Progress in Aerospace Sciences 158 (2025) 101143

Contents lists available at ScienceDirect

Progress in Aerospace Sciences

journal homepage: www.elsevier.com/locate/paerosci

Regulatory pathways to certifiable condition based maintenance solutions in aviation: A comprehensive review

Robert Meissner [a,*], Ahmad Ali Pohya [a], Oliver Weiss [b], David Piotrowski [c], Gerko Wende [a]

[a] German Aerospace Center (DLR), Hamburg, Germany
[b] Airbus Operations GmbH, Hamburg, Germany
[c] Delta Air Lines, Technical Operations, Atlanta, GA, United States

ARTICLE INFO

Keywords:
Condition Based Maintenance
Aircraft maintenance
Aircraft Health Management
Preventive maintenance
OSA CBM
Automated Condition Monitoring
MSG-X

ABSTRACT

Aircraft maintenance is characterized by strict regulations to ensure airworthiness throughout an air vehicle's complete lifetime. While the traditional maintenance approach of regular inspections and functional checks has led to extraordinary high levels of safety, it is also cause for considerable maintenance-related downtimes and substantial operating cost contributions. At the same time, the vast majority of performed inspections does not reveal any defects and will leave the aircraft's condition unchanged. Therefore, the promise of substantial cost savings pushes manufacturers and operators constantly towards replacement of those tasks by automated Condition Monitoring (CM) and Health Management (HM) systems. However, regulatory guidance for the development of certifiable HM solutions to substitute manual scheduled maintenance tasks by Condition Based Maintenance (CBM) approaches is sparse. Consequently, the introduction of these technologies remains slow while current use cases are limited to non-critical maintenance tasks or the avoidance of unscheduled maintenance events due to system breakdowns. With this work, we will provide an in-depth review of existing guidelines from (a) regulatory authorities, (b) institutions such as SAE and the International Organization for Standardization (ISO), and (c) academic publications. Using these insights, we will derive a holistic framework that provides HM experts a guiding document to support their development of certifiable technical solutions. As a result, this guidance will help to exploit the existing technical capabilities for a continuous CM to determine airworthiness statuses and to replace scheduled preventive maintenance tasks by automation.

Contents

* Corresponding author.
E-mail address: robert.meissner@dlr.de (R. Meissner).

https://doi.org/10.1016/j.paerosci.2025.101143
Received 10 May 2025; Received in revised form 27 August 2025; Accepted 30 September 2025
Available online 15 October 2025

1. Introduction

In order to ensure a continued airworthiness, commercially operated aircraft need to be maintained regularly [1]. At the same time, aircraft maintenance is heavily regulated and an important contributor to the high safety standards the industry has achieved to date. However, with these levels of safety and the corresponding frequent checks, aircraft maintenance also contributes up to 20% to an operator's Direct Operating Cost (DOC) [2] and decreases the asset's operational availability through regular downtimes [3]. Although these can be caused by unscheduled (corrective) and scheduled (preventative) maintenance tasks, especially the latter significantly drive an operator's Direct Maintenance Cost (DMC) [4]. Simultaneously, a majority of these preventatively executed scheduled maintenance tasks do not result in any restorative action and consume limited resources unnecessarily [4].

As a consequence, new maintenance concepts – such as CBM in general [5] and predictive [6] or prescriptive maintenance [7] in particular – have been developed throughout the years. Their aim is to reduce the need of repetitive manual inspections and to project imminent system failures in order to avoid extensive maintenance downtimes [8]. In the past decade, extensive research on the development of automated CM and HM technologies has resulted in an abundance of scientific publications. Although these terminologies are often used synonymously, CM is the pure act of measuring system performance parameters, while HM extends the capabilities further by incorporating diagnostics and prognostics insights as well as providing a decision support function [9]. As the sheer quantity of publications on CM and HM technology development extends beyond the scope of this paper, interested readers are kindly referred to exemplary review papers on the topic of automated diagnostics and prognostics techniques, e.g., Jardine et al. [10], Lei et al. [11] and Zio [12]. Furthermore, a large variety of synonyms for HM technologies has been developed throughout the years (see Table 1). Even though these abbreviations emphasize slightly different connotations, for consistency and to avoid confusion, we will exclusively use the generic term HM and the aviation-specific term AHM in the context of this paper.

In addition to academic advances, first aviation-related industrial use cases of HM technologies have been introduced, e.g., an SHM system [23] and HM solutions for the engine bleed air system [24] and for the Integrated Drive Generator (IDG) [25], respectively. However, despite extensive developments, first industrial use cases, and the promise of substantial economic benefits [26–36], the broad introduction into service for HM solutions continues to be slow [37,38]. While many of the technologies can support avoiding Aircraft on Ground (AOG) scenarios, i.e., unplanned operational interruptions due to technical failures, they often focus exclusively on either non-critical systems or introduce additional maintenance tasks that have not been covered by traditional scheduled maintenance [39].

One reason for the slow adoption rate of HM technologies is the lack of regulatory guidelines for their integration into an aircraft Maintenance Program (MP) [40,41]. While a set of guidelines for the development of certifiable HM solutions for rotorcraft applications exists [18,42], there is no equivalent regulation for large airplanes as of CS 25 design specifications [43]. However, the application of HM technologies for airworthiness determinations must have prior approval by a competent aviation authority [21]. Consequently, regulatory guidance material is needed to drive development efforts in a harmonized level across all aviation jurisdictions [21] for the subsequent elimination of regular inspections through HM usage [44]. Otherwise, the developed solutions will continue to offer limited flexibility to drive maintenance actions based on health indicators and failure predictions [39]. Besides the regulatory aspects, IATA [21] also emphasizes that any developed HM solution, and consequently the corresponding certification framework, will need to offer the possibility for operators to opt-out of its usage. Therefore, there need to be fall-back solutions in place, e.g., traditional scheduled preventative maintenance tasks, that ensure an assets airworthiness without the mandatory requirement to use some sort of prescribed HM solution [45].

Based on these limitations, this paper aims at reviewing existing HM development guidelines to derive a framework that will support certification of HM solutions and ultimately allow the automation of current scheduled maintenance tasks. Therefore, it will contribute to the following three aspects.

1. Analysis of the current aviation regulation concept and the definition of conventional scheduled aircraft maintenance tasks
2. Identification of key characteristics of an automated CBM concept and the corresponding development objectives
3. Derivation of a suitable framework for certification based on the identified objectives and established industry practices

The remainder of the paper is structured as follows. After a brief overview of the aviation regulatory environment (Section 2), we will introduce the CBM principle in the aviation context and define development objectives for certifiable HM solutions (Section 3). Based on these insights, we will then present the baseline for our framework (Section 4), followed by in-depth discussion of the respective developmental steps for each stage of the process (Sections 5 to 9). In order to allow the application of our framework for subsequent industry use cases, we will summarize the essential findings, identify existing limitations of our study, and provide an outlook in Section 10.

2. Maintenance-related regulations in civil aviation

Since the development and establishment of AHM systems touches upon various aspects and entities of civil aviation regulations, we first need to review the existing regulatory framework and the associated established industry process. This will include the examination of the general maintenance-related regulatory tree and the determination process for a scheduled MP in civil aviation.

Table 1
Definition for different health management technologies and their characteristics.

Abbreviation	This definition incorporates...	Ref.
Integrated Vehicle Health Management (IVHM)	... technologies for anomaly detection, diagnosis, and prognosis – integrated across different subsystems and industries.	[13–15]
Prognostics and Health Management (PHM)	... techniques that enable maintenance actions on products and processes based on need – determined by the current system condition via diagnostic analyses and/or the expected future condition through prognostic methods.	[16]
Health and Usage Monitoring System (HUMS)/Vibration Health Monitoring (VHM)	... monitoring capabilities of critical components of the propulsion system for rotorcrafts – to detect degradation and fatigue, and to prevent failures while increasing the availability.	[17–19]
Structural Health Monitoring (SHM)	... the observation of a structure or mechanical system over time – using periodically spaced measurements, the extraction of damage-sensitive features, and their statistical analysis to determine the current state of system health.	[20]
Aircraft Health Management (AHM)/Integrated Aircraft Health Management (IAHM)	... the capability of using health monitoring of aircraft structures and systems (including propulsion systems) – in order to control the scheduling of needed maintenance actions.	[21]
Integrated System Health Management (ISHM)	... the design, development, operation, and life cycle management of components, subsystems and vehicles – with the aim of maintaining the nominal system behavior and ensuring operational safety and performance under non-nominal conditions.	[22]

A central terminology that is often referred to is airworthiness. While different nuances of its definition exist, airworthiness can be generally defined as the *capability* of an air system configuration to *safely* complete flight missions within *approved limits* [46,47]. Further examining the respective dimensions reveals that [47]

- *Safely* refers to the normal course and satisfactory conclusion of the flight without any conditions that can cause death, injury, the loss of equipment, or damage to the environment,
- *Capability* describes the aircraft's conformity to established design and manufacturing criteria, and
- *Approved limits* refer to the flight envelope an aircraft is designed for operating within, mainly depending on speed and structural load factors.

2.1. General regulatory framework

The civil aviation industry is characterized by strict legal processes that shall ensure safe and reliable operations. The accompanying documents can be subdivided into three parts (see Fig. 1),

- General regulations governing all aspects of civil aviation [48],
- Rules that are relevant to the (initial) air vehicle design [49] with their respective Certification Specification (CS) [18,43], and
- Rules that shall ensure a continuous airworthiness [50].

While initial airworthiness ensures that Type Certificate Holders (TCHs) adhere to established industry practices with their aircraft design and manufacturing processes, continued airworthiness regulations are designed for aging and degrading aircraft to provide a minimum safety level throughout their lifetime – with regular maintenance performed by qualified personnel.

These implementing rules are further being supplemented by designated annexes that provide specifications for different organizational aspects, e.g., qualification requirements for certifying staff.

These regulatory documents are supported by a number of Acceptable Means of Compliance (AMC) that, if followed closely, ensure compliance with the regulatory intent. One of the more prominent examples in the field of aircraft MP development is the MSG-3 methodology to comply with the continuing airworthiness requirements of CS 25 [51] and CS 29 [52], respectively (see Section 2.2.2).

The exceptional importance of operational safety can already be seen in Regulation (EU) 2018/1139 [48, Annex II] where the design requirements for a continued airworthiness of a product are highlighted. These requirements explicitly incorporate the establishment of Instructions for Continued Airworthiness (ICA) and their availability for operators to ensure the airworthiness standard for the aircraft

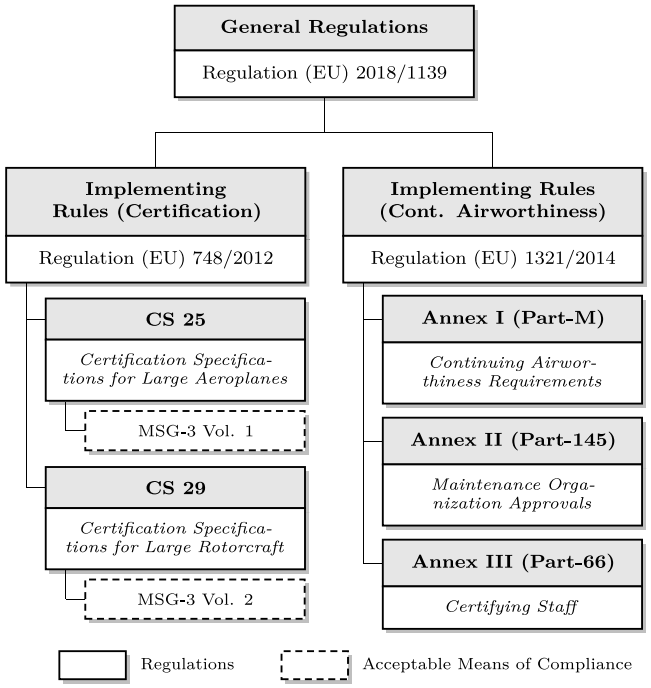

Fig. 1. Maintenance-related civil aviation regulations in the European Union. *Source:* Based on [1,2].

and all its associated parts is maintained throughout its lifetime; furthermore, these ICA must contain manuals that cover maintenance instructions and procedures as well as servicing and trouble-shooting information [49, 21.A.61], [43, 25.1529 & Appdx. H], [50, M.A.302].

Additionally, maintenance can only be performed by qualified maintenance organizations [50, Annex II], which are responsible to ensure, only qualified personnel with appropriate levels of knowledge, skill, and experience are performing the maintenance tasks [50, 145.A.30]. Since the associated skills and formal qualification change for different types of maintenance activities, a number of corresponding qualification requirements (e.g., mathematics, physics, and electrical fundamentals) exist, that need to be fulfilled to obtain the respective maintenance licenses [50, Annex III].

2.2. Current process of developing a scheduled aircraft maintenance program

With the general regulatory concept explained, we will now discuss the current process of determining legacy scheduled maintenance tasks in accordance with the certification specifications for large airplanes (CS 25) [43].

2.2.1. General requirements – Instructions for continued airworthiness

As mentioned before, aircraft need to be safe throughout all phases of their life cycle. Therefore, an aircraft owner is required to ensure that their asset maintains an airworthy condition [50, M.A.201]. However, this does not imply that owners need to execute the maintenance tasks themselves as they can subcontract this execution to approved maintenance organization [50, Annex II].

Consequently, TCHs are instructed to provide descriptive maintenance-related data and maintenance accomplishment instructions that have been prepared in accordance with the type certificate basis and consider the effect of aging structures [49, 21.A.61], [43, 25.1529 &25.1729]. Furthermore, Appendix H of CS 25 [43] specifies that the ICA shall be prepared in the form of manuals. Among others, these manuals have to include the following information.

Aircraft Maintenance Manual (AMM). A document containing information on (a) aircraft features and data that are relevant for (preventive) maintenance measures, (b) installed systems, appliances and engines, (c) how the aircraft's components and systems are controlled and operated (including special procedures and limitations), and (d) servicing information (e.g., access points for lubrication, inspection and servicing, reservoir capacities, servicing fluid and lubricant specifications, applicable pressures, and Ground Support Equipment (GSE)) [43].

Maintenance instructions. These instructions have to provide scheduling information for each part of the aircraft (i.e., required frequencies for cleaning, inspection, adjustment, tests or lubrication) together with corresponding degrees of wear tolerances from the respective equipment manufacturer. Furthermore, if applicable, they have to contain recommended overhaul periods with information on replacement procedures and general procedural instructions (e.g., ground test processes or storage limitations). Lastly, troubleshooting information for probable malfunctions, their diagnosis and appropriate restorative actions need to be provided [43].

Miscellanea. Additionally, the ICA shall contain (a) diagrams of structural access plates for inspection, (b) details for necessary special inspection techniques (e.g., radiographic or ultrasonic testing), (c) information for protective treatments to the structure after inspection, (d) data relevant to structural fasteners (e.g., means of identification, discard recommendations and torque values) and (e) a list of special tools that are required [43].

Airworthiness limitation section. Lastly, a section dedicated to airworthiness limitations needs to be provided. It needs to provide information for maintenance procedures and intervals of safety-critical items, structures, the fuel tank system, Electrical Wiring Interconnect System (EWIS) components, Certification Maintenance Requirements (CMRs), and lightning protection. Furthermore, the total permissible number of accumulated flight cycles or flight hours for the aircraft structure needs to be provided [43].

