F_{ji}, S_i\right) \log_b \frac{P\left(F_{ii}, F_{ji}|S_i\right)}{P\left(F_{ii}|S_i\right) P\left(F_{ji}|S_i\right)} \tag{1}$$

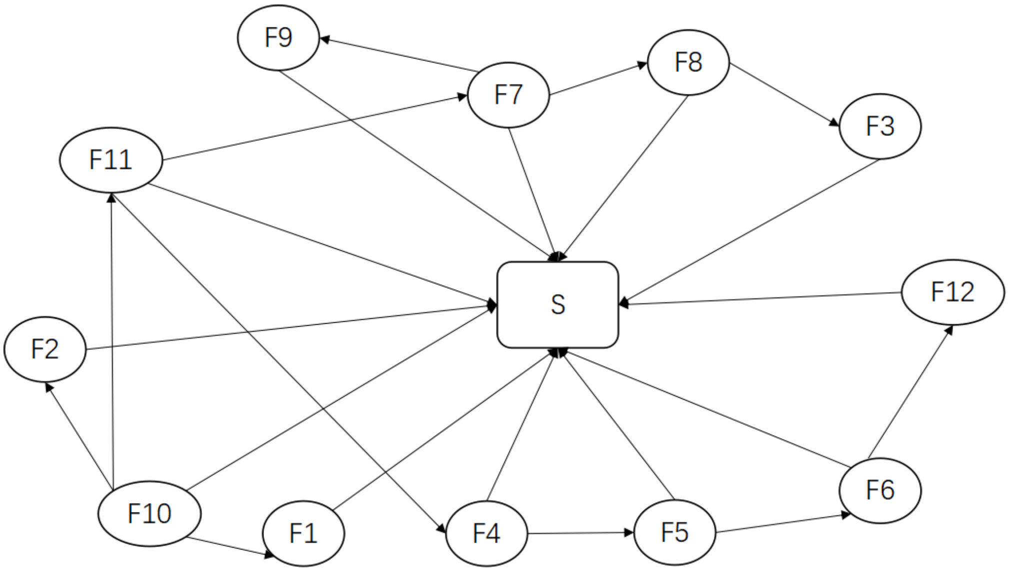

Figure 1. Tree structure of risk factors and event formation.

$$\prod I_i = \begin{cases} \{I, E_{\pi(i)}\}, \pi(i) > 0 \\ \{I\} \qquad \pi(i) = 0 \end{cases} \tag{2}$$

Where S_{CMI} represents the conditional mutual information between two risk factor variables, F_{ii} represents the i-th state of the risk factor variable F_i, F_{ji} represents the i-th state of the risk factor variable F_j, S_i represents the i-th state of "the probability of maintenance event occurrence", and usually b is set to 2.

2.2.2. Parameter learning

For the resulting tree-structured network, the conditional probability table for each risk factor node is parameterized using a counting learning algorithm, with the counting algorithm formula defined in Equation (3). By traversing the dataset to identify the value of risk factor F and the parent node value S, the count for each parent node is determined, and the conditional probability table (CPT) is estimated for each F and S using the formula in Equation (4). To address potential zero-count scenarios arising from sparse observations, Laplace smoothing is applied during parameter estimation. Specifically, a small positive constant is added to each frequency count before normalization, ensuring that no conditional probability is assigned a zero value and improving the numerical stability and generalization ability of the learned CPTs. This yields the conditional probability distribution characteristics of each node's value under different combinations of parent node states, ultimately forming the TAN-BN structure. This step is implemented using Python programming.

$$P(F_1, \cdots, F_n, S) = P(I) \cdot \prod_{i=1}^{n} P(\mathrm{F}_i|S) \tag{3}$$

In the formula: $P(F_1, \cdots, F_n, S)$ represents the joint probability distribution of the network; "the probability of maintenance event occurrence" I is regarded as the only parent variable of the risk factor variable set $F = \{F_1, \cdots, F_i$ ($\Pi = \{S\}$, $1 \le i \le 12$), and S has no parent variables ($\Pi = \varnothing$).

$$\hat{P}_{\text{MLE}}(F = f|P_a(F) = s) = \frac{N[f, s]}{N[s]} \tag{4}$$

In the formula, $N[s]$ = the number of samples in the data where $P_a(F) = s$; $N[f, s]$ = the number of samples in the data that simultaneously satisfy $F = f$ and $P_a(F) = s$. The parent node set of F is $P_a(F)$.

Import the resulting TAN-BN model into the GeNIe software. This model employs a framework centered on the maintenance operation risk system, "organizational factors, unsafe supervision, preconditions for unsafe behavior, and unsafe behavior," and incorporates 37 key risk nodes, as illustrated in **Figure 2**.

