



Cyber Risk Maturity Assessment Model Based on SeMS for Aircraft Operations: A Systematic Literature Review

Agoes Soebagio¹, Dwi Afriyanto², Girinoto³, Arizal⁴, Jeckson Sidabutar⁵

¹²Department of Airport Operation, Indonesia Aviation Polytechnic Curug, Tangerang, Indonesia

³⁴⁵Department of Cryptography and Cybersecurity, National Cyber and Cryptography Polytechnic, Bogor, Indonesia

Email Address: dwiafriyanto@ppicurug.ac.id², girinoto@poltekssn.ac.id³, arizal@poltekssn.ac.id⁴, jeckson.sidabutar@poltekssn.ac.id⁵

Email Correspondence: soebagioagoes@gmail.com¹

ARTICLE INFO

Article History

Received: 17 September 2026

Revised: 21 September 2026

Accepted: 29 September 2026

Keywords

Aircraft operations, Aviation cybersecurity, CRMAM, Cyber risk maturity, SeMS



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright ©2025 by Author. Published by Akademi Penerbangan Indonesia.

ABSTRACT

Aviation cybersecurity governance rests on a three-document RTCA/EUROCAE airworthiness security family; DO-355A/ED-204A provides operational-phase compliance guidance, but no validated, scored maturity layer aligned with Security Management System (SeMS) principles exists above it. This study mapped evidence on cyber risk maturity frameworks in aviation operations, identified methodologies for developing a Cyber Risk Maturity Assessment Model (CRMAM), and characterized certification- and operational-phase gaps. A literature review followed Kitchenham's methodology and PRISMA 2020 standards, searching Google Scholar, IEEE Xplore, collections, and grey literature, with quality assessed with a modified Design Science Research Methodology rubric. From 889 records, 164 studies included after deduplication and screening. Across five questions, no aviation-specific, SeMS-integrated maturity model was identified; DO-326A governed certification only; no Security Performance Indicators library existed for aircraft operators; 97% of cross-regulatory studies addressed one regulatory body; and combined Delphi-AHP validation had aviation precedent but lacked application to cybersecurity maturity assessment. Findings indicate that aviation cybersecurity governance research has matured technically but not institutionally, confirming a multi-dimensional governance gap and supporting the proposed CRMAM design.

1. INTRODUCTION

Civil aviation cybersecurity governance sits at the intersection of physical safety and digital infrastructure. The RTCA/EUROCAE airworthiness security family comprises three complementary standards: DO-326A/ED-202A (RTCA, 2014) specifies the certification-phase security process required at type approval; DO-356A/ED-203A (RTCA, 2018) supplies supporting methods for that process; and DO-355A/ED-204A (RTCA & EUROCAE, 2020) guides continuing airworthiness: the operation, support, maintenance, and decommissioning of aircraft already in service. DO-355A/ED-204A therefore already addresses the operational phase at the level of compliance guidance, but offers guidance rather than a scored, tiered maturity-assessment instrument integrated with Security Management System (SeMS) performance management. This review confirms the gap as the absence of a validated maturity-assessment layer built on top of existing guidance, not the absence of operational-phase regulatory guidance itself.

SeMS, embedded in ICAO Annex 17 (ICAO, 2022) and operationalized by IATA's SeMS Manual (IATA, 2023), extends the Plan-Do-Check-Act structure of ICAO Annex 19 Safety Management Systems to security threats, including cybersecurity: hazard identification becomes threat identification, safety risk assessment becomes cyber risk assessment, and safety performance indicators become Security Performance Indicators (SePIs). EASA Part-IS, Regulation (EU) 2023/203 (European Commission, 2023), now requires commercial operators to establish an Information Security Management System, effectively operationalizing SeMS for EU-certified carriers. Despite this regulatory momentum, the literature has not produced a validated, aviation-specific instrument for assessing cyber risk management maturity at the operator level. Generic frameworks such as CMMI (CMMI Institute, 2018), C2M2 (U.S. Department of Energy, 2022), and COBIT (ISACA, 2019) are applied by analogy, as this review's RQ1 findings confirm (Table 1), without domain mappings calibrated to the avionics system lifecycle or to SeMS performance-management principles.

This systematic literature review is Phase 1 of a multi-phase program to design and validate a Cyber Risk Maturity Assessment Model for Aircraft Operations Based on SeMS (CRMAM), targeted at Indonesian commercial air carriers. Phase 2 will construct and validate the instrument through a combined Delphi-AHP design; Phase 3 will pilot it with an Indonesian commercial carrier. Five research questions structure the review: which cyber risk maturity frameworks exist in aviation operations and how far they integrate SeMS principles (RQ1); how existing standards bridge the DO-326A certification gap and operational governance under SeMS and EASA Part-IS (RQ2); which cyber-specific SePIs or risk metrics the literature proposes for operational contexts (RQ3); how frameworks address cross-regulatory alignment across ICAO, EASA, FAA, and IATA (RQ4); and which methodologies have validated cyber risk maturity instruments in aviation and other safety-critical industries (RQ5). In the Indonesian context, CRMAM aligns with national aviation cybersecurity regulation KM 39/2024 (Kementerian Perhubungan Republik Indonesia, 2024) and BSSN guidance (Badan Siber dan Sandi Negara, 2024), a regulatory context absent from the peer-reviewed corpus identified here and one that gives the review a national contribution alongside its international regulatory mapping. Closing this gap matters in practice: absent a validated instrument, Indonesian and international regulators currently assess operator cyber risk maturity through ad hoc or imported means rather than a framework calibrated to aviation's certification-to-operations lifecycle.