2.2.2. Definition of system maintenance tasks for the maintenance review board report

With the basic regulatory requirements presented, we now want to discuss how TCHs define the necessary maintenance work for the continued airworthiness of their products. The results of this definition are subsequently summarized in the Maintenance Review Board Report (MRBR) and build the foundation for operators to develop their maintenance schedule accordingly. One of the most established means for that development process is the MSG-3 analysis [51] – a risk-based maintenance approach that ensures compliance with the regulatory requirements presented in Section 2.1.

The general idea of current preventive aircraft maintenance is based on the concept of Reliability Centered Maintenance (RCM) as described by Nowlan and Heap [53]. That is, it employs a decision process to identify appropriate maintenance measures for the management of failure modes that could otherwise result in severe functional failures with operating safety implications [54]. RCM further assumes that an item's reliability is the result of its design and built quality [55]. Since the main objective of RCM is not the complete avoidance of failures but their management with respect to operational consequences [55], the development of a cost-effective preventive MP requires (a) an understanding of influencing factors for functional failures, (b) the analysis of failure consequences, and (c) the definition of preventive measures [53,56].

For the definition of aircraft system maintenance requirements, the process start with the results from an analysis in the style of a Preliminary System Safety Assessment (PSSA), where all critical failure conditions and their effects on an aircraft are identified top-down [57]. Following that, the MSG-3 methodology subsequently requires the completion of two steps [51]:

1. An allocation of possible system failures to predefined Failure Effect Categories (FECs). These can range from evident/hidden failure with implications for the safe completion of a flight mission to mere economic implications.
2. The identification of applicable and effective *manual* maintenance tasks from a predefined task set and the definition of suitable maintenance intervals.

Based on this analysis, one (or a combination) of the following maintenance tasks can be chosen.

Lubrication or servicing. These tasks represent the lightest form of preventive maintenance and intend to maintain an item's inherent design capabilities. Examples for servicing tasks can be cleaning of items or checking and replenishing fluid levels [51].

Operational or visual check. These tasks serve as failure-finding tasks with obvious pass/fail criteria to detect defects for hidden system functions, e.g., protective equipment. The (incorrect) function of those systems is unknown to the operating crew during execution of their normal duty. Therefore, an operational or visual check determines if an item fulfills its intended purpose or is in its intended state without measuring quantitative tolerances [51,53,58].

Inspection or functional check. A functional check is defined as a quantitative check to determine if an item performs with its functions within specified limits [51]. Furthermore, the inspection is an examination of specific items or areas and can be subdivided into [51]

- General Visual Inspection (GVI), i.e., a visual examination – made from within touching distance under normally available lighting conditions and without specialized equipment – to detect obvious damage or irregularities,
- Detailed Inspection (DET), i.e., an intensive examination – enhanced by (a) tactile assessment to check for tightness, (b) a direct source of good lighting at an intensity deemed appropriate or (c) inspection aids (e.g., magnifying lenses) – to detect damage or irregularities, and
- Special Detailed Inspection (SDI), i.e., an extensive examination – performed with specialized techniques (e.g., Nondestructive Testing (NDT)) or equipment (e.g., boroscope) – to detect damage or irregularities.

For these tasks to be effective, it has to be possible to (a) reliably detect an item's reduced failure resistance for a specific failure mode, (b) define a failure threshold that can be detected and (c) provide sufficient time between the initial detectability of a potential failure and its ultimate occurrence [58].

Restoration. With a restoration task, a system will be reworked to return the item to a specific standard. These tasks will be issued in fixed time intervals and can vary from cleaning or replacement of single parts up to a complete overhaul. For it to be effective, systems need to possess a clearly identifiable age limit after which the failure rate significantly increases and the majority of components need to reliably operate up to that limit [51,58].

Discard. A discard task is similar to scheduled restorations, as items will also be removed from service at a specified life limit. However, after the item's removal, systems will not be reworked but replaced with new parts. These maintenance tasks are usually applied to systems with high operational criticality and comparably low replacement costs [51,58].

Run-to-failure. Outside of the scope of preventive maintenance, it can be the deliberate decision to operate a system up to the point of failure. Prerequisite of this decision is that a system is not critical for safe operations and its economic impact on the occurrence of a malfunction is acceptable. Especially for systems with failure conditions that are evident to the flight crew, this option is often advantageous over regular preventive maintenance tasks [51,58].

3. Automating preventive scheduled aircraft maintenance

In the following section, we will focus on the aviation industry's efforts to replace the previously discussed, manual maintenance tasks by automation. Therefore, it is important to gain a general understanding and characterization of task automation levels, the definition of CBM, and the regulatory implications and challenges when automating scheduled manual preventive aircraft maintenance.

3.1. Characteristics of task assistance & task automation

First, before we can discuss the characteristics of CBM, we need to examine and differentiate the different levels of automation. Throughout the years, a substantial number of (conceptual) CM solutions have been developed that range widely in their level of task autonomy, e.g., from purely visual augmentation with drones for GVIs [59] to advanced Machine Learning (ML) algorithms for automated fault diagnosis and failure prognosis [60]. Consequently, as the associated requirements for their development vary, it is important to understand their autonomy levels and to identify legal responsibilities for airworthiness decisions.

In general, automation can be defined as the automatically controlled operation of processes or systems by mechanical or electronic devices [61]. Here, it has to be noted that an automated system cannot have one overall level of automation as these levels always refer to a specific function being supported. Consequently, since an automated system can support more than one function, it can potentially result in different levels of automation [62].

In one of the earliest work on categorizing levels of automation, Sheridan and Verplank [63] have proposed a 10-step scale to differentiate between low automation levels with manually performed tasks and full automation with fully autonomous computers. While this scale offered a first categorization of autonomy levels, it neglected different dimensions of the automation process. Consequently, since the degree of automation typically varies in terms of desired autonomy and functional capabilities [64], the approach has been extended by incorporating the following dimensions to represent the individual functions for a task and their corresponding autonomy levels [65,66].

Information acquisition. First, all aspects related to the sensing and registration of input data need to be addressed. Therefore, this category is equivalent to the human information processing stage and intended to support human sensory processes. At the lowest level, the automation may consist of strategies for mechanically moving sensors in order to scan and observe. Furthermore, moderate levels of automation may involve the organization and highlighting of incoming information in accordance to predefined criteria while preserving the original information. In contrast, for more complex operation and high levels of automation, data may be filtered so that certain items of information are exclusively selected and brought to the operator's attention – potentially leading to differing human performance consequences [66].

Information analysis. The automation of information analysis involves cognitive functions, e.g., data processing for generating insights. At low levels, algorithms can be applied to incoming data for their extrapolation over time, e.g., trend analysis and event projection. A higher level of automation would involve the integration of several input variables into a single value in order to generate insights. In a highly automated level, information will be processed autonomously and only summaries of data will be provided to the user [66].

Decision and action selection. The third dimension covers the selection from decision alternatives. Here, automation involves varying levels of augmentation or replacement of human selection of decision options with machine decision-making. The different levels of automation at this stage progress from systems that merely recommend courses of action to those that execute these actions – with or without the possibility of intervention [63,66].

Action implementation. The final dimension refers to the actual execution of the action choice. The corresponding levels of automation are defined as manual-to-automated-activity ratio for the execution of responses. Action automation includes capabilities to track user interaction with the computer system and allow automatic initiation of certain sub-tasks in a contextually-appropriate manner [66].

Lastly and most recently, based on these insights, the European Union Aviation Safety Agency (EASA) [67] has published their categorization of automation levels with specific focus on AI/ML applications:

- Human augmentation (Level 1A) with automation support for data acquisition and analysis and full end user authority,
- Human assistance (Level 1B) with automation support for decision-making and full end user authority,
- Human-AI cooperation (Level 2A) with directed decision and automatic action implementation and full end user authority,
- Human-AI collaboration (Level 2B) with supervised decision and action implementation and partial end user authority,
- Safeguarded advanced automation (Level 3A) with the end user's capability to oversee and override the operations of the AI-based system (for selected decisions and actions) upon alerting, and
- Non-supervised advanced automation (Level 3B) without any end user's involvement into the decision and action implementation and no capability to override the AI-based system's operations.

3.2. Condition based maintenance in aviation

Taking these different levels of task automation into account, we now want to present and discuss the different aspect of CBM in aviation. Starting with a definition and characterization of a CBM approach, we subsequently review what practices of CM already exist. This section will close with some examples of industry use case for an automated CBM approach.

3.2.1. Definition & characteristics of condition based maintenance

In a very basic definition, CBM is described as preventive maintenance that includes a combination of condition monitoring, inspection, testing, and analysis – to determine subsequent maintenance actions [68].

Arguably, one of the more comprehensive publications on the subject of CBM has been developed by the Department of Defense [69]. In their document, the authors refer to CBM as CBM$^+$ to explicitly incorporate its predictive capabilities. They define CBM as maintenance that (a) is only performed when needed, (b) is based on observations in alignment with RCM analysis and other integrated technologies, and (c) aims to improve a system's reliability and maintenance effectiveness. Therefore, it uses a systems engineering approach to collect data, enable analyses, and support decision-making processes [69].

Consequently, CBM and RCM are closely interlinked. While RCM analysis helps to identify failure modes and their criticality to define appropriate maintenance actions (see Section 2.2.2), CBM approaches build on these insights to improve the system's safety, reliability, and affordability by incorporating automating technologies [69,70], such as

- Sensing and data acquisition hardware,
- Signal processing and transmission interfaces,
- Equipment condition and health assessment methods,
- Failure prognostics algorithms, and
- Decision support systems and human system interfaces.

Thus, it is essential to extend the maintenance perspective beyond the asset itself and raise awareness of imminent failures for system operators and maintenance support teams [69].

In the context of CBM, a system's End-to-End (E2E) capability is often emphasized; for example, the certification of the HUMS demands consideration of the complete process – from data acquisition to the ultimate intervention action [42]. Although there is no universally agreed upon definition of E2E capability, we found the explanation provided by the Federal Aviation Administration (FAA) [42] in the context of the HUMS development fitting for our purpose. They define it as a process starting with the airborne data acquisition and ending once a meaningful result for an identified failure mode without the need of further processing has been obtained. Since this E2E process heavily relies on the utilization of automation technologies that often incorporate a variety of Commercial Off-The-Shelf (COTS) hardware and software [70], their application needs to be certified to ensure hazard-free operations for the aircraft (see Section 6). In the context of our work, COTS can be defined as equipment that has originally not been qualified in adherence to aircraft standards, e.g., consumer computer devices and standard operating software [42].

3.2.2. Existing approaches for automating condition based maintenance

Even though the aviation industry is continuously shifting towards tailored CBM approaches [71], the fundamental philosophical principles of current MP definition are still based on decades-old system architectures and capabilities. Therefore, regulatory authorities and industry standard committees, such as the ISO and SAE, have published guidance material on how to incorporate HM technologies into the MP development. A good overview of their respective publications is provided by Vogl et al. [16] (for ISO norms) and SAE [72] (for relevant SAE publications). In the following, we want to examine a selection of relevant Advisory Circular (AC) and AMC documents as well as Aerospace Recommended Practices (ARPs) in a bit more detail.

FAA AC 29-2C MG.15 & EASA AMC 29.1465. As part of the design certification specifications for transport category rotorcraft, the FAA [42] and the EASA [18] provide guidance on airworthiness approval of an HUMS systems.[1] They each present extensive guidelines of different development aspects that will need to be addressed when designing an effective HUMS system. Furthermore, a definition of a flow chart for the installation and validation of HUMS – including necessary interfaces to a ground-based system with COTS hardware and software – is given. While these guidelines are extensive in their support for the HUMS development, they are rather restrictive for monitoring safety-critical functionalities. For example, the certification of HM capabilities for potentially catastrophic system functions is explicitly excluded (see *O1.3* in Section 5 for an explanation of the failure criticality levels). Based on this limitation, SAE [73] provides an overview of exemplary use cases for HUMS data to extend fixed Times Between Overhauls (TBOs) for rotorcraft power train transmissions [18,42].

FAA AC 43-218. More recently, the FAA [17] has published this guidance material for the operational authorization of IAHM systems. Here, a variety of aspects are listed that would need to be addressed for a certified-for-credit[2] AHM system, e.g., minimum requirements for safe data transmission methods, minimum performance standards for ground-based equipment, or necessary qualification standards for ground personnel. However, despite listing these categories, the FAA refers to Design Approval Holders (DAHs) for provision of the necessary information without further specification on how to achieve compliance with the regulatory intent. In consequence, there still is a lack of specific aspects a DAH has to demonstrate to show effectiveness of a developed AHM solution.

IMRBPB issue papers 180 & 197. The International MRB Policy Board (IMRBPB) presented in their Issue Papers (IPs) 180 [75] and 197 [76] a potential approach to integrate AHM technologies into the current MSG-3 workflow. The basic idea is to extend the analysis to allow identification of AHM candidates as alternative to conventional scheduled maintenance tasks. While the presented logic diagram addresses the aspects of lead times for issued alerts and the effectiveness of AHM systems, it explicitly excludes the aspects of qualification of 'on-board' and 'on-ground' segments in terms of hardware and software [75]. Furthermore, the amended logic merely requires AHM alternatives to legacy scheduled maintenance to be effective without providing any specifics on how to determine this effectiveness, e.g., through appropriate AHM performance requirements.

SAE ARPs 5987 & 7122. In addition to these regulatory efforts, SAE [74] has published guidelines for the use of E2E HM systems as Alternative Means of Compliance (AMOC) for scheduled aircraft engines maintenance in their ARP 5987. A more comprehensive framework that will be applicable for the entire aircraft is currently under revision for publication as ARP 7122 [77]. In these documents, SAE provides a ten-step process to demonstrate compliance with certification by automating traditional interval-based maintenance approaches. Starting with the determination of failure criticality, they emphasizes the definition of a suitable AHM system design, its implications for conventional scheduled maintenance tasks, and the cost-effectiveness of the task automation. Finally, they propose a feedback loop in case of future technological improvements of the AHM system. However, while they include a process step for benchmarking the automated system's performance against conventional human inspection performance and provide insights through two examples, these SAE guidelines do not elaborate on how an AHM performance shall actually be determined (see Section 8.3 for a selection of possible HM performance parameters).

In addition to these guideline documents, there have been some industry initiatives to achieve certification or regulatory approval through Supplemental Type Certificates (STCs) or Service Bulletins

[1] The EASA uses in their document the terminology of VHM system but describes essentially identical system functionalities.

[2] So called maintenance credits testify regulatory approval for intended intervention with current scheduled preventive maintenance tasks by changing their scope, interval, or extent [42,74].

(SBs) – predominantly for SHM applications. Examples for that are the integration of SHM into (a) the ICA of WiFi antennas, (b) a SB for the Boeing 737 Wing Center Section Shear Fittings, and (c) an AMC for application on the Boeing 737 Aft Pressure Bulkhead [25,78–80]. Although the gained insights are valuable for the involved parties, the technical details remain undisclosed and cannot serve as a general industry-wide blueprint for future HM applications. However, the lessons learned of these pilot use cases are currently compiled with the intention of publication [81].

3.3. Challenges for regulatory-compliant condition based maintenance process

After we have reviewed the existing CBM maintenance approaches and their shortcomings, we want to focus in the following on listing requirements to comply with regulation for the certification of automated CBM approaches. While there is only a few number of publications on these requirements, regulatory challenges for the application field of Certification by Analysis (CBA), i.e., simulation-based, fully-digital aircraft certification processes, have been discussed already. Since the challenges are comparable to the development of certifiable CBM solutions, we will base the list of requirements on those CBA-related observations.

