

A system-based safety analysis of AI-enabled UAV operations for runway FOD detection: application of STPA and PROMETHEE[☆]

Nafissa Abdouramane Attou Bounou^a, Gulsum Kubra Kaya^{a,*}, Fanny Camelia^b

^a Cranfield University, Faculty of Engineering and Applied Science, Safety and Accident Investigation Centre, Cranfield, UK

^b Cranfield University, Faculty of Engineering and Applied Science, Cranfield Defence and Security, Cranfield, UK

ARTICLE INFO

Keywords:

STPA
Hazard analysis
System safety
PROMETHEE
FOD detection
Runway inspection

ABSTRACT

Unmanned Aerial Vehicles (UAVs) supported by Artificial Intelligence (AI) offer a promising capability for detecting Foreign Object Debris (FOD) on runways in non-segregated airport operations. However, they are also introducing new operational and safety challenges. This study conducts a systems-based hazard analysis for such operations using the System-Theoretic Process Analysis (STPA) method. The safety recommendations generated from STPA findings are then ranked to support sequencing and tiering using the PROMETHEE method as decision support, providing a transparent basis for implementation under practical constraints without compromising comprehensive hazard coverage. Recommendations were ranked using 15 criteria, 11 of which were operationalised from STPA artefacts (hazards, unsafe control actions and loss scenarios) to quantify risk contribution and traceability. This study identified 8 hazards, 149 unsafe control actions, 188 loss scenarios and 161 safety recommendations. The combined STPA and PROMETHEE analysis emphasises the importance of human-automation coordination, operator training, situational awareness and system-level integration in ensuring the safe deployment of FOD detection UAVs. Overall, the study demonstrates an STPA-grounded multi-criteria decision-making prioritisation workflow in which recommendation scores are derived from, and traceable to, STPA artefacts, enabling transparent near-term implementation planning under practical resource and operational constraints for AI-enabled UAV runway inspection.

1. Introduction

Despite continuous improvements in aviation safety, runway safety remains one of the most critical concerns (International Civil Aviation Organization [ICAO], 2018b). It is vulnerable to a range of hazards and operational risks on the runway surface and in the runway environment, including runway incursions and excursions, degraded surface conditions, and obstacles that threaten aircraft during take-off and landing. One critical operational process in runway safety is the detection, classification and timely removal of runway Foreign Object Debris (FOD), which can include aircraft parts, pavement fragments, wildlife, environmental debris and even personal belongings (Fizza et al., 2021).

Traditionally, FOD detection has relied on visual inspections carried out by ground personnel. Although generally effective, these inspections can be compromised in poor visibility or extreme weather, limiting accurate localisation of detected debris (Fizza et al., 2021; Parker et al., 2022). To address these challenges, recent work has explored

integrating Unmanned Aerial Vehicles (UAVs) with Artificial Intelligence (AI), as a complementary approach to FOD detection and localisation (Papadopoulos & Gonzalez, 2021; Parker et al., 2022). In proof-of-concept studies, UAVs can execute pre-planned waypoint flight paths to collect high-resolution runway imagery, while AI automates detection and supports geolocation (Shan et al., 2025). However, their implementation also introduces new operational challenges and safety concerns, therefore, requiring additional safety and regulatory measures (M'Sila et al., 2022; Nakamura & Kajikawa, 2018; Scott & Kurtupinar, 2026; Shan et al., 2025). Accordingly, airport operators should conduct a structured risk assessment prior to deploying the UAVs (Civil Air Navigation Services Organization [CANSO], 2011).

Airport operations constitute a complex socio-technical system, in which multiple actors, such as ATC, airside operations, maintenance, contractors and flight crews, must coordinate effectively under time pressure and variable environmental conditions (Chikha & Skorupski, 2022; Liou et al., 2026; Studic, 2015). In such systems, assuring runway

[☆] This article is part of a special issue entitled: 'STAMP Frontiers and Tools' published in Safety Science.

* Corresponding author at: College Road, Wharley End, Martell House, MK43 0AL, Cranfield University, Faculty of Engineering and Applied Science, Cranfield, UK.
E-mail address: kubra.kaya@cranfield.ac.uk (G.K. Kaya).

safety for AI-enabled UAV operations cannot rely solely on component-level reliability. Unsafe control actions, mismatched assumptions, inadequate coordination or deficiencies in feedback and oversight can lead to hazardous states even when no physical component has failed (Leveson, 2011; Leveson & Thomas, 2018). While traditional techniques such as Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis (FMEA) remain useful in analysing component reliability, their largely linear failure-centric framing can limit their ability to capture emergent behaviour and organisational, coordination or human-system interactions that are central to modern, tightly coupled airport environments (Kaya et al., 2025; Studic et al., 2017).

For these reasons, this study adopts a systems-thinking perspective and applies System-Theoretic Process Analysis (STPA) to identify hazards and derive safety recommendations for an AI-enabled UAV for FOD detection in the airport context. STPA treats safety as a control problem and examines how design flaws in the system, such as inadequate control and feedback loops, can lead to losses (Leveson, 2011). Although its application in AI-enabled UAVs for FOD detection is limited due to the system’s recent development, STPA has been widely applied to analyse hazards in UAV operations and AI-enabled and safety-critical automation (Patriarca et al., 2022; Charalampidou et al., 2020, 2024; Chatzimichailidou et al., 2017; Dghaym et al., 2021; Plioutsias et al., 2018; Rothmund et al., 2025; Chen et al., 2015; Hobbs et al., 2018; Pennington et al., 2026) supporting its suitability for this case study.

As STPA can generate many safety requirements or measures, organisations often require a transparent way to sequence and tier their implementation under resource and operational constraints, without implying that lower-ranked measures are unnecessary. Kim et al. (2021) note that, under limited resources and large findings, implementation focus may shift to less critical items, potentially leaving higher-risk items under-addressed. In many organisations, prioritisation is performed through risk management processes based on risk matrices and scoring mechanisms. Yet, the literature highlights limitations of risk matrices and risk scoring, particularly their potential to mislead in mid-range risk categories, thereby affecting the prioritisation of resulting actions (Baybutt, 2015; Cox, 2008; Duijm, 2015; Kaya et al., 2019). Moreover, prioritisation based solely on risk level can be insufficient, as decisions typically involve additional criteria that may conflict, such as implementation cost and duration.

To overcome this limitation, Multi-Criteria Decision-Making (MCDM) methods provide a structured approach for prioritising actions across multiple criteria (Aslantas et al., 2025; Cheraghi et al., 2022; Gul et al., 2022; Senić et al., 2025; Siraj et al., 2023). Accordingly, this study uses Preference Ranking Organization METHod for Enrichment Evaluations (PROMETHEE) to prioritise the safety recommendations using criteria operationalised from and traceable to STPA artefacts (e.g., hazards, unsafe control actions and loss scenarios). PROMETHEE involves comparing each alternative against every other alternative in pairs, providing a ranking of the alternatives based on their relative preferences (Brans et al., 1986; Brans & Mareschal, 2005; Brans & Vincke, 1985). This method was considered suitable due to its ability to perform pairwise comparisons and prioritise safety recommendations, as well as its well-known use in prioritisation problems (Coquelet et al., 2025; da Cunha et al., 2022; Dejaegere & De Smet, 2023; Mareschal, 2013). While there is no single, universally ‘best’ method for prioritisation, PROMETHEE provides a transparent outranking approach for the STPA-grounded criteria used in this study. The process of applying PROMETHEE to the STPA-grounded criteria can also be implemented using other prioritisation-focused MCDM methods to support methodological comparison or organisational preference. Moreover, the STPA-PROMETHEE workflow methodology can be transferred to any safety-critical process or industry.

This study aims to apply STPA to the runway FOD detection service using AI-enabled UAVs, conduct a hazard analysis, and use PROMETHEE to support transparent near-term implementation sequencing and tiering of the STPA-driven recommendations under practical

constraints. Specifically, this study operationalises STPA artefacts as MCDM inputs by scoring each recommendation using criteria derived from its linked STPA evidence: loss-scenario attributes (impact, likelihood and detectability) and the UCA (conditional to loss scenario)-to-hazard propagation potential based on the UCA-hazard mapping, alongside feasibility criteria (Table 2). When a recommendation is linked to multiple loss scenarios, risk-contribution criteria use a worst-case (maximum) inheritance rule across the linked loss scenarios to avoid masking high-consequence pathways during implementation sequencing. Hazard coverage remains represented through the hazard-propagation criteria. This preserves traceability to safety-relevant pathways without excluding broader hazard coverage, while supporting staged implementation planning. The primary contribution is a traceable STPA-to-MCDM workflow that converts a large STPA output set into an auditable implementation-priority list, while preserving explicit links to the underlying STPA evidence. In doing so, this study addresses a common practical challenge in STPA applications: converting large recommendation sets into a transparent, resource-aware implementation order while preserving links to STPA outputs.

2. Methods

This research design comprises two stages: (1) STPA-based hazard analysis and (2) PROMETHEE-based prioritisation of the resulting safety recommendations using assessment criteria operationalised from STPA artefacts. Because STPA is inherently iterative, PROMETHEE prioritisation is performed only after completing the STPA iterations for the target concept. The resulting hazards, UCAs and loss scenarios from the completed iteration provide the baseline for the evaluation matrix and implementation tiers.

2.1. STPA

2.1.1. STPA steps

STPA was applied in four key steps, as described by Leveson and Thomas (2018). The analysis was conducted iteratively and was informed throughout by the review of multiple document types, including academic literature, regulatory and guidance material (ICAO, 2015, 2018a, 2020b; EASA, 2024; FAA, 2021; FAA, 2024), operational and industry documents (CANSO, 2011; ACI, 2020), and available UAV incidents and accidents data (ATSB, 2025; UK CAA, 2023). Document review was not treated as a separate preliminary phase. It was embedded within each STPA step to ground assumptions about operational context, roles and responsibilities, and information and design flows.

For Step 1, the reviewed sources were utilised to establish the operational context and identify losses and hazards relevant to the integration of AI-enabled UAVs in the airport environment. The analysed system is a runway FOD detection service delivered by an AI-enabled UAV operating within non-segregated airport operations. The service comprises mission request and authorisation, flight planning, UAV launch and runway inspection, sensing and AI-based detection and localisation, alerting and communication to stakeholders and

Table 1
Subject matter expert (SME) interviews.

SME ID	Experience	Length of interview (min)	STPA validation in
1	ATS instructor and Head of operations	50 (online)	Step 1 and 2
2	Drone operator	40 (online)	Step 4
3	Military drone pilot	70 (online)	Step 1, 2 and 3
4	ATC with PhD in computer vision	70 (online)	Step 3 and 4
5	ATC and Runway inspection team	40 (online)	Step 1 and 2
6	Manager for AI solutions with several projects for object detection	65 (online)	Step 1, 2 and 3

Table 2

Assessment criteria and scale.

Criteria	Criteria Explanation	Assessment Scale	Scenario 1 Weights	Scenario 2 Weights
C1. Loss Scenario Impact	Severity of the linked loss scenario(s) mitigated by the recommendation	(1) Negligible, (2) Minor, (3) Moderate, (4) Major, (5) Catastrophic	0.067	0.1
C2. Loss Scenario Likelihood	Likelihood of the linked loss scenario(s) under the operational concept	(1) Very Unlikely, (2) Unlikely, (3) Possible, (4) Likely, (5) Almost Certain	0.067	0.05
C3– C10. The possibility of UCA (conditional to LS) leading to hazard 1 to 8.	For each hazard, likelihood that the linked UCA (under the LS context) escalates to the hazard	(1) Very Unlikely, (2) Unlikely, (3) Possible, (4) Likely, (5) Almost Certain	0.067	0.05
C11. LS detectability	Difficulty of detecting the linked loss scenario/ precursor before the occurrence of loss.	(1) Very easy to detect, (2) Easy, (3) Moderate, (4) Difficult, (5) Almost Impossible	0.067	0.05
C12. Cost of Recommendation	Relative implementation cost of the recommendation	(5) Very low, (4) Low, (3) Moderate, (2) High, (1) Very high	0.067	0.1
C13. Duration of the implementation of the recommendation	Relative implementation time and effort	(5) Very low, (4) Low, (3) Moderate, (2) High, (1) Very high	0.067	0.1
C14. Ease of implementing the recommendation	Practical ease of implementation	(5) Very easy, (4) Easy, (3) Moderate, (2) Difficult, (1) Very difficult	0.067	0.1
C15. Potential effectiveness of the recommendation in mitigating LS	Expected effectiveness in preventing or mitigating the linked loss scenario(s)	(5) Very effective, (4) Effective, (3) Moderate, (2) Ineffective, (1) Very ineffective	0.067	0.1

coordination of runway access and response actions. The boundary of analysis is the runway FOD detection service controller (Fig. 1), treated as the integration point where autonomy and perception outputs are translated into operational decisions under ATC clearances and airside coordination constraints. Key assumptions follow the architecture proposed by Gutiérrez (2023): a DJI Matrice 300 equipped with a camera conducts runway overflights within radio line-of-sight and streams footage to a Ground Base Station (GBS) for processing. A flight plan is uploaded before each inspection to enable autonomous flight, with remote piloting available when required. When FOD is detected, the runway inspection team is dispatched to remove it. In this study, the detection model is assumed to be offline-trained and deployed as a fixed version during a given operational period (i.e., no online learning during missions). Model updates, such as retraining or database expansion, are

treated as controlled configuration changes subject to governance, validation and operational approval before redeployment. This assumption limits the STPA scope to operational use and managed lifecycle updates, rather than to autonomous on-the-fly adaptation.