2. METHODS

Research Design

This review follows the Kitchenham methodology for software-engineering systematic reviews (Kitchenham & Charters, 2007), supplemented by PRISMA 2020 reporting standards (Page et al., 2021), with the protocol fixed before database searches began. Records qualified for inclusion if they addressed cybersecurity, information security, or cyber risk in a civil aviation context; concerned aircraft operators, airlines, maintenance organizations, or aircraft-system-level security; addressed maturity assessment, governance, framework, policy, regulatory compliance, or SeMS/SMS; were published between 2014, the year DO-326A was released, and 2026; appeared in English or Indonesian; and were peer-reviewed journal articles, conference papers, technical reports, standards, or recognized grey literature. Records were excluded if they addressed military aviation exclusively; concerned air traffic management without aircraft-operator relevance; offered purely technical vulnerability analysis without a governance dimension; addressed generic IT/OT security without an aviation application; addressed physical security only; appeared in another language; had no retrievable full text; or duplicated another record. We chose Kitchenham's protocol-driven method over PRISMA alone because reporting guidance does not by itself specify a search-and-screening protocol, and over a formal scoping-review design (JBI or Arksey-O'Malley) because the review tests five explicit, pre-specified research questions rather than open-endedly mapping a literature base, even though most included sources are conceptual, framework-proposing, or standards-analytic rather than empirical. PRISMA 2020 was layered on top of Kitchenham for reporting transparency. The review team fixed the protocol internally before database searches began; it was not registered on PROSPERO or OSF, a limitation noted in Section 2.5.

Population and Sample

Five information sources were searched during May 1-31, 2026, with no top-up search run between that search and manuscript submission: Google Scholar, accessed via Publish or Perish and substituted for Scopus owing to unavailable institutional access; IEEE Xplore; manual collection; grey literature from ICAO, EASA, FAA, IATA, BSSN, and KM 39 documents; and, unsuccessfully, Web of Science and the ACM Digital Library, which returned zero accessible records. Four Google Scholar query batches combined aviation terms with cybersecurity and maturity terms and yielded 687 DOI-filtered records. An unrestricted all-metadata search on IEEE Xplore returned roughly 15,000 results, too broad for meaningful screening, so we restricted an IEEE Xplore Command Search to aviation terms in the title field and cybersecurity/maturity terms in the abstract field; that search returned 53 exportable records. This title-field restriction is a directional search bias. Manual collection contributed 134 records, assembled by the review team through backward and forward citation snowballing from the IEEE Xplore records and from ICAO, EASA, FAA, and IATA guidance documents, and supplemented by the authors' own institutional libraries and by sources recommended by aviation-cybersecurity practitioners at the authors' institutions; manually collected records were screened against the same inclusion and exclusion criteria as database-sourced records. Grey literature contributed a further 15 records, for 889 records identified overall. The resulting screening population spanned publications from 2014-2026 across journal articles, conference papers, technical reports/standards, and grey literature.

Data Collection Technique

Deduplication removed 88 records: 73 by DOI-based exact matching and 15 by title fuzzy matching, leaving 801 records for title and abstract screening. A single screener applied the inclusion and exclusion criteria as explicit rules and routed any record that could not be confidently classified from the title and abstract alone to an uncertain category rather than to exclusion. The same screener resolved every uncertain record at full-text retrieval by re-applying the eligibility criteria to the full text rather than by consensus with a second rater; this documents the internal auditing procedure used during screening, but it remains a single-rater resolution process, not an independent dual-screening check. Title and abstract screening excluded 526 records, leaving 275 for full-text assessment; full-text screening excluded a further 111, leaving 164 studies in the synthesis (Figure 1). We extracted data into six structured files covering bibliographic and thematic core fields and research-question-specific fields for framework characteristics (RQ1), DO-326A and operational-phase coverage (RQ2), proposed metrics (RQ3), regulatory frameworks (RQ4), and validation methods (RQ5). Because a study could inform more than one research question, the sum of studies reported across RQ1-RQ5 ($126 + 11 + 17 + 31 + 16 = 201$) exceeds the 164 unique studies in the corpus; no study was double-counted within a single research question. Three studies were re-scored after full-text reading because the DSRM problem-definition or validation-evidence items could not be reliably rated from the abstract alone; rescoring updated only the extracted DSRM item values, and all three studies remained included with unchanged eligibility status. Quality was assessed with a modified Design Science Research Methodology (DSRM) rubric adapted for aviation cybersecurity governance studies, comprising eight items rated 0 (absent), 1 (partial), or 2 (present) for a maximum score of 16, covering problem definition, design methodology, domain justification, validation evidence, aviation-context specificity, regulatory mapping, limitations, and generalizability. Because DSRM assumes a design-science or empirical structure, it evaluates a source's utility as a structural design input for CRMAM rather than the source's own standalone scientific rigor, so a low DSRM score for a formal standard or technical report reflects a mismatch between the rubric and that document type, not a judgement on the standard's regulatory authority. We also assessed grey literature with the AACODS checklist and systematic reviews with the Critical Appraisal Skills Program (CASP) checklist.

Data Analysis Technique

Narrative synthesis followed Popay et al. (2006): five research-question-specific synthesis documents summarized evidence strength, key findings, CRMAM design implications, and research gaps; we did not conduct a meta-analysis, given the heterogeneity in study designs and outcome measures. We preferred DSRM over standard tools such as the Mixed Methods Appraisal Tool or JBI checklists because most included studies are conceptual, framework-proposing, or standards-analytic rather than empirical; DSRM instead evaluates whether a source provides a well-justified problem definition, a defensible design method, and validation evidence, properties that make a study usable as design input for CRMAM regardless of whether its own method was empirical or conceptual. The single-screener design is the primary internal-validity threat: conservative-bias routing protects against false exclusion but not against a systematic misapplication of a rule across many records, a risk dual screening with Cohen's Kappa calibration would ordinarily catch. No inter-rater reliability statistic is available, and the same single-rater risk applies to DSRM quality scoring. Google Scholar via Publish or Perish has documented lower precision than curated bibliographic

databases for topic-specific engineering literature, and the unconfirmed access to Web of Science and the ACM Digital Library is a coverage gap of unknown direction and magnitude, a material risk given that the review's headline claims are absence claims (e.g., that no SeMS-integrated maturity model exists), which are more sensitive to database coverage than presence claims. A targeted triangulation search across the review's core query strings, using the publicly accessible abstract-level indices of Scopus, Web of Science, Dimensions, Lens.org, and OpenAlex, had not been completed at the time of this review; it is planned before the Phase 2 protocol is finalized, to verify that no additional eligible records were missed. The IEEE Xplore title-field restriction (Section 2.2) is a further directional, nonrandom bias: it systematically excludes records whose title is generic and whose aviation application appears only in the abstract or body. The English- and Indonesian-language restrictions may underrepresent aviation cybersecurity governance practices in non-Anglophone jurisdictions. Finally, the review protocol was fixed internally rather than registered on PROSPERO or OSF, so its pre-specification cannot be independently verified from a public record.

