Flight Validation in Constrained Economic Environments: Optimizing Compliance, Efficiency, and Quality Through Intelligent Process Optimization

Fabrizio Maracich
Sr. Captain (BSc, MSc, MBA)
ENAV S.p.A.
Rome, Italy
E-mail: fabrizio.maracich@enav.it

ABSTRACT

In an increasingly constrained economic environment, flight validation activities are challenged to maintain safety and regulatory integrity while controlling operational costs. This paper examines how a smart, risk-based application of regulatory requirements, grounded in ICAO DOC 9906, Volume V, Second Edition, can preserve validation quality without unnecessary resource expenditure. Emphasis is placed on distinguishing mandatory safety objectives from prescriptive process elements, enabling more efficient allocation of validation effort. The integration of artificial intelligence within data analysis, anomaly detection, and preparatory validation phases is explored as a force multiplier rather than a substitute for human expertise. By combining regulatory intelligence, process optimization, and selective AI deployment, the proposed framework enhances consistency, traceability, and decision-making quality. The result is a sustainable flight validation model that aligns safety outcomes with economic realism.

INTRODUCTION AND CONTEXT

The validation of instrument flight procedures is a critical component of modern air navigation systems, ensuring compliance with design criteria and operational suitability. Increasing procedural complexity, driven by performance-based navigation, advanced avionics, airspace constraints, and environmental considerations, has expanded the scope and effort required for validation activities.

In parallel, flight validation organizations operate under increasingly constrained economic conditions. Budgetary pressure, limited access to specialized resources, and growing expectations in terms of delivery timelines and output quality have become structural features of the operating environment. These constraints do not alter safety or regulatory obligations but expose inefficiencies in traditional, resource-intensive validation models.

This paper addresses how flight validation activities can be organized and executed to preserve full regulatory compliance while improving efficiency, predictability, and sustainability. The focus is on optimizing the application of existing requirements and guidance rather than redefining them.

To define this context in the light of ICAO DOC 9906, Volume V, Second Edition, which was finally published at the end of 2025, the new process flow chart is presented in Figure 1.

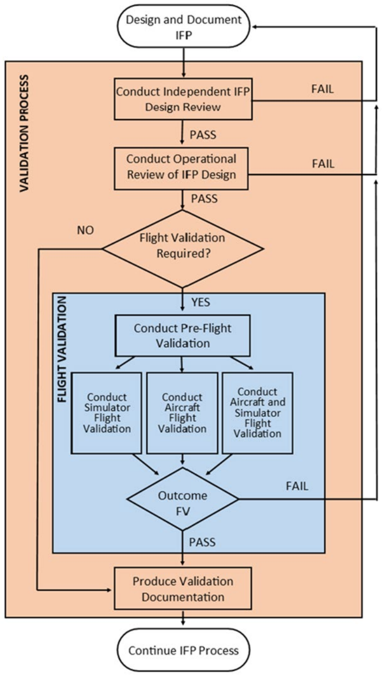

Figure 1 - ICAO DOC 9906 Vol. V Second Ed. process flow chart

REGULATORY AND STANDARDS FRAMEWORK

Flight validation operates within a layered framework of ICAO SARPs, PANS, and associated guidance material. SARPs and PANS define mandatory safety objectives, while guidance material—particularly ICAO Doc 9906, Volume V, Second Edition—provides acceptable means for achieving them.

Validation is a structured quality assurance activity preceding approval and publication, and States retain ultimate responsibility regardless of execution arrangements. Flexibility in implementation is therefore limited to methods that preserve safety intent, traceability, and evidentiary integrity.

FLIGHT VALIDATION: ROLE, SCOPE, AND STAKEHOLDERS

Flight validation ensures that a procedure, as designed and coded, is flyable, operationally suitable, and fit for publication. It is distinct from procedure design and flight inspection.

The process combines technical analysis, simulator-based evaluation, and aircraft validation where required. Key stakeholders include the State authority, procedure design organization, Flight Validation Organization, and the Flight Validation Pilot (FVP), with clearly defined roles and decision authority. Given the expert audience, detailed procedural descriptions are omitted in favor of functional clarity.

PROBLEM STATEMENT: CONSTRAINTS, INEFFICIENCIES, AND RISK DRIVERS

Despite a mature regulatory framework, flight validation execution remains affected by structural inefficiencies and variability that are primarily operational rather than regulatory in origin.

Economic and resource constraints

Flight validation is inherently resource-intensive, requiring specialized assets, highly qualified personnel, and extensive documentation. Under constrained budgets and limited staffing flexibility, inefficiencies become more visible, particularly in aircraft utilization, simulator access, and workload distribution.