For Step 2, the same sources were used to determine controllers, controlled processes and control and feedback loops. The control structure was then developed in draw.io and refined as additional information emerged during reviews or subsequent steps. For Step 3, Unsafe Control Actions (UCAs) were generated by systematically applying the four UCA types (e.g., not providing, providing, providing but too early, too late, or in the wrong order, and stopped too soon or applied too long) to each identified control action. Although UCAs were analytically generated from the control actions, the reviewed documents influenced the interpretation of operational conditions, constraints and

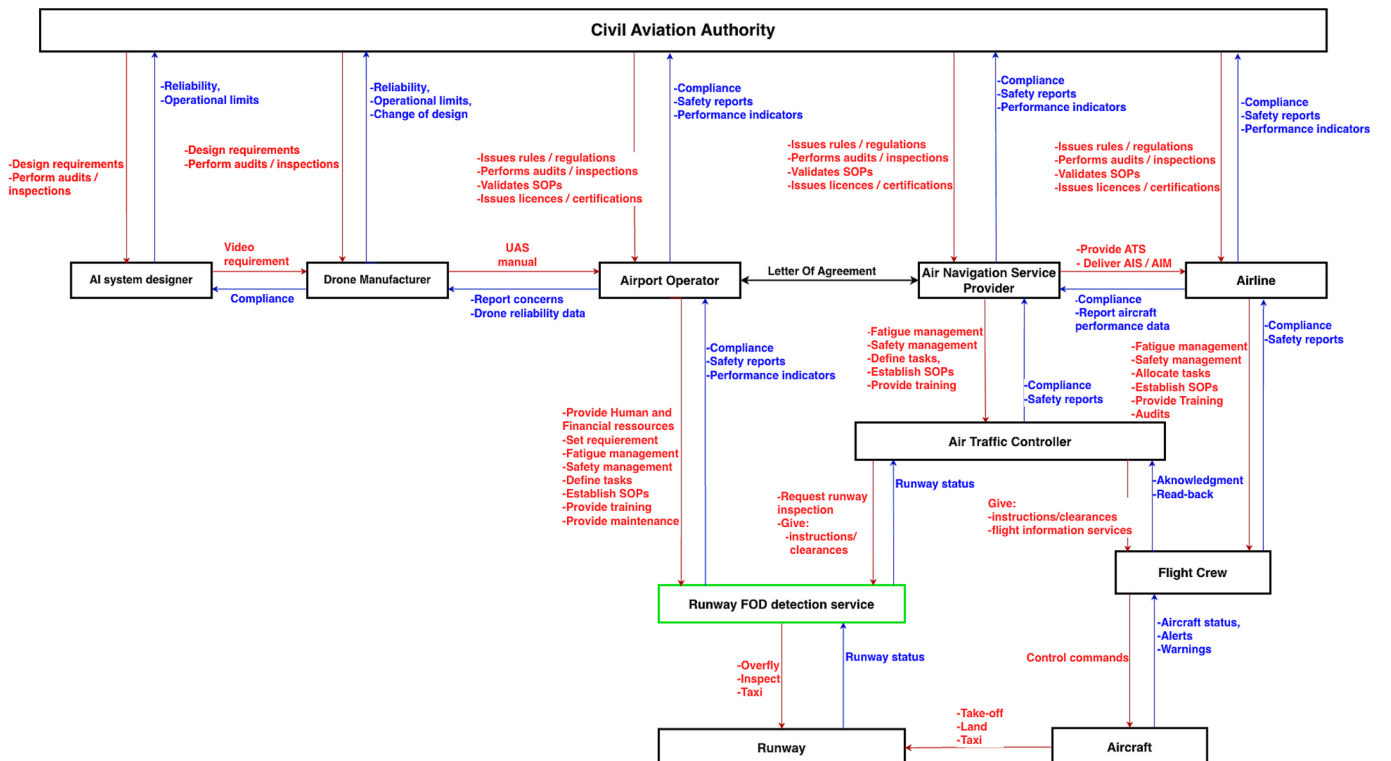


Fig. 1. The high-level control structure for the runway FOD detection process.

typical interactions. In addition, for each controller, where relevant, the interpretation of UCAs and the development of loss scenarios considered the controller's assumptions and beliefs about the system and environmental states (i.e., process models). This includes examining potential discrepancies between the controller's internal representation and the actual system or environmental state. For example, for CA-1 ("Follow mission trajectory"), UCA-8 captures the UAV initiating the mission trajectory before ATC clearance is granted. This reflects a potential process-model mismatch in which the drone pilot interprets an ATC request for runway inspection as implied clearance, while the actual operational state is that no explicit clearance has been issued. This mismatch is reflected in LS13 and the associated recommendation SR21. Representative contextualised UCA examples are illustrated in Table 4, and the detailed linked findings are documented in Appendix 2. For Step 4, Loss Scenarios (LS) were developed by revisiting the control structure and analysing how and why UCAs could occur and lead to hazards. Scenario development was informed by the reviewed sources, where available, and complemented by structured analyst reasoning to identify plausible human, procedural, organisational and technical contributors, such as process model flaws, inadequate feedback and communication and coordination breakdowns.

Finally, safety recommendations were derived by tracing causal factors documented in each LS to interventions that either prevent the associated UCA(s) (by addressing the identified causal mechanisms) or mitigate their consequences (by reducing the escalation of hazards and losses). Recommendations were coded by primary intervention type: technical (T), organisational (O) or human (H). All outputs were reviewed by the authors and revised iteratively to improve

completeness, internal consistency and traceability across STPA steps.

2.1.2. STPA validation

The STPA validation was conducted using a validation framework proposed by Sadeghi & Goerlandt (2023a). This framework supports evaluating the correctness and completeness of STPA steps, including the use of convergent, content and face validation methods. Six Subject Matter Experts (SMEs) participated in this process, providing feedback to build confidence in the analysis and thereby address the risk of analyst bias (Pitchforth & Mengersen, 2013). SMEs reviewed selected STPA artefacts for each step's logical consistency, accuracy and relevance to real-world operations, with iterations based on their feedback. STPA Steps 1 and 2 were validated through convergent validation, using a prior study of UAV systems (Charalampidou et al., 2020, 2024; Chatzimichailidou et al., 2017; d'Ambrosio et al., 2025; Dghaym et al., 2021; Marliere et al., 2025; Pan et al., 2023; Plioutsias et al., 2018; Rothmund et al., 2025). SMEs participated in content and structure validation by reviewing documentation, discussing UCAs and LSs, and ensuring consistency with operational standards.

Table 1 details the involvement of SMEs across each validation step. Specifically, the Step 3 findings were subject to face validation through interviews with three SMEs, focusing on the logical consistency, accuracy and completeness of the identified UCAs. Meanwhile, the Step 4 findings were validated through face validation involving two SMEs. Moreover, historical validation, utilising the contributing factors of previous incidents and accidents involving drones (ATSB, 2025; UK CAA, 2023), was also employed to validate Step 4. Feedback from the validation activities resulted in minor refinements. Examples included refining controller role boundaries (e.g., ATC and runway inspection team responsibilities), clarifying decision-making controllers and the associated sensing and feedback pathways in the control structure (e.g., the localisation function and FPV camera feedback stream), and rewording selected UCAs to align with STPA UCA categories. These refinements improved the interpretability and operational realism of the artefacts. Finally, the authors reviewed the STPA outputs for completeness and internal consistency and incorporated these refinements into the final reported results.

2.2. PROMETHEE

2.2.1. Criteria, assessment scale and preference functions

In this study, the methodological contribution lies in the construction of criteria and scoring for STPA artefacts, and in linking them back to those artefacts. A total of 15 criteria were identified to prioritise safety recommendations (see Table 2). Of these, 11 criteria were grounded in STPA artefacts, using the LSs, UCAs and hazard mappings to characterise risk contribution, specifically, analyst-assigned ratings of LS impact, likelihood and detectability, and the UCA (conditional to LS)-to-hazard propagation potential (i.e., the extent to which a given UCA occurring under an LS could escalate to each system-level hazard). Each safety recommendation is traceable to one or more LSs, their associated UCA (s), and the hazard(s). The remaining 4 criteria (C12- C15) were related to the feasibility of implementing the safety recommendations. This supports conservative implementation sequencing while maintaining traceability to hazard coverage via the propagation criteria. Similar risk level and feasibility-related evaluation criteria were identified by others (Cheraghi et al., 2022; Gul, 2018; Senić et al., 2025).

Where a recommendation mitigated multiple LSs, criterion scores C1-C11 were assigned using a conservative aggregation rule: the recommendation inherits the maximum (worst-case) impact, likelihood, low-detectability and hazard-propagation scores across its linked LSs. For example, if a recommendation is linked to LS-1 (impact (C1) = 3 and likelihood (C2) = 2) and LS-2 (C1 = 5 and C2 = 1), it inherits C1 = 5 and C2 = 2 (and likewise the maximum low-detectability and hazard-propagation scores across linked loss scenarios). This aggregation does not exclude or diminish lower-severity loss scenarios: all linked loss

Table 3
System-level hazards and constraints.

System-level hazards (H) and constraints (SC)	Linked Losses
H-1. UAV violates the safe separation from other flying aircraft during operations SC-1. UAV must maintain a minimum separation distance from other flying aircraft during operations, as specified in operational guidelines.	L1, L2, L4, L5, L6, L7
H-2. Vehicles (UAV or inspection vehicle) violate the safe separation from terrain or objects on the ground during operations SC-2. Vehicles (UAV and inspection vehicle) must maintain a safe separation distance from terrain, infrastructure, and objects on the ground.	L2, L3, L5, L6, L7
H-3. Detection system is unable to detect FOD on the runway SC-3. The detection system must reliably identify all Foreign Object Debris (FOD) on the runway, with minimal false negatives or false positives.	L1, L2, L4, L5, L6, L7
H-4. Safety-critical operational data is missing, stale or corrupted such that control and coordination decisions may be made using incorrect or untrusted information. SC-4. Operational decisions must be based only on validated and available safety-critical data.	L1- L7
H-5. UAV and GBS operated in a degraded or out-of-configuration condition SC-5. UAV and GBS equipment must be subject to regular maintenance to ensure operational reliability and safety.	L1-L7
H-6. Runway inspection is interrupted or incomplete such that runway FOD status is unknown while the runway may be used for aircraft operations. SC-6. If inspection is interrupted, the system must manage runway status explicitly.	L4, L5, L7
H-7. UAV loss of control SC-7. The UAV must maintain its airframe integrity and control at all times, with built-in fail-safes that activate in case of loss of control.	L2, L3, L5, L6, L7
H-8. Inadequate communication or coordination between drone pilot, ATC and runway inspection team SC-8. Communication and coordination between the drone pilot, ATC and runway inspection team must be seamless and occur according to established protocols to ensure safe and efficient operations.	L1, L2, L4, L5, L6, L7

Table 4

Representative, contextualised STPA findings (full findings are provided in Appendix 2).