3. RESULTS

Study Selection, Characteristics, and Quality

From 889 records identified across all sources, we removed 88 duplicates, yielding 801 records for title and abstract screening. Screening excluded 526 records: 431 for generic information-technology content, 62 for purely technical analysis, 18 for military-only context, 8 for air-traffic-management-only context, 6 for physical-security-only content, and 1 duplicate, leaving 275 records for full-text assessment. Full-text review excluded a further 111 records: 105 for generic information-technology content, 4 for military or air-traffic-management contexts, and 2 for purely technical analyses, yielding 164 included studies, summarized in the PRISMA 2020 flow diagram in Figure 1.

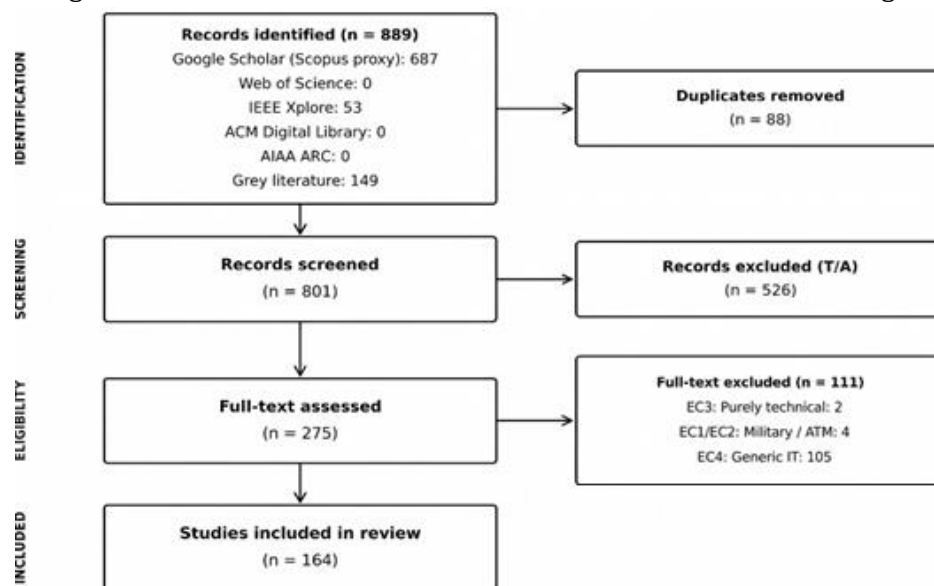


Figure 1. PRISMA 2020 flow diagram

Included studies span 2014 to 2026, with publication volume increasing markedly across the period, reflecting growing regulatory attention following EASA Part-IS's 2023 promulgation and ICAO Doc 10099 updates. Journal articles are the largest study-type category (72 studies, 43.9%),

followed by conference papers (45, 27.4%), technical reports and standards (39, 23.8%), and grey literature (8, 4.9%). Country of origin was reported in 46 of 164 studies (28%); among these, the United States, the European Union or multiple countries, and the United Kingdom dominate. The most frequent research themes are aviation cybersecurity, threats and vulnerabilities, frameworks, governance, and maturity models; airport-specific convergence of information technology, operational technology, and physical security is also well represented. SeMS integration, by contrast, appears in only one study, and EASA Part-IS in only two.

DSRM quality scores ranged from 4 to 15, with a mean of 7.95 (Figure 2). Seven studies (4.3%) reached the high-quality tier (12-16); 122 studies (74.4%) were moderate (7-11); and 35 studies (21.3%) were low (0-6), predominantly manual-collection records without accessible abstracts that we scored conservatively. Grey literature achieved the highest mean quality (10.62), reflecting the institutional backing of ICAO, EASA, and FAA source documents in that subset; conference papers averaged 8.69 and journal articles 8.65, while technical reports and standards, despite their formal provenance, scored lowest on this design-oriented rubric (5.23). DSRM is structurally oriented toward design-science and empirical reporting conventions, so it systematically penalizes standards documents, which are not written to report a problem definition or validation evidence in DSRM's terms; the 5.23 figure should be read as an artifact of the instrument's fit to that document type, not as a quality judgment on RTCA/EUROCAE standards themselves. The seven high-quality studies (Pemerintah Republik Indonesia, 2022; Ryon & Rice, 2018; Werthwein & Annighöfer, 2024; Uludağ et al., 2025; Federal Aviation Administration, 2018; Transportation Research Board, n.d.; Strohmeier et al., 2023) all address aviation cybersecurity governance with explicit regulatory or institutional backing, and three of them contribute to RQ2, making it the research question with the most focused, highest-quality evidence base.

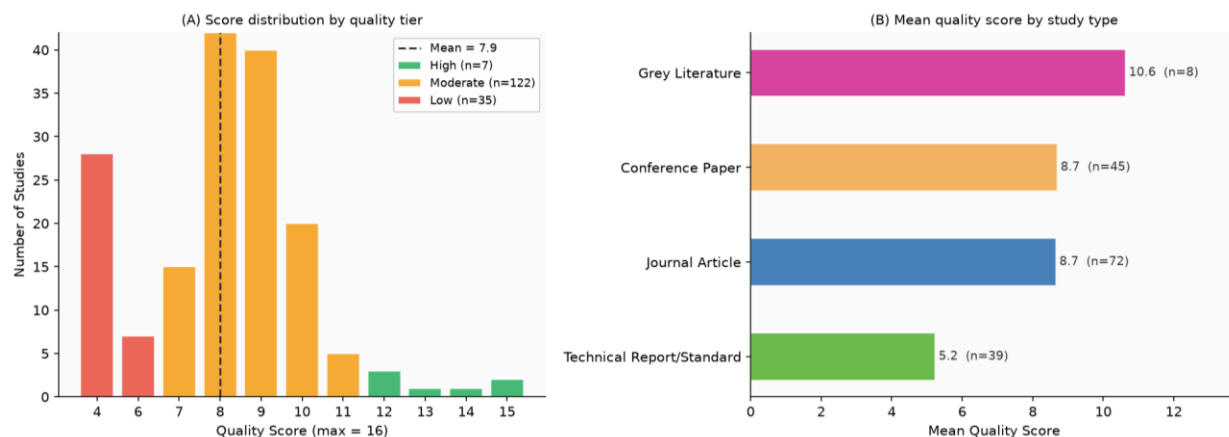


Figure 2. Quality score distribution (max = 16)