The main areas of concern for a successful CBA are, among others, the appropriate consideration of errors and uncertainties, and the verification and validation of any developed methods [82,83]. Consequently, key challenges that need to be addressed during a CBA process are (a) the development and application of comprehensive verification tools, (b) the availability of validation data for full scale applications, and (c) the quantification and efficient management of uncertainties [84].

This is in line with the FAA [42] who highlighted in their certification guidelines for HUMS that the application of a CM system will need to be validated, if it is intended as an alternative to a conventional scheduled maintenance task. Therefore, compliance with that intent requires (a) description of the application, (b) thorough understanding of the underlying physics, (c) comprehensive definitions of suitable validation methodologies, and (d) plans for a controlled introduction into service (e.g., as discussed in *O2.2* in Section 6). Additionally, developed HM capabilities need to be incorporated in the corresponding ICA.

3.3.1. Aircraft health management verification & validation

As seen above, an essential prerequisite for demonstrating regulatory compliance is the Verification & Validation (V&V) of the developed AHM solution [18,83]. As emphasized by SAE [85], the identification of verification and validation steps is essential for developing IVHM requirements. This requirement has also been recognized by the EASA [18, AMC 29.1465] as they require developers to demonstrate that VHM systems provide acceptable fault detection performances.

Since the exact meaning of the terminologies *verification* and *validation* is necessary to understand their individual importance, we will define and characterize them first.

Verification. In general, verification examines if a technical (AHM) solution is being developed correctly [82,85,86], i.e., if a model represents the underlying mathematics and physics sufficiently accurate [87–89]. In other words, verification processes assure that a system functions according to the defined requirements [85]. IEEE [90] extends the scope of the verification process to also ensure compliance with (a) requirements for correctness, completeness, and consistency and (b) standards, practices, and established conventions during each life cycle phase.

Validation. The process of validation addresses the questions if the right technical (AHM) solution is being built [82,86]. It examines in particular the degree to which a model is an accurate representation of the real world – with the boundaries of the intended use [82,87,89]. According to the IEEE [90], this process provides supporting evidence that the solution satisfies its allocated system requirements, e.g., from the Concept of Operations (ConOps) document (see Section 5), while solving the right problem. Therefore, validation determines if defined requirements meet the stakeholders' needs [85].

The verification process is typically executed bottom-up from a unit level up to component and system levels. It involves the use of analyses, simulations, models, or other means to test products and demonstrate their correct implementation. This demonstration also enables the identification of important areas and weak spots. The testing process should be based on representative test scenarios, e.g., experimental fault injections representing actual faults encountered during operation or determined relevant as part of an FMEA. However, since it is often very challenging to characterize relevant but hard-to-observe degradation mechanisms at the time of requirements definition and system design, the fidelity, granularity, and performance of IVHM systems strongly correlates with the designer's understanding of the fault modes [85].

Once the component and model verification is complete, the models are then used for validation through additional tests, analyses, inspections, or demonstrations that ensure the compliance of developed technical solutions with functional requirements from the ConOps document. This includes the deployment and testing in realistic environments under various scenarios. For each application, it needs to be proven that the involved physics is understood and the monitoring technique, rejection criteria and associated intervention actions have been chosen appropriately [42]. Furthermore, a thoroughly validated model will be essential for establishing a certifiable automated condition monitoring framework and obtaining maintenance credits [42]. However, since a thorough experimental validation of IVHM functions can be impractical to impossible, the process often relies heavily on simulation-based testing. Other proposed approaches for IVHM validation take an incremental approach to build trust by introducing IVHM solutions in a controlled loop, e.g., as presented by Piotrowski [23] for a SHM solution.

Examples for the associated technical challenges in the V&V process are

- Extending maintenance intervals without the need to wait for the duration of the new interval or
- Monitoring for a fault through IVHM with a certain detection and false alarm rate without seeding a safety-critical monitored system with actual faults.

Therefore, it is typically assumed if a model could adequately predict some related instances of the intended use with available experimental data, it will also be validated for predictions beyond the experimental data for the intended use. Consequently, the correct V&V of any developed AHM solution is of paramount importance [18,85,90].

3.3.2. Uncertainty management

Uncertainty is a foundational concept in engineering and decision-making, yet it remains inconsistently defined across various disciplines [91]. Consequently, despite its widely acknowledged relevance – particularly in contexts where systems rely on data, models, or measurements – there is neither a universally accepted taxonomy for uncertainty description nor uncertainty management methods. However, insights from established frameworks in other domains can offer valuable guidance for a systematic approach. For example, the Guide to the Expression of Uncertainty in Measurement (GUM) [92] outlines a structured approach of Uncertainty Quantification (UQ) in measurement processes. Most importantly, it distinguishes the type of uncertainty into those that (a) arise from statistical variation and (b)

are based on incomplete knowledge. Especially for cases where measurements serve as the foundation for further processing or inference, this distinction enables a comprehensive expression of confidence in sensor readings.

Additionally, in one widely cited approach, Walker et al. [93] organize uncertainty along the following three dimensions.

Location. This dimension focuses on the origin and manifestation of uncertainties in the system. In the context of CBM, sources of uncertainty range across physical and computational layers. At the physical interface, noise, faulty calibration, degradation, or a limited resolution may result in sensor input uncertainties that can potentially propagate downstream – with sensor data typically being the system's primary observation of reality. Additionally, latency, packet loss, or synchronization issues during data transmission may compromise the temporal accuracy of incoming signals. Once received, data is often subject to filtering, rounding, or aggregation, each of which can introduce subtle but meaningful distortions. Computational models may add another layer of uncertainty, e.g., when they rely on fixed assumptions such as constant environmental conditions or static baselines that rarely hold in practice. Furthermore, maintenance action triggers – especially when based on rigid thresholds or heuristics – may insufficiently reflect the variability of real-world conditions and lead to increased false alarm or missed detection rates. Finally, as diagnostic algorithms and ML models become more prevalent in CBM, their inherent uncertainties (e.g., stemming from limited training data or generalized models) must also be considered. It has to be noted, these locations of uncertainty are not isolated as they can interact and compound if initial measurement imprecision cascades through assumptions, processing, and decision logic.

Level. This dimension refers to the degree or magnitude of uncertainty. In order to express the spread or reliability of a given estimate, uncertainty is often quantified statistically – using indicators such as variance, standard deviation, or confidence intervals. However, since meaningful UQ typically relies on statistical assumptions such as normally distributed errors or sufficiently large sample sizes, it can be underpinned by principles like the central limit theorem and the law of large numbers [91]. In the context of CBM in aviation, such assumptions are frequently challenged by sparse, variable, or non-representative data, particularly due to differences in mission profiles or limited failure occurrences. As a consequence, assessing the level of uncertainty becomes difficult and prone to misinterpretation. An understated level may lead to high but unjustified trust in automated decisions, while overstated levels can result in unnecessary caution or inefficiencies – both undermining confidence in a system's output. Because people's discomfort with uncertainty affects their perception of a technology's usefulness and influences its acceptance, understanding and communicating inherent uncertainties is essential for successful technology introduction [94,95].

Nature. This dimension characterizes the type of uncertainty present. A common distinction is made between (a) *epistemic uncertainty* that stems from incomplete knowledge and is potentially reducible and (b) *aleatory uncertainty* that arises from inherent variability and is considered irreducible. However, this dichotomy may be insufficient for practical use in CBM. As Sankararaman [96] points out, the classification of uncertainty is not always clear-cut – especially in prognostic contexts where boundaries between the incomplete understanding of degradation mechanisms and inherent variability can blur. Therefore, a shift towards interpretation of uncertainty in terms of its effects on predictions and decisions may be more useful than strictly categorizing its origin. For example, while a change in measured vibration levels could result from environmental fluctuations or unmodeled system behavior, the key question remains whether and how the uncertainty can be reduced, bounded, or otherwise incorporated into robust decision-making.

Complementing this distinction, Oberkampf and Roy [97] have developed a framework that focuses on uncertainty in computational modeling – with direct relevance to algorithm-driven systems. Their work discusses the aforementioned V&V, and presents methods for assessing model credibility through statistical comparisons with experimental data. Moreover, it addresses different approaches to UQ in order to offer tools for updating beliefs, analyze parameter sensitivity, and deal with non-probabilistic forms of uncertainty. These considerations are of high relevance when models are used not just to analyze but also to support decisions under uncertainty – a point emphasized in their treatment of risk-informed design.

In a broader organizational and operational context, the ISO 31000 standard [98] provides a high-level framework for managing risk and uncertainty systematically. It emphasizes principles such as transparency, integration into decision processes, and continual improvement – making it especially relevant for institutions aiming to align technical uncertainty management with enterprise-level governance and accountability.

In addition to these foundational frameworks, several academic contributions have addressed the treatment of uncertainty in the context of condition-based maintenance. These range from discussions on challenges of predicting Remaining Useful Lifetimes (RULs) under uncertainty [96], to examinations of broader difficulties to consistently manage uncertainty in limited data setting [39], to recommendations on the inclusion of confidence levels in model outputs for safety-relevant decision making [12]. Additionally, Thacker [99] and the EASA [100] provide structured guidance for dealing with uncertainty during simulation-based certification. Their respective approaches emphasize the need to define sources of uncertainty explicitly, evaluate predictive accuracy, and establish traceability across validation hierarchies.

While all these frameworks provide foundational perspectives that help anchoring the complex analysis and management of uncertainty in practice, applying these principles to automated CBM remains a significant challenge [39]. The layered and interconnected nature of technical, algorithmic, and operational uncertainties in such systems makes straightforward classification or mitigation difficult. Data scarcity – particularly of failure data – limits model accuracy and validation [101]. Combined with implicit assumptions about the quality of the monitoring system, it becomes clear that recognition of the complexity of uncertainty management is essential for building CBM systems that remain reliable under real-world conditions.

4. Baseline model for a certifiable end-to-end condition based maintenance process

With the key characteristics of an E2E CBM system defined and the essential challenges for regulatory compliance identified, we now want to develop a corresponding process that complies with the regulatory intent and supports in the development of automated HM systems. Based on prior publications [70,85] and through iterations with industrial partners, we have defined and categorized a set of development objectives in the logic of our baseline model. We will further highlight them in the respective sections and figures – in order to guide readers throughout the rest of this document. The listed objectives and their respective contents should be followed to justify issuance of maintenance credits for developed HM solutions.

System design (see Section 5)
- O1.1 Identify the intended AHM application
- O1.2 Identify failure modes of the monitored system
- O1.3 Determine the failure criticality

Data acquisition (see Section 6)
- O2.1 Define properties and key characteristics of the sensor network
- O2.2 Qualify sensors according to the required Assurance Level (AL)
- O2.3 Assess uncertainties & their causes

Data processing & transfer (see Section 7)
- O3.1 Identify data quantities & define off-aircraft transmission frequencies
- O3.2 Validate the data processing software according to the required AL
- O3.3 Determine appropriate cybersecurity requirements

Data analysis (see Section 8)
- O4.1 Define suitable alert & alarm levels for the system condition
- O4.2 Define an appropriate human-machine interface
- O4.3 Demonstrate sufficient fault diagnosis & failure prognosis performance

Advisory generation (see Section 9)
- O5.1 Define displayed information for maintenance decision making
- O5.2 Determine necessary skills & training for maintenance personnel

As stated in Section 3.2.1, an essential ingredient of CBM is a system's E2E capability. One of the more prominent models to ensure this capacity for CBM applications is the OSA CBM [70]. Alternative but similar E2E workflows have been developed by SAE [102] with their Sense, Acquire, Transfer, Analyze, Act (SATAA) framework and IEEE [104] in their standard with the five-step approach of Sense, Acquire, Analyze, Advice, Act. While the OSA CBM model has originally focused on the development and demonstration of a software architecture that facilitates interoperability of CBM modules, it has become a blueprint for E2E automated condition monitoring and diagnostics of machines (see also SAE ARP 5987 in Section 3.2.2) [103,105]. Essentially, the workflow consists of the following steps as depicted in Fig. 2.

AHM system design. Although the design development process is no designated part of the OSA CBM model, an effective AHM requires to obtain a sound technical understanding of the system, its characteristics, and possible failure modes [106]. Therefore, this step would translate to the exploratory/concept stage of the system life cycle process [107]. Consequently, a range of existing guidelines [18,85,108,109] emphasizes to preface the development of CM capabilities by performing system design assessments, e.g., a Functional Hazard Assessment (FHA) or a System Safety Assessment (SSA) to determine the AHM's criticality and the required AL [42].

Data acquisition. The first 'true' step of the original OSA CBM process is the acquisition of digitized sensor data by converting a transducer output to a digital parameter. Consequently, the data represents a physical quantity and potentially related information, such as time stamps, calibration, or sensor configuration. As a result, the data acquisition module is essentially a server of calibrated digitized sensor data records [103,110,111].

Data processing. In a next step, the acquired signals and data from the sensor module are collected and processed to extract features (e.g., through signal analysis, computation of meaningful descriptors, or derivation of virtual sensor readings from raw measurements). Subsequently, the output of this step includes digitally filtered sensor data, frequency spectra, virtual sensor signals and other CBM features [103,110,111].

State detection/condition monitoring. Based on the acquired input data from sensors, signal processors, and other condition monitoring devices, this step will (a) compare the extracted features against expected values or operational limits and (b) output condition indicators to detect system abnormalities. With the resulting findings and preset operational limits, alerts or alarms of imminent system deficiencies will be issued – if necessary [70,103,110,111].

Health assessment/diagnosis. Next, the health status of a monitored system, subsystem, or component is determined, using the condition monitoring input data. Therefore, this step considers trends in the health history, operational status, and maintenance records to derive degradation analyses. Finally, the degradation progression is examined to derive possible evolving fault conditions with associated confidence levels [70,103,110,111].

Prognostic assessment. After the current degradation status of the system has been obtained, the primary function of the prognostic assessment is to project the anticipated future health state. In addition to that, the prognostics layer provides RUL information, i.e., the expected operable time until a functional breakdown occurrence – given the observed usage profile [70,103,110,111].

Advisory generation. Lastly, based on the degradation diagnosis and prognostic assessment, recommended maintenance and operational actions need to be provided. Furthermore, these shall indicate the respective implications for each action towards the intended mission objectives or an optimized system life cycle. Consequently, the advisory generation layer needs to incorporate information about the operational history, current and future mission profiles, high-level unit objectives, and resource constraints. In order to support the decision making process, the developed advisories – together with supplemental data – and alerts need to be displayed to qualified personnel. That explicitly includes the ability to access source data when anomalies are reported to identify, confirm, or understand their nature [70,103,110,111].

5. Development of a condition monitoring system design

The overarching goal for any kind of system automation development is to ensure trustworthiness through design assurance, which is integral to be approved for safety-related applications. Therefore, the corresponding analysis starts with a characterization of the application, followed by safety and security assessments as key elements of the trustworthiness analysis concept[3] [112]. The trustworthiness analysis process further considers the following aspects.

Assurance concept. The first step addresses the need for explainability of automatically generated result; that is, the capability to provide human users with understandable, reliable, and relevant information in the appropriate level of details on how results have been produced. Furthermore, it includes the data recording capabilities for continuous monitoring of the automated system's safety and for supporting incident or accident investigations [112].

Human factors. Next, guidance to account for specific human needs linked with the application of automation technologies is introduced. Among other aspects, this includes the definition of appropriate information levels for end users to reliably conclude necessary actions based on provided insights. In addition, the concept of automation level categorization (see Section 3.1) is addressed to ensure adequate cooperation or collaboration between human end users and (AI-based) automation systems [112].