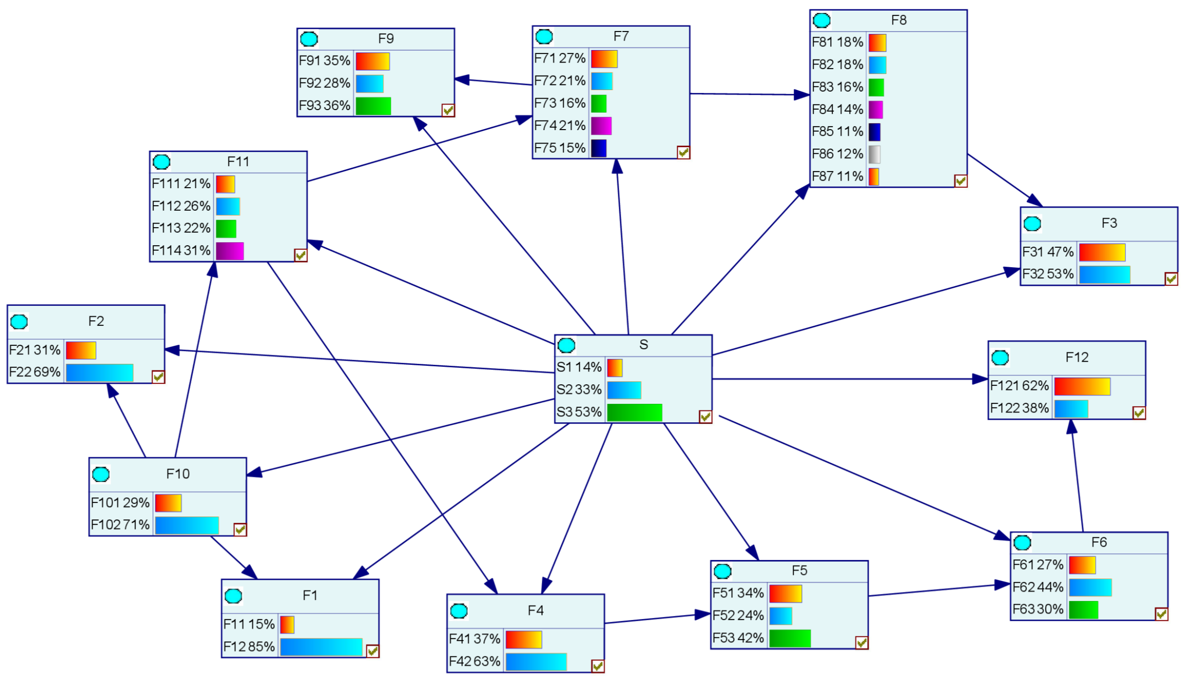

Figure 2. Tree Augmented Naïve Bayesian (TAN-BN) structure for civil aviation maintenance risk analysis.

2.3. Decision-making layer of system dynamics

Construct a maintenance management environment based on the risk factors extracted by the risk analysis layer, and use system dynamics for simulation decision analysis. The implementation steps of decision analysis are as follows:

(a) Problem definition and goal setting. Covering the entire process of maintenance organization from maintenance plan formulation, execution, inspection, acceptance, operation, and supervision, involving maintenance personnel factors and management factors.

(b) System modeling. Set according to the result indicators of risk factors established by the data analysis layer.

(c) Parameter setting and calibration. Establish system functions and parameter settings based on actual maintenance operation conditions and the expert scoring method.

(d) Scenario simulation and decision analysis. Simulate a certain maintenance management environment to obtain the optimal decision-making scheme through simulation.

2.3.1. Causal loop

A system dynamics model is constructed based on nine risk factors (F4–F12) to simulate the evolution process leading to the occurrence of maintenance events (I). The core concept is that organizational-level deficiencies (F10, F11) foster an unsafe working environment, leading to problems at the process and execution levels (F7, F8, and F9). This, in turn, triggers individual errors (F4, F5, and F6), and ultimately, these errors breach defenses, resulting in the occurrence of the event (I), as shown in **Figure 3**.

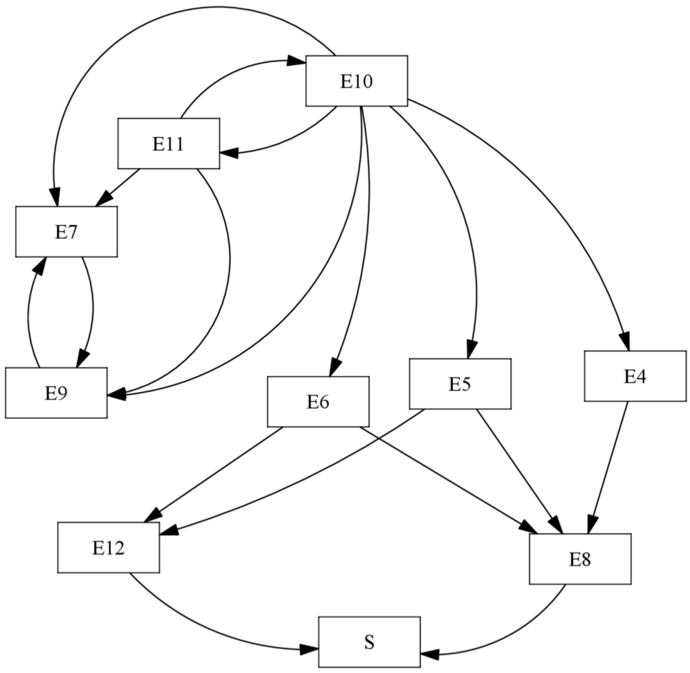

Figure 3. Causal loop.

The two core feedback loops are the vicious cycle and the process failure cycle. The vicious cycle, comprising F10 and F11 mutual exacerbation, forms a continuously deteriorating loop, which is the most fundamental driving loop in the entire model. The process failure cycle involves the mutual influence of F7 and F9, leading to the continuous degradation of the entire management and process foundation.

2.3.2. Stock-flow diagrams

To quantify the causal relationships in **Figure 2**, three core stock variables are defined to represent the accumulation of risks: Systemic Risk Foundation (SRF), Operational Process Risk (OPR), and Immediate Error Risk (IER). The SRF represents long-standing fundamental issues at the organizational and management level, primarily driven by F10 and F11. OPR represents the degree of imperfection in daily workflows and systems, mainly driven by E7 and E9. IER represents the level of impending or existing errors among frontline staff, primarily driven by F4, F5, F6, F8, and F12. The Initial values are SRF = 10 Points, OPR = 5 Points, ECR = 5 Points.

The expression can be described as:

$$\text{SRF}(t) = \text{SDR}(t-dt) - \text{SIR}(t-dt) \tag{5}$$

$$\text{OPR}(t) = \text{PDR}(t-dt) - \text{POR}(t-dt) \tag{6}$$

$$\text{IER}(t) = \text{EGR}(t-dt) - \text{ECR}(t-dt) \tag{7}$$

Systemic Deterioration Rate (SDR) and the Systemic Improvement Rate (SIR) describe the changes in the SRF, Process Degradation Rate (PDR) and Process Optimization Rate (POR) describe the changes in the OPR, the Error Generation Rate (EGR) and the Error Correction Rate (ECR) describe the changes in IER, which can be described mathematically as:

$$\text{SDR}(t) = 0.01 * \text{SRF}(t) * (1 - \text{SRF}(t)/100) \tag{8}$$

$$\text{SIR}(t) = 0.02 * \text{MI}(t) * \text{SRF}(t) \tag{9}$$

$$\text{PDR}(t) = 0.05 * \text{SRF}(t) * (1 - \text{SRF}(t)/100) + 0.1 * \text{IER}(t) \tag{10}$$