Controller/ Control- relevant Functions	Control Action	Unsafe Control Action and Controller Constraints	Loss Scenario	Recommendation
UAV control system	CA-1: Follow mission trajectory	UCA-8: UAV follows mission trajectory before ATC clearance is granted while runway-entry and inspection require explicit ATC clearance and the runway movement area is active or may be used by aircraft, (H-1, H-2, H-6) <i>CC-8: UAV must only follow the mission trajectory after ATC clearance granted.</i>	LS13: UCA-8 might occur if the drone pilot assumes implied clearance following an ATC request for runway inspection.	SR21: Coordination protocol developed by airport management must include a positive clearance confirmation from ATC before UAV mission initiation (O).
UAV automation	CA-4 Control commands	UCA-21: UAV automation sends inappropriate commands to UAV subsystems when the UAV is operating in the runway movement area or within defined safety buffers (H-1, H-2, H-6) <i>CC-21: Appropriate commands must be sent to UAV subsystems by the UAV automation.</i>	LS32: UCA-21 might occur if the drone pilot filled the flight plan with incorrect data (e.g. wrong route coordinate, inappropriate flight level and speed changes).	SR20: Airfield maintenance department must create and save pre-filled flight plans in the system for each runway of the airport, which only necessitate a few changes from the drone pilot before each operation (O, T).
Flight Controller (FC)	CA-5: Execute pilot commands	UCA-26: UAV FC executes commands from unknown third party during any operational state in which command authenticity and integrity are required for safe control (H-1, H-2, H-3, H-4, H-6, H-7) <i>CC-26: UAV FC must be prevented from executing commands from an unknown third party.</i>	LS43: UCA-26 might occur due to signal spoofing or an unsecured C2 link.	SR14: Airport management must implement security protocols such as encrypted communication protocols with strong authentication between the UAV and GBS, secure data storage, and use only secured devices (O, T).
Payload camera system	CA-7: Capture and transfer live video/images	UCA-36: Payload camera system captures a low-quality runway video during runway inspection (H-3) <i>CC-36: The payload camera system must capture high-quality runway video.</i>	LS55: UCA-36 might occur due to low-resolution camera settings.	SR69: Drone manufacturer must provide a camera system with high resolution standards for FOD detection (T).
Detection controller	CA-9: Detect objects	UCA-47: Algorithm is unable to detect an FOD which is present in the video feed during an active runway inspection (H-3) <i>CC-47: Algorithm must detect FOD present in the video feed.</i>	LS66: UCA-47 might occur if the FOD was not within the objects database used to train the algorithm.	SR79: AI system designer must use a good-quality and robust object database to train the algorithm (T).
Localisation function	CA-12: Localise detected FOD	UCA-57: Geolocation is not provided for detected FOD during runway inspection when a location estimate is required to enable timely and safe dispatch of the runway inspection team (H-3) <i>CC-57: Localisation software program must provide the geolocation of the detected FOD.</i>	LS76.1: UCA-57 might occur if the detection algorithm detects FOD, but the localisation function is not triggered due to failure to read the detection file output.	SR90: AI system designer must develop a robust localisation software system which is able to read and interpret data from the detection algorithm (T).
Drone pilot	CA-13: Fill and submit the flight plan	UCA-62: Flight plan includes unsafe altitude or route, or overlaps with aircraft operations when the runway movement area is active and the UAV route and altitude must remain within approved corridors, geofences and separation constraint (H-1, H-2, H-3, H-6) <i>CC-62: Drone pilot must not include unsafe altitude or route in the flight plan.</i>	LS81: UCA-62 might occur if the drone pilot made a data entry error due to fatigue from increased working hours.	SR97: Airport management must implement a fatigue risk management system
Air Traffic Control (ATC)	CA-28: Request runway inspection	UCA-111: Request comes too late when an arriving or departing aircraft is already committed to runway use (H-3, H-6) <i>CC-111: ATC must request runway inspection at TBD time before arriving or departing aircraft.</i>	LS142: UCA-111 might occur if ATC is managing much traffic at the same time, and due to workload, omits to request a runway inspection before the departing aircraft starts taxi.	SR139: ANSP must schedule at least two ATCs per shift and specify in SOPs that the actions of the ATC executive must be monitored by the ATC assistant (O, H).
Runway inspection team	CA-32: Remove FOD	UCA-125: Inspection team takes too much time to localise the reported FOD when timely localisation and removal are required to restore runway availability or prevent aircraft exposure to debris (H-2) <i>CC-125: Inspection team must localise the FOD.</i>	LS160: UCA-125 might occur if inaccurate FOD coordinates were sent to the inspection team.	SR154: AI system designer must use a robust system (e.g. accurate drone geolocation) which can provide FOD coordinates with an accuracy (T).
FPV camera system	CA-38: Capture and transfer live video	UCA-145: FPV camera stops recording and streaming during manual piloting when continuous visual feedback is required for safe control and separation (H-1, H-2, H-7) <i>CC-145: FPV camera must not stop recording in the middle of manual piloting of the UAV.</i>	LS181: UCA-145 might occur if the UAV loses power in the middle of the flight.	SR22: Drone pilot must check the UAV battery status before the flight and ensure it is sufficient to complete the runway inspection (H).

Note: H, T and O denote the recommendation category: Human-related (H), Technology-related (T), and Organisation-related (O).

scenarios and hazards remain represented in the evaluation matrix via the LS-based criteria and the UCA-to-hazard propagation criteria. This worst-case inheritance is a conservative sequencing choice that avoids diluting recommendations linked to at least one high-consequence pathway, while still capturing hazard coverage through the UCA-to-hazard propagation criteria.

All criteria were assessed using a 5-point qualitative scale (see Table 2). For “lower is better” criteria (e.g., cost and implementation duration), the scale was defined in reverse so that higher scores consistently represent more preferred performance. A 5-point ordinal scale was selected to support consistent expert judgement and to provide a midpoint with clear differentiation between levels (Likert, 1932), as used in other MCDM applications (da Cunha et al., 2022; Kaya, 2020). The numerical coding is intended for rank-order comparisons and PROMETHEE preference modelling and does not imply equal-interval distances between adjacent levels. If quantitative data is available, the same approach can be used with the quantitative values.

Preference functions were specified as linear functions with indifference and preference thresholds, implemented in Visual PROMETHEE using the Preference Function Assistant. Under this function, score differences are treated as negligible up to the indifference threshold q , preference then increases linearly for differences between q and the preference threshold p , and full preference is reached for differences $\geq p$. Using the 5-point scale, thresholds were set consistently across criteria at $q = 1$ and $p = 2$, reflecting the assumption that a one-level difference may be within judgement noise while differences of two or more levels indicate clear preference. This choice ensures consistent treatment of one-level versus two-level differences across all criteria. Criterion-specific thresholds could be elicited from decision-makers in operational use, but common thresholds were used in this study for consistency, given the shared scoring scale.

Following the definition of criteria, scales and preference functions, an evaluation matrix was compiled in Excel that contains all safety recommendations and their criterion scores. The matrix was imported into Visual PROMETHEE software (Mareschal, 2014), and the software was used to compute ϕ^+ , ϕ^- , and ϕ values, generate the PROMETHEE I network visualisation and produce the PROMETHEE II complete ranking.

2.2.2. Managing subjectivity in criteria scoring

The PROMETHEE evaluation matrix relies on analyst-assigned qualitative ratings derived from traceable STPA evidence. Subjectivity is therefore inherent and is managed through four measures. First, all criteria were defined a priori using an explicit assessment scale (Table 2) and a consistent preference model (a linear function with common thresholds $q = 1$ and $p = 2$) to reduce scale drift across raters and criteria. Second, scoring followed a traceability-first rule: each rating was justified by the most safety-relevant linked STPA artefact(s), with the “worst-case inheritance” rule applied for recommendations linked to multiple LSs to avoid dilution of high-consequence pathways without excluding lower-severity loss scenarios, which remain represented through the LS and hazard-propagation criteria. Third, the evaluation matrix was populated using an iterative consensus procedure. One author performed the initial scoring of all recommendations using the predefined criterion definitions and linked STPA evidence. A second author then reviewed the full matrix in two rounds of revision. Disagreements were resolved through discussion and by referring back to the STPA traceability links, resulting in a final consensus score for each recommendation assessment entry. Fourth, the prioritisation results are interpreted as decision support rather than a deterministic “correct” ordering. Rankings are preference-dependent and are, therefore, reported together with robustness checks to support transparent and revisable implementation planning. While these measures improve consistency and traceability, they do not eliminate the role of expert judgement.

2.2.3. Robustness and sensitivity analyses

To examine the robustness of the prioritisation to plausible changes in criterion weighting (w_i), two weighting scenarios were analysed (see Table 2). Scenario 1 assigns equal weights to all criteria ($w_i = 1/15$). Meanwhile, Scenario 2 assigns greater importance to impact and feasibility criteria while reducing the weights assigned to the remaining risk-propagation criteria, reflecting a preference structure that emphasises consequences and implementability. In other words, Scenario 1 represents a neutral ‘no-preference’ baseline, whereas Scenario 2 approximates an operational stakeholder preference that emphasises consequence and implementability. In both scenarios, weights were normalised such that $\sum_i w_i = 1$, and PROMETHEE II net flows and complete rankings were computed.

Since this study includes a large alternative set (161 recommendations), robustness was assessed using complementary rank-stability indicators that capture both overall agreement and decision-relevant stability. Rank-order agreement between Scenario 1 and Scenario 2 was quantified using Spearman’s ρ and Kendall’s τ_b applied to the two rank vectors. The magnitude of reordering was assessed via a rank-shift analysis, calculating the change in ranking position for each recommendation between scenarios and summarising absolute shifts (e.g., mean, median and maximum). To focus on stability of the highest-priority recommendations, a top- k overlap curve was computed by comparing the Scenario 1 and Scenario 2 top- k sets across a range of k values (e.g., $k = 10, 15, 20, 30 \dots$ up to the full set of 161). This provides an interpretable measure of stability of priority subsets and avoids over-interpreting small ranking differences in the mid-ranked region, where net flows are often closely spaced.

In addition to scenario-based robustness comparisons, a weight sensitivity analysis was performed using the Weight Stability Intervals (WSI) tool in Visual PROMETHEE. For each criterion, its weight was varied while the remaining weights were proportionally re-normalised, and PROMETHEE II net flows and rankings were recalculated. The tool reports the WSI [Lower%, Upper%] for each criterion, defined as the range of that criterion’s weight for which a specified stability condition is preserved under the chosen baseline weighting scheme. In this study, WSIs were computed at stability level 161, corresponding to the stability of the complete ranking across all 161 recommendations. Under this strict definition, narrow WSIs indicate that even small weight perturbations can induce at least one rank reversal somewhere in the ranking, which is an outcome that is expected for large alternative sets with many near-ties. Importantly, such full-rank instability does not necessarily undermine decision-relevant stability: the highest-priority set can remain stable even if some lower-ranked items swap positions. Wider WSIs indicate greater tolerance to complete ordering of weight variation. WSIs were generated for all 15 criteria and were interpreted alongside the scenario-based robustness indicators to distinguish full-order sensitivity from decision-relevant stability of the highest-priority recommendations.

These robustness and sensitivity checks were used to evaluate the stability of the prioritisation under plausible changes in preferences and weights (Brans & Mareschal, 2005; Dweiri et al., 2016; Kabassi & Martinis, 2021; X. Liu & Liu, 2024; Senthil et al., 2018).

3. Results

3.1. STPA findings

Operational stakeholders for AI-enabled UAV for runway FOD detection include the Civil Aviation Authority (CAA), Air Navigation Service Provider (ANSP), airport operator, UAS operator, airlines, drone manufacturer and AI system designer. In the detailed control structure, these stakeholders are represented through decision-making and supervisory controllers (e.g., the drone pilot, flight plan management, automation and flight controller functions, the detection controller and the runway inspection team), together with the communication

channels and sensing services that mediate safety-critical control and feedback (e.g., the C2 link, payload/FPV camera streams and weather data provision).

Seven losses identified, which are: L1: Injury or loss of human life, L2: Loss of or damage to vehicles, L3: Damage to infrastructure, L4: Loss of public and customer confidence and reputational harm, L5: Loss of mission, L6: Financial loss and L7: Delay or disruption to airport operations. [Table 3](#) presents eight system-level hazards and constraints that can lead to losses.

The high-level control structure contains nine controllers and two controlled processes (see Fig. 1). In Fig. 1, red arrows denote control actions and blue arrows denote feedback. This study focuses on the ‘runway FOD detection service’ controller, shown in the green box in Fig. 1, because it is the integration point where AI-enabled UAV autonomy and perception outputs are translated into operational decisions, and where coordination with ATC and airspace actors constrains runway access and separation. This boundary captures the dominant system-level control and feedback loops that govern safe runway inspection in non-segregated airport operations. At this level of abstraction, the UAV physical platform is treated as part of the overall Runway

FOD detection service. In the detailed control structure, it is represented explicitly as a separate controlled-process element.