Safety risk mitigation. The last step focuses on residual risks that may need to be addressed due to the inherent uncertainty of (AI-based) automation technologies. For example, it may inherently be impossible to avoid AI black box solutions; therefore, the risks associated with the uncertainty will need to be minimized to an acceptable level for the intended task [112].

[3] Although these steps have been developed specifically for AI applications, the aspects of this approach can be generalized for any kind of automation process that involves data processing.

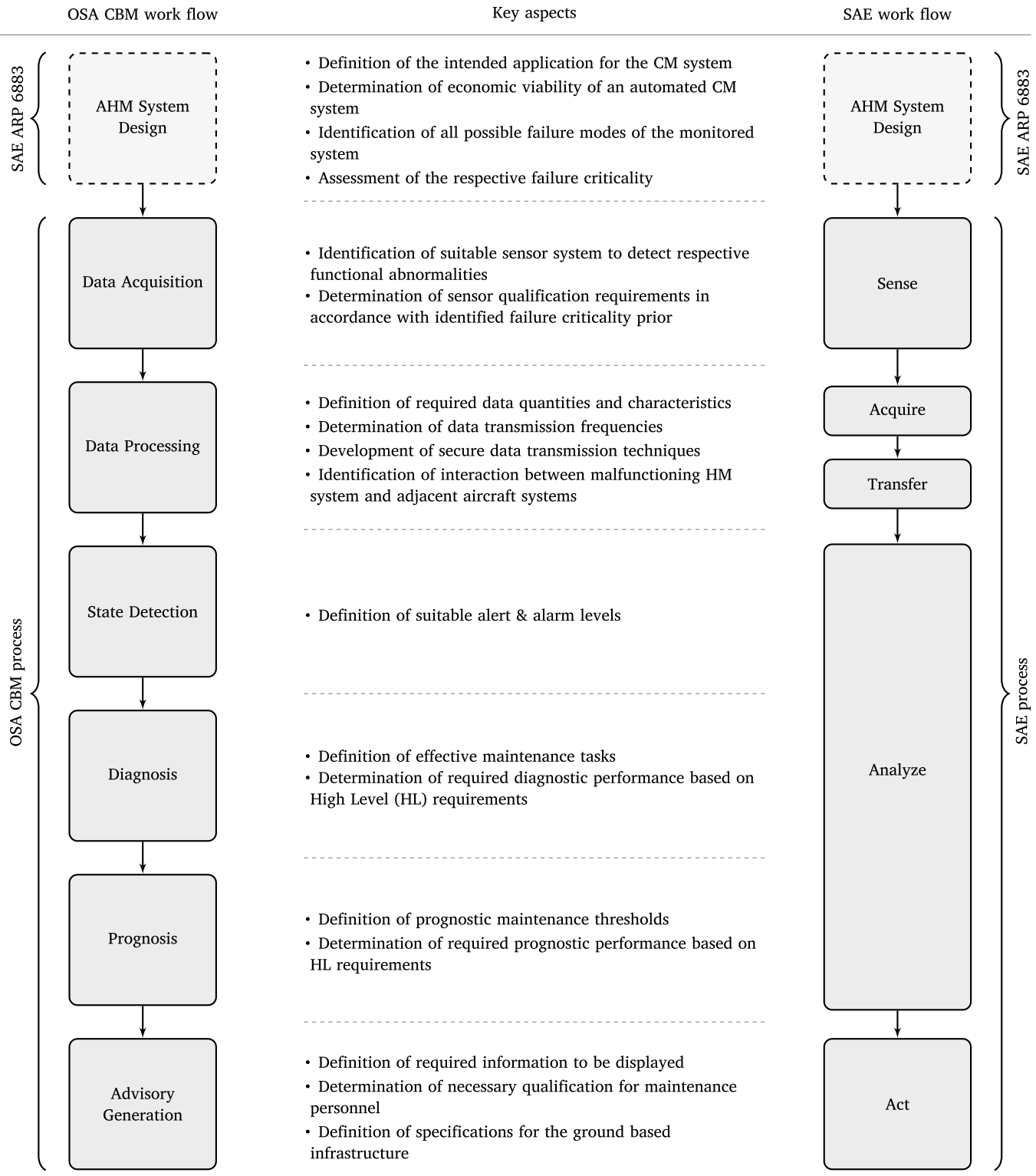

Fig. 2. End-to-End CBM processes with their respective inputs [70,85,102,103].

Additionally, especially for AI applications, the EASA [112] emphasizes to perform an ethics-based assessment to align with the AI ethical guidelines developed by the European Commission [113].

While a reliable system – performing as designed in an operational environment over time – is a primary focus during the system design and architecture development process, the underlying analyses consider trade-offs between system performance and life-cycle cost to maximize the technical effectiveness and affordability [69]. Building on that and the essential steps mentioned prior, we propose a process work flow as shown in Fig. 3.

O1.1: Identify the intended AHM application

In order to enable any kind of deeper technical analysis of functional failure mechanisms and to derive concrete specifications for prognostics requirements, it is important to gain a thorough understanding of the

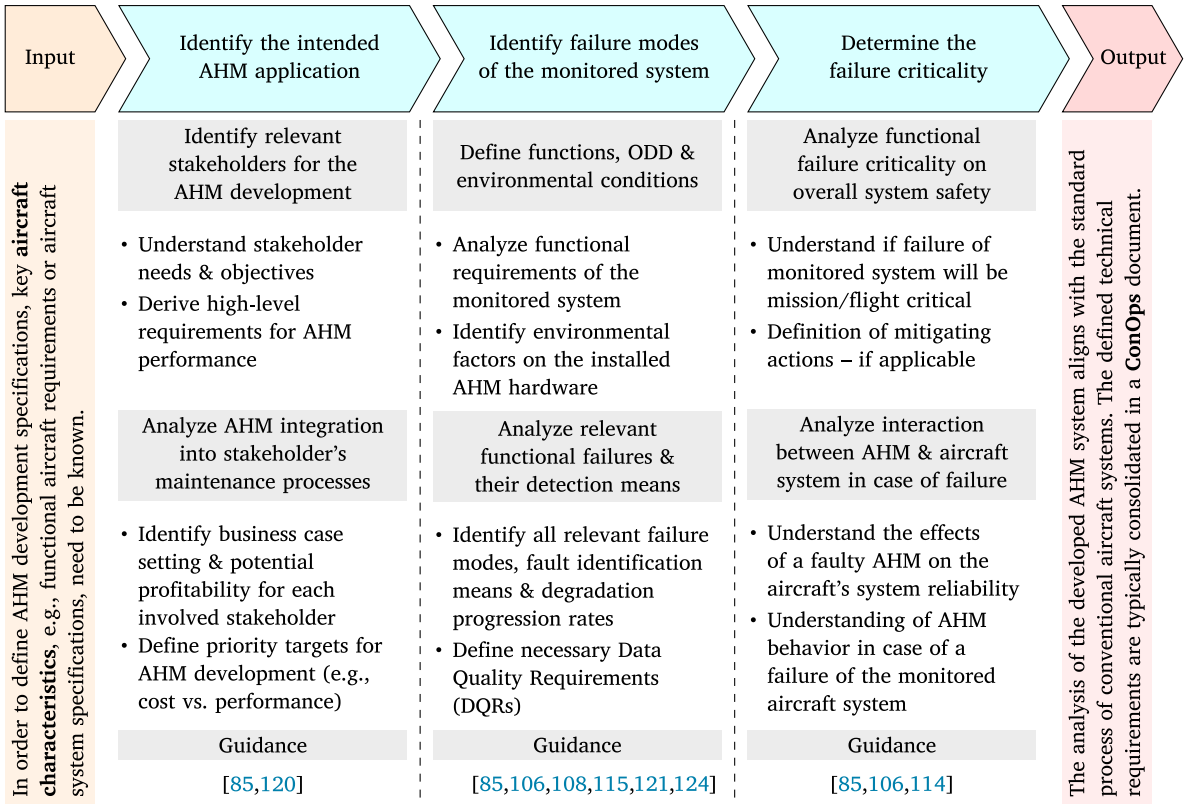

Fig. 3. AHM system design process step.

AHM's system requirements and the operating concept, such as the intended system functions and expected environmental conditions [114,115]. Especially for AI and ML applications, it is imperative to define the so called Operational Design Domain (ODD) to identify edge cases [116], i.e., operating conditions under which a given automated system is specifically designed to function – including environmental, geographical, and time-of-day restrictions [117,118].

As of this step, the objectives are (a) to identify all relevant aspects that could potentially influence an AHM's effectiveness [108], (b) to define the scope of AHM functions for satisfying relevant stakeholder requirements [85], and (c) to derive top-level functional requirements based on the defined goals [115]. It has to be noted that typical process flows start with an economic assessment of the AHM technology's potential to ensure the validity of underlying business cases [108]. However, we focus for this work exclusively on safety-relevant factors. Thus, readers who are interested in the economic evaluation of HM technologies are kindly referred to SAE ARP 6275 [119] or Meissner et al. [7]. The results of this stage are typically summarized in a ConOps document that [67,85,120]

- Describes the characteristics of an AHM system from a user's point of view,
- Contains quantitative as well as qualitative system characteristics for all involved stakeholders, and
- Defines the ODD with specific operational limitations and assumptions.

O1.2: Identify failure modes of the monitored system

As pointed out by SAE [85], based on the ConOps document information, any AHM development shall be prefaced by the definition of the scope for AHM functionalities to (a) satisfy all relevant stakeholder requirements, (b) focus on prioritized fault modes, and (c) limit development efforts. Complex designs can consist of numerous subsystems and components with their own fault modes that are infeasible or impractical to be tracked by AHM solutions. Therefore, the focus should be placed on failures with the highest criticality from the overall system's perspective, especially with respect to the organizational integration of AHM capabilities and the associated budget restrictions [85].

Furthermore, it needs to be ensured that all limiting failure modes and the corresponding failure progression rates with their driving parameters are sufficiently understood [73]. A practical implementation of this recommendation can be performed through a FHA [42,115], where all functions are listed and evaluated with respect to impact factors that can influence the vehicle's safety [115]. In addition to an in-depth analysis of potential failure mechanisms and their respective means of detection, this step includes a breakdown of the top-level system-specific requirements into AHM-specific requirements – as shown in Fig. 4. Since the predominant objective here is to ensure the vehicle's health is maintained, these requirements can include the necessary AHM coverage, required diagnostic and prognostic performance (see Section 8.3), and applicable mitigation actions [115,120]. Especially for AI and ML applications, Data Quality Requirements (DQRs) need to be defined in accordance with the functional failure criticality and should focus on [67,121]

- Data relevance to support the intended use,
- Format, accuracy and resolution of the data,
- Traceability of data from their origin through the whole pipeline of operations,
- Mechanisms ensuring that information will not be corrupted while stored, processed, or transmitted over a communication network,
- Completeness and representativeness of the data sets, and
- Level of independence between the training, validation and test data sets.

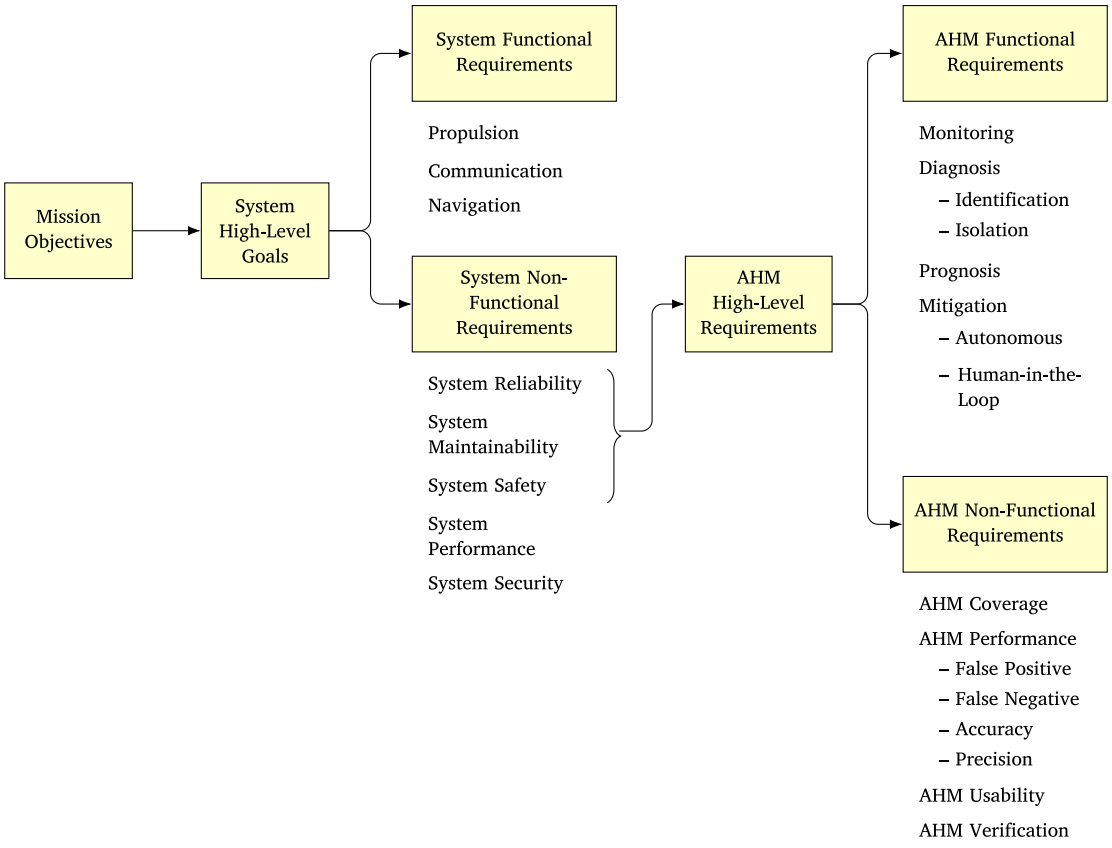

Fig. 4. Identification of AHM requirements [120].

O1.3: Determine the failure criticality

To determine the required integrity level of the AHM system, the identified functional failures of the monitored system need to be evaluated towards their criticality, e.g., through an FMEA [46,108] or Fault Tree Analysis (FTA) [69,115]. In the context of our analysis, criticality describes the severity of an AHM application failure or malfunction [42]. Based on the obtained FHA insights from the prior step, system level functions (e.g., aviate, navigate, communicate, mitigate hazard, etc.) are decomposed into lower level basic functions [115]. Subsequently, possible failures that are associated with these basic functions are identified and categorized into severity levels [115]. An example developed by Tobon-Mejia et al. [122] and Vogl et al. [123] of such a FMEA is shown in Table 4. One possible approach is to categorize the failure criticalities as [43]

(A) Catastrophic, i.e., a failure condition that would result in multiple fatalities – typically with the complete loss of the vehicle,
(B) Hazardous, i.e., a failure condition that would largely reduce the airplane's safety margins and functional capabilities or the crew's task reliability and ability to cope with adverse operating by exposing them to excessive workload, or cause serious to fatal injuries to a small number of occupants,
(C) Major, i.e., a failure condition that would significantly reduce the airplane's safety margins and functional capabilities or increase the crew's workload and limit their ability to cope by impairing their efficiency, or cause physical distress and injuries to occupants,
(D) Minor, i.e., a failure condition that would slightly reduce the aircraft's safety margins and functional capabilities or increase the crew's workload well within their capabilities, or cause physical discomfort to occupants, or
(E) No Safety Effect (NSE), i.e., a failure condition that does not affect the operational capability and safety of the aircraft, or the crew's workload.