$$\text{OPR}(t) = 0.1 * \text{MI}(t) * \text{OPR}(t) \tag{11}$$

$$\text{EGR}(t) = (0.3 * \text{OPR}(t) + 0.2 * \text{SRF}(t)) * (1 - \text{IER}(t)/50) + \text{EP}(t) \tag{12}$$

$$\text{ECR}(t) = 0.4 * \text{IER}(t) \tag{13}$$

The rate parameters in the model were determined based on both system dynamics conventions and practical considerations. Specifically, the coefficient 0.01 in SDR(t) reflects the slow accumulation of systemic risk, while 0.02 in SIR(t) represents the mitigating effect of management input, which is higher than the natural deterioration rate. The coefficient 0.05 in PDR(t) captures the amplifying effect of systemic risk on process degradation, and 0.1 in OPR(t) reflects a moderate-strength management intervention at the operational level. Finally, 0.4 in ECR(t) indicates that once errors occur, corrective actions are implemented relatively quickly. These parameter values were selected to ensure numerical stability, realistic time delays, and non-explosive system behavior, consistent with prior system dynamics–based safety risk studies. Where MI represents Management Input, with a value range of a random number within [1]; EP represents external pressure, with a value range of a random number within [1]. The initial values, MI = 0.5 Dmnl, EP = 0.5 Points/Week.

Specifically, a panel of 5 aviation maintenance safety experts was invited to evaluate the relative influence strength between variables. Including maintenance managers, safety inspectors, and senior engineers, each with more than 5 years of experience. To quantify the relative influence strength between variables, a five-point Likert scale was adopted in the expert elicitation process, where 1 denotes very weak influence, 2 denotes weak influence, 3 denotes moderate influence, 4 denotes strong influence, and 5 denotes very strong influence.

Let $r_{ij}^{(k)}$ denote the score assigned by the k-th expert ($k = 1, 2, \cdots, N$) to the influence of factor *i* on factor *j*. The expert scoring matrix can thus be expressed as:

$$R_{ij} = \left\{ r_{ij}^{(1)}, r_{ij}^{(2)}, \cdots, r_{ij}^{(N)} \right\} \tag{14}$$

To obtain a consolidated evaluation of influence strength, the arithmetic mean of expert scores was calculated as:

$$\bar{r}_{ij} = \frac{1}{N} \sum_{k=1}^{N} r_{ij}^{(k)} \tag{15}$$

To translate expert judgment into quantitative model parameters, a linear normalization and interval mapping strategy was employed. Specifically, the averaged expert score was normalized into the interval [0,1] according to:

$$\mathrm{w_{ij}} = \frac{\mathrm{\bar{r}_{ij}} - 1}{4} \tag{16}$$

Subsequently, the normalized weights were mapped to predefined coefficient intervals commonly used in system dynamics modeling to represent different levels of causal influence. Weak, moderate, and strong influences were respectively mapped to the intervals 0.1–0.3, 0.3–0.5, and 0.5–0.7. These interval ranges were selected based on empirical conventions widely adopted in system dynamics and socio-technical risk modeling studies, where causal coefficients are typically constrained to moderate magnitudes to avoid excessive amplification or numerical instability. Specifically, values below 0.1 are generally considered to have negligible dynamic impact, while coefficients exceeding 0.7 may lead to unrealistic feedback dominance. The final coefficients within each interval were selected based on a combination of representativeness and preliminary simulation trials. In general, a coefficient near the midpoint of the mapped interval was preferred to reflect the average causal influence, while adjustments were made where necessary to ensure model stability and realistic dynamic behavior, as shown in **Table 3**.

Table 3. Expert scoring results and parameter mapping for model coefficients.

Causal relationship	Mean expert score (1–5)	Normalized weight	Mapping interval	Final coefficient
SRF→E10	4.2	0.80	strong (0.5–0.7)	0.5
SRF→E11	4.2	0.80	strong (0.5–0.7)	0.5
OPR→E7	4.5	0.88	strong (0.5–0.7)	0.7
OPR→E9	3.2	0.55	moderate (0.3–0.5)	0.3
E10→E4	3.0	0.50	moderate (0.3–0.5)	0.3
E10→E5	3.8	0.70	moderate (0.3–0.5)	0.4
E10→E6	3.1	0.53	moderate (0.3–0.5)	0.3
E7→E6	3.4	0.60	moderate (0.3–0.5)	0.3
E4→E8	3.6	0.65	moderate (0.3–0.5)	0.4
E5→E8	3.1	0.53	moderate (0.3–0.5)	0.3
E6→E8	3.0	0.50	moderate (0.3–0.5)	0.3
E5→E12	4.0	0.75	strong (0.5–0.7)	0.5
E6→E12	4.0	0.75	strong (0.5–0.7)	0.5
E8→S	4.5	0.88	strong (0.5–0.7)	0.6
E12→S	3.8	0.7	strong (0.5–0.7)	0.4

The functional relationships of the dynamic variable S with F4–F10 are described as follows:

$$\mathrm{E10}(t) = 0.5 * \mathrm{SRF}(t) \tag{17}$$

$$\mathrm{E11}(t) = 0.5 * \mathrm{SRF}(t) \tag{18}$$

$$\mathrm{E7}(t) = 0.7 * \mathrm{OPR}(t) \tag{19}$$

$$\mathrm{E9}(t) = 0.3 * \mathrm{OPR}(t) \tag{20}$$

$$\mathrm{E4}(t) = 0.3 * \mathrm{E10}(t) \tag{21}$$

$$\mathrm{E5}(t) = 0.4 * \mathrm{E10}(t) \tag{22}$$

$$\mathrm{E6}(t) = 0.3 * \mathrm{E10}(t) + 0.3 * \mathrm{E7}(t) \tag{23}$$

$$\mathrm{E8}(t) = 0.4 * \mathrm{E4}(t) + 0.3 * \mathrm{E5}(t) + 0.3 * \mathrm{E6}(t) \tag{24}$$

$$\mathrm{E12}(t) = 0.5 * \mathrm{E5}(t) + 0.5 * \mathrm{E6}(t) \tag{25}$$

$$\mathrm{S}(t) = (0.6 * \mathrm{E8}(t) + 0.4 * \mathrm{E12}(t)) * 10 \tag{26}$$

Based on the functional relationships among flow, stock, and variable, a flow-stock diagram was constructed to simulate the maintenance risk control process, and modeling and simulation were carried out using AnyLogic software, as shown in **Figure 4**.