Operational inefficiencies

Fragmented planning of simulator and aircraft activities leads to duplicated trajectories, unnecessary repositioning, and inefficient use of flight time. Weather dependency further introduces schedule volatility and reactive replanning.

In parallel, IFP package analysis—often conducted manually—extends preparation cycles and delays detection of inconsistencies between textual descriptions, chart depictions, and coding intent.

Risk of over-validation and under-validation

Organizations face a dual risk: excessive conservatism leading to inefficient resource use, and excessive compression of effort introducing safety and compliance exposure. Sustainable validation requires proportional allocation of effort based on procedural complexity and risk.

Management and governance challenges

Limited process visibility and fragmented toolsets lead to reactive decision-making, reduced predictability, and weak cost control.

Need for an optimized operating model

These constraints demonstrate that incremental improvements are insufficient. A structured approach is required, integrating optimized planning, platform allocation, and technology-enabled support while preserving human expertise.

METHODOLOGY AND OPERATING MODEL

The proposed framework maintains full compliance with ICAO SARPs, PANS, and Doc 9906 guidance, preserving all validation steps and accountability structures. Optimization is achieved through improved task allocation, integrated planning, and selective use of AI to support non-judgmental activities.

Two operating models are defined: a baseline ICAO-compliant model and an optimized AI-augmented model.

Baseline operating model

The validation process comprises five sequential steps:

- Independent IFP design review;
- Flight operational review;
- Pre-flight validation;
- Simulator and/or aircraft validation;
- Validation reporting

While fully compliant, this model often exhibits inefficiencies due to fragmented planning, duplicated effort, and manual reporting.

Optimized operating model

The optimized model preserves the same structure while enhancing execution:

- IFP package analysis (pre-validation): AI supports detection of inconsistencies between textual descriptions, charts, and design intent.
- Operational review: AI may support analysis, but the FVP retains full responsibility for flyability, workload, and human factors.
- Pre-flight validation: AI-assisted trajectory planning reduces overlap, repositioning, and inefficiencies.
- Execution phase: An integrated simulator–aircraft strategy allocates validation objectives to the most appropriate platform.
- Validation evidence package (post-validation): Data collected during simulator and/or aircraft activities is consolidated with AI-supported processing and validated by the FVP.
- Reporting: AI supports drafting, standardization, and traceability; human approval remains mandatory.

Outcomes

The optimized model delivers reduced cycle time, improved resource utilization, enhanced traceability, and better cost predictability, without compromising compliance.

PROCESS OPTIMIZATION THROUGH SIMULATOR AND AIRCRAFT INTEGRATION

Optimization is achieved through an access-based integration of simulators and aircraft rather than asset ownership. Simulators, typically accessed through rental arrangements, offer structural advantages including weather independence, higher dispatch availability, and repeatability.

Aircraft validation remains essential for real-world assessment, including obstacle proximity, infrastructure verification, and operational integration.

An optimized approach:

- Allocates validation objectives to the most suitable platform;
- Applies a simulator-first strategy where permitted;
- Uses AI-assisted trajectory planning to reduce overlap and repositioning.

This integrated model improves schedule robustness, efficiency, and cost predictability while preserving evidentiary completeness.

ECONOMIC EFFICIENCY, AI INTEGRATION, AND MANAGERIAL CONTROL

Efficiency is achieved through system-level optimization rather than isolated cost reduction.

Efficiency drivers

- Reduced airborne hours through optimized platform allocation;
- Accelerated IFP package analysis through AI;
- Improved schedule predictability.

AI vs human allocation

AI provides decisive advantages in:

- Bulk document analysis;
- Consistency verification;
- Repetitive, rule-based tasks.

Human expertise remains essential for:

- Operational judgement;
- Human factors assessment;
- Safety-critical decision-making.

Performance differences in analytical tasks can be orders of magnitude, justifying AI integration without altering decision authority.

Managerial control

An integrated IT/QMS environment ensures:

- End-to-end traceability;
- Real-time process visibility;
- Structured governance.

AI supports analysis and reporting but does not perform approval functions.

QUALITY ASSURANCE AND COMPLIANCE ASSURANCE

Quality assurance ensures process consistency and reliability, while compliance assurance demonstrates adherence to ICAO requirements.

A robust framework includes:

- Standardized processes and templates;
- Independent review functions;
- Full traceability of decisions and outputs.

AI enhances QA through consistency checks, workflow monitoring, and early detection of discrepancies.

The validation evidence package consolidates post-validation data from simulator and aircraft activities, combining AI-supported processing with FVP validation.