Once the AHM system has been fully developed, it needs to demonstrate compliance with these criticality level that have been established through the assessment techniques mentioned before [42,73].

Although not within the scope of our work, it has to be noted that the functional requirements can extend beyond pure safety concerns. An example for these non-regulatory-related demands can be the use of AHM information for an intelligent battery usage management to extend its overall life expectancy by controlling the depth-of-discharge. With these specifications feeding into the requirement tree, the ultimate definition of an AHM system can substantially be influenced [115].

6. Data acquisition infrastructure

After the general requirements for the AHM system have been defined, we now need to determine the necessary sensing infrastructure for the identified failure modes. The corresponding workflow is shown in Fig. 5.

O2.1: Define properties and key characteristics of the sensor network

In general, as emphasized in AMC 29.1465 [18], installed HM systems have to be designed and manufactured in an appropriate way and be reliable in accordance with their intended function. Therefore, the corresponding hardware that measures the system state should provide a reliable signal with an appropriate performance. Furthermore, with the position and installation of a sensor being as critical as its performance, the sensor selection, positioning and installation has to enable the analysis of processed signals to identify failure modes

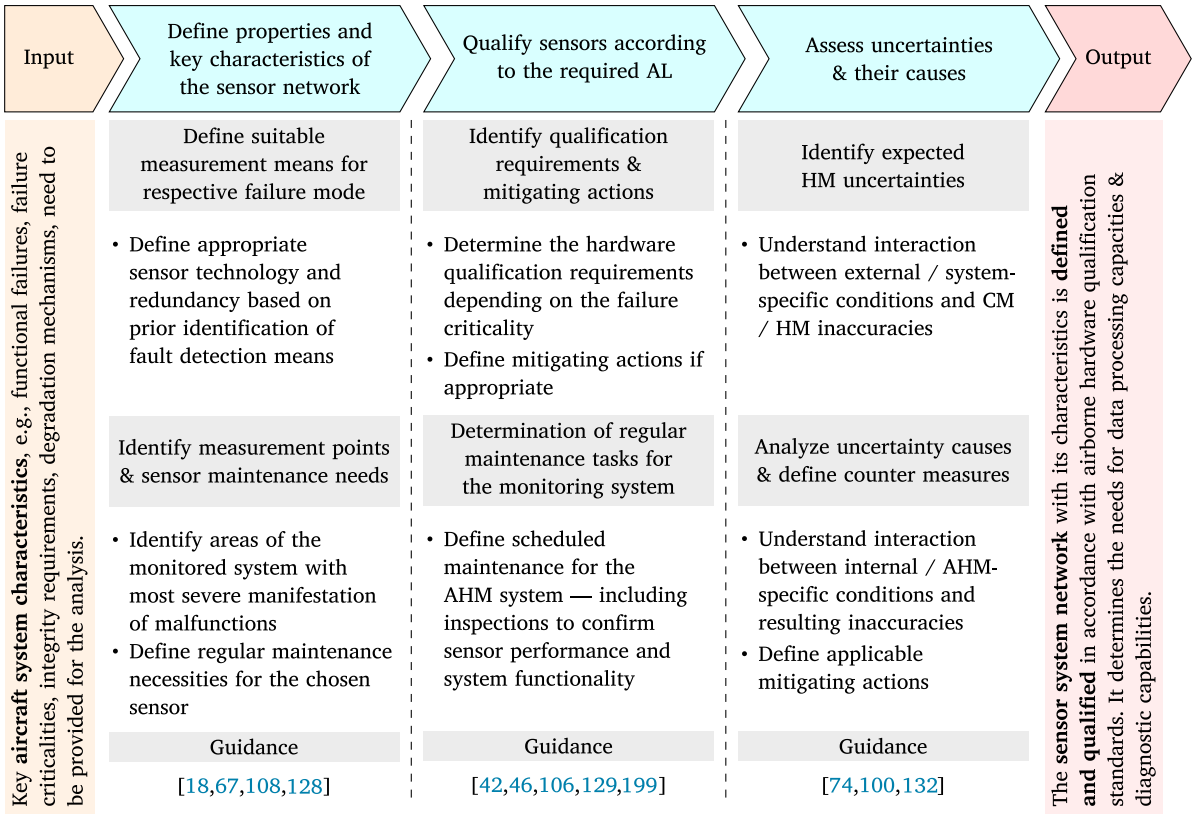

Fig. 5. Data acquisition work flow.

of the monitored component. Since it is also possible that acquired sensor data will be sensitive to the current flight condition, it may be desirable to focus data acquisition – through the Operational Domain (OD) definition – to particular operating conditions or phases of a flight. However, when defining this limitation, the following factors need to be considered [18].

- The signal sampling rate needs to be sufficient for the required bandwidth and to avoid aliasing.
- The data should be automatically gathered in specifically defined regimes, at an appropriate rate and quantity, for the signal processing to produce robust data for defect detection.
- The acquisition cycle should be designed in such a way that all relevant components and their defects are monitored with an adequate sensing frequency – irrespective of interruptions due to different operational profiles.
- The most likely operating scenario and data acquisition adequacy needs to be identified to enable suitable alert and alarm processing (see Section 8.1).

Furthermore, in order to be useful for any AI and ML applications, the acquired data needs to fulfill the following set of requirements [124–126].

Data completeness. In general, data completeness expresses the degree to which data associated with an entity has values for all expected attributes and related instances in a certain context of use [125]. A more specific definition is given by the MLEAP Consortium [124] who define a data set as complete if it sufficiently (in accordance with the predefined DQRs) covers the entire space of the ODD for the intended application. However, one of the major difficulties in assessing completeness of data sets is to obtain reliable information on distributions of phenomena within the ODD [67]. Therefore, the assessment has to be performed on a case-by-case basis with extensive expert work and judgment. The EASA [67, AMOC DM-07-1] proposes multiple techniques to evaluate the completeness of data sets. One such technique is the Principal Component Analysis (PCA), i.e., an approach for data set analysis used to gain visual insight on the completeness. By plotting the (projected) data in lower dimensions and visually checking the distribution's homogeneity within the entire plot, any appearing cluster or empty space might be indicative of some form of lack of completeness. In addition, a sample-wise similarity analysis allows the comparison between data sets, ensuring that the characteristics of data are preserved across different sets. Lastly, Emran [127] also provides in his work definitions and methods for computing data completeness, linking four different types of missing values: (a) null-based missing values with nulls representing missing data, (b) tuple-based missing values with absent attribute-value tuples, (c) schema-based missing values with missing attributes and entities from the schema, and (d) population-based missing values with missing data compared to a reference population.

Data representativeness. Data representativeness is defined as the degree to which a data set represents the population under study [126]. Therefore, a data acquisition system needs to ensure that any measured data is being independently sampled from the possible input space according to its distribution [67, AMOC DM-07-2]. The EASA [67] further provides multiple examples for statistical methods to verify the representativeness of derived or low-dimensional-feature data sets (e.g., Z-, Chi-Square-, or Kolmogorov–Smirnov-Tests).

Data accuracy and correctness. Data accuracy describes the degree to which acquired data has attributes that correctly represent the true value of an event in a specific context of use [125]. Consequently, the EASA highlights different types of errors and biases that need to be identified [67, AMOC DM-07-3], such as

- Errors introduced by sensors,
- Errors introduced by collecting data from a single source,

- Errors introduced by sampling, or
- Errors introduced when performing data cleaning or removal of supposed outliers.

With these influencing factors identified, we finally need to determine appropriate measurement points and suitable sampling rates. In order to ensure steady-state conditions, the data acquisition rate should be high enough to capture a complete set of measurement data before the system's condition changes [108]. Consequently, that may result in the requirement of higher sampling rates for fast-evolving transient conditions. Furthermore, in addition to the actual monitored parameters, it is recommended to record of at least [108]

- Essential information describing the monitored system,
- Operating conditions,
- Measurement position,
- Measured variable, unit and performed processing, and
- Date and time of the measurement.

These information may also be amended by details of the measurement system and its accuracy to support subsequent data analyses [108]. The specific points for the sensor placement should ensure the best possible damage detection and need to incorporate factors such as safety, high sensitivity to changes in degradation, low sensitivity to background noise, repeatability of the measurements, accessibility, and costs [108]. Depending on the rate of failure progression and its criticality towards operating safety, ISO 13373-1 [128] provides the following overview of different sensing system options for the example of a vibration condition monitoring.

Permanently installed sensing systems. In this type of system, the whole measurement unit, consisting of transducers, signal conditioning and data processing/storage units, is permanently installed on the aircraft. Therefore, it is also referred to as on-line system. The measurement itself can be issued either continuously or periodically. Due to their installation costs and additional weight, permanently installed sensing systems are usually only applicable for expensive and safety-critical system failures or aircraft systems with complex monitoring requirements [128].

Semi-permanently installed sensing systems. These sensing systems are a mixture of permanent and mobile systems. With this system architecture, the transducers are usually permanently installed, while the electronic data processing/recording devices are only connected intermittently [128].

Mobile sensing systems. Mobile sensing systems are used to manually record measurement data at predetermined measuring points in periodic time intervals, e.g., weekly or monthly. The measurement data is usually recorded and stored with a mobile data collector. While a preliminary condition assessment can be carried out immediately, measurement data needs to be transferred to a suitable computer with necessary analysis software for detailed data insights [128].

Furthermore, in its Annex A, ISO 13371-1 [128] highlights specific recommendations of measurement installations for different kind of equipment. It has to be noted that the type of sensing system is strongly interconnected with the intended monitoring strategy (see *O3.1* in Section 7).

O2.2: Qualify sensors according to the required AL

In general, for the qualification of the E2E infrastructure, we need to examine the declared intent.

If the development shall obtain maintenance credits, the E2E criticality for such an application has to be determined (see Section 5) and used as an input to establish the equipment's integrity criteria. However, it may be possible to modify the integrity requirement for certification by introducing mitigating actions. Those are autonomous and continuing compensating factors that become part of the certification requirements and have to be performed during normal operations. Two examples of mitigating actions are (a) the pilots' comparison of airborne HM data with aircraft instrument data and (b) the coverage of the failure mode by an existing Flight Warning System (FWS) [42].

In contrast, if the declared intent is for non-credit developments, it may be sufficient to demonstrate that the equipment installation will not result in a hazardous condition to the aircraft by adversely affecting the proper function of any other system or equipment [48, Annex II - 1.3.2].

In order to qualify an E2E CBM process that is supposed to be certified for maintenance credits, we need to ensure that all relevant equipment – airborne and on-ground – is qualified [42]. Consequently, an installation approval must cover systems and equipment that acquire, store, process, and display HM data, including aspects such as the reliable supply of electrical power, human factors for affected operations, and the non-interference of installed HM systems with an otherwise undamaged aircraft system [42,46]. Therefore, for qualification of airborne equipment in accordance with the required AL, the FAA [42] emphasizes to distinguish as follows.

Direct evidence. This approach is prescribed to be performed for a *hazardous* and *major* failure criticality. It has to be noted that the guidelines for the HUMS development [42] specifically exclude the application of automated condition management technologies for catastrophic failures. In order to give consistent alerts, the collection of direct evidence aims to establish that the HM application is sensitive to and obeys predicted response rules for the damage type. It can be gathered from (a) actual service experience on HM-equipped vehicles, (b) 'seeded tests', i.e., tests where a defect or deterioration is introduced, allowed to develop, and the response is verified, and (c) on-aircraft trials, i.e., investigating causes and effects to calibrate the HM response. The FAA emphasizes explicitly the requirement for representativeness of the (on-ground or rig-based seeded) test conditions to reflect the normal flight regime [42].

Indirect evidence. When the functional failure is classified not higher than *minor* or the probability of undetected failures for otherwise on-condition maintenance actions shall be lowered, HM technologies can be qualified through the collection of indirect evidences. Since it may be impracticable to generate direct evidence for each failure mode, the qualification can be achieved through monitoring of a high number of potential failure modes to collectively determine the probability of undetected failures. For that, analytical methods may be combined with sound engineering judgment to provide derived maintenance criteria – supported by validation tests. Furthermore, it may also be suitable to validate model-based damage progression methods through analogy with evidence generated for other aircraft types or equipment [42].

For the HM-related airborne equipment qualification, the FAA [42] prescribes to use the same procedures as for any other conventional airborne equipment. That is, equipment qualification has to consider environmental factors [129], High Intensity Radiated Fields (HIRF) and lightning conditions, and software development standards [130]. Furthermore, there should be signal independence to the extent that acquisition of HM signals should not compromise the level of safety or reliability of functions provided by other equipment as a result of signal sharing [42].

Besides the determination of initial airworthiness through appropriate design standards, the continuing airworthiness of the AHM equipment has to be ensured through ICA activities [42] (see Section 3.3). Since the equipment and tools needed for those regular maintenance activities have to be controlled and calibrated [50, M.A.608(b)], any AHM-related equipment will need to be subjected to regular maintenance tasks as well. Additionally, as emphasized by the EASA [18], Built-In Test (BIT) capabilities may be necessary to determine the correct functioning of sensor infrastructure. Therefore, ICA should incorporate the corresponding maintenance procedures that ensure the correct function of sensors.

Table 2
Uncertainty mitigation based on its source [131,132].

Source of uncertainty	Recommended countermeasures
Sensor noise	Band-pass filtering, spectral analysis, sensor redundancy
Calibration drift	Regular recalibration schedules, self-diagnostic routines
Environmental interference	Shielding, environmental compensation algorithms
Quantization and aliasing	Oversampling, anti-aliasing filters
Data loss during transmission	Buffering, time-stamping, integrity checks
Installation effects	Standardized procedures, interface documentation, alignment verification

O2.3: Assess uncertainties & their causes

To ensure the credibility and safety of sensor-based CM systems, multiple regulatory and industry guidelines stress the importance of understanding, quantifying, and mitigating measurement uncertainties. According to AMC 29.1465 [18], it is essential to identify and classify uncertainty sources, propagate them through the decision-making chain, and design mitigation strategies aligned with system criticality and certification requirements as follows.

Identification of uncertainty sources. Measurement uncertainty arises from hardware and processing components. Typical sources include sensor noise, drift due to environmental influences, digitization errors, loss or corruption during data transmission, and signal filtering artifacts [131]. Furthermore, AMC 29.1465 [18] highlights that additional sources may include installation effects, human errors on interpretation, and insufficient calibration histories. These sources should be systematically analyzed and mapped for each measurement pathway during system design.

Classification of uncertainty types. The classification of uncertainty informs how it can be addressed. SAE [132] recommends a taxonomy based on uncertainty entry points into the system, i.e., input, model structure, numerical approximation, or implementation. Additionally, uncertainty may be described as systematic (bias) or random (variability), each requiring different mitigation approaches [18],. Aleatory and epistemic uncertainties (see Section 3.3.2) may co-exist within the same sensor reading, e.g., a vibration signal can be affected by material tolerance (aleatory) and insufficient environmental modeling (epistemic).

Quantification of input uncertainty. In order to use sensor data for predictive maintenance applications, it is of paramount importance to understand the measurement bounds and confidence levels [74]. For that, AMC 29.1465 [18] offers practical guidance by specifying that total measurement uncertainty should be expressed as the combination of systematic and random components, ideally represented with 95% confidence intervals. For critical measurements, applicants are expected to follow the principles laid out in the GUM [92]. If a full uncertainty budget cannot be developed, qualitative justifications and prior experience may be used to estimate bounds – but only if they are explicitly justified.