Fig. 2 presents the detailed control structure, further elaborated through the key controllers, supporting functions, and their associated responsibilities (Appendix 1). Key controllers include the drone pilot, flight plan management, automation and flight controller functions, the detection controller, and the runway inspection team. The UAV vehicle is represented as part of the controlled process, receiving actuator commands from the flight controller and returning sensor feedback to the flight controller within the closed-loop control structure. Supporting channels and sensing services include the C2 link, weather data provision and the payload/FPV camera streams. The C2 link is modelled explicitly because it carries safety-critical control and feedback flows (e. g., commands, telemetry and status, and inspection video and data) and is treated as part of the control and feedback path rather than a decision-making controller. Similarly, payload/FPV camera streams and weather data provision are represented explicitly where they mediate safety-critical information and feedback, but they are not treated as decision-making controllers in the STPA model. For readability, internal control algorithms and procedures and representative process model

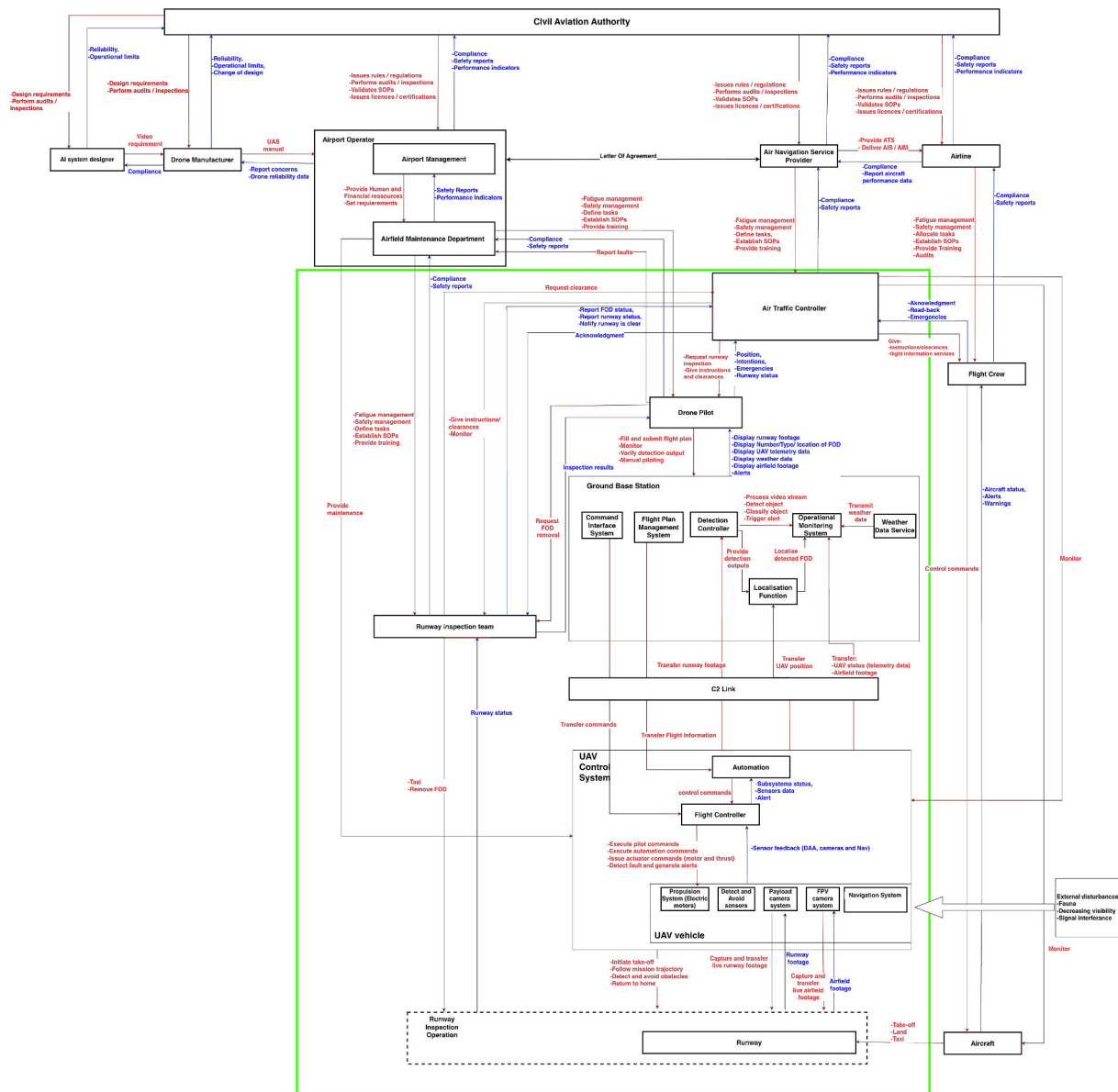


Fig. 2. The detailed control structure of the runway FOD detection process.

variables are not shown in Fig. 2, but were considered in UCA interpretation and loss-scenario development; examples are provided in Table 4, and the full linked findings in Appendix 2.

A total of 149 UCAs were identified, considering the green box in Fig. 2. Based on these UCAs, 188 LSs were identified, explaining the causality of the UCAs and 161 safety recommendations were provided. Appendix 2 provides the full traceable STPA outputs linking each control action to its UCAs, associated constraints, LSs and derived recommendations. Table 4 provides a partial list of UCAs, one UCA for each controller, along with the associated controller constraints, loss scenarios and safety recommendations. Safety recommendations were categorised as being human (H), technology (T) or organisation-related (O).

3.2. PROMETHEE findings

A total of 161 recommendations were derived from the STPA analysis, and these recommendations were ranked to support implementation sequencing and tiering using 15 evaluation criteria (Appendix 3 provides the evaluation matrix). Figs. 3 and 4 illustrate the PROMETHEE network model for Scenario 1 (equal weights) and Scenario 2 (impact and feasibility emphasised weights, see Table 2). Across both scenarios, action 54 (SR54. Airport management must ensure the GBS is placed within the UAV's maximum communication range (15 km for the DJI Matrice 300)) achieved the highest overall priority, indicating that C2-range assurance is consistently the most influential near-term safety measure when balancing risk reduction and deployability across the full criteria set.

Table 5 reports the PROMETHEE I and II outputs for the top 15 recommendations under both scenarios, with the complete list provided in Appendix 4. The PROMETHEE outputs should be read as a structured decision-support prioritisation (preference- and judgement-dependent), not as a deterministic prescription that lower-ranked recommendations are unnecessary. The results show a high degree of agreement at the decision-relevant level. Twelve of the top 15 recommendations are common across Scenarios 1 and 2, demonstrating that the highest-

priority set is largely preserved despite differences in criterion weights. The scenario comparison using PROMETHEE II net flows (see Fig. 5) further shows that most variability occurs in the mid-ranked region, where many recommendations have similar net-flow values and therefore small changes in weights can reorder alternatives without substantially changing their overall performance category. In Fig. 5, recommendations on the positive (green) side exhibit positive net flows and can be interpreted as more desirable, whereas those on the negative (red) side have negative net flows and are less desirable relative to the full set.

To quantify the robustness of the prioritisation to changes in criterion weights, complementary rank-stability indicators were computed between Scenario 1 and Scenario 2 ranking vectors. The overall rank-order agreement was strong (Spearman's $\rho = 0.944$ and Kendall's $\tau\text{-b} = 0.812$, with $p < 0.001$), indicating that the two scenarios produce highly consistent global ordering despite local reassignments. Rank-shift analysis shows that reordering is concentrated in the central part of the list. The median absolute rank shift was 7 positions (mean 10.6), while large shifts were relatively uncommon, such as $\approx 14\%$ moved more than 20 positions and $\approx 5\%$ more than 30 positions. Consistent with the top-15 comparison, the top-k overlap curve confirms that decision-relevant stability improves with k: overlap was 0.70 for the top-10 set and 0.80 for the top-15 set, rising to 0.88 for the top-50 and exceeding 0.90 beyond $k \approx 60$. Collectively, these results indicate that scenario-based changes affect mainly boundary cases around the mid-ranks, while the high-priority subset remains comparatively stable. In practical terms, this robustness supports confidence that the “implement-first” tier is not an artefact of a particular weighting choice, and that stakeholders can use the top-tier recommendations as a stable basis for near-term implementation planning.

Sensitivity analysis further examined the effect of varying individual criterion weights using Visual PROMETHEE WSI at stability level 161. Under Scenario 1 (equal weights, $w = 6.67\%$), WSIs were point intervals for 14 of the 15 criteria ($\text{WSI} = [6.67\%, 6.67\%]$), indicating that any deviation from the baseline weight led to at least one rank reversal somewhere in the full ordering. Only criterion C7 showed a small

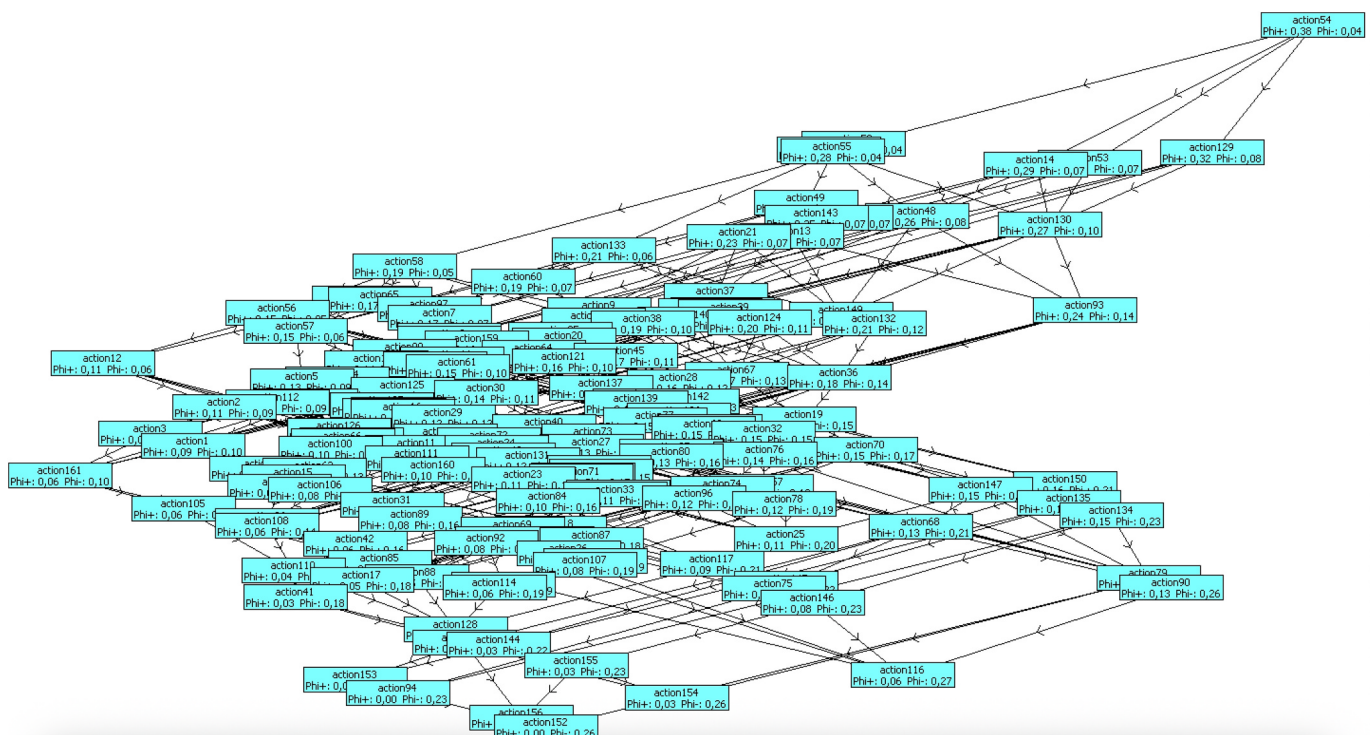


Fig. 3. PROMETHEE network model for scenario 1.

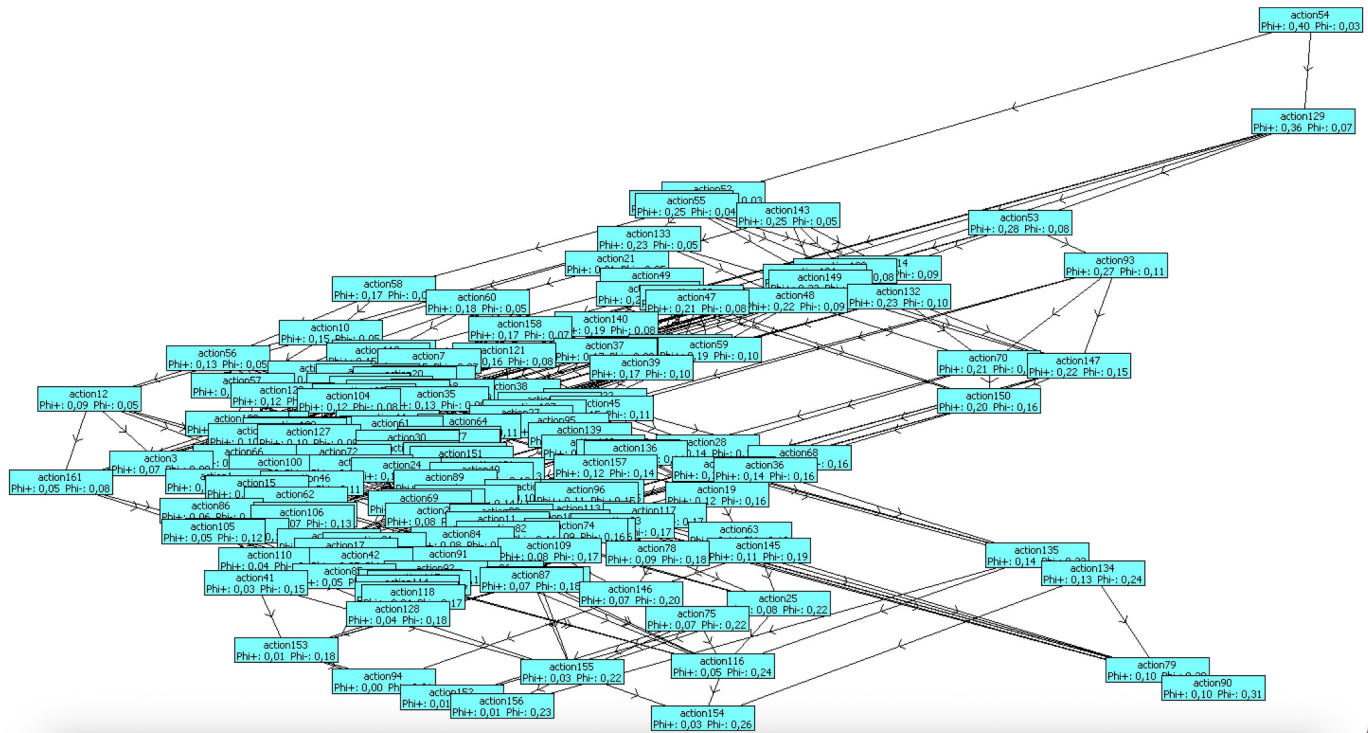


Fig. 4. PROMETHEE network model for scenario 2.

tolerance to downward adjustment ($WSI = [6.08\%, 6.67\%]$). Under Scenario 2, WSIs were likewise point intervals for all criteria (e.g., $[5\%, 5\%]$ or $[10\%, 10\%]$). These results indicate that the fine-grained complete ordering is highly sensitive to small weight perturbations, an expected outcome in large alternative sets with many near-ties, while the scenario-based robustness metrics above demonstrate that decision-relevant priorities (top-15 ranked) remain comparatively stable.