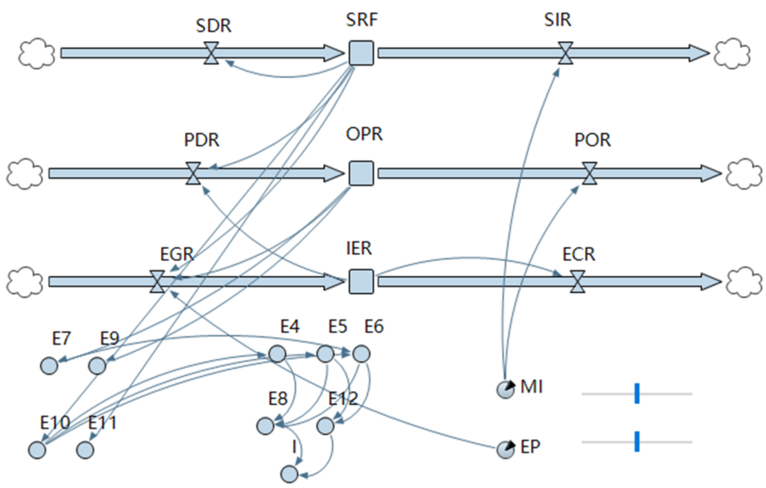

Figure 4. Stock–flow diagram of the system dynamics model.

3. Results and discussion

3.1. Bayesian network risk analysis results

After the BN parameters are determined, the BN structure and the conditional probability of each node are used to set the risk factors node of the event as the target variable, and the target value is set to 100% of the occurrence probability value. Simulation is performed using software to inversely deduce and solve the posterior probability of each node when maintenance events of different severity levels occur, and a comparative analysis of the prior and posterior probability results of each network node is conducted.

Bayesian network model simulation

The probability of event severity is shown in **Table 4**.

Table 4. Probability of event severity.

Event severity	Probability
S1	14%
S2	33%
S3	53%

In the maintenance operation environment of a certain airline, most risk consequences are relatively minor (S3). In **Table 4**, the value of S3 is the highest, indicating that under current Maintenance Scenarios, the losses and impacts caused by most potential risk events are at a low level. This reflects that the overall maintenance work is currently in a relatively controllable state. However, we must not let our guard down because of this, as there are still a small number of risks that may lead to serious consequences. Therefore, in terms of maintenance resource allocation, more resources can be appropriately invested in measures to prevent and respond to minor risks, so as to improve maintenance efficiency and quality.

In the field of risk assessment, the difference between the prior probability and the posterior probability can be used as a quantitative indicator to describe the degree of change of risk factors. **Table 5** shows the risk change values corresponding to each risk factor.

Table 5. Probability of risk factors.

Symbol	Initial	S1/%	S2/%	S3/%	Symbol	Initial	S1/%	S2/%	S3/%
F11	15	27	10	15	F81	18	19	18	18
F12	85	73	90	85	F82	18	17	19	17
F21	31	24	35	30	F83	16	20	13	17
F22	69	76	65	70	F84	14	8	14	15
F31	47	44	47	48	F85	11	8	14	10
F32	53	56	53	52	F86	12	17	10	12
F41	37	32	47	32	F87	11	10	11	10
F42	63	68	53	68	F91	35	34	31	38
F51	34	36	41	29	F92	28	29	35	24
F52	24	24	24	23	F93	36	37	34	37
F53	42	40	35	47	F101	29	36	35	24
F61	27	34	20	29	F102	71	64	65	76
F62	44	38	41	47	F111	12	41	13	20
F63	30	28	39	24	F112	26	22	19	32
F71	27	33	28	24	F113	22	21	31	16
F72	21	15	24	22	F114	31	16	36	32
F73	16	12	17	17	F121	62	47	60	68
F74	21	22	19	21	F122	38	53	40	32
F75	15	17	13	16					

In the field of risk assessment, the difference between the prior probability and the posterior probability can be used as a quantitative indicator to describe the degree of change of risk factors. **Table 6** shows the risk change values corresponding to each risk factor. The risk factor variation formula is:

$$TRI(FI_i) = \frac{\sum_{i=1} (P(FI_i) - P(F_i))^2}{\sum_{i=1} P(FI_i) - P(F_i)} \tag{27}$$

Table 6. Risk change values corresponding to each risk factor.

Risk factor	S1	S2	S3	Risk factor	S1	S2	S3
F1	12	5	0	F7	4.84	0.33	1.83
F2	7	4	1	F8	4.24	2.56	1
F3	3	0	1	F9	1	5.31	3.25
F4	5	10	5	F10	7	6	5
F5	2	7	4.64	F11	16.05	11.9	5.29
F6	5.93	7.32	4.45	F12	15	2	6

The Regulatory and Rectification Mechanism Failures (F11) can be regarded as a fundamental influencing factor in the maintenance risk system. As shown in **Table 6**, the value of F11 is the highest, which clearly reveals that during the actual maintenance process, multiple risk factors are intertwined and interact together, leading to a significant deviation of the risk situation of supervision and rectification mechanism failure from the initial estimation. In response to this situation, a multi-level supervision network should be established to clarify responsibilities at each level and strengthen daily inspections and irregular spot checks. Meanwhile, a tracking and feedback mechanism for rectification effects should be established to ensure the implementation of rectification measures. In addition, regular professional training should be provided to supervision and rectification personnel to improve their professional capabilities and sensitivity in risk identification.

(a) S1 analysis

Figure 5 is a probability graph of the risk factor when the risk probability of S1.