RQ1: Cyber Risk Maturity Frameworks in Aviation

126 studies contributed to RQ1, with moderate evidence strength. The corpus is large but fragmented: it characterizes threat landscapes and proposes frameworks conceptually, while validated, aviation-specific maturity instruments are almost entirely absent. Only three studies name an explicit maturity framework applied in aviation: COBIT in Adhikari (2020), C2M2 in Perrett and Wilson (2023), and CMMI in Ibanez (2025), each applied by analogy without aviation-specific domain mapping or regulatory calibration. Two further studies propose five-level maturity architectures

without referencing an existing framework (Stastny, 2022; Chen, 2024). Table 1 summarizes the structured maturity approaches identified in the corpus; CRMAM's own proposed design is presented separately in the Design Implications subsection of the Discussion, since it is this study's proposal rather than a study identified by the search.

Table 1. Maturity Frameworks Named or Proposed in RQ1 Studies

Framework	Study	Validation Status
COBIT (applied)	Adhikari (2020)	Not validated
C2M2 (applied)	Perrett & Wilson (2023)	Case study (OT)
CMMI (adapted)	Ibanez (2025)	Not validated
Five-level model	Stastny (2022)	Conceptual only
Five-level, project mgmt.	Chen (2024)	AHP-weighted

Only one study demonstrates even partial integration of SeMS principles into a cybersecurity governance framework (Wagner, 2023). Seven further studies independently identify the absence of an aviation-specific maturity framework as a significant research gap, and three name the SeMS-alignment gap specifically (Axon et al., 2022; Rugo et al., 2022; Pascarella et al., 2025). Table 2 shows assessment-domain coverage across the 126 RQ1 studies; the modal domain is threat and vulnerability management, reflecting that the literature better characterizes which threats exist than assesses organizational readiness to govern them. SeMS integration is the least-covered domain; along with the near-absent cross-regulatory alignment in RQ4 and the missing validated SePI library in RQ3, it is one of several evidence gaps CRMAM's design must close without direct literature support.

Table 2. Assessment Domain Frequency Across RQ1 Studies (n = 126).

Domain	Freq.	Notes
Threat/vulnerability mgmt.	36	Modal theme; technical, not governance
Governance & policy	25	Regulatory compliance, organizational policy
Maturity model design	23	Conceptual frameworks; rarely validated
Cyber risk management	14	Risk quantification, prioritization
Avionics security	13	DO-326A/DO-178C domain; system-level
Airport physical-cyber convergence	9	Airport IT/OT/physical convergence
Resilience	6	Cyber-physical resilience; links to SMS
Airline operator security	5	Aircraft operator as unit; least developed
SeMS integration	1	Only Wagner (2023), partial

Note: Frequencies sum to 132 assignments across 126 studies because a study could be coded to more than one domain; the sum is not a computational error.

Of the 126 RQ1 studies, 57 (45%) address civil aviation broadly, 48 (38%) address safety-critical industry generically, 10 address air-traffic management, 7 address airports, and only 4 explicitly address the aircraft operator or cargo carrier as the unit of analysis. The four validated instruments identified in the corpus all address industrial operational technology or cyber-physical resilience rather than aviation governance.

RQ2: DO-326A Certification vs. Operational SeMS Gap

Eleven studies contributed to RQ2, including the three highest-quality studies in the corpus, which supports the central finding despite the small study count. Five address DO-326A explicitly, two reference it, and four do not mention it, but all eleven confirm DO-326A as a design- and certification-phase standard, and none identifies an operational maturity-scoring mechanism. DO-355A/ED-204A, the RTCA/EUROCAE companion standard that already addresses the continuing-airworthiness operational phase at the compliance-guidance level, does not appear in the extraction data for any of the eleven studies, and none proposes a maturity-scoring or SeMS-aligned performance-monitoring mechanism for that phase; EASA Part-IS alignment is absent from all eleven, consistent with its 2024 effective date. Uludağ et al. (2025), the highest-quality study in the corpus, gives the most systematic characterization, showing that DO-326A addresses only the identification and design phases and leaves operation and maintenance unaddressed. Pascarella et al. (2025) name the SeMS-alignment gap specifically, calling for bridging the safety-management-system and cybersecurity-governance traditions that form CRMAM's architectural premise. Although DO-355A/ED-204A is the primary operational airworthiness security standard, the included literature has not turned it into a scored maturity framework: that gap is exactly what CRMAM's operational-layer contribution is meant to fill.

RQ3: Security Performance Indicators and Metrics

Seventeen studies contributed to RQ3, with weak evidence strength. The evidence reveals two dominant archetypes, benchmark-based indicators and quantitative security-process metrics, alongside an emerging pattern of maturity-level self-assessment rubrics; proposed metrics range from compliance benchmark indices and computational avionics risk scores to case-study-validated assessment scores and resilience indicators. Field validation appears in four studies, but Omido (2025) is the only one reporting operational field deployment of security metrics in an aviation-adjacent context, a step beyond field validation alone. Its metric definitions are accordingly a high-priority input for CRMAM's indicator design. No study in the corpus aligns proposed metrics with SeMS performance-management principles, under which Security Performance Indicators must be measurable, linked to security objectives, monitored at defined intervals, and tied to corrective-action thresholds; none of the 17 RQ3 studies meet all four criteria, and all draw exclusively on literature rather than operational airline data.