4. Discussion

4.1. Safety recommendations

In this study, STPA was applied to analyse hazards associated with AI-enabled UAV operations for runway FOD detection, identifying 8 hazards, 149 unsafe control actions, 188 loss scenarios and 161 safety recommendations. Table 5 presents the top 15 prioritised safety recommendations to illustrate a stable “implement-first” set for staged implementation planning. Lower-ranked recommendations remain relevant and are interpreted as subsequent implementation tiers rather than unnecessary measures. Since PROMETHEE was applied directly to STPA-derived recommendations using criteria largely operationalised from STPA artefacts and complemented by feasibility criteria, the resulting ranking should be interpreted as the measures that best balance safety benefit with system deployability for near-term integration.

The prioritised set consistently highlights recommendations that strengthen operational resilience and reliability, with particular emphasis on C2 link resilience and robust handling of degraded states (e.g., SR49, SR52, SR53 and SR54). Complementary measures enhance the pilot’s detectability of abnormal conditions via aural and, where appropriate, multimodal warnings and support timely escalation to emergency action when link-loss thresholds are exceeded (e.g., SR48, SR51 and SR133). This aligns with evidence that UAV operations depend on wireless communication and remain vulnerable to noise, fading and attenuation, with breakdowns potentially leading to mission loss and crashes (Jeelani & Gheisari, 2021; Xing & Johnson, 2023). The autonomous return-to-home procedure (SR47) further ensures that the

UAV can respond to critical C2 link failures without requiring human intervention. These recommendations align with CANSO (2011) guidelines, which state that ANSPs should review contingency and emergency procedures based on airport operations contingency plans to mitigate collision risk following C2 loss (e.g., SR37 and SR38). Indeed, regulatory reviews highlight that telecommunications, including C2 links, are a core domain in emerging drone integration frameworks (ICAO, 2015; Scott & Kurtpinar, 2026).

Beyond communications, component reliability remains relevant because failures in critical systems such as sensors, power sources and flight controllers can terminate missions when redundancy is limited (e.g., LS14, LS16 and LS53). Indeed, Ghasri and Maghrebi (2021) reported that approximately 60% of UAV accidents are due to hardware issues. Broader UAV studies emphasise the combined contribution of mechanical and human factors to UAV safety risks, including issues linked to bandwidth and latency as well as sensory isolation (Luongo Salvatore and Di Gregorio 2019; Mu, Yue, and Chen 2022). To address this, preventive maintenance and inspection practices (e.g., SR7, SR10, SR67 and SR68) can be strengthened by structured defect or anomaly reporting to surface emerging failure modes (e.g., SR129), together with fault alerting and pilot alarm monitoring during abnormal conditions (e.g., SR48 and SR51). Relatedly, C2 loss and degradation appear across several UCAs (e.g., UCA-24, UCA-28, UCA-126 and UCA-130), reinforcing the value of combining technical resilience with disciplined monitoring and response.

Cybersecurity concerns (e.g., LS5 and LS43) are tightly coupled with C2 integrity and data integrity hazards, where unsecured channels could expose the UAV to hijacking or Global Positioning System (GPS) spoofing. The broader literature shows that UAV vulnerabilities are increasingly exploited through GPS spoofing and hijacking (d’Ambrosio et al., 2025; Kerns et al., 2014; Seo et al., 2015; Yaacoub et al., 2020), and that many commercial drones lack baseline protections such as encrypted communications and secure storage, with ICAO (2015) further advising biometric access control and anti-spoofing safeguards, which align with SR14. Notably, data payloads (e.g., images and videos) may serve as vectors for malware targeting ground stations, meaning

Table 5

PROMETHEE I and II findings for the top 15 recommendations (illustrative implement-first set).

Scenario 1					Scenario 2				
Rank	Action (SR) Description	Phi	Phi+	Phi-	Rank	Action (SR) Description	Phi	Phi+	Phi-
1	SR54. Airport management must ensure the GBS is placed within UAV's maximum communication range (15 km for the DJI Matrice 300) (T).	0.3421	0.3796	0.0375	1	SR54. Airport management must ensure the GBS is placed within UAV's maximum communication range (15 km for the DJI Matrice 300) (T).	0.3716	0.3997	0.0281
2	SR52. Airport management must allocate an adequate frequency range to the UAV to avoid interference (T).	0.2412	0.2825	0.0413	2	SR129. Airport management must provide reporting forms which are structured to support effective defect identification and should include a designated area to note any unforeseen issues or anomalies (O).	0.2841	0.3550	0.0709
3	SR50. Airport management must define in SOPs the minimum flight altitude below which the UAV cannot operate safely, assuming the UAV C2 link is operated via radio line of sight (O).	0.2362	0.2775	0.0413	3	SR52. Airport management must allocate an adequate frequency range to the UAV to avoid interference (T).	0.2203	0.2541	0.0338
4	SR55. A confirmation window must be displayed for any input in the system by the drone pilot (T).	0.2346	0.2771	0.0425	4	SR50. Airport management must define in SOPs the minimum flight altitude below which the UAV cannot operate safely, assuming the UAV C2 link is operated via radio line of sight (O).	0.2128	0.2466	0.0338
5	SR129. Airport management must provide reporting forms which are structured to support effective defect identification and should include a designated area to note any unforeseen issues or anomalies (O).	0.2342	0.3154	0.0812	5	SR55. A confirmation window must be displayed for any input in the system by the drone pilot (T).	0.2103	0.2459	0.0356
6	SR51. Drone pilot must monitor the aural and visual alarms from GBS (H).	0.2250	0.2983	0.0733	6	SR143. ANSP must include in SOPs the ATC responsibility to inform the drone pilot of any severe change in the weather during the runway inspection (O, H).	0.2022	0.2537	0.0516
6	SR53. UAV must be operated by taking into account the effect of external electromagnetic fields (T, O).	0.2250	0.2983	0.0733	7	SR51. Drone pilot must monitor the aural and visual alarms from GBS (H).	0.1959	0.2778	0.0819
8	SR14. Airport management must implement security protocols such as encrypted communication protocols with strong authentication between the UAV and GBS, secure data storage, and use only secured devices (O, T).	0.2233	0.2921	0.0688	7	SR53. UAV must be operated by taking into account the effect of external electromagnetic fields (T, O).	0.1959	0.2778	0.0819
9	SR49. Airport management must ensure the GBS is placed in a direct line of sight with the UAV's trajectory, without any obstacles between the UAV and GBS (O, T).	0.1917	0.2529	0.0612	9	SR133. SOPs must require immediate emergency action upon C2 link loss over predefined thresholds (O).	0.1816	0.2272	0.0456
10	SR48. An aural warning must be generated to the drone pilot when the C2 link is lost (T).	0.1800	0.2583	0.0783	10	SR21. Coordination protocol developed by airport management must include a positive clearance confirmation from ATC before UAV mission initiation (O).	0.1603	0.2128	0.0525
11	SR47. UAV must execute the return to home (RTH) procedure in the case of loss of C2 signal (T).	0.1787	0.2500	0.0712	11	SR93. Ensure time-stamped positional data is updated at frame level (T).	0.1591	0.2706	0.1116
11	SR143. ANSP must include in SOPs the ATC responsibility to inform the drone pilot of any severe change in the weather during the runway inspection (O, H).	0.1787	0.2475	0.0688	12	SR14. Airport management must implement security protocols such as encrypted communication protocols with strong authentication between the UAV and GBS, secure data storage, and use only secured devices (O, T).	0.1566	0.2428	0.0862
13	SR130. SOPs must include a predefined reaction timeline for each critical fault class (O).	0.1729	0.2683	0.0954	13	SR130. SOPs must include a predefined reaction timeline for each critical fault class (O).	0.1544	0.2366	0.0822
14	SR13. Airport management and ANSP must include in the SOPs an appropriate coordination protocol between all staff involved in the runway inspection (O).	0.1629	0.2371	0.0742	14	SR124. SOPs must instruct the pilot to resume to automated mode in degraded visibility conditions for manual control (O, H).	0.1491	0.2303	0.0813
15	SR21. Coordination protocol developed by airport management must include a positive clearance confirmation from ATC before UAV mission initiation (O).	0.1621	0.2313	0.0692	15	SR49. Airport management must ensure the GBS is placed in a direct line of sight with the UAV's trajectory, without any obstacles between the UAV and GBS (O, T).	0.1459	0.2097	0.0638

cybersecurity controls also protect the integrity of safety-relevant information used by operators and organisations. At the governance level, recent analyses of drone regulation emphasise that safety and security are increasingly treated as interdependent domains in emerging frameworks (Scott & Kurtpinar, 2026), supporting SR14-type controls as operational safety measures.

Recommendations also target the human-automation interface to preserve Situational Awareness (SA) and reduce errors under time pressure. This includes interface design controls such as confirmation windows for pilot inputs (SR55), alarm-driven monitoring and alerting during anomalies (e.g., SR48 and SR51), and data and traceability controls that support reliable monitoring and diagnosis (e.g., SR93).

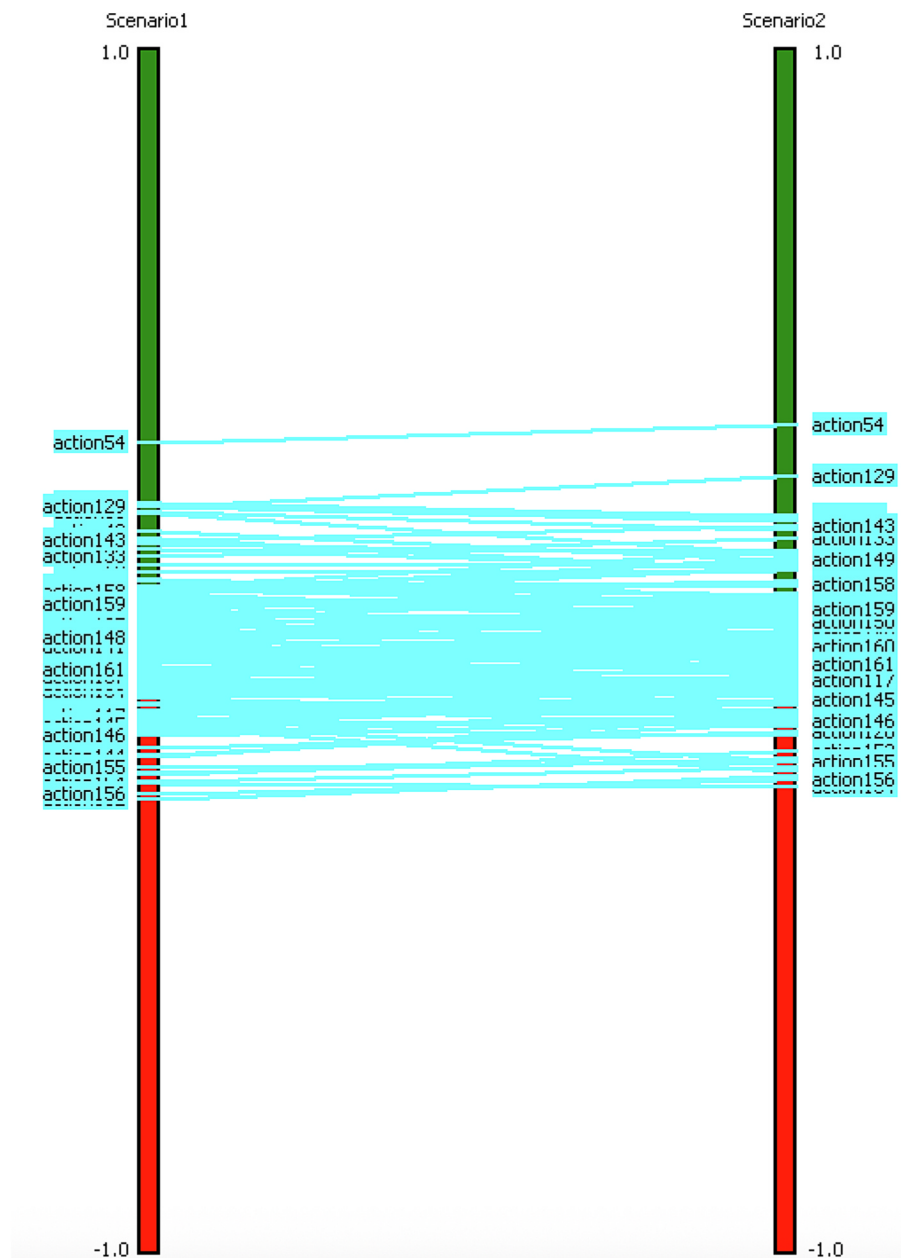


Fig. 5. Scenario comparisons based on PROMETHEE II.