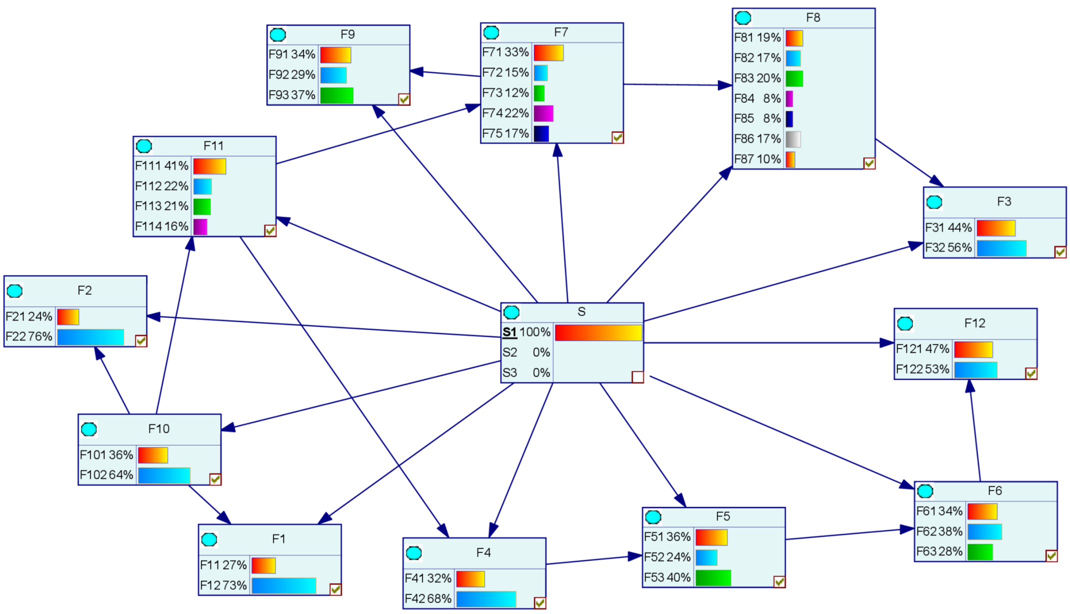

Figure 5. Posterior probability distribution of maintenance risk factors under Hazardous (S1) event conditions.

Hazardous-severity maintenance incidents are typically not triggered by a single operational error, but rather stem from the amplified effect of adverse working conditions coupled with systemic organizational deficiencies. When an incident severity is rated as dangerous (S1), severe weather (F1), night operations (F2), safety culture and resource shortages (F10), and improper release (F12) significantly emerge as critical risk factors. Adverse weather and night operations substantially increase operational uncertainty and cognitive load on personnel, making critical errors more likely to occur when safety resources and managerial support are inadequate. Safety culture and resource shortages reflect long-term organizational imbalances in time pressure and resource allocation, weakening risk identification, decision-making, and defense capabilities. Unauthorized release represents the final breach of safety barriers under multiple pressure transmissions. This characteristic indicates that the essence of S1 risks lies in systemic failures at the organizational decision-making and safety defense levels.

(b) S2 analysis

Compared to hazardous incidents, major incidents more often reflect deficiencies in personnel skills and cognition. When an incident severity is classified as significant (S2), skill and knowledge gaps (F4), unsafe behaviors and attitudes (F5), and cognitive and judgmental biases (F6) become the dominant risk factors. In scenarios where organizational and institutional systems have not yet experienced comprehensive failure, mismatched professional capabilities among maintenance personnel, lax safety attitudes, and cognitive biases are more likely to evolve into incidents with significant consequences through non-standard operations or judgment errors. This indicates that risks at this severity level primarily concentrate at the "human-machine-procedure" interface, necessitating governance priorities focused on competency development, behavioral standardization, and cognitive intervention.

Figure 6 is a probability graph of the risk factor when the risk probability of S2.

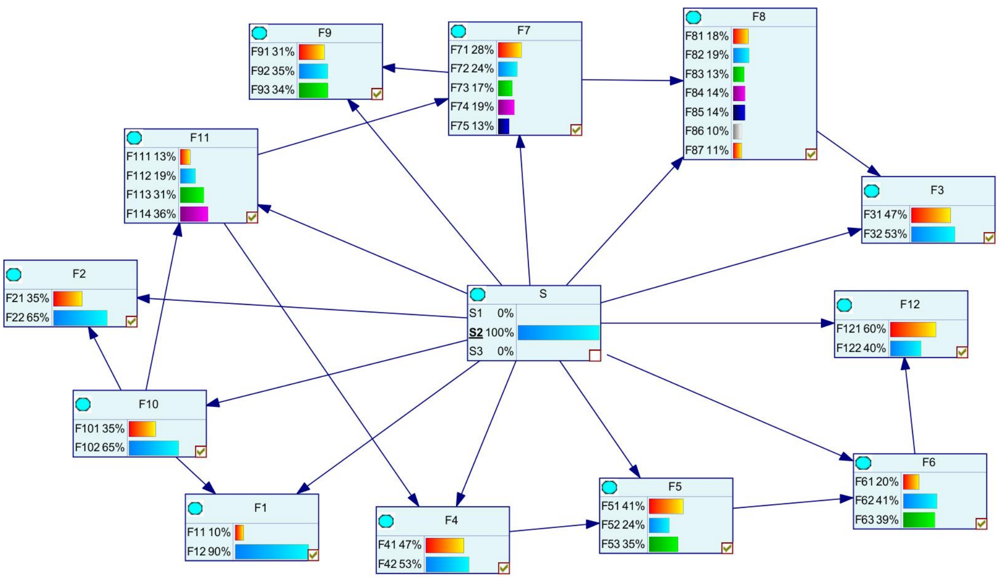

Figure 6. Posterior probability distribution of maintenance risk factors under Major (S2) event conditions.

(c) S3 analysis

Minor incidents typically stem from the gradual accumulation of slow-moving risks during routine operations. These incidents reflect both inadequate personnel capabilities and suboptimal compliance with regulations, as well as the organization's long-term neglect of low-severity risks. When an incident is classified as General Severity (S3), the key risk factors combine deficiencies in skills and knowledge (F04), shortcomings in safety culture and resources (F10), and unauthorized release (F12). Although such risks typically do not immediately lead to severe consequences, their persistent presence erodes the foundational safety of systems, creating conditions for higher severity incidents. Therefore, S3 risk governance should focus on process standardization, continuous capability enhancement, and proactive control of low-level violations and release actions.

Figure 7 is a probability graph of the risk factor when the risk probability of S3.