RQ4: Cross-Regulatory Alignment

Thirty-one studies contributed to RQ4, with weak evidence strength. The dominant pattern is single-framework focus: 30 of 31 studies (97%) address only one regulatory body or cite generic, unspecified standards, and only one study attempts any cross-regulatory comparison. Figure 3 makes the structural gap self-evident: the frameworks most important to CRMAM's operational governance scope show zero coverage, while general/unspecified references dominate. Specifically, EASA Part-IS, ICAO Doc 10099, and the IATA SeMS Manual show zero coverage in the peer-reviewed studies identified; Indonesian regulatory instruments KM 39/2024 (Kementerian Perhubungan Republik Indonesia, 2024) and BSSN guidance (Badan Siber dan Sandi Negara, 2024) appear only in grey literature, and no international peer-reviewed study addresses the Indonesian aviation cybersecurity regulatory framework. NIST CSF is coded in only 1 of the 31 RQ4 studies (3%), on a par with single-country FAA advisory-circular coverage and pre-Part-IS generic EASA references, and is not itself a dominant framework in this corpus. The single cross-regulatory synthesis identified

(Schlager et al., 2023) offers a partial EASA-FAA comparison at the objective level that does not produce a usable mapping matrix, and no study addresses the four-way ICAO-EASA-FAA-IATA harmonization that CRMAM requires.

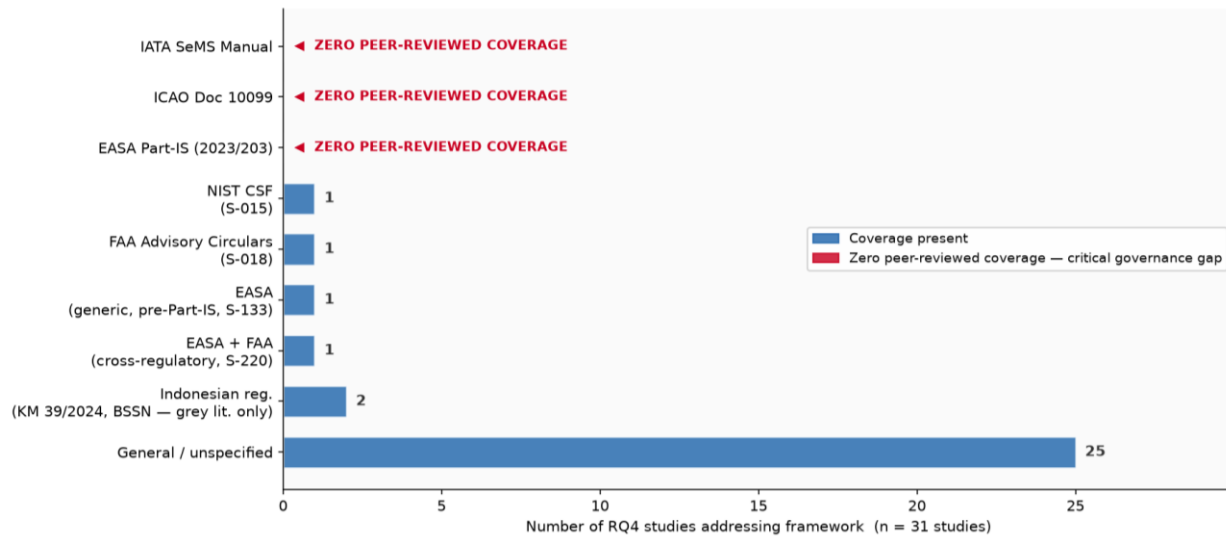


Figure 3. Regulatory framework coverage across RQ4 studies (n = 31)

RQ5: Validation Methodologies

Sixteen studies contributed to RQ5, with moderate evidence strength. Table 3 summarizes the validation methods used; conceptual and theoretical approaches dominate (8 studies, 50%), followed by case studies (4 studies) and surveys (2 studies). AHP-based methods appear in two studies: Mizrak and Turhan (2026) apply AHP alone, while Chen (2024) combines Delphi-derived indicator weighting with AHP-based scoring for aviation project-management capability rather than cybersecurity risk. This is the corpus's only combined Delphi-and-AHP design and the strongest methodological precedent identified for CRMAM's own validation design; no study in the corpus used Delphi in isolation. Perrett and Wilson (2023) further establish that maturity-instrument validation requires a real organizational deployment rather than expert review alone. RQ1 dominates the evidence base overall, while RQ2 and RQ5, the questions most central to SeMS and EASA Part-IS integration, have the smallest evidence bases.

Table 3. Validation Methods in RQ5 Studies (n = 16)

Method	N	Notes
Conceptual/theoretical	8	Predominantly expert-opinion or literature-based reasoning
Case study/pilot	4	Includes Perrett & Wilson (2023)
Survey	2	Includes Saleem (2024)
AHP (alone)	1	Mizrak & Turhan (2026); Consistency Ratio < 0.10
Delphi + AHP (combined)	1	Chen (2024), project management, not cyber risk; CR < 0.10
Delphi (alone)	0	Not identified in this corpus

4. DISCUSSION

Interpretation of Findings Against the Research Questions

The evidence assembled across five research questions converges on a single structural conclusion: aviation cybersecurity governance research has matured technically but not institutionally. For RQ1, the assessment-domain distribution reflects a broader bifurcation between technical and governance-oriented research strands. Studies addressing threat and vulnerability characterization and avionics security cluster tightly around shared methods, reflecting a mature sub-literature focused on attack-surface characterization informed by DO-326A's certification-phase vocabulary. By contrast, studies addressing governance, cyber risk management, and maturity-model design form a looser, more conceptually varied cluster with fewer shared methodological conventions, consistent with governance research occupying an earlier, less standardized stage relative to the technical sub-literature. SeMS integration and EASA Part-IS, the two themes most central to CRMAM's scope, sit at the periphery of both clusters instead of bridging them: a sign that the SeMS-alignment gap has not yet become a shared reference point for either research community. The near-absence of the aircraft operator as the unit of analysis (4 of 126 RQ1 studies) further indicates that the corpus's technical maturity has not translated into operator-level governance tooling.