Operationally, the prioritised set also emphasises structured fault response through Standard Operating Procedure (SOP)-level definitions of a predefined reaction timeline for critical fault classes (SR130) and procedures to support safe mode management during degraded visibility (SR124). These controls align with recent studies on human-factors evidence on drones, which suggests that the physical separation between pilot and vehicle removes important sensory cues and increases cognitive load, making SA disproportionately dependent on how information is selected, filtered and presented in the interface (Lochner et al., 2019; Luongo Salvatore and Di Gregorio, 2019; Ramallo-Luna et al., 2025). Consistent with STPA-driven recommendations (e.g., SR8, SR48, SR51 and SR101), haptic-enabled multimodal displays can enhance SA and strengthen anomaly alerting, such as C2 degradation, by pairing auditory alarms with salient visual warnings and, where feasible, haptic cues (Malik et al., 2020; Roberts et al., 2017; Wu et al., 2025; Xia et al., 2024).

Since runway inspection is conducted in a tightly coupled operational environment with limited tolerance for delay or miscoordination,

procedural controls and clear role boundaries appear prominently in the top-ranked recommendations. For example, SR21 requires positive ATC clearance confirmation before mission initiation, directly addressing implied-clearance and timing and coordination failure pathways (e.g., UCA-8 and LS13). In parallel, SR50 requires airport SOPs to define minimum safe flight altitudes under the radio line-of-sight assumptions used for the operation, helping to constrain low-altitude operations where interactions with manned aircraft, vehicles or static infrastructure remain credible outcomes (e.g., LS35). These control recommendations are consistent with broader aerodrome and remotely piloted aircraft system integration guidance that emphasises explicit clearances, runway occupancy awareness, disciplined monitoring and robust coordination procedures in mixed operations (CANSO, 2011; ICAO, 2015, 2020a; Scott & Kurtpinar, 2026).

Relatedly, the STPA-driven scenarios (e.g., LS4, LS13 and LS151) highlight the need for shared safety-related information across airspace stakeholders to prevent unsafe overlaps between UAV inspection activities and other movement-area operations. This supports embedding

explicit runway-inspection coordination and information-sharing requirements in SOPs (e.g., SR13 and SR21), consistent with UK CAA guidance on shared safety-related information standards (CAA, 2018). To support ATC oversight, it is further recommended to equip the UAV with an Automatic Dependent Surveillance-Broadcast (ADS-B) transponder to enhance detectability on the surveillance system (SR134), in line with ICAO (2015). Also, where necessary, implement aerodrome ground radar to detect vehicles in the movement area (SR135), alongside runway-occupancy and alerting support measures (e.g., SR137 and SR148).

Recommendations also include organisational measures that sustain safe performance across shifts and over time, including recurrent Air Traffic Control Officer (ATCO) training (SR138) and workload-aware scheduling (SR140). Fatigue-related scenarios can contribute to data-entry errors and degraded decision-making (e.g., LS81 and LS113), consistent with evidence that fatigue is a major contributor to human error (Griffith & Mahadevan, 2011). Accordingly, recommendations highlight scheduling practices (SR140) and the implementation of the Fatigue Risk Management System (FRMS) (SR97), aligning with ICAO FRMS guidance (ICAO, 2020a). Organisational learning is also reinforced through structured defect and anomaly reporting, which surfaces emerging failure modes as operations mature (e.g., SR129). This aligns with evidence that safe UAV operations depend on learning mechanisms and systematic transfer of lessons learned in dynamic contexts (Moorkamp et al., 2016).

Environmental conditions were repeatedly identified as precursors to unsafe events in the STPA scenarios (e.g., LS1, LS6, LS30, LS116 and LS138), particularly when decisions are made without up-to-date environmental information. SR143 explicitly assigns ATC responsibility to inform the drone pilot of severe weather changes during inspection. The STPA findings link this to broader weather controls outside the top-ranked recommendations, including defining operating minima (e.g., SR1 and SR2), continuous meteorological updates (SR3) and real-time pilot monitoring (SR4), consistent with ICAO Doc 10,019 guidance (ICAO, 2015). This is especially important for small UAVs, which are more susceptible to performance degradation in adverse weather (Altawy & Youssef, 2016), and aligns with empirical UAV safety findings that high winds, precipitation, poor lighting, dust and particulate exposure can degrade system response and sensor performance (Parker et al., 2022). Environmental effects also interact with sensing and AI performance. Visibility conditions, such as fog and rain, can reduce image clarity and detection reliability. The associated optical camera degradation scenarios (e.g., LS56) therefore motivate camera readiness and quality assurance measures (e.g., SR67, SR68 and SR69), together with operational controls that constrain when inspections can be conducted. More generally, adverse-weather scenarios motivate redundancy and resilience across UAV sensing and critical subsystems (e.g., multi-sensor detect-and-avoid in SR29 and redundant sensors for critical modules in SR63).

In parallel, detection effectiveness depends on the coverage of the AI training data. For example, LS66 highlights that objects absent from the training set may be missed. This motivates robust dataset governance (e.g., SR79) and explicit lifecycle feedback loops that capture missed FOD and use it to update the model and database (e.g., SR80). Where feasible, residual uncertainty can be managed via human-in-the-loop verification for low-confidence outputs (e.g., SR113 and SR115). This is consistent with recent autonomous-systems assurance studies emphasising monitoring and update loops under operational uncertainty (Buyse et al., 2025; Parker et al., 2022b). Beyond training-set coverage, AI-enabled inspection also faces lifecycle risks, including dataset shift across seasons and lighting conditions, performance degradation due to sensor ageing or operational changes, calibration drift affecting confidence thresholds, and change-control risks during retraining and redeployment (Shan et al., 2025). Although the present STPA focuses on operational control and the immediate causal pathways captured in the control structure, these lifecycle issues motivate additional assurance

measures, such as performance monitoring, revalidation triggers and retraining governance, that can be modelled as higher-level control loops in future STPA iterations.

The analysis showed that, out of 161 recommendations, most were technological (90), organisational (62) and human-related (35). Considering the highest-priority recommendations (Table 4), organisational measures appear most frequently (11), followed by technology-related (9) and human-related measures (3). This reflects both the high degree of automation in the operational concept and STPA's ability to surface organisational and system-design contributors alongside technical issues (Elizebeth et al., 2025; Kaya, 2021). This distribution is consistent with evidence that organisational design and coordination materially shape the emergence of safety in dynamic UAV operations (Moorkamp et al., 2016), and with broader UAV safety research indicating that risks extend beyond technical performance to include regulatory, organisational and operational factors (Mu et al., 2022).

4.2. Prioritising safety recommendations

While STPA enables the identification of a comprehensive list of safety recommendations, it does not prescribe an implementation order. Organisations decide what to implement first under resource and operational constraints, balancing anticipated risk reduction with factors such as cost, implementation duration, ease of deployment, and operational disruption. This is inherently a multi-criteria decision problem. In practice, recommendations may also be grouped by lifecycle function (e.g., design and engineering changes, verification activities, operational procedures, and organisational controls) and prioritised within each group. The same STPA-PROMETHEE logic applies, while criteria weights can be tailored to stakeholder priorities for each group. MCDM techniques are increasingly used in safety and risk management to structure expert judgement (Cheraghi et al., 2022; Di Graziano & Marchetta, 2021; Gul, 2018; Liu et al., 2019).

In this study, PROMETHEE II was used to rank 161 STPA-derived safety recommendations. PROMETHEE can support the prioritisation when decision-makers must balance heterogeneous criteria and when full compensation between criteria is undesirable (Brans et al., 1986; Brans & Mareschal, 2005; Brans & Vincke, 1985). The criteria captured both safety contribution, operationalised from STPA artefacts, and the practical implementability of the recommendations. Two weighting scenarios and a sensitivity analysis were used to test whether the highest-priority set remains stable under plausible stakeholder preferences. Compared with simplified proxies (e.g., single-criterion rankings, cost-benefit ratios or risk matrices), PROMETHEE makes preferences explicit via weights and preference functions and provides an auditable outranking-based ordering across multiple criteria.

The same “what should we implement first?” challenge is recognised in aviation regulatory practice. The UK CAA's STPA-based eVTOL study (CAP 3141) introduces an explicit triage extension that prioritises UCAs to decide which are taken forward into Step 4, and then prioritises the resulting mitigations, using expert scoring supported by Monte Carlo simulation and a five-level, colour-banded matrix format for communicating priorities (UK CAA, 2025). This approach is useful for rapid screening and stakeholder communication; however, discretised banding can reintroduce known limitations of risk-matrix style tools, including under-specified ordering within bands and boundary discontinuities that may violate weak consistency and lead to ranking anomalies (Cox, 2008; Duijm, 2015; Kaya et al., 2019). In this study, the PROMETHEE-based approach aims to achieve the same objective: transparent prioritisation of a large set of controls. It prioritises recommendations using explicit multi-criteria trade-offs, with safety-contribution criteria operationalised from STPA outputs (hazards, UCAs and LSs) and complemented by feasibility criteria. Robustness is then examined through weighting scenarios and sensitivity analysis.

The combined workflow strengthens practical decision support in two ways. First, STPA provides a causal and system-theoretic basis for

deriving recommendations traceable to hazards, UCAs and LSs, thereby ensuring that implementation sequencing remains traceable to the STPA safety argument rather than treating recommendations as isolated risk items. Second, PROMETHEE provides a transparent mechanism for ordering a large recommendation set when resources are constrained, while making trade-offs explicit through weights and preference functions and enabling robustness checks under alternative stakeholder priorities. A key lesson is that PROMETHEE should be used to identify a stable “implement first” set and to form priority tiers, rather than to overinterpret fine-grained differences in the mid-ranked region, where recommendations are often near-ties and preference-sensitive.

In practice, organisations can use the ranking to implement safety controls in phases: (i) select a stable top-k “implement first” package (e. g., top-10/15/20, chosen based on available resources and the demonstrated top-k stability), (ii) cluster the remaining recommendations by control strategy or owner, and (iii) translate each cluster into a staged implementation plan with accountable roles, verification activities and operational readiness criteria. Importantly, lower-ranked recommendations are not “optional”; rather, the ranking provides a transparent sequencing mechanism for incremental implementation and assurance planning when resources constrain immediate adoption of all measures.

PROMETHEE also makes an important practical point explicit, as rankings are preference dependent. Different decision makers may reasonably prioritise different criteria. For instance, while one might consider the cost criterion the top decision-making indicator, another might prefer the implementation duration. So, there is rarely a single “correct” ordering independent of organisational strategy and constraints. In this study, although the mid-ranked region was more sensitive to weight changes (see Fig. 5), the top of the ranking was relatively robust, with 12 of the 15 top recommendations remaining unchanged across the two scenarios. This concentration of reordering in the middle is expected because many recommendations have similar net-flow values, so modest weight changes can swap their relative positions without materially affecting their overall comparative performance.

Finally, although MCDM provides a more transparent and defensible basis for prioritisation than common heuristics, it is still important not to over-interpret fine-grained differences, especially in the mid-range, where rankings can be less stable (Kaya, 2020). As a consequence, it is recommended to use MCDM to classify recommendations into priority tiers and then prioritise within each tier to guide staged implementation while preserving traceability to STPA results.

4.3. Study limitations

This study has limitations that need to be mentioned. First, the STPA control structure was developed for a specific AI-enabled UAV runway inspection architecture concept. Alternative configurations, such as relocating AI processing from a ground-based system to an onboard UAV, would change control and feedback loops and could lead to different hazards, UCAs, LSs and safety recommendations. This limits direct generalisation across all potential AI-enabled UAV runway inspection designs. Accordingly, some recommendations are concept specific. Nevertheless, several system-level insights are transferable across UAV types and airport contexts, including ATC and UAS operator coordination and clearance management, robust C2 assurance and contingency handling, operational minima and weather-triggered constraints, human-automation coordination for detection uncertainty and escalation and organisational learning and reporting mechanisms. The STPA-PROMETHEE workflow is also platform-agnostic and can be re-instantiated by updating the control structure, control actions and traceability links while retaining the same criteria-construction logic. For Beyond Visual Line of Sight (BVLOS) drones or higher-autonomy concepts, this re-instantiation would extend the control structure to include additional controllers and feedback loops, which would typically increase the prominence of hazards related to separation assurance, contingency management and assurance of autonomous decision

logic. The prioritisation step remains unchanged, although criteria weights may reasonably shift to reflect stakeholder emphasis on separation and contingency robustness under reduced human oversight.

Second, STPA is inherently knowledge and judgment intensive. The completeness and reliability of the identified UCAs and LSs depend on the quality of available information, the analyst’s understanding of the operation and the breadth and experience of the consulted SMEs (Kaya et al., 2025; Neto et al., 2024). As a result, it is not possible to claim that all plausible UCAs and LSs were exhaustively captured. To mitigate this, the study applied a structured STPA validation framework and iterated on the results based on expert feedback (Sadeghi & Goerlandt, 2023b, 2023a), but residual dependence on the analyst and SMEs remains.