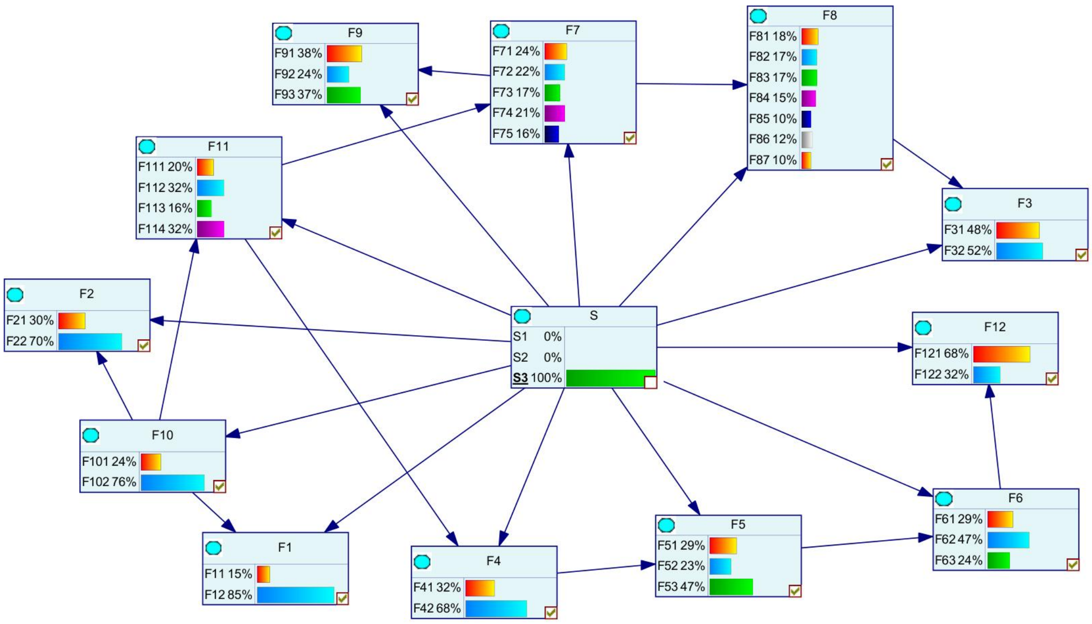

Figure 7. Posterior probability distribution of maintenance risk factors under Minor (S3) event conditions.

Overall, the combination of key risk factors across different severity levels exhibits a layered pattern: dominated by organizational and safety barriers (S1), personnel capabilities and behaviors (S2), and routine operations and slow-moving variables (S3). This indicates that maintenance safety risk governance requires tiered, differentiated management strategies rather than a uniform approach.

3.2. System dynamics decision analysis results

3.2.1. Analysis of simulation results for the reference scenario

The baseline scenario serves as a control condition with no additional pressure or management intervention, and all risks remain as they are. At this time, the impacts of SRF, OPR, and IER are shown in **Figure 8**.

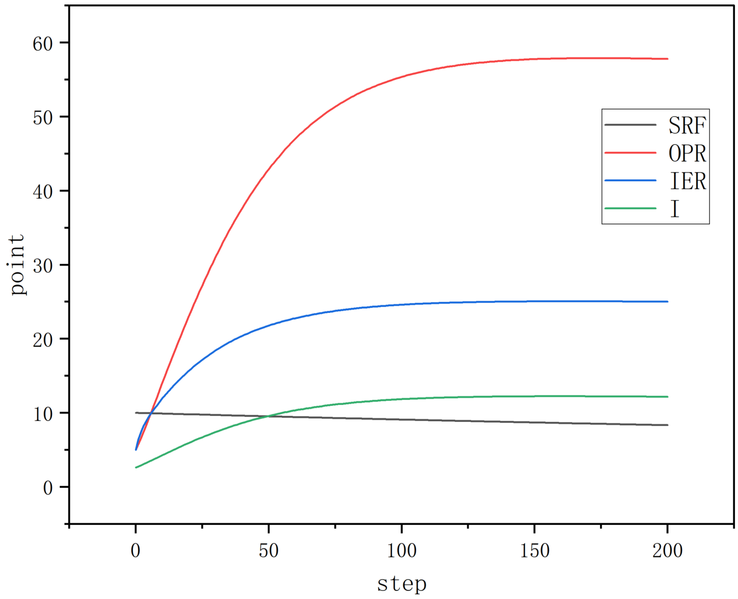

Figure 8. Simulation results of the reference scenario without management intervention.

From the perspective of the Reason Model, the dynamic evolution of the simulated risk variables reflects the progressive interaction between latent organizational conditions, degraded safety defenses, and active failures at the frontline.

In this study, the stock variable Systemic Risk Foundation (SRF) represents long-term latent deficiencies at the organizational and management levels, such as inadequate safety culture, insufficient resources, and ineffective supervision mechanisms. The slow decline of SRF in the baseline scenario indicates that organizational-level weaknesses tend to persist over extended periods and cannot be rapidly eliminated without sustained management intervention, which is consistent with the concept of latent conditions in the Swiss Cheese Model.

The Operational Process Risk (OPR) corresponds to the integrity of procedural and supervisory safety barriers. The rapid increase of OPR observed in the reference scenario suggests that when latent organizational deficiencies remain unaddressed, procedural defenses are progressively weakened, resulting in an increasing number of "holes" in the safety barriers. This behavior reflects the mechanism by which latent conditions translate into unsafe preconditions within operational processes.

The Immediate Error Risk (IER) represents the likelihood of active failures by frontline personnel. The sharp rise and subsequent high-level stabilization of IER indicate that once procedural defenses are compromised, frontline operators are exposed to sustained high-risk working conditions, under which human errors become more frequent and difficult to eliminate. This phenomenon aligns with Reason's assertion that active failures are often the visible manifestations of deeper systemic issues rather than isolated individual shortcomings.

The increase in the incident occurrence probability (I) reflects the cumulative alignment of multiple degraded safety barriers. As SRF, OPR, and IER simultaneously deteriorate, the probability that the "holes" in different layers of defense align increases, leading to a higher likelihood of incident occurrence.

Overall, the simulation results demonstrate that the system dynamics model reproduces the core causal logic of the Reason Model, namely that accidents emerge from the interaction and accumulation of organizational latent conditions, weakened defenses, and frontline active failures rather than from single-point failures.