For RQ2, the interpretation is that DO-355A/ED-204A's operational-phase compliance guidance and SeMS's performance-management architecture have evolved along parallel but unconnected tracks. None of the eleven RQ2 studies proposes the maturity-scoring bridge between them, even though Uludağ et al.'s (2025) systematic characterization and Pascarella et al.'s (2025) explicit call for SeMS bridging show that domain experts recognize the gap without yet closing it.

For RQ3, the absence of any metric library meeting SeMS's four performance-indicator criteria (measurable, objective-linked, periodically monitored, corrective-action-linked) means CRMAM cannot adopt an existing SePIs catalog wholesale; it must be developed from first principles, most plausibly through a three-tier architecture of compliance benchmarks, process metrics, and maturity-level descriptors that accommodates operators at varying data-collection maturity, informed most directly by Omido's (2025) field-deployed metrics. the gap without yet closing it.

For RQ4, the dominance of single-framework, generic citations over the specific instruments CRMAM must harmonize (EASA Part-IS, ICAO Doc 10099, IATA SeMS Manual) suggests that cross-regulatory harmonization is not merely under-researched but has not yet been framed as a research problem in its own right. NIST CSF itself is coded in only 1 of the 31 RQ4 studies (Results), so its selection as CRMAM's cross-regulatory backbone is a design judgment, not a review finding: no study in this corpus documents NIST CSF's adoption specifically within aviation cybersecurity governance, so the choice rests on NIST CSF's familiarity as a general-purpose security-function taxonomy across other critical-infrastructure sectors, rather than on any evidence of aviation-sector uptake found in this review. CRMAM's domain structure is designed to map explicitly to its five functions (identify, protect, detect, respond, and recover) for interoperability, while its Indonesian national regulatory layer is built directly from primary sources rather than secondary literature, given the complete absence of peer-reviewed treatment.

For RQ5, the single aviation-sector precedent for combined Delphi-and-AHP validation (Chen, 2024), applied to project-management capability rather than cybersecurity, is direct evidence that the combined design is feasible and produces interpretable relative domain importance in an

aviation setting, even though it was not built for CRMAM's purpose. Expert profiles across the RQ5 studies lean toward academia, with regulators absent from every documented expert panel: face, content, and regulatory validity cannot be assumed here and must be deliberately engineered into CRMAM's own validation design.

Design Implications for CRMAM

Table 1 situates CRMAM among the maturity approaches identified in RQ1: unlike COBIT, C2M2, and CMMI, which the included studies apply to aviation by analogy without domain-specific mapping, and unlike the two unreferenced five-level proposals in Stastny (2022) and Chen (2024), CRMAM is proposed here as a five-level, SeMS-explicit architecture with its own planned Delphi-and-AHP validation (Phase 2). This is a proposed design, not an included record, and its features below are design parameters derived from the Phase 1 findings rather than empirically validated claims. Taken together, these interpretations support six CRMAM design decisions.

First, a five-level maturity architecture, consistent with Stastny (2022), Chen (2024), and the CMMI/NIST CSF convention, ranging from Level 1 (ad hoc) to Level 5 (continuous improvement) and aligned with SeMS management-review cycles. Second, five core domains, governance and policy, risk management, incident response, supply chain security, and regulatory compliance, drawn from the highest-frequency assessment domains in Table 2 and calibrated to ICAO Annex 17 and EASA Part-IS requirement clusters, plus a sixth domain, continuous airworthiness security, that extends DO-355A/ED-204A's existing compliance guidance into a scored maturity structure covering electronic flight bag update logs, avionics line-replaceable-unit configuration histories, and maintenance-system connectivity records.

Third, a three-tier Security Performance Indicator architecture (compliance benchmarks, process metrics, and maturity-level descriptors) populating those domains with measurable indicators, since no existing SePIs catalog could be adopted wholesale (RQ3). Fourth, NIST CSF's five functions as a cross-regulatory backbone for mapping ICAO, EASA, FAA, IATA, and KM 39/2024 requirements, adopted as an external design judgment rather than a review finding (see Results, RQ4).

Fifth, a two-stage Delphi-then-AHP validation sequence with 15-20 experts drawn from operators, regulators, and academics, recruited deliberately across all three groups to close the expert-panel gap identified in RQ5, followed by pilot deployment with an Indonesian carrier, using Delphi consensus with Kendall's W of at least 0.70 and AHP weighting with a Consistency Ratio below 0.10. Sixth, an explicit Indonesian national regulatory layer built on KM 39/2024 and BSSN guidance, alongside the international ICAO, EASA, FAA, and IATA layers. All six remain proposed design parameters pending the Phase 2 Delphi-AHP construction and validation phase and the Phase 3 pilot deployment; none should be read as an empirically validated feature of CRMAM at this stage.

Comparison with Prior Reviews

To the authors' knowledge, this is among the first systematic reviews to examine cyber risk maturity assessment instruments in civil aviation operations under a SeMS framing. Prior aviation cybersecurity reviews address broader territory without engaging the maturity-governance or SeMS-integration dimensions to the same extent: Gnatyuk et al. (2023) review threats and countermeasures with a focus on technical attack surfaces, and Mizrak and Akkartal (2024) approach the domain from a systems-safety and multi-criteria prioritization perspective; neither addresses the DO-326A certification-operations boundary, EASA Part-IS operationalization, or SeMS-based

performance indicators as a central focus. In adjacent safety-critical sectors, maturity-model applications to industrial operational technology and aerospace digital transformation show that maturity assessment is feasible without engaging aviation's distinct DO-326A, DO-355A, and EASA Part-IS regulatory interface.