Validation of uncertainty models and assumptions. Assumptions made during UQ must be examined – since insufficiently validated assumptions can lead to overconfident or misleading results. SAE [132] and AMC 29.1465 [18] recommend validating underlying assumptions such as distribution shapes, independence between variables, and linearity. Consequently, model developers should use residual analysis, goodness-of-fit testing, and empirical data comparisons to support the assumptions made during UQ and propagation activities.

Uncertainty propagation. Any uncertainty in the input propagates through the CBM pipeline – from state estimation and diagnostics to prognostics and maintenance advisories. Therefore, models must reflect not only average behavior but also the variability induced by uncertain inputs [74]. Consequently, the validation of diagnostic and prognostic models has to be based on accuracy and robustness of these models under uncertain input conditions [132]. This includes the use of Monte Carlo methods or bounding analysis to demonstrate how uncertainty impacts RUL predictions and decision thresholds.

Uncertainty in decision logic. Given that maintenance decisions may depend on thresholds or confidence levels, SAE [132] and AMC 29.1465 [18] recommend that uncertainty be explicitly integrated into the alerting logic. Thresholds should be designed not as fixed values but as uncertainty-aware boundaries that account for sensor variability and model limitations. For example, a predictive alert may be triggered only if the lower bound of a confidence interval exceeds a fault threshold; thus, minimizing false alarms while preserving safety.

Sensitivity analysis. To prioritize uncertainty reduction efforts, it is also recommended to perform sensitivity analyses that identify parameters with the highest effect on system outputs [18,132]. These analyses can be local (e.g., one-at-a-time) or global (e.g., Sobol indices) [133]. AMC 29.1465 [18] addresses this approach by encouraging deterministic parameter sweeps for safety-relevant components to ensure robustness. Key metrics may include variance contributions or gradient-based impact scores, with results used to guide both design improvements and data quality requirements.

Mitigation strategies. As mitigation depends on the nature of the uncertainty, SAE [131,132] provides a list of recommended countermeasures (see Table 2). Each mitigation action should be justified in terms of expected impact, supported by validation evidence, and aligned with the criticality of the measurement. In systems using COTS equipment, mitigation may also include architectural segregation or software safeguards, as noted in AMC 29.1465 [18].

Continuous learning and data feedback. Many sources of uncertainty – particularly those classified as epistemic – can be reduced over time through exposure to operational data. Therefore, system developers are encouraged to implement feedback mechanisms that allow in-service experience to refine uncertainty estimates and update model calibrations [132]. With these validated and appropriately documented updates, CBM systems are enabled to become more confident and less conservative over their life cycle.

Documentation and traceability. Throughout the process, uncertainty management activities must be clearly documented. SAE [74] and the EASA [100, CM-S-014] emphasize the need for traceability of uncertainty analyses. Consequently, identified sources, assumptions, propagation results, and mitigation actions should be recorded in a validation plan and linked to the system architecture. As a result, this documentation not only supports certification but also facilitates system updates and stakeholder communication.

Integration with verification and validation. In summary, uncertainty management is inseparable from V&V planning. Therefore, it is required that all known sources of uncertainty need to be documented and traced through the validation process [18,74]. Furthermore, SAE [132] outlines a tiered V&V approach where uncertainty quantification is used to define acceptance criteria and model applicability limits. AMC 29.1465 [18] provides additional context for aligning analytical model outputs with observed field behavior when measurement uncertainty is non-negligible. Lastly, Thacker [99] emphasizes that predictive credibility must be assessed not only through accuracy but also through a rigorous understanding of uncertainty sources, their propagation, and possible mitigation strategies.

All regulatory and industry guidance highlight the importance to systematically address uncertainties in measurement and modeling to

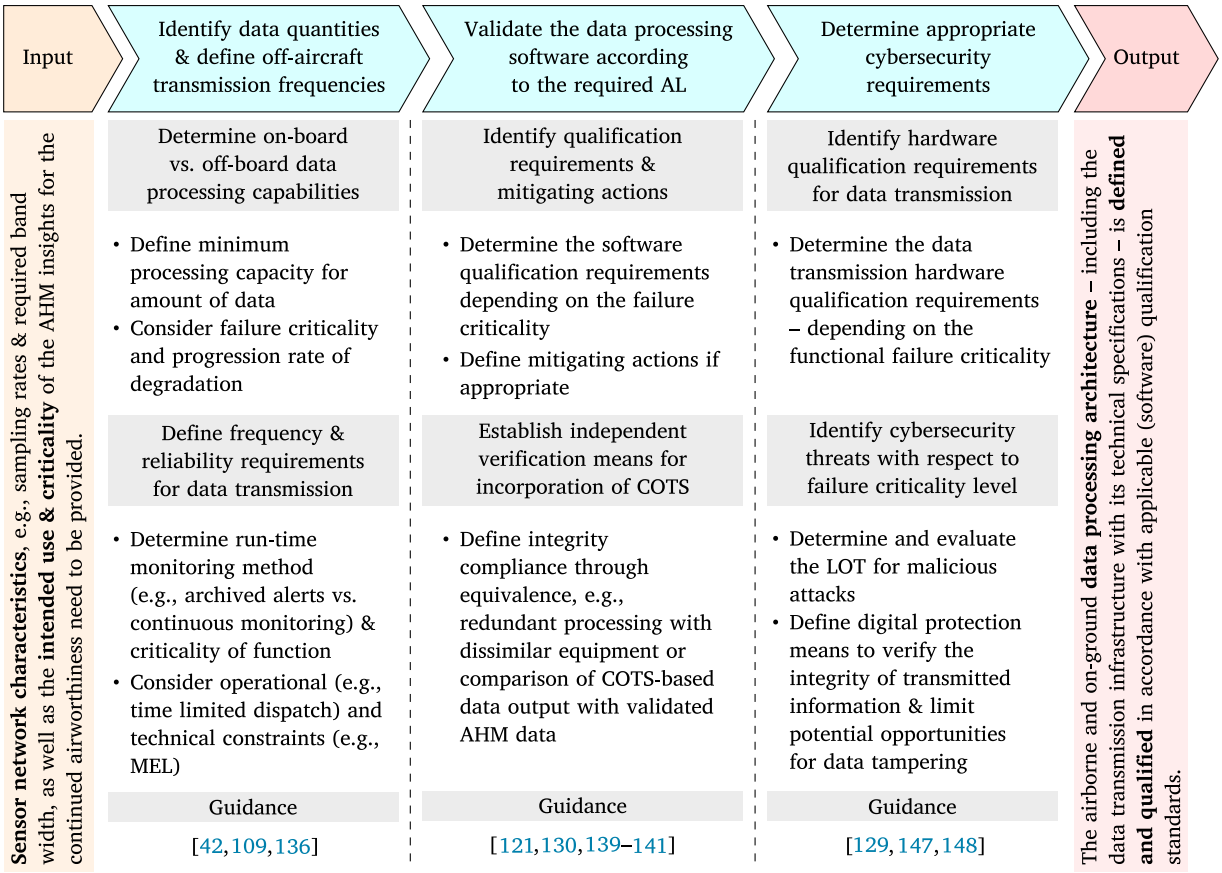

Fig. 6. Data processing & data transmission.

support the credibility of automated maintenance systems. Using structured processes – ranging from uncertainty identification and classification to propagation, validation, and continuous refinement – developers can ensure that model outputs are both accurate and trustworthy under real-world conditions. Lastly, continuous refinement requires that underlying assumptions be periodically reassessed, as an assumption's initial validity may no longer hold under changing operational conditions.

7. Data processing & transmission network

With the data acquisition steps for the AHM development addressed, we will now focus on aspects for the subsequent data processing and secure off-board data transmission (see Fig. 6).

O3.1: Identify data quantities & define off-aircraft transmission frequencies

In general, the FAA [42] emphasizes that data processing equipment and software has to provide the capacity to process the amount of required HM data with processing speed not being limited to an unacceptable rate by hardware or software. While the acceptability of processing speed depends on (a) the amount of data to be processed and (b) the specified performance for HM data processing, the speed should be "reasonable to accomplish data processing in a reasonable time" [42, p. MG 15-12].

Although there are no specific thresholds for the (on-board) data processing speed provided, it is apparent that on-board legacy computing power is insufficient to handle large volumes of data that would be required for data-driven ML applications. At the same time, the continuous downstream of unprocessed raw data becomes less viable with increasing data volumes generated by sensors [134] and the limited capacity of (existing) communication systems, e.g., Very High Frequency (VHF) or Satellite Communication (SATCOM) [135].

In order to define the frequency and reliability of data transmission technologies, we first need to determine the monitoring strategy, i.e., identifying functions that have to be performed during runtime and those that could be performed off-line [109]. SAE [109] distinguishes the following monitoring methods:

- Active monitoring, where information is requested actively from the monitoring unit,
- Passive monitoring, where the monitoring unit receives the information that is sent to the system,
- Live/Online monitoring, where the condition is monitored while the system is in operation, and
- Offline monitoring, where information is archived during operation of the system for subsequent (retrospective) analysis.

Furthermore, we need to define the fundamental philosophy for the sensing architecture. Here, SAE [109] provides an overview of three distinct options.

Centralized data acquisition/processing. This architecture follows the principle of a centralized Health-ready Component (HRC) – both in location and functionality. In accordance with SAE [136], HRCs are defined as components that provide all IVHM functionalities to allow them to be easily integrated into the overall vehicle. In particular, these functionalities include features such as (a) sensor mechanisms for continuous monitoring of critical component functions, (b) raw sensor data processing units to produce data related to the state of health, (c) state detection and health assessment functions to synthesize health state information, (d) prognostic algorithms to predict RULs based on performance and usage data, and (e) communication interfaces to

transmit raw, health state, condition and prognostics data from the component to a higher-level data subsystem or vehicle [136]. This type of an architecture is typically chosen when the monitored system requires outputs of the health monitoring to perform its intended function, e.g., for flight controls. Therefore, the concept would translate to the *active monitoring* run-time mode. While this monitoring concept eases the supply of integrated processing power for data handling, any additional IVHM functionalities or bandwidth would necessitate extensive modifications to the HRC [109].

Distributed data acquisition, shared processing. This architecture consists of so called smart sensors, which individually acquire sensing data and share a common data processing capability among multiple sensors for each monitored component. While the data processing and state detection capabilities are co-located at the sensing units, the health assessment, prognosis, and alert generation will be performed in a distributed manner among multiple aircraft component systems – however, still on-board the air vehicle. Therefore, this architecture concept offers the advantage of local data processing with higher redundancy and fewer single points of failure. The main disadvantage of this approach is the requirement of component modifications and data processing units in multiple locations – implying a reduced processing power [109].

On-board and off-board processing. With this concept, the data acquisition and processing takes place on-board with the component, while the subsequent data analysis for health assessment, prognostics, and advisory generation to the maintenance personnel happens off-aircraft. Therefore, it enables time-sensitive or less complex information to be processed during flight with less time-critical or more complex data analysis to be performed after landing. However, that working principle implies a more complex infrastructure to enable the combination of on-board and delayed off-board processing [109].

In addition to the processing speed, the architecture also has to provide resistance against the introduction of errors or out-of-specification inaccuracies for any parameter [42]. Since a developed AHM system is likely to interface with numerous other systems, the definition of these interfaces are likely to provide constraints on the CM system design [109]. Therefore, as highlighted by SAE [109], an AHM architecture will need to adhere to the following constraints.

- Data availability through limitations of adding new sensors
- Data movement through limitations of sampling, transfer or status reporting rates with shared sensors
- Data processing and storage limitations through limited allocation of memory or processing capability within the monitored system
- Allowable functional criticality through differences between available ALs of a data generator (by another legacy system) and the required ALs for the subsequent usage by the AHM system
- Desired usage of COTS equipment

Especially for the integration of COTS items in conjunction with hazardous or major criticality applications, the FAA [42] emphasizes to either isolate COTS elements on designated sub-system or demonstrate adequate protection for higher-level processing on the same equipment.

O3.2: Validate the data processing software according to the required AL

In order to determine qualification requirements for the data processing software, we utilize the corresponding application intent based on the approach for qualifying airborne AHM hardware (see *O2.2* in Section 6) and the determined integrity level (see *O1.3* in Section 5) [42].

In accordance with AMC 29.1465 [18], signal processing capabilities for an helicopter VHM need to address the complexity of the monitored mechanical elements and signal transmission pathways. Furthermore, the capabilities have to be demonstrated as appropriate for the potential failure modes. Consequently, the corresponding software should be developed in accordance with the determined E2E criticality and the defined integrity level as part of the system design process (see Section 5) [42].

Since the objective of processing sampled data is to produce HM indicators that correlate to degradation characteristics of the monitored components, it is essential to enhance the signal-to-noise ratio through appropriate data processing techniques (e.g., vibration signal averaging for gears, and signal band-pass filtering and enveloping for bearings) [18,137]. Since data processing is an essential part of every data analysis procedure [138], a range of guidelines for different applications exist already, e.g., ISO 13373-2 [139] for vibration signals or Cremona and Santos [140] for SHM applications.

With the data processing primarily being software-driven, RTCA [130] provides guidelines for the certification of airborne software[4] – extended by additional guidance from RTCA [141] for on-ground software applications. Furthermore, RTCA [121] also lists requirements for data processing standards in order to ensure integrity and reliability of the data itself – with special focus on data quality aspects as defined by the DQRs (see *O1.2* in Section 5) and categorized through data integrity levels. In addition, it defines a Data Process Assurance Level (DPAL) – the integrity level that is required to protect against erroneous data processing depending on the corresponding failure severity due to faulty data insights [121].

However, especially for proprietary software products, e.g., Microsoft Excel for data analysis applications, it may be infeasible to apply the standard software qualification procedure as laid out by RTCA [130]. As a remedy, the FAA [42] proposes so called *independent verification means*, i.e., an independent process to verify the correct functionality of a HM application that utilizes COTS (see *O5.1* in Section 9). Consequently, the intent of independent verification is to gain additional confidence in the operational reliability of COTS and may be discontinued once sufficient confidence in the application has been achieved [42].

Besides the software qualification aspect, consideration must be given to the data transfer requirements as they will help to determine applicable run-time options. Depending on the urgency and priority of the processed information, the method of communication should be determined – potentially ranging from utilizing existing data buses up to the development of dedicated HM communication networks. Furthermore, the importance of the data availability will need to be reflected in the definition of dispatch limitations, e.g., by incorporation into the Minimum Equipment List (MEL) [109].

O3.3: Determine appropriate cybersecurity requirements

The European Union has introduced two new legislation acts [142,143] that lay down organizational requirements for the management of information security risks for aviation safety. While one applies, among others, to production and design organizations [142], the other is focused on maintenance organizations, Continuing Airworthiness Management Organizations (CAMOs), and operators [143]. In essence, these regulations require the corresponding organizations to implement and maintain an Information Security Management System (ISMS)[5] to [142,143]

(a) Establish a policy on information security with regard to the potential impact on aviation safety,
(b) Identify and review information security risks,

[4] As emphasized by Bello et al. [116], since these guidelines do not entirely cover the challenges of AI-enabled systems, the corresponding aspects such as AI assurance and trustworthiness have been addressed by the EASA [67].

[5] The FAA [144] has published the counterpart with a strong focus on aircraft-related information security and guidelines for the development of an Aircraft Network Security Program (ANSP).

Table 3
Acceptability of risk treatment [149].