3.2.2. Policy simulation

The decision to increase management input (MI) in the policy simulation is directly driven by the risk identification results of the BN model. The BN analysis identifies failures in supervision and rectification mechanisms at the organizational level as the dominant root risk factor, exhibiting the largest deviation between prior and posterior probabilities across severity scenarios. Within the SMS framework, such organizational deficiencies correspond primarily to weaknesses in Safety Policy and Objectives and Safety Assurance, which emphasize management commitment, allocation of adequate resources, and continuous monitoring and improvement. Increasing MI, therefore, represents a targeted response to the BN-identified root causes, operationalizing SMS principles by strengthening supervision intensity, corrective action effectiveness, and resource support for safety-critical activities. A baseline value of MI = 0.5 represents a routine operational state in which safety management practices meet standard regulatory requirements and organizational norms. The increase of MI to 0.8 simulates a reinforced management scenario commonly observed in aviation maintenance practice, such as intensified internal audits, increased supervisory staffing, enhanced procedural control, and focused rectification efforts following safety assessments or regulatory oversight. This range captures a realistic shift from routine compliance to proactive and strengthened safety governance, while remaining within feasible organizational limits. When simulating to the 100th step, the management input is increased from 0.5 to 0.8, as shown in **Figure 9**.

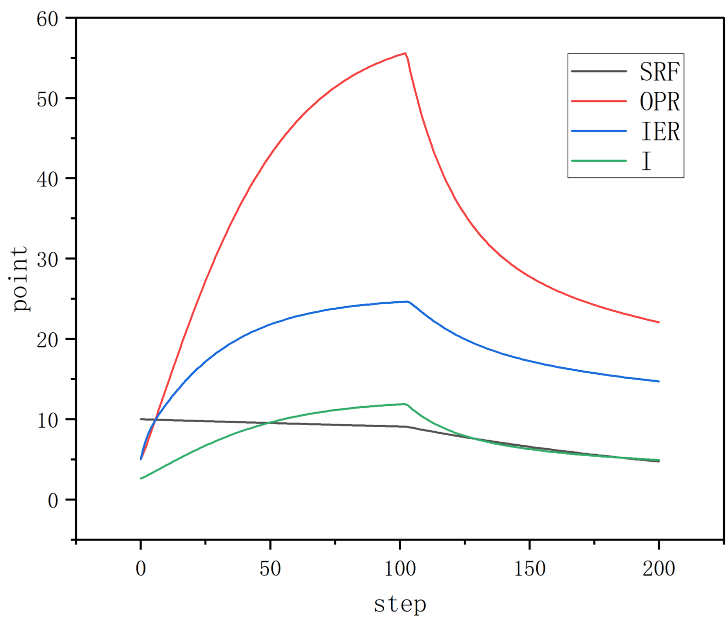

Figure 9. Simulation results of the policy intervention scenario with increased management input.

The simulation results indicate that enhanced management input leads to a rapid reduction in Operational Process Risk (OPR), which is consistent with practical SMS experience. Under increased management commitment, organizations typically intensify internal audits, improve maintenance procedures, enhance documentation quality, and allocate additional supervisory resources. These actions directly strengthen procedural defenses and improve compliance, thereby yielding an immediate and observable effect on process-related risks. This finding aligns with the SMS principle that proactive safety assurance measures are effective in identifying and correcting operational deficiencies before they propagate.

In contrast, the improvement in Systemic Risk Foundation (SRF) occurs at a slower rate following the increase in management input. This reflects the fact that organizational culture, leadership attitudes, and long-standing management practices cannot be transformed instantaneously through short-term interventions. According to SMS theory, the development of a positive safety culture and robust organizational capability requires sustained commitment, continuous feedback, and long-term investment. The delayed response of SRF in the simulation, therefore, mirrors the time-lag effect commonly observed in real-world SMS implementation.

Furthermore, although Immediate Error Risk (IER) decreases after the increase in management input, it does not immediately return to a low-risk state. This result highlights a fundamental SMS insight: while management actions can rapidly improve systems and processes, human performance risks are inherently dynamic and require ongoing training, monitoring, and learning mechanisms. The persistence of a residual error risk underscores the necessity of continuous safety promotion and learning, which are core elements of the SMS framework.

Overall, the policy simulation demonstrates that increasing management input is an effective risk mitigation strategy because it operationalizes key SMS principles within the maintenance system. By strengthening safety policy enforcement, enhancing safety assurance activities, and sustaining safety promotion efforts, increased management input reduces risk propagation across organizational, process, and human levels, thereby lowering the overall probability of incident occurrence.

3.3. Validation of BN and SD models

3.3.1. BN model

The conditional probability tables of the BN were learned from real maintenance data using counting-based parameter estimation. Parameter reasonableness was evaluated through predictive consistency and sensitivity analysis.

The inferred probability distribution of event severity shows that minor events account for the majority of outcomes, while hazardous events occur with relatively low probability. This distribution aligns with historical maintenance event characteristics observed in actual airline operations, where low-severity events dominate. Furthermore, sensitivity analysis was performed by setting the event node to a given severity state and observing posterior probability changes of parent risk factors. The results indicate that key organizational and management-related factors exhibit larger probability shifts, demonstrating that the BN effectively distinguishes critical risk contributors. These results confirm that the BN parameters are reasonable and that the model possesses meaningful diagnostic capability.

3.3.2. SD model

Behavioral validation was conducted by comparing simulation outcomes under different management scenarios. In the reference scenario without additional intervention, systemic, process, and frontline error risks gradually accumulate and stabilize at a high-risk level, reflecting typical maintenance system behavior under insufficient management input.

In contrast, the policy intervention scenario with increased management input demonstrates a clear reduction in process-related risks and a delayed but sustained decrease in frontline error risks. This delayed response is consistent with practical experience, as improvements in human performance and safety culture typically require longer time horizons than procedural corrections. The simulated system behavior, therefore, reproduces realistic risk evolution patterns observed in maintenance operations, supporting the behavioral validity of the SD model.

4. Conclusion

A novel data-driven approach for constructing a risk analysis and decision-making model for civil aviation maintenance incidents is proposed. The risk analysis layer extracts key risk factors, identifying that the existing maintenance management system possesses a certain risk buffer capacity under routine operating conditions, effectively suppressing the evolution of risks toward high severity. It reveals that combinations of key risk factors at different severity levels exhibit hierarchical characteristics: dominated by organizational and safety barriers at lower severity, personnel capability and behavior at medium severity, and daily operations and slow-moving variables at high severity. This indicates that maintenance safety risk governance requires tiered, differentiated management strategies. The decision layer simulates the long-term effects of management strategies. Increasing management investment significantly reduces process risks, while systemic risks and frontline errors require sustained, long-term investment and continuous improvement.