Limitations

Across most research questions, more than half of the studies rely on conceptual validation only, and only four of the seventeen RQ3 studies achieve field validation, which limits certainty about what works in practice. EASA Part-IS's 2024 effective date means the included corpus cannot yet reflect substantial practitioner compliance experience, so the zero-coverage finding for Part-IS is scoped strictly to the 164 studies included in this review's May 2026 search and should not be read as a claim about the wider, continuously evolving literature. The single-screener design means exclusion errors from inconsistent rule application cannot be fully excluded. Because a second, independent screener was not employed, no inter-rater reliability statistic (e.g., Cohen's kappa) could be computed or reported; this absence should be read as a limitation of the current search, not as evidence of screening consistency. The review protocol was not registered on PROSPERO or OSF before screening, which limits external verification that the eligibility criteria were fixed in advance rather than refined during the review. The IEEE Xplore search was restricted to the title field rather than title, abstract, and keywords, which likely under-retrieved studies that discuss cyber risk maturity, SeMS, or related constructs only in their abstracts. This field restriction, together with the Google Scholar-for-Scopus substitution and unconfirmed Web of Science/ACM Digital Library access, are coverage gaps of unknown direction and magnitude that the planned Phase 2 multi-database triangulation is intended to address.

Practical Implications

For aircraft operators, the intended end users of CRMAM, a tiered maturity score is designed to support at least four concrete uses once the instrument is validated. First, self-assessment: an operator could use the CRMAM rubric to locate its own cybersecurity governance on a five-level scale without waiting for an external audit finding to reveal the gap. Second, gap-remediation prioritization: a score of Level 2, for example, would indicate that basic controls exist but are not yet consistently measured or managed, pointing the operator toward SePI instrumentation and process-metric collection as the next investment rather than more advanced capabilities the model would place at Level 4 or 5. Third, audit and certification readiness: because CRMAM is designed to sit above DO-326A/DO-355A certification-phase compliance, an operator could use its score to anticipate operational-phase evidentiary expectations before a Part-IS or KM 39/2024 audit rather than during one. Fourth, board and executive reporting: a single tiered score, together with its underlying domain-level breakdown, is designed to give operator leadership a governance metric comparable in form to the maturity scores already used for safety management systems. For Indonesian aviation stakeholders more broadly, these findings carry additional practical implications. The national regulator responsible for KM 39/2024 currently has no peer-reviewed instrument to benchmark commercial operators' cyber risk maturity, a gap CRMAM is designed to close through its planned Phase 3 pilot deployment with Indonesian carriers. For BSSN, which holds a national cybersecurity coordination mandate, CRMAM's Indonesian regulatory layer offers a mapping between international airworthiness security standards and domestic cybersecurity regulation that has not previously been constructed from primary sources. Empirical validation of CRMAM itself, through

the planned Delphi and AHP phases and subsequent pilot deployment, is accordingly the most immediate research priority, since a design proposal grounded in literature synthesis alone cannot establish construct validity.

5. CONCLUSION

This systematic literature review of 164 studies indicates that the literature searched lacks a validated, aviation-specific, SeMS-integrated cyber risk maturity assessment instrument for aircraft operators. Five research questions produce convergent and mutually reinforcing findings: no SeMS-aligned maturity model was identified; the DO-326A certification-to-validated-operational-maturity gap is structural and unaddressed in all eleven relevant studies, notwithstanding DO-355A/ED-204A's existing non-maturity compliance guidance for the operational phase; no standardized Security Performance Indicators library was identified; cross-regulatory harmonization across ICAO, EASA, FAA, and IATA was not identified in the literature reviewed; and combined Delphi-and-AHP validation has an aviation precedent but was not identified as applied to cybersecurity maturity assessment specifically. The proposed CRMAM addresses all five gaps simultaneously, building on rather than duplicating existing certification-phase and continuing-airworthiness regulatory guidance. Its proposed design, a Delphi-and-AHP instrument-construction phase with 15 to 20 experts drawn from operators, regulators, and academics, and its planned validation, field testing with Indonesian commercial operators, are grounded in the evidence base this review produces, but both remain to be carried out; no component of CRMAM has yet been empirically validated. Recommendations for future work are fourfold. First, conduct the Delphi-AHP validation phase with a balanced panel of operator, regulator, and academic experts. Second, pilot CRMAM with at least one Indonesian commercial carrier to establish construct validity. Third, extend the cross-regulatory mapping to a formal ICAO-EASA-FAA-IATA harmonization matrix. Fourth, for civil aviation authorities specifically, the Directorate General of Civil Aviation (DGCA) under Kementerian Perhubungan (Kemenhub) and the Badan Siber dan Sandi Negara (BSSN), future work should pilot the use of tiered CRMAM scores as a supplement to, rather than a replacement for, existing SeMS compliance audits. A validated CRMAM score could give DGCA/Kemenhub an early, standardized signal of an operator's cybersecurity maturity between full audit cycles, while BSSN could use operator-level tiered scores in aggregate to identify sector-wide governance gaps against KM 39/2024 that a compliance-only audit checklist does not surface. Establishing how CRMAM scores would be reported to, and acted on by, these authorities alongside existing audit instruments is itself a task for the Phase 3 pilot, not a claim made by this review.

6. REFERENCES

- Adhikari, S. (2020). An analysis of AIAA aviation cybersecurity framework in relation to NIST, COBIT, and DHS frameworks. In *Proceedings of the AIAA AVIATION 2020 Forum*. <https://doi.org/10.2514/6.2020-2930>
- Axon, L., Fletcher, K., Scott, A. S., Stolz, M., Hannigan, R., El Kaafarani, A., Goldsmith, M., & Creese, S. (2022). Emerging cybersecurity capability gaps in the industrial internet of things: Overview and research agenda. *Digital Threats: Research and Practice*, 3(4). <https://doi.org/10.1145/3503920>
- Badan Siber dan Sandi Negara. (2024). *Peraturan BSSN tentang Keamanan Siber*. BSSN.