Third, the prioritisation results should be interpreted as a decision-support demonstration rather than a definitive implementation plan. The PROMETHEE evaluation matrix and the weighting scenarios were populated by the authors based on their domain understanding, rather than being elicited directly from the organisational decision-makers who would ultimately own the trade-offs. Accordingly, the reported ranking should be interpreted as an illustration of how the proposed STPA-grounded prioritisation can be operationalised. Nevertheless, the study discusses the top-ranked set because it is comparatively stable across the two weighting scenarios, thereby supporting interpretations of priority tiers under plausible preference structures. The workflow is designed to be rerun with stakeholder-provided scores and weights and to compare how priorities change across different decision-maker views. Different stakeholders may reasonably assign different scores and weights, which can change ranking, particularly in the mid-range. In addition, other MCDM methods could be applied using the same STPA-derived criteria and traceability mappings to test method-dependence of the prioritised set. In applied MCDM studies, this is commonly addressed through group decision-making by aggregating multiple expert inputs and fuzzy and uncertainty-aware variants to represent linguistic assessments and reduce subjective bias (Chin et al., 2008; Gul, 2018; Gul et al., 2019; Purba et al., 2014). Future work should therefore elicit ratings and weights from a broader stakeholder panel and, where feasible, incorporate uncertainty treatment, including fuzzy evaluations or probabilistic aggregation.

Furthermore, despite the identified loss scenarios and safety recommendations, the absence of testing in a real airport environment limits the study’s ability to utilise empirical data, such as incident reports or operational performance records. Despite these limitations, the study provides a structured, traceable set of socio-technical risks and mitigations for AI-enabled UAV runway FOD detection and offers a transparent prioritisation approach to support staged implementation.

5. Conclusion

This study applied STPA to analyse an AI-enabled UAV concept for runway FOD detection in non-segregated airport operations. The analysis shows that losses may emerge not only from component failures but also from software and automation behaviour, inadequate process models, procedural design, and coordination and communication breakdowns under operational constraints. The STPA perspective makes these socio-technical dependencies explicit by linking safety constraints and responsibilities across the airport operator, ANSP, ATC, UAV operator, system and AI developers, providing a structured basis for controlling hazardous system behaviour across the end-to-end operation.

The analysis identified 8 hazards, 149 unsafe control actions, 188 loss scenarios and 161 safety recommendations, spanning the technical, procedural and human-organisational domains. To support resource-bounded implementation planning, these recommendations were prioritised using PROMETHEE II across 15 criteria and two weighting scenarios. The prioritisation consistently highlighted measures that strengthen command-and-control and control and feedback communications under degraded modes, improve operator monitoring and alerting with timely escalation, and reinforce structured, closed-loop

coordination and clearance protocols with ATC. Scenario comparison indicates that the highest-priority set is relatively stable under plausible preference changes, whereas the mid-ranked ordering is more sensitive and should therefore be interpreted as priority tiers rather than a strict sequence.

The contribution of this work is twofold. First, it provides a traceable STPA-based hazard analysis for an AI-enabled UAV runway inspection concept, resulting in a structured set of hazards, UCAs, LSs, and safety constraints relevant to non-segregated airport operations. Second, it demonstrates a practical STPA-to-MCDM workflow that translates a large STPA-derived recommendation set into an auditable priority ranking to support staged implementation while preserving traceability back to UCAs and LSs. The prioritisation should be treated as decision support because weights and ratings are preference-dependent and may vary across stakeholders. Future work should elicit weights and ratings from a broader stakeholder panel, incorporate uncertainty treatment where appropriate, and validate prioritised measures against operational evidence as the concept and deployment context mature.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author(s) used ChatGPT 5.2 in order to improve the language, clarity and readability of the manuscript, including editing and rewriting text for concision and consistency. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

CRedit authorship contribution statement

Nafissa Abdouramane Attou Bounou: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gulsum Kubra Kaya:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fanny Camelia:** Writing – review & editing, Validation, Methodology.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssci.2026.107280>.

Data availability

Data is provided as an Appendix

References

- ACI, 2020. Drones in the airport environment: Concept of operations and industry guide. Altawy, R., Youssef, A.M., 2016. Security, privacy, and safety aspects of civilian drones: a survey. *ACM Trans. Cyber-Phys. Syst.* 1 (2), 1–25. <https://doi.org/10.1145/3001836>.
- Aslantas, M., Kutlu Gündoğdu, F., Moslem, S., 2025. Evaluating the potential risks posed by autonomous vehicles by using a decomposed fuzzy multi-criteria decision-making model. *Transp. Eng.* 21. <https://doi.org/10.1016/j.treng.2025.100372>.
- ATSB, 2025. Control issues and ditching involving RPA swarm of 500 Damoda Newton 2.2 RPA.
- Baybutt, P., 2015. Designing risk matrices to avoid risk ranking reversal error. *Process Saf. Prog.* 35 (1), 41–46. <https://doi.org/10.1002/prs.11768>.
- Brans, J., Mareschal, B., 2005. PROMETHEE methods. In *International Series in Operations Research and Management Science* (Vol. 78, pp. 187–219). Springer. Doi: 10.1007/978-1-4939-3094-4_6.
- Brans, J., Vincke, P., 1985. A preference ranking organisation method. *Manag. Sci.* 31 (6), 647–656. <https://doi.org/10.1287/mnsc.31.6.647>.
- Brans, J., Vincke, P., Mareschal, B., 1986. How to select and how to rank projects: the PROMETHEE method. *Eur. J. Oper. Res.* 24, 228–238. [https://doi.org/10.1016/0377-2217\(86\)90044-5](https://doi.org/10.1016/0377-2217(86)90044-5).
- Buyse, L., Habli, I., Vanoost, D., Pissort, D., 2025. Safe autonomous systems in a changing world: operationalising dynamic safety cases. *Saf. Sci.* 191 (106965), 1–16. <https://doi.org/10.1016/j.ssci.2025.106965>.
- CAA, 2018. Airside safety management CAP 642.
- CANSO, 2011. ANSP considerations for unmanned aircraft systems (UAS) operations. Charalampidou, S., Lygouras, E., Dokas, I., Gasteratos, A., Zacharopoulou, A., 2020. A sociotechnical approach to UAV safety for search and rescue missions. *International Conference on Unmanned Aircraft Systems*. September 1.
- Charalampidou, S., Zeleskidis, A., Dokas, I.M., 2024. Hazard analysis in the era of AI: assessing the usefulness of ChatGPT4 in STPA hazard analysis. *Saf. Sci.* 178, 106608. <https://doi.org/10.1016/j.ssci.2024.106608>.
- Chatzimichailidou, M.M., Karanikas, N., Plioutsias, A., 2017. Application of STPA on small drone operations: A benchmarking approach. 4th European STAMP Workshop 2016, 179, 13–22. Doi: 10.1016/j.proeng.2017.03.091.
- Chen, J., Zhang, S., Lu, Y., Tang, P., 2015. STPA-based hazard analysis of a complex UAV system in take-off. In: *The 3rd International Conference on Transportation Information and Safety*, p. 774.
- Cheraghi, M., Eslami Baladeh, A., Khakzad, N., 2022. Optimal selection of safety recommendations: a hybrid fuzzy multi-criteria decision-making approach to HAZOP. *J. Loss Prev. Process Ind.* 74. <https://doi.org/10.1016/j.jlp.2021.104654>.
- Chikha, P., Skorupski, J., 2022. The risk of an airport traffic accident in the context of the ground handling personnel performance. *J. Air Transp. Manag.* 105. <https://doi.org/10.1016/j.jairtraman.2022.102295>.
- Chin, K.S., Chan, A., Yang, J.B., 2008. Development of a fuzzy FMEA based product design system. *Int. J. Adv. Manuf. Technol.* 36 (7–8), 633–649. <https://doi.org/10.1007/s00170-006-0898-3>.
- Coquelet, B., Dejaegere, G., De Smet, Y., 2025. Evaluating the PROMETHEE II ranking quality. *Algorithms* 18 (597). <https://doi.org/10.3390/a18100597>.
- Cox, L.A., 2008. What's wrong with risk matrices? *Risk Anal.* 28 (2), 497–512. <https://doi.org/10.1111/j.1539-6924.2008.01030.x>.
- da Cunha, R.A., Rangel, L.A.D., Rudolf, C.A., dos Santos, L., 2022. A decision support approach employing the PROMETHEE method and risk factors for critical supply assessment in large-scale projects. *Oper. Res. Perspect.* 9. <https://doi.org/10.1016/j.orp.2022.100238>.
- d'Ambrosio, N., Perrone, G., Romano, S.P., Urraro, A., 2025. A cyber-resilient open architecture for drone control. *Comput. Security* 150. <https://doi.org/10.1016/j.cose.2024.104205>.
- Dejaegere, G., De Smet, Y., 2023. Promethee γ : a new Promethee based method for partial ranking based on valued coalitions of monocriterion net flow scores. *J. Multi-Criteria Decis. Anal.* 30 (3–4), 147–160. <https://doi.org/10.1002/mcda.1805>.
- Dghaym, D., Hoang, T.S., Turnock, S.R., Butler, M., Downes, J., Pritchard, B., 2021. An STPA-based formal composition framework for trustworthy autonomous maritime systems. *Saf. Sci.* 136. <https://doi.org/10.1016/j.ssci.2020.105139>.
- Di Graziano, A., Marchetta, V., 2021. A risk-based decision support system in local railways management. *J. Rail Transp. Plann. Manage.* 20 (September), 100284. <https://doi.org/10.1016/j.jrtpm.2021.100284>.
- Duijm, N.J., 2015. Recommendations on the use and design of risk matrices. *Saf. Sci.* 76, 21–31. <https://doi.org/10.1016/j.ssci.2015.02.014>.
- Dweiri, F., Kumar, S., Khan, S.A., Jain, V., 2016. Designing an integrated AHP based decision support system for supplier selection in automotive industry. *Expert Syst. Appl.* 62, 273–283. <https://doi.org/10.1016/j.eswa.2016.06.030>.
- EASA, 2024. Easy access rules for unmanned aircraft systems.
- Elizebeth, M.J., Khastgir, S., Jennings, P., 2025. Hazard analysis of an automated lane keeping system using systems-theoretic process analysis. *Accid. Anal. Prev.* 221 (March), 108171. <https://doi.org/10.1016/j.aap.2025.108171>.
- FAA, 2021. Small unmanned aircraft system (Small UAS)- AC: 107- 2A.
- FAA, 2024. Small unmanned aircraft-based foreign object debris detection on airport surfaces- DOT/FAA/TC-24/29.
- Fizza, G., Idrus, S.M., Iqbal, F., Hassan, W.H.W., Ahmad, R., Meerah, Z., Kanno, A., Kawanish, T., Shibagaki, N., Kashima, K., 2021. Review on foreign object debris detection technologies and advancement for airport safety and surveillance. *Turkish J. Comput. Math. Educ.* 12 (3), 1431–1436.
- Ghasri, M., Maghrebi, M., 2021. Factors affecting unmanned aerial vehicles' safety: a post-occurrence exploratory data analysis of drones' accidents and incidents in Australia. *Saf. Sci.* 139. <https://doi.org/10.1016/j.ssci.2021.105273>.
- Griffith, C.D., Mahadevan, S., 2011. Inclusion of fatigue effects in human reliability analysis. *Reliab. Eng. Syst. Saf.* 96, 1437–1447. <https://doi.org/10.1016/j.res.2011.06.005>.
- Gul, M., 2018. A review of occupational health and safety risk assessment approaches based on multi-criteria decision-making methods and their fuzzy versions. *Hum. Ecol. Risk Assess.* 24 (7), 1723–1760. <https://doi.org/10.1080/10807039.2018.1424531>.
- Gul, M., Ak, M.F., Guneri, A.F., 2019. Pythagorean fuzzy VIKOR-based approach for safety risk assessment in mine industry. *J. Saf. Res.* 69, 135–153. <https://doi.org/10.1016/j.jsr.2019.03.005>.
- Gul, M., Yucesan, M., Ak, M.F., 2022. Control measure prioritization in Fine-Kinney-based risk assessment: a Bayesian BWM-Fuzzy VIKOR combined approach in an oil station. *Environ. Sci. Pollut. Res.* 29 (39), 59385–59402. <https://doi.org/10.1007/s11356-022-19454-x>.
- Gutiérrez, G.B., 2023. Foreign object debris detection on runways by computer vision from drone. *Universitat Politècnica de Catalunya*.