Author contributions: Conceptualization, JD and MC; methodology, JD and MC; software, JD; validation, JD; formal analysis, JD; investigation, JD; resources, JD; data curation, JD; writing—original draft preparation, JD; writing—review and editing, JD and MC; visualization, JD; supervision, MC; project administration, MC; funding acquisition, MC. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Safety Special Project—Civil Aviation Safety Supervision Capacity Building Project (HA202511) and National Key Research and Development Program of China (2023YFB4302901).

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Data availability statement: The data used in this study are available from the corresponding author upon reasonable request.

Conflict of interest: The authors declare no conflict of interest.

References

1. Strengthened profitability expected in 2025 even as supply chain issues persist. Available online: https://www.iata.org/en/pressroom/2024-releases/2024-12-10-01/ (accessed on 16 November 2025).
2. Esteban A, Zafra A, Ventura S. Data mining in predictive maintenance systems: A taxonomy and systematic review. WIREs Data Mining and Knowledge Discovery. 2022; 12(5): e1471.
3. Stanton I, Munir K, Ikram A, et al. Predictive maintenance analytics and implementation for aircraft: Challenges and opportunities. Systems Engineering. 2023; 26(2): 216–37.
4. Bauranov A, Rakas J. Bayesian network model of aviation safety: Impact of new communication technologies on mid-air collisions. Reliability Engineering & System Safety. 2024; 243: 109905.
5. Wang L, Wang Y, Ding Y. Risk analysis and resilience of humanitarian aviation supply chains: A Bayesian network approach. Applied Sciences. 2025; 15(19): 10508.
6. Pan W, Li Y, Jiang Y, et al. Analysis of controller-caused aviation accidents based on association rule algorithm and Bayesian network. Applied Sciences. 2025; 15(17): 9690.
7. Zhang X, Mahadevan S. Bayesian network modeling of accident investigation reports for aviation safety assessment. Reliability Engineering & System Safety. 2021; 209: 107371.
8. Zhou Z, Yu X, Zhu Z, et al. Development and application of a Bayesian network-based model for systematically reducing safety risks in the commercial air transportation system. Safety Science. 2023; 157: 105942.
9. Guo R, Wu J, Ji F, et al. Analysis of traffic safety in airport aircraft activity areas based on Bayesian networks and fault trees. Digital Transportation and Safety. 2024; 3(1): 8–18.
10. Liu H, Hu M, Yang L. A new risk level identification model for aviation safety. Engineering Applications of Artificial Intelligence. 2024; 136(Part A): 108901.
11. Dang X, Shao Y, Liu H, et al. Risk Analysis Method of Aviation Critical System Based on Bayesian Networks and Empirical Information Fusion. Electronics. 2025; 14(12): 2496.
12. Chen YH, Lin JJ. Determinants of flight delays at East Asian airports from an airport, route and network perspective. Journal of Air Transport Management. 2021; 94: 102064.
13. Barry DJ. Estimating runway veer-off risk using a Bayesian network with flight data. Transportation Research Part C: Emerging Technologies. 2021; 128: 103180.
14. Qi X, Liu C, Fan D, et al. Causal modeling and risk assessment of runway excursion events via interpretable machine learning. Reliability Engineering & System Safety. 2025; 269: 112055.
15. Zhao Y, Hua S, Ren X. Relevance research of threat/error and undesired states in air traffic management based on Bayesian network model. Journal of Air Transport Management. 2017; 60: 45–48.
16. Arnaldo Valdés RM, Gómez Comendador VF, Perez Sanz L, et al. Prediction of aircraft safety incidents using Bayesian inference and hierarchical structures. Safety Science. 2018; 104: 216–230.
17. Ifelebuegu AO, Awotu-Ukiri EO, Theophilus SC, et al. The application of Bayesian—Layer of Protection Analysis method for risk assessment of critical subsea gas compression systems. Process Safety and Environmental Protection. 2018; 113: 305–18.
18. He S, Shi S, Lin Z, et al. Risk assessment of coal mine gas explosion based on cloud model and Bayesian network. Reliability Engineering & System Safety. 2025: 111798.
19. Wang N, Yuen KF, Li D, et al. A bayesian network approach to ship safety assessment: Integrating machine learning and expert opinions. Reliability Engineering & System Safety. 2025; 111659.
20. Zhao L, Gao Q, Dong Y, et al. Vision-guided dynamic risk assessment for long-span PC continuous rigid-frame bridge construction through DEMATEL–ISM–DBN modelling. Buildings. 2025; 15(24): 4543.
21. Li X, Wang G, Gui Y. Enhancing subway fire safety with a symmetric framework: From fault tree analysis to dynamic Bayesian network inference. Symmetry. 2025; 17(12): 2090.
22. Huang X, Wang Q, Li X, et al. Navigation risk assessment of Arctic shipping routes based on Bayesian networks. Journal of Marine Science and Engineering. 2025; 13(12): 2306.

23. Zhang J, Wu L, Zhang T, et al. Systemic and dynamic risk analysis of drilling construction based on Bayesian network and system dynamics model. Scientific Reports. 2025; 15(1): 27989.
24. Bakhtiari S, Najafi MR, Goda K, et al. A dynamic Bayesian network approach to characterize multi-hazard risks and resilience in interconnected critical infrastructures. Reliability Engineering & System Safety. 2025; 257(Part A): 110815.
25. You Q, Guo J, Zeng S, Che H. A dynamic Bayesian network based reliability assessment method for short-term multi-round situation awareness considering round dependencies. Reliability Engineering & System Safety. 2024; 243: 109838.
26. Li X, Chen R, Ren Y, Jim CY. Intrinsic drivers of urban flood disasters from the resilience perspective in China. Reliability Engineering & System Safety. 2025; 265(Part B): 111596.
27. Yuan Z, Cai L, Xie Z, et al. Dynamic evolution and scenario-based prediction of urban flood resilience: A system dynamics modeling approach in Kunming, China. Journal of Environmental Management. 2025; 395: 127740.
28. Reason JT. Human Error. Cambridge University Press; 1990.