- Chen, Z. (2024). Evaluation of civil aircraft project management capability based on improved analytic hierarchy process algorithm. *Proceedings of SPIE – Computer Network Security and Communication Engineering*, 13175. <https://doi.org/10.1117/12.3032036>
- CMMI Institute. (2018). *CMMI for development (Version 2.0)*. CMMI Institute.
- European Commission. (2023). Commission Implementing Regulation (EU) 2023/203 of 27 January 2023 laying down rules for the application of Regulation (EU) 2018/1139. *Official Journal of the European Union*.
- Federal Aviation Administration. (2018). Cybersecurity for global aviation: A trust framework enabling global secure aviation interoperability. *Proceedings of the 2018 Integrated Communications, Navigation, Surveillance Conference (ICNS)*, 1–19. <https://doi.org/10.1109/ICNSURV.2018.8385007>
- Gnatyuk, S., Sydorenko, V., & Polozhentsev, A. (2023). Method for cybersecurity level evaluation in the civil aviation critical infrastructure. *Proceedings of the International Workshop on Cyber Security, Privacy and Trust*, 203–214. https://doi.org/10.1007/978-3-031-38082-2_16
- Ibanez, R. (2025). *Toward a framework dedicated to the design and development of information assurance for the Internet of Things*. Academia Engineering. <https://doi.org/10.20935/AcadEng7731>
- International Air Transport Association. (2023). *Security Management System (SeMS) manual* (2nd ed.). IATA.
- International Civil Aviation Organization. (2022). *Annex 17 to the Convention on International Civil Aviation: Security* (12th ed.). ICAO.
- ISACA. (2019). *COBIT 2019 framework*. ISACA.
- Kementerian Perhubungan Republik Indonesia. (2024). *Peraturan Menteri Perhubungan Nomor KM 39 Tahun 2024 tentang Keamanan Siber di Bidang Penerbangan*. Kemenhub RI.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering (Technical Report EBSE-2007-01)*. Keele University.
- Mizrak, F., & Akkartal, G. R. (2024). Prioritizing cybersecurity initiatives in aviation: A DEMATEL-QSFS methodology. *Heliyon*, 10(15), e35487. <https://doi.org/10.1016/j.heliyon.2024.e35487>
- Mizrak, F., & Turhan, K. (2026). *Threat-responsive prioritization of cybersecurity in aviation: A fuzzy decision framework integrating AHP, CFA, and K-means clustering*. Security and Privacy. <https://doi.org/10.1002/spy2.70182>
- Omido, T. (2025). *Industry 4.0 in commercial airlines: A bibliometric analysis*. Frontiers in Future Transportation. <https://doi.org/10.3389/ffutr.2025.1630011>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pascarella, D., Gigante, G., Lanzi, P., Spiller, E., & Fornaciari, E. (2025). Gaps and challenges in automation assessment to support human-centric aviation certification. In *Engineering Psychology and Cognitive Ergonomics (HCII 2025) (LNCS Vol. 15777)*. Springer. https://doi.org/10.1007/978-3-031-93721-7_9
- Pemerintah Republik Indonesia. (2022). *Peraturan Presiden Nomor 82 Tahun 2022 tentang Pelindungan Infrastruktur Informasi Vital*. Sekretariat Negara.
- Perrett, K., & Wilson, A. (2023). A cyber resilience analysis case study of an industrial operational technology environment. *Environment Systems and Decisions*, 43(2), 178–190. <https://doi.org/10.1007/s10669-023-09895-1>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., et al. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. ESRC Methods Programme.
- RTCA, & EUROCAE. (2020). DO-355A/ED-204A: *Information security guidance for continuing airworthiness*. RTCA Inc. / EUROCAE.
- RTCA. (2014). *DO-326A: Airworthiness security process specification*. RTCA Inc.

- RTCA. (2018). *DO-356A: Airworthiness security methods and considerations*. RTCA Inc.
- Rugo, A., Ardagna, C. A., & El Ioini, N. (2022). A security review in the UAVNet era: Threats, countermeasures, and gap analysis. *ACM Computing Surveys*, 55(1). <https://doi.org/10.1145/3485272>
- Ryon, L., & Rice, G. (2018). A safety-focused security risk assessment of commercial aircraft avionics. *Proceedings of the 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. <https://doi.org/10.1109/DASC.2018.8569314>
- Saleem, H. (2024). A survey of cybersecurity laws, regulations, and policies in technologically advanced nations: A case study of Pakistan to bridge the gap. *International Cybersecurity Law Review*. <https://doi.org/10.1365/s43439-024-00128-y>
- Schlager, C., Macher, G., Messnarz, R., & Brenner, E. (2023). Development of embedded systems for mobility: A narrative review. *Proceedings of the International Conference on Intelligent Systems and Applications (LNNS Vol. 817)*, 177–187. https://doi.org/10.1007/978-981-99-7886-1_16
- Stastny, V. (2022). Protecting aviation safety against cybersecurity threats. *IOP Conference Series: Materials Science and Engineering*, 1226(1), 012025. <https://doi.org/10.1088/1757-899X/1226/1/012025>
- Strohmeier, M., Leonardi, M., Markochev, S., Ricciato, F., Schäfer, M., & Lenders, V. (2023). In pursuit of aviation cybersecurity: Experiences and lessons from a competitive approach. *IEEE Security & Privacy*, 21, 2–14. <https://doi.org/10.1109/MSEC.2023.3265523>
- Transportation Research Board. (n.d.). *ACRP Report 140: Guidebook on best practices for airport cybersecurity*. National Academies Press.
- U.S. Department of Energy. (2022). *Cybersecurity capability maturity model (C2M2) (Version 2.1)*. U.S. Department of Energy.
- Uludağ, M. F., Bayoğlu, B., Candan, F., & Yilmaz, T. (2025). Enhancing safety and security assessment process in IMA platforms: A model-based approach aligned with ARP-4754B. *Proceedings of the 2025 AIAA/IEEE 44th Digital Avionics Systems Conference (DASC)*. <https://doi.org/10.1109/DASC66011.2025.11257420>
- Wagner, S. (2023). Analysis of safety and cybersecurity management systems interaction in decision-making including correlation for automated driving systems. *Proceedings of the International Conference on Computer Safety, Reliability, and Security*. https://doi.org/10.1007/978-3-031-64748-2_16
- Werthwein, M., & Annighöfer, B. (2024). Model-based avionics cybersecurity framework for identification of risk and evaluation. *Proceedings of the 2024 AIAA/IEEE 43rd Digital Avionics Systems Conference (DASC)*. <https://doi.org/10.1109/DASC62030.2024.10749689>