Level of Threat	Severity of threat condition effect				
	NSE	Minor	Major	Haz.	Cat.
Very High	Unacceptable	Unacceptable	Unacceptable	Unacceptable	Unacceptable
High	Acceptable	Unacceptable	Unacceptable	Unacceptable	Unacceptable
Moderate	Acceptable	Acceptable	Unacceptable	Unacceptable	Unacceptable
Low	Acceptable	Acceptable	Acceptable	Two independent security measures	Unacceptable
Extr. Low	Acceptable	Acceptable	Acceptable	Acceptable	Two independent security measures

NSE ... No Safety Effect | Haz. ... Hazardous | Cat. ... Catastrophic | Extr. Low ... Extremely Low

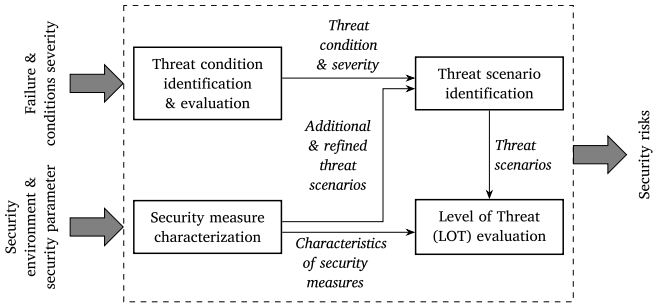

Fig. 7. Workflow to identify and assess security threats in accordance with RTCA DO-356 [149].

(c) Define and introduce risk treatment measures,
(d) Implement an internal reporting scheme,
(e) Detect and responds to information security incidents with a potential impact on aviation safety,
(f) Comply with personnel and record-keeping requirements, and
(g) Monitor compliance of the organization and – in case of any findings – to provide feedback to the accountable manager/head of the design organization.

In order to support a development that complies with the intent of these regulations, the EASA [145, Annex I&II] also provides AMC and Guidance Material (GM). While these guidelines are specifically tailored for the needs of aviation-related applications, their content consistently aligns with established standards (e.g., ISO 27001 [146], RTCA DO-326 [147], RTCA DO-355 [148], and RTCA DO-356 [149]).

Originally developed for avionics applications, the RTCA guidelines in particular provide useful guidance for the development of AHM systems with respect to a secure information management system design. Especially, RTCA DO-326 [147], in conjunction with RTCA DO-356 [149], describes extensively the procedure to perform security analyses as part of the system design process. It differentiates from RTCA DO-178 [130] by providing a broader context beyond the mere software development procedure with strong focus on the interaction of software security and effects on aircraft safety [150]. Consequently, the central aspect is the determination and evaluation of a so called Level of Threat (LOT) to determine the acceptability for the combination of residual risk of successful malicious attacks and their respective severity (see Table 3) [147].

In order to determine the severity of threat conditions, RTCA [147] suggests to perform a security assessment – in parallel to the system development process of SAE [151] with its proposed methods (see Section 5). Simultaneously, to evaluate the LOT, it is necessary to examine the effectiveness of protection, i.e., the resistance of security measures against attacks. This effectiveness will subsequently be rated – with higher resulting numbers indicating superior effectiveness. Furthermore, every security measure is evaluated by three criteria [149].

- Preparation means, i.e., how much system-specific knowledge and special hardware is necessary for an attack
- Window of opportunity, i.e., how often are suitable opportunities present to launch an attack
- Execution means, i.e., how skilled and proficient does an individual need to be for the development of an attack

The corresponding workflow for the resulting evaluation process of security risks is shown in Fig. 7.

In addition to these guidelines for initial system designs, RTCA [148, p. 21] provides measures that shall be established throughout the life cycle of the air vehicle to ensure a continuing airworthiness. It distinguishes between responsibilities of the DAH and operator, respectively. Furthermore, RTCA [148, pp. 25–28] specifically devotes a section to Ground Support Information Systems (GSIS), i.e., ground systems that shall ensure the secure data distribution from the air vehicle to the destination (e.g., a Maintenance Control Center (MCC) or airborne data storage). It recommends to use digital signature methods (e.g., in accordance with ARINC 827 or ISO 9797 [152–154]) for the verification of authenticity and integrity of the exchanged data, and to limit any opportunities for tampering. Lastly, the DAH is responsible to provide guidance to the operator so that all information security requirements that have been specified in accordance with RTCA [147] can be followed [148]. The information may be provided through aircraft security ICA and can include

> *"activities ranging from scheduled data integrity and software conformity checks to aircraft assigned maintenance laptop/GSE restoration and updates".* [144, p. 6]

However, in addition to that, the operator needs to have policies and procedures in place that ensure the secure handling and management of GSIS, including log data for incident detection and documentation, risk assessment for their operations, and prevention of unauthorized access [144,148,155].

8. Data analysis

After the acquired data has been processed to allow for an effective derivation of HM indicators, we will now focus on the necessary aspects for detecting the current degradation status of the monitored system – as shown in Fig. 8.

8.1. State detection

As a first step for an automated CBM system, we need to define and characterize alerts and alarms. While an AHM-based alert solely indicates the requirement for further processing or investigation to determine the need of corrective maintenance actions [122], an alarm signals the necessity of such a restorative intervention [18]. Consequently, the alert value is typically set at higher remaining system conditions than the alarm value – to issue earlier indications [122].

O4.1: Define suitable alert & alarm levels for the system condition

As emphasized by ISO 17359 [108], preliminary alert and alarm criteria should be defined in a way to allow an early indication of faults or failures. Since abrupt changes in a system's state within given limits may indicate the need for further investigation [128], it can be helpful to define multiple thresholds with different levels of intervention. It is important to note that alert and alarm criteria should be optimized with increasing operating experience over time in an iterative process [108].

Any developed solution to (automatically) detect a system's condition needs to comply with the safety analyses performed initially in the AHM system design phase (see Section 5). This includes the detection and isolation of all mission- or safety-critical malfunctions [156].

In order to define suitable alert values, the following aspects must be considered [157]:

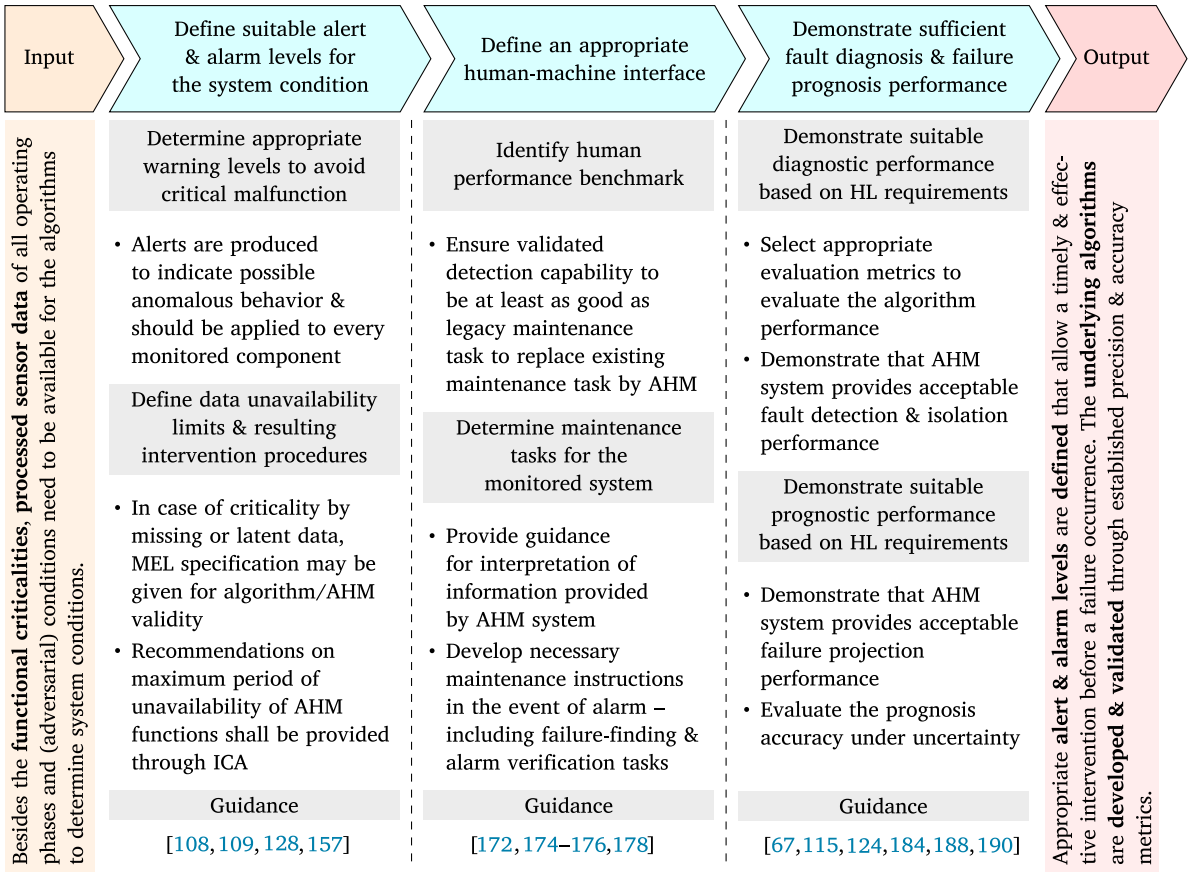

Fig. 8. State detection, fault diagnosis & failure prognosis.

- Confidence level of the prognostics
- Future operating requirements
- Lead times of spare parts
- Required maintenance planning
- Work required to rectify the faults
- Trend extrapolation and projection

An exemplary combination of the resulting thresholds with an FMEA of the system design process (see Section 5) is shown in Table 4.

8.2. Human performance benchmarks in aircraft maintenance

As described in Section 2.2, the current maintenance paradigm is built on the philosophy of manual task execution. Therefore, for a fair comparison between an automated CBM approach and the manual task execution, we need to identify a realistic human performance baseline.

O4.2: Define an appropriate human-machine interface

Over the decades, a significant body of literature on human factors in aircraft maintenance has been published. A study performed by Marais and Robichaud [158] has concluded that roughly 7% of FAA-captured commercial aviation incidents between 1999 and 2008 were maintenance-related – with a majority of incidents involving mechanical malfunctions (i.e., components operating incorrectly) or failures (i.e., component break-downs). In addition, incorrect servicing represents the second most frequent type of maintenance error – only behind faulty installations [159]. These observations confirm a Boeing study from 1994 [160], who examined 86 incidents and found for roughly 12% that undiscovered degradation, inappropriate testing, forgotten material, or skipped necessary servicing were contributing factors. Furthermore, Hobbs and Williamson [161] have found in their study that 4% to 9% of maintenance-related safety occurrence are linked to a lack of required servicing or missed degradation. Simultaneously, maintenance faces the challenge that errors in the task execution are likely to result in consequences that are not immediately obvious. However, this delayed feedback substantially reduces the ability to learn from errors, their causes and contributing factors – and hinders the establishment of prevention policies [162].

Consequently, several studies have analyzed common human errors in aviation maintenance, such as [162–167]

- Omissions (e.g., access panels not closed or unsecured),
- Incorrect installations (e.g., usage of worn or degraded components),
- Wrong parts (e.g., superseded part numbers),
- Inadequate execution (e.g., insufficient lubrication, introduction of functional defects by over-torquing, or system contamination),
- Mishandling (e.g., foreign object impact or left tools).

While the majority of airline technicians always adhere to implemented policies [168,169], the predominant drivers for non-compliance are (a) the lack of practicality of existing procedures that would prevent a timely job completion, (b) the strong reliance on own skills and expertise, and (c) the inherent assumption to follow a procedure [169, 170].

Furthermore, ambient condition for the task execution have to be considered as these environmental aspects can affect maintenance performance, e.g., [162,166,168,171]

- Extremely cold conditions that may require technicians to wear gloves complicating manipulation and reducing tactile sensation,
- Hydrocarbons in fuel tanks and cleaning agents creating noxious atmospheres,

Table 4
FMEA example for a condition monitoring system [122].

System	Sub-System	Component	Failure mode	Potential effect	Means of detection	Detection values		Recommended actions
						Alarm	Alert	
Water pump	Electric motor	Bearing	Outer race broken	1. Increasing vibrations of the shaft, risking involuntary contact between shaft and casing 2. Jamming of shaft and ceasing of pump motor	Accelerometer			Bearing replacement
		Stator	Short circuit	Damaged isolation layer could allow contact between rotor bars and coils	Resistance measurement			Replacement of degraded insulation layers
	Centrifugal pump	Impeller	Cavitation	1. Vibration from cavitation can increase degradation level in the pump bearings. 2. Vibration can go through the shaft into the hydrostatic and motor bearings, increasing their degradation.	Energy consumption			Pump check & impeller replacement

- Inadequate lighting conditions for inspection tasks, and
- Postural hazards in restricted spaces resulting in decreased attentiveness for inspection tasks.

Besides these qualitative description of influencing factors, we need to estimate quantitatively the likelihood of erroneous inspections to determine a human performance baseline.

One such qualitative method has been developed by Williams [172] who uses the so called Human Error Assessment and Reduction Technique (HEART) in combination with proposed nominal human unreliability values to estimate the likelihood of an error occurrence for different conditions and settings. For example, for a "routine, highly-practiced, rapid task involving relatively low level of skill" [172, p. 439], the nominal unreliability can be expected to vary between 0.7% and 4.5%; that is, even under perfect outside conditions the error likelihood can be as high as 4.5%. If the ambient conditions change to a less favorable environment, the human error rate is likely to increase and will need to be adjusted accordingly.

Another approach is to use the structural integrity characteristics for deteriorating assets. As part of the regular design process, structures need to be developed so that they can withstand structural loads in the presence of so called Barely Visible Impact Damages (BVIDs). These are dents with a maximum depth that ensure a Probability of Detection (POD) of 90% during scheduled inspection [173].

To determine the POD for structural defects, Boeing [174] has developed reliability curves for different levels of inspection, i.e., GVI, zonal inspections, and DET (see Section 2.2.2). While their approach appears to be qualitative in nature, an experimental study by Spencer [175] has found that a POD of 90% can be achieved for damage sizes as little as 0.91".

The influence of varying ambient conditions (lighting and cleanliness) on the detectability of dents has been examined by Erhart et al. [176]. Unsurprisingly, they have found that the detectability of dents increases with (a) the cleanliness of the structure and (b) the brightness of the inspection area. Furthermore, their experiments showed that the POD ranges from about 70% for small dents on dirty surfaces with poorly lit environments up to 96% for larger dents on clean surfaces in well lit settings. However, it has to be noted that for extreme condition, there are deviation from this range. Additionally, the data quality seemed to be insufficient to provide a conclusive statement for these cases.

Lastly, Williams and Borow [177] have examined how the probability of detecting visual anomalies changes with varying inspection speed, defect density, and examination direction. They could show that the POD (a) is highest for lower speeds, (b) decreases with increasing visual distractions, and (c) is higher for horizontal inspection directions. Their findings indicate that the POD ranges between 70% in a worst case scenario and 98% under ideal conditions. Based on these results, the POD can be approximated by [178]

$$P = 1 - \exp\left(-\frac{t}{\bar{t}}\right)$$

and

$$\bar{t} = \frac{t_o A}{apn}$$

with t_o as the average inspection time for one area, A as the size of the area, a as the area of the visual lobe, and p as the probability for a detected imperfection.[6]

Given that an increased task time seems beneficial for the POD of defects, Drury [179] has examined if there are possible adversarial implications with these extended task times. Therefore, he analyzed how the probability of false alarms (i.e., indicating a defect despite acceptable conditions) and missed alarms (i.e., missing existing faulty conditions) changes for visual inspection with respect to the available task time. His results indicate, with increasing available inspection times, (a) the missed rate decreases exponentially, but (b) simultaneously, the false alarm rate increases linearly. Consequently, a stopping time for the inspection task should be defined that minimizes missed and false alarms while also accounting for the cost of the inspector's time [178].