Quality gates ensure controlled progression through the process. The Flight Validation Organization internal oversight is fundamental to the credibility of these gates. Proper management independence ensures that commercial, scheduling, or resource pressures do not override safety and compliance considerations. The entire process is aligned with the State Regulatory Framework as per ICAO DOC 10068.

CONCLUSIONS AND FUTURE PERSPECTIVES

Conclusions

This paper proposes a structured framework enabling flight validation activities to be conducted efficiently and sustainably in constrained economic environments, without compromising safety, regulatory compliance, or output quality.

Optimization is achieved through integrated simulator–aircraft use, AI-supported analysis, and focused deployment of human expertise.

Perspectives

The model improves cost predictability, process control, and output consistency while maintaining regulatory alignment. Further benefits may arise from continued digitalization and integration of AI within validation systems.

FINAL REMARKS

Flight validation remains a cornerstone of instrument flight procedure design and implementation. Safety and efficiency are not competing objectives but can be mutually reinforcing when supported by disciplined governance and intelligent process design.

It is of paramount importance for the Flight Validation Organization not to fall into the (tempting) trap of allocating lower tier job to the AI only, because this will destroy the knowledge base rapidly, eroding the expertise and precluding younger and/or less experienced FVP to build the required exposure to multiple validation scenarios, thus impeding them to truly reach senior status in terms of overall IFP Validation competence.

REFERENCE MATERIAL

ICAO DOC 8168, vol. II, 7th Ed.
ICAO DOC 9613, 5th Ed.
ICAO DOC 9905, 3rd Ed.
ICAO DOC 9906, vol. V, 2nd Ed.
ICAO DOC 10068.

APPENDIX

Empirical Comparison of Human and AI Performance in IFP Package Analysis

A.1 PURPOSE AND CONTEXT

This appendix provides empirical support to the main paper on Flight Validation (FV) by comparing human expert performance and AI-assisted analytical performance in the specific context of Instrument Flight Procedure (IFP) package evaluation.

The analysis is intentionally restricted to IFP package review activities, as these represent a core, recurring component of FV workflows and are directly relevant to the technical scope and conclusions of the main paper. The appendix does not address generic document analysis or tasks outside the FV domain.

A.2 EMPIRICAL BASIS AND HISTORICAL DATASET

The assumptions, workload parameters, and time estimates used throughout this appendix are derived from historical operational data comprising approximately 1,500 completed validation processes.

These validation processes:

- Span multiple years of operational activity;
- Include both standard and non-standard IFP designs;
- Reflect typical iteration cycles between design, review, and correction.

The historical dataset is used exclusively to establish representative averages and conservative workload assumptions. No individual validation outcome is analyzed, reproduced, or disclosed.

A.3 SCOPE AND BOUNDARY CONDITIONS

The analysis focuses on the desktop analytical phase of IFP package evaluation, including:

- Review of procedure design documentation;
- Consistency and completeness checks;
- Identification of potential non-compliances or ambiguities;
- Preparation of structured analytical findings.

The following activities are explicitly excluded:

- Flight inspection execution;
- Operational flight checks;
- Formal regulatory approval acts.

This delimitation ensures alignment with the analytical support role evaluated for AI systems.

A.4 METHODOLOGICAL FRAMING

The comparison is conducted under a human-in-the-loop model:

- AI systems operate solely in a non-decisional, advisory role;
- Human experts retain full responsibility for interpretation, professional judgment, decision making, and regulatory accountability.

Any efficiency gains are therefore interpreted strictly as reallocation of expert time, not as delegation of authority or workforce reduction.

A.5 PRIMARY ANALYTICAL TASK: IFP PACKAGE EVALUATION

A.5.1 Task Definition

The primary task analyzed is the review of a complete IFP package, consisting on average of 30 pages, including procedure design material, coding data, and supporting notes and references. Identical inputs and constraints are provided to both human analysts and AI systems.

A.5.2 Evaluation Dimensions

Performance is assessed across the following dimensions:

- Completeness: identification of relevant issues;

- Correctness: avoidance of false positives;
- Traceability: linkage of findings to applicable standards;
- Internal coherence: consistency of observations and conclusions.

Human outputs are produced by qualified FV practitioners. AI outputs are generated under fixed prompting constraints and are subject to a single clarification pass, mirroring standard human rereading.

A.5.3 Comparative Observations

The comparison shows that AI systems are effective in first-pass screening of documentation, detection of internal inconsistencies, and structured extraction of potential issues. Human experts remain essential for contextual interpretation, resolution of ambiguities, acceptance or rejection of observations, and final decision making. AI assistance does not replace expert judgment but reduces preparatory analytical effort.