- Hobbs, K.L., Cargal, C., Feron, E., Scott Burns, R., 2018. Early safety analysis of manned-unmanned team system. AIAA Inform. Syst.-AIAA Infotech at Aerospace 2018. <https://doi.org/10.2514/6.2018-1984>.
- ICAO, 2015. Manual on remotely piloted aircraft systems (PRAS)- Doc: 10019 AN/507. ICAO, 2018a. Annex 14 to the convention on international civil aviation- Aerodromes. ICAO, 2018b. Safety Report. ICAO Safety Report 2018 International Civil Aviation Organization <https://www.icao.int> \ ICAO SR 2018.30082018.
- ICAO, 2020a. Doc 9966 - Manual for the oversight of fatigue management approaches. ICAO, 2020b. Doc 9981: Procedures for air navigation services aerodromes.
- Jeelani, I., Gheisari, M., 2021. Safety challenges of UAV integration in construction: conceptual analysis and future research roadmap. Saf. Sci. 144 (105473), 1–16. <https://doi.org/10.1016/j.ssci.2021.105473>.
- Kabassi, K., Martinis, A., 2021. Sensitivity analysis of PROMETHEE II for the evaluation of environmental websites. Appl. Sci. 11 (19), 9215. <https://doi.org/10.3390/app11199215>.
- Kaya, G.K., 2020. The use of multi-criteria decision-making methods to support risk prioritisation. 5th International Conference on Industrial Engineering and Operations Management.
- Kaya, G.K., 2021. A system safety approach to assessing risks in the sepsis treatment process. Appl. Ergon. 94, 103408. <https://doi.org/10.1016/j.apergo.2021.103408>.
- Kaya, G.K., Humphreys, M., Camelia, F., Chatzimichailidou, M., 2025. Integrating causal analysis based on system theory with network modelling to enhance accident analysis. Ergonomics 1–28. <https://doi.org/10.1080/00140139.2025.2516060>.
- Kaya, G.K., Ward, J.R., Clarkson, P.J., 2019. A review of risk matrices used in acute hospitals in England. Risk Anal. 39 (5), 1060–1070. <https://doi.org/10.1111/risa.13221>.
- Kerns, A.J., Shepard, D.P., Bhatti, J.A., Humphreys, T.E., 2014. Unmanned aircraft capture and control via GPS spoofing. J. Field Rob. 31 (4), 617–636. <https://doi.org/10.1002/rob.21513>.
- Kim, H., Lundteigen, M.A., Hafver, A., Pedersen, F.B., 2021. Utilization of risk priority number to systems-theoretic process analysis: a practical solution to manage a large number of unsafe control actions and loss scenarios. Proc. Inst. Mech. Eng., Part O: J. Risk and Reliab. 235 (1), 92–107. <https://doi.org/10.1177/1748006X20939717>.
- Leveson, N., 2011. Engineering a safer world: systems thinking applied to safety. The MIT Press.
- Leveson, N., Thomas, J., 2018. STPA handbook. http://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf.
- Likert, R., 1932. A technique for the measurement of attitudes. Arch. Psychol. 22, 5–55.
- Liou, J.J.H., Chien, C.W., Chang, F.C., Chuang, Y.C., 2026. Integrating FMEA and HFACS to assess the risk to airport operations. J. Air Transp. Manag. 131. <https://doi.org/10.1016/j.jairtraman.2025.102902>.
- Liu, H.C., Chen, X.Q., Duan, C.Y., Wang, Y.M., 2019. Failure mode and effect analysis using multi-criteria decision making methods: a systematic literature review. Comput. Ind. Eng. 135, 881–897. <https://doi.org/10.1016/j.cie.2019.06.055>.
- Liu, X., Liu, Y., 2024. Sensitivity analysis of the parameters for preference functions and rank reversal analysis in the PROMETHEE II method. Omega 128. <https://doi.org/10.1016/j.omega.2024.103116>.
- Lochner, M., Duenser, A., Sarker, S., 2019. Trust and cognitive load in semi-automated UAV operation. ACM International Conference Proceeding Series 437–441. <https://doi.org/10.1145/3369457.3369509>.
- Luongo Salvatore and Di Gregorio, M. and V. G. and V. A., 2019. Human machine interface issues for drone fleet management. In W. and T. R. Ahram Tareq and Karwowski (Ed.), Human Systems Engineering and Design (pp. 791–796). Springer International Publishing.
- Malik, H.A., Rasool, S., Maqsood, A., Riaz, R., 2020. Effect of haptic feedback on pilot/operator performance during flight simulation. Appl. Sci. 10 (11), 3877. <https://doi.org/10.3390/app10113877>.
- Mareschal, B., 2013. Visual PROMETHEE manual. 1–192.
- Mareschal, B., 2014. Visual PROMETHEE. <https://bertrand.mareschal.web.ulb.be/promethee.html>.
- Marliere, T.A., de Azevedo Castro Cesar, C., Hirata, C.M., 2025. Extending the STPA to model the control structure with finite state machine. J. Saf. Sci. Resilience 6 (4). <https://doi.org/10.1016/j.jnlssr.2025.04.004>.
- Moorkamp, M., Wybo, J.L., Kramer, E.H., 2016. Pioneering with UAVs at the battlefield: the influence of organizational design on self-organization and the emergence of safety. Saf. Sci. 88, 251–260. <https://doi.org/10.1016/j.ssci.2015.09.029>.
- M'Sila, C., Ayad, R., Ait-Oufroukh, N., 2022. Automated Foreign Object Debris Detection System based on UAV. In: 2022 IEEE International Conference on Networking, Sensing and Control (ICNSC), pp. 1–7. <https://doi.org/10.1109/ICNSC55942.2022.10004050>.
- Mu, D., Yue, C., Chen, A., 2022. Are we working on the safety of UAVs? an LDA-based study of UAV safety technology trends. Saf. Sci. 152 (105767), 1–9. <https://doi.org/10.1016/j.ssci.2022.105767>.
- Nakamura, H., Kajikawa, Y., 2018. Regulation and innovation: how should small unmanned aerial vehicles be regulated? Technol. Forecast. Soc. Chang. 128, 262–274. <https://doi.org/10.1016/j.techfore.2017.06.015>.
- Neto, C.R.V., Rose, R.L., Thomas, J., 2024. Application of causal analysis based on systems theory (CAST) to regulatory decision-making: a case study of the Sikorsky S92A. Aersp. Sci. Technol. 155 (P3), 109708. <https://doi.org/10.1016/j.ast.2024.109708>.
- Pan, C., Gao, Y., Gu, A., 2023. Enhancing reliability analysis of AI systems using STPA: a case study on the DroneResponse UAV system. 43. Doi: 10.1117/12.3011516.
- Papadopoulos, E., & Gonzalez, F., 2021. UAV and AI Application for Runway Foreign Object Debris (FOD) Detection. 2021 IEEE Aerospace Conference (50100), 1–8. Doi: 10.1109/AERO50100.2021.9438489.
- Parker, A., Gonzalez, F., Trotter, P., 2022. Live detection of foreign object debris on runways detection using drones and AI. IEEE Aerospace Conference Proceedings, 2022-March. Doi: 10.1109/AERO53065.2022.9843697.
- Patriarca, R., Chatzimichailidou, M., Karanikas, N., Gravio, G.D., 2022. The past and present of system-theoretic accident model and processes (STAMP) and its associated techniques: a scoping review. Saf. Sci. 146, 105566. <https://doi.org/10.1016/j.ssci.2021.105566>.
- Pennington, E.S., Johnson, K.E., Hobbs, K.L., Colombi, J.M., 2026. Engineering safe human-autonomy teaming using STPA-coordination. Saf. Sci. 193. <https://doi.org/10.1016/j.ssci.2025.107011>.
- Pitchforth, J., Mengersen, K., 2013. A proposed validation framework for expert elicited Bayesian Networks. Expert Syst. Appl. 40 (1), 162–167. <https://doi.org/10.1016/j.eswa.2012.07.026>.
- Plioutsias, A., Karanikas, N., Chatzimichailidou, M.M., 2018. Hazard analysis and safety requirements for small drone operations: to what extent do popular drones embed safety? Risk Anal. 38 (3), 562–584. <https://doi.org/10.1111/risa.12867>.
- Purba, J.H., Lu, J., Zhang, G., Pedrycz, W., 2014. A fuzzy reliability assessment of basic events of fault trees through qualitative data processing. Fuzzy Set. Syst. 243, 50–69. <https://doi.org/10.1016/j.fss.2013.06.009>.
- Ramallo-Luna, M.A., Gonzalez-Torre, S., Rodríguez-Mora, Á., de la Torre, G.G., 2025. Neurocognitive factors of new drone pilots: identifying candidates with expert potential. Comput. Hum. Behav. Rep. 19 (100705), 1–11. <https://doi.org/10.1016/j.chbr.2025.100705>.
- Roberts, R., Barajas, M., Rodriguez-Leal, E., Gordillo, J.L., 2017. Haptic feedback and visual servoing of teleoperated unmanned aerial vehicle for obstacle awareness and avoidance. Int. J. Adv. Rob. Syst. 14 (4), 1–13. <https://doi.org/10.1177/1729881417716365>.
- Rothmund, S.V., Thieme, C.A., Utne, I.B., Johansen, T.A., 2025. Supervisory risk control with application to industrial drone inspection. Eng. Appl. Artif. Intel. 157. <https://doi.org/10.1016/j.engappai.2025.110970>.
- Sadeghi, R., Goerlandt, F., 2023a. A proposed validation framework for the system theoretic process analysis (STPA) technique. Saf. Sci. 162 (February), 106080. <https://doi.org/10.1016/j.ssci.2023.106080>.
- Sadeghi, R., Goerlandt, F., 2023b. Reasonableness of a proposed System Theoretic Process Analysis (STPA) validation framework: an interview study. J. Loss Prev. Process Ind. 83, 105064. <https://doi.org/10.1016/j.jlp.2023.105064>.
- Scott, B.I., Kurtpinar, Ö., 2026. Safety first, safety's choice: regulating unmanned aircraft and an illusion of choice. Saf. Sci. 193 (107029), 1. <https://doi.org/10.1016/j.ssci.2025.107029>.
- Senić, A., Ivanović, M., Dobrodolac, M., Stojadinović, Z., 2025. Prioritization of preventive measures: a multi-criteria approach to risk mitigation in road infrastructure projects. Mathematics 13 (2), 278. <https://doi.org/10.3390/math13020278>.
- Senthil, S., Murugananthan, K., Ramesh, A., 2018. Analysis and prioritisation of risks in a reverse logistics network using hybrid multi-criteria decision making methods. J. Clean. Prod. 179, 716–730. <https://doi.org/10.1016/j.jclepro.2017.12.095>.
- Seo, S.-H., Lee, B.-H., Im, S.-H., Jee, G.-I., 2015. Effect of spoofing on unmanned aerial vehicle using counterfeited GPS signal. J. Positioning, Navigation, and Timing 4 (2), 57–65. <https://doi.org/10.11003/jpnt.2015.4.2.057>.
- Shan, J., Miccinesi, L., Beni, A., Pagnini, L., Cioncolini, A., Pieraccini, M., 2025. A review of foreign object debris detection on airport runways: sensors and algorithms. Remote Sens. (Basel) 17 (225), 133. <https://doi.org/10.3390/rs17020225>.
- Siraj, M.T., Debnath, B., Payel, S.B., Bari, A.B.M.M., Islam, A.R.M.T., 2023. Analysis of the fire risks and mitigation approaches in the apparel manufacturing industry: Implications toward operational safety and sustainability. Heliyon 9 (9), e20312. <https://doi.org/10.1016/j.heliyon.2023.e20312>.
- Studic, M., 2015. Developing a framework for Total Apron Safety Management (Number October) [PhD]. Imperial College London.
- Studic, M., Majumdar, A., Schuster, W., Ochieng, W.Y., 2017. A systemic modelling of ground handling services using the functional resonance analysis method. Transp. Res. Part C Emerging Technol. 74, 245–260. <https://doi.org/10.1016/j.trc.2016.11.004>.
- UK CAA, 2023. Preventing technical failures- Drone educational and safety articles.
- UK CAA, 2025. STPA-based safety analysis of eVTOL operations.
- Wu, J., Sun, B., You, H., Du, J., 2025. Enhancing human-AI perceptual alignment through visual-haptic feedback system for autonomous drones. Int. J. Ind. Ergon. 109. <https://doi.org/10.1016/j.ergon.2025.103780>.
- Xia, P., Zhou, T., Ye, Y., Du, J., 2024. Human autonomy teaming for ROV shared control. J. Comput. Civ. Eng. 38 (4), 4024015. <https://doi.org/10.1061/JCEES.CPENG-5756>.
- Xing, L., Johnson, B.W., 2023. Reliability theory and practice for unmanned aerial vehicles. IEEE Internet Things J. 10 (4), 3548–3566. <https://doi.org/10.1109/JIOT.2022.3218491>.
- Yaacoub, J.P., Noura, H., Salman, O., Chehab, A., 2020. Security analysis of drones systems: attacks, limitations, and recommendations. Internet of Thing 11 (100218). <https://doi.org/10.1016/j.iot.2020.100218>.



DSPACE

<https://dspace.org/>

A system-based safety analysis of AI-enabled UAV operations for runway FOD detection: application of STPA and PROMETHEE

Abdouramane Attou Bounou, Nafissa; Kaya, Gulsum Kubra; Camelia, Fanny

2026-10

Elsevier

<https://doi.org/10.1016/j.ssci.2026.107280>

Abdouramane Attou Bounou N, Kaya GK, Camelia F. (2026) A system-based safety analysis of AI-enabled UAV operations for runway FOD detection: application of STPA and PROMETHEE. Safety Science, Volume 202, October 2026, Article number 107280