A.6 ANNUAL MAN-HOUR IMPACT ASSESSMENT

A.6.1 Workload Assumptions

Based on historical data and conservative normalization:

- IFP packages per Flight Validation Pilot (FVP) per year: 50;
- Average package size: 30 pages;
- Average review cycles per package: 1.5.

A.6.2 Human-Only Baseline

Estimated effort per IFP package is 5–7 hours, including review, standards cross-checking, documentation, and partial re-review. Annual effort per FVP is estimated at 250–350 hours.

A.6.3 AI-Assisted Scenario

AI systems are used for first-pass analysis and consistency checks, with mandatory human validation. Estimated effort per package includes 0.25–0.4 hours for AI processing and setup and 2–2.5 hours for human validation and decision making, for a total of 2.25–2.9 hours. Annual effort per FVP is estimated at 113–145 hours.

A.6.4 Net Effect

Scenario	Annual Hours
Human-only	250–350
AI-assisted	113–145
Freed human effort	≈100–200

A reference value of approximately 150 hours per year per FVP is used for contextual analysis. No workforce reduction is implied.

A.7 HUMAN CAPITAL REALLOCATION

The freed capacity is fragmented across the year and cannot be converted into headcount reduction without loss of safety margin. Instead, it enables reallocation toward activities requiring human expertise, including:

- Deeper analysis of complex or non-standard procedures;
- Cross-procedure consistency checks at aerodrome level;
- Improved preparation of regulatory justifications;
- Expanded peer review and internal quality assurance

AI assistance reduces repetitive analytical effort, while humans retain full responsibility for decision making and accountability.

A.8 EXECUTIVE SYNTHESIS

The empirical evidence presented in this appendix supports the following conclusions:

- AI-assisted analytical workflows can materially improve efficiency in IFP package evaluation;
- Benefits persist under conservative workload assumptions;
- Human expertise remains central and irreplaceable;
- Efficiency gains translate into better use of expert time, not automation of authority.

Independent academic (1)(2)(3) and governmental research (4) provides additional context for these findings. University-led and peer-reviewed studies comparing AI systems and human experts in document reading, comprehension, and summarization tasks consistently show that, even when using generalist large language models, AI assistance can significantly reduce the time required for first-pass analysis, coverage checks, and structured synthesis of technical documentation, while maintaining performance comparable to human reviewers on defined metrics. These studies also consistently emphasize that human oversight remains essential for interpretation, contextual judgment, and final decision making.

In particular, comparative work on administrative, educational, and clinical documentation demonstrates that generalist AI systems already achieve measurable efficiency gains in structured analytical tasks, despite not being specifically trained for the target domain. A policy-grade study conducted by the UK Government (4) reports an average 56% reduction in time for AI-assisted evidence reviews compared to human-only workflows, with humans retaining full responsibility for conclusions and accountability. This magnitude of efficiency gain is consistent with the results obtained in the present analysis of IFP package evaluation, which identifies a reduction of approximately 40–55% in analytical effort under conservative assumptions.

The convergence between independent academic literature, government evidence, and the empirical results reported here suggests that the observed benefits are not contingent on the use of specialised or domain-trained AI systems. Rather, they arise from the ability of even generalist AI models to support repetitive, high-volume analytical steps. It follows that a specialised AI system, explicitly trained on IFP design and applicable regulatory material, could reasonably be expected to deliver equal or greater efficiency gains, while remaining subject to the same human-in-the-loop governance and accountability constraints described in this appendix.

When implemented within clearly defined procedural boundaries, AI therefore functions as an enabler of human performance in FV-related analytical activities, reinforcing safety assurance and regulatory robustness rather than undermining them.

A.9 REFERENCE MATERIAL FOR THE APPENDIX

(1) Ripoll Y Schmitz, L. M., & Sonnleitner, P. (2025). Evaluating AI-generated vs. human-written reading comprehension passages: An expert SWOT analysis and comparative study. Large-scale Assessments in Education (Springer Nature). DOI: 10.1186/s40536-025-00255-w.

(2) Russodivito, M., Ganfi, V., Fiorentino, G., & Oliveto, R. (2024). AI vs. Human: Effectiveness of Large Language Models in simplifying administrative documents. Proceedings of the 10th Italian Conference on Computational Linguistics (CLiC-it 2024), Italian Association for Computational Linguistics, Pisa, Italy.

(3) Van Veen, D., et al. (2024). Adapted large language models for clinical text summarization. Nature Medicine, Nature Publishing Group.

(4) UK Government. (2025). AI-assisted vs human-only evidence review: Results from a comparative study. UK Cabinet Office – Government Digital Service. Published on GOV.UK. Cabinet Office – Government Digital Service. Published on GOV.UK.