8.3. Health assessment & failure prognosis

In the following section, we will shift our focus to the development of fault diagnosis and failure prognosis approaches. Since there is a broad variety of papers addressing the technical challenges of algorithm developments [10,12,180–182], we will limit our work on the determination of performance outputs with respect to the underlying functional failure criticality. Furthermore, a particular challenge for an algorithm V&V is the data availability. SAE [72] and the FAA [42] discuss at length the possibilities for an effective validation, e.g., through historical fault data, seeded tests, or in-service experience. For this paper, we assume that the required data is available and an appropriate technique has been chosen. Thus, we will focus on how to determine a suitable performance to demonstrate an AHM system's effectiveness. As of this paper, we will align our definition of effectiveness with the MSG-3 approach [52]; they define maintenance tasks to be effective if they reduce the risk of failure to an acceptable level with the intent to assure safe aircraft operation.

First, we want to highlight that maintenance organizations are required to provide tools – specified in the maintenance instructions of an aircraft or their verified equivalents – for day-to-day maintenance [50, M.A.608]. Consequently, in order to satisfy especially the latter part of this requirement, any developed AHM solution will need to possess an interface for independent developers to allow verification of their solutions. A similar aspect has been envisioned already by

[6] This depends on how a is defined. It is often defined such that p = 1/2, i.e., an area with a 50% chance of detecting an imperfection [178].

Table 5
Fault diagnosis metrics [190].

Diagnostic category	Performance category	Metric
Detection	Response time	Time to detect
	Accuracy	False positive detection rate False negative detection rate Fault detection rate False detection accuracy
	Sensitivity	Detection sensitivity factor
	Stability	Detection stability factor
Isolation	Response time	Time to isolate
	Accuracy	Isolation classification rate Isolation misclassification rate
	Resolution	Size of isolation set
	Stability	Isolation stability factor

the FAA [42, MG.15-11] with their illustration of COTS inclusion into the HM development process. They propose an interface between data outputs from a certified system and its COTS counterpart. While the FAA [42] only specifies that COTS shall be verified by independent means in terms of their accuracy and integrity, such an interface could be established through comparison of dedicated performance metrics. That is, if a COTS-based system delivers a measurable performance at least as good as the qualified verification means, it should also be considered verified.

O4.3: Demonstrate sufficient fault diagnosis & failure prognosis performance

Before we discuss how to measure diagnostic and prognostic performance, the corresponding baseline needs to be defined. Since the goal of HM applications is to identify hazards that can contribute to anomalous behavior, it is necessary (a) to quantitatively describe safety margins for the targeted operation and (b) to predict safety margin effects in real time [40]. Consequently, as part of the model evaluation process, minimum E2E performance criteria have to be defined that are consistent with the application's intended use and are reflected in corresponding indicators for the developed algorithms. The performance criteria should at least consider aspect such as accuracy, timing/sampling, resolution, event recognition, and consistency. With the intended use and potential functional failure criticality already identified as part of the AHM system design process (see Section 5), Saxena et al. [115] propose a workflow to translate those top level specifications into actionable AHM performance requirements [42,46,124].

Subsequently, in order to measure the criteria, a significant body of literature on the topic of performance metrics exists [183–188]. Thus, since it would be impractical to list all possible metrics in this work, we will focus on their underlying concept. Interested readers are kindly referred to the publication by SAE [184] and Saxena and Roemer [188] who present a rather extensive list of possible metrics. In the following, despite overlaps, we distinguish between diagnostic and prognostic metrics. Furthermore, we discuss the challenges for measuring the performance of artificial intelligence algorithms.

8.3.1. Diagnostic metrics

Diagnosis can be defined as a capability to detect and determine the root cause of a symptom [32]. Therefore, its objective is to detect and isolate faults in a timely and accurate manner with sufficient resolution to identify the specific faulty component [32,189]. In order to measure these capabilities, corresponding detection and isolation metrics have been developed (see Table 5) [189–191].

Table 6
Decision matrix for fault detection measurement [184].

	Fault	No Fault	*Total*
Detected	Detected faults (a)	False alarms (b)	***All alarms (a+b)***
Not detected	Missed faults (c)	Correct rejections (d)	***All non-alarms (c+d)***
Total	***Faults (a+c)***	***Fault-free cases (b+d)***	***All cases (a+b+c+d)***

Fault detection. Kurtoglu et al. [190] defines fault detection as indication of a malfunctioning system. Since fault detection is classified on a binary scale of the system state (faulty or non-faulty), the majority of metrics are constructed around decision matrices [184,190]. A decision matrix is a binary classification matrix that represents the distribution of predicted and actual states of faulty and non-faulty cases (see Table 6). Therefore, by calculating the ratios of correctly classified system conditions, we can calculate the detection accuracy factors of Table 5. In addition to these metrics, the time to detect refers to the time interval from the beginning of a fault injection to the moment of the first reliable detection signal. Furthermore, since the strength of a detection signals may vary over time, the detection sensitivity factor represents the required strength of a present fault before an indication is triggered. Once a detection signal has been sent, the detection stability metric indicates its presence over time – starting from the time of fault occurrence until the ultimate restoration task [190].

Fault isolation. In contrast to the binary fault detection category, fault isolation determines the exact fault mode and location within a system [190]. Therefore, isolation is a multi-state problem with multiple candidates for fault modes and locations – typically represented in a confusion matrix (see Table 7) [184,190]. Consequently, benchmarking isolation functionality is a far more challenging task compared to assessing the detection capability. Similar to the detection metrics, the accuracy parameters for fault isolation performances are also based on the ratios of correct classifications in Table 7. Since diagnostic systems typically do not converge on a single isolated fault but a number of possible faults with corresponding probabilities, the size of the isolation set indicates the level of ambiguity [188]. For the overall classification performance evaluation of the confusion matrix, SAE [184] proposes an Kappa coefficient that calculates the ratio of correctly classified faults divided by the total number, both corrected by those fault classifications that are expected by chance [192]. Thus, the coefficient's formula is

$$\kappa = \frac{N_o - N_e}{N_t - N_e}$$

with N_o as the number of correct classifications, N_e as the number of expected correct classifications by chance, and N_t the total number of observed faults. Additionally, the time to isolate similarly represents the time interval from the beginning of a fault injection to the moment of the first reliable fault classification. Finally, the isolation stability factor can similarly be interpreted to the detection stability, i.e., it measures the presence of a positive fault signal over time [190].

In addition to these detection and isolation metrics, SAE [184] proposes the use of a so called 'Metrics FMEA' for the evaluation of the diagnostics algorithm's performance. Similar to a conventional FMEA for system design purposes (as discussed in Section 5), this Metrics FMEA evaluates the different functions (e.g., the probability of detection) towards their failure mechanisms, effects, and counter measures.

8.3.2. Prognostic metrics

Prognostics can be defined as the process of predicting a system's RUL by estimating the time when the system will no longer perform its intended function within desired specifications [157,188]. However, projecting the future with limited information and inherent uncertainties (e.g., future operating loads and environments, model inaccuracies, data noise, and observer faults) is not a trivial exercise and leads to

Table 7
Confusion matrix for fault isolation measurement [184].

Diagnosed failure mode	Observed failure mode					Total
	Fault 1	Fault 2	...	Fault i	No Fault	
Fault 1	n_{11}	n_{12}	...	n_{1i}	n_{10}	$\sum_{k=0}^{i} n_{1k}$
Fault 2	n_{21}	n_{22}	...	n_{2i}	n_{20}	$\sum_{k=0}^{i} n_{2k}$
...	...	...	...	...	...	...
Fault j	n_{j1}	n_{j2}	...	n_{ji}	n_{j0}	$\sum_{k=0}^{i} n_{jk}$
No Fault	n_{01}	n_{02}	...	n_{0i}	n_{00}	$\sum_{k=0}^{i} n_{0k}$
Total	$\sum_{k=0}^{j} n_{k1}$	$\sum_{k=0}^{j} n_{k2}$	...	$\sum_{k=0}^{j} n_{ki}$	$\sum_{k=0}^{j} n_{k0}$	$\sum_{k,l=0}^{j,i} n_{kl}$

Table 8
Failure prognosis metrics [188].

Prognostic category	Metric	Description
Offline	Prognostic horizon	Maximum advance warning possible with desired confidence
	$\alpha - \lambda$ accuracy	Improvement of algorithm's performance with respect to time interval until EOL
	Prognostic false alarm rate	Extension from classical diagnostic metric with early prediction as false positive and late prediction as false negative
	Relative accuracy	RUL error normalized by actual RUL for any given time
	Convergence rate	Rate of prognostic performance improvement with updated predictions over time
	Sensitivity	Measure of robustness for prognostic algorithm against input changes or external disturbances
Online	RUL online precision index	Quantifies precision of predicted RUL distributions by assessing 95% confidence bounds with respect to predicted RUL over time
	Dynamic standard deviation	Quantifies stability of predictions by assessing the variance between individual predictions within a given time window
	Critical-α performance measure	Assesses the critical percentile of a RUL distribution that would allow a just-in-time maintenance action

uncertainties in the predictions (see Section 3.3.2) [187,188]. Consequently, any prognostic model will need to be validated thoroughly before it can be certified for critical applications [187]. While important, the assessment of an prognostics algorithm's 'live' performance would require knowledge of a system's true End of Life (EOL), a future event that is clearly not possible to know in advance [115,188]. Therefore, most prognostics performance metrics have been developed for so called offline (retrospective) assessments with relatively few for online applications (see Table 8 for an excerpt of possible metrics). Readers who are interested to learn more about possible prognosis metrics are kindly referred to SAE [184] and Saxena and Roemer [188].

Offline assessment. An essential prerequisite for a performance assessment of the prediction is the availability of reliable EOL information ('ground truth'). Since these information are only available after the predicted event has actually taken place, offline metrics are designed for a retrospective evaluation of failures from the past. Thus, it inherently assumes that if an algorithm conforms to specified performance requirements for failure events in the past, it will also do so for future failure projections. The offline prognostic performance metrics shown in Table 8 follow a systematic progression in terms of the provided information. While the prognostic horizon identifies if an algorithm can predict within a specified error margin yielding a sufficient advanced failure warning, the $\alpha - \lambda$ performance assesses how the error margins develop over time to evaluate their applicability to a desired use case. In a next step, the accuracy of the projection for given times throughout a system's lifetime are analyzed, i.e., how close to the actual EOL has the predicted EOL been for every time instance. As emphasized by Saxena et al. [115], it has to be noted that inaccurate predictions may result in different criticalities. For example, while a premature failure prediction mainly results in economic penalties and operational restrictions, a late prediction after the actual failure occurrence may result in catastrophic scenarios. Lastly, the convergence rate quantifies how fast the algorithm satisfies the prior metrics, i.e., how much operating time needs to pass before the algorithm can sufficiently predict the RUL [115,186,188].

Online assessment. In contrast to the retrospective offline assessments, these metrics will evaluate the performance of prognostic algorithms without knowledge of the true EOL, i.e., during normal operations prior to the actual failure occurrence. As mentioned before, the research field is much newer and there are comparably few corresponding metrics available [188]. Furthermore, indicators that do exist focus exclusively on the prediction process quality (e.g., stability and precision), since the evaluation of accuracy aspects requires the knowledge of ground truth EOL information [188]. Consequently, online metrics can only compare the RUL prediction trend observations to evaluate an algorithm's performance.

8.3.3. AI metrics

Especially for the application of AI-based algorithms, the following categories for the performance evaluation need to be considered.

Generalizability. The generalization of ML models describes their capacity to keep an acceptable level of performance on unseen input data (during the training phase) from within the ODD. It can be compromised by the data science phenomena of *overfitting* and *underfitting*. Model overfitting occurs when a statistical model fits exactly the training data but fails to perform accurately against unseen data from within the ODD. Therefore, an overfitted model fails to generalize since it has learned the patterns of the training data with the underlying noise to a degree where it negatively impacts the model's performance on new data. In contrast, while an underfitted model also cannot generalize well to new data, that is due to its lack of ability to create a mapping between input and target variables [124].

Stability & robustness. While stability of an ML algorithm ensures that the produced model does not substantially change with perturbations of the training data set, robustness describes a system's ability to maintain its performance level under all foreseeable (adversarial) conditions [124]. Consequently, the EASA [67] defines in their concept paper objectives to ensure the performance of trained ML models. Essentially, these objectives require a developer to demonstrate the algorithm's stability and robustness through

- Analyses of existing perturbations in the development phase due to fluctuations in the training data set (e.g., replacement of data points or labeling errors) and their resulting effect of an algorithm's instability [67, Objective LM-11],

Table 9
Corner case scenarios for AI applications.
Source: Based on [193].

Corner cases	Description	Examples
Pixel Level (Perceived) errors in data.	**Global Outlier** *All or many data points fall outside of the expected range of measurement.*	• Lighting conditions • External vibration interference
	Local Outlier *One or few data points fall outside of the expected range of measurement.*	• Pixel errors (dead pixels) • Dirt accumulation on sensor
Domain Level World model fails to explain observations.	**Domain level Shift** *A large, constant shift in appearance, but not in semantics.*	• Abnormal ambient conditions • Different generations of monitored components
Object Level Instances that have not been seen before.	**Single-Point Anomaly (Novelty)** *An unknown object.*	• Lost objects • Loose parts
Scene Level Non-conformity with expected patterns in a single image.	**Contextual Anomaly** *A known object, but in an unusual location*	• Faulty part installation
	Collective Anomaly *Multiple known data points, but in an unseen quantity*	• Sensor interference
Scenario Level Patterns are observed over the course of an image sequences. Recognition requires scene understanding.	**Risky Scenario** *Pattern that was observed during the training process, but still contains potential for accident.*	• System abnormality with short reaction time to failure
	Novel Scenario *Pattern that was not observed during the training process, but does not increase the potential for accident.*	• System abnormality with installed redundancy
	Anomalous Scenario *Pattern that was not observed during the training process and has high potential for accident.*	• Insufficiently understood system interactions with novel failure mechanisms • Unauthorized (external) interference

increasing complexity

- Verification of the stability through case studies that address anticipated perturbations in the operational phase due to fluctuations in the data input (e.g., added noise) within the whole ODD – including nominal and corner cases [67, Objective LM-12], and
- Verification of the model's robustness in adverse conditions (e.g., extreme weather or poor visibility for computer vision) [67, Objective LM-13].

In order for the algorithm to achieve the desired performance, Bello et al. [116] highlights the importance of a sufficiently defined OD and ODD – since these can greatly affect the completeness and representativeness of the data set selection and help in the identification of corner cases. This is in line with the development objective LM-12 from the EASA [67] concept paper that requires developers to verify the stability of the trained model for the whole ODD.

Therefore, we briefly present a definition of corner cases and introduce a potential classification. It has to be noted, there are several terminologies used in the literature related to corner cases [193–196]; however, since detecting corner cases can be difficult with a lack of common understanding, we will use the characterization as presented by the MLEAP Consortium [124]. Here, corner cases may result from (a) coinciding normal situations representing a rare case or scene [194], (b) entirely new situation and not just combinations of already known ones [124], or (c) an anomaly expressed in nonconform behavior or patterns [193]. Furthermore, corner cases can potentially comprise data samples that exhibit erroneous and unforeseen behaviors, e.g., adversarial data on boundaries and misclassified data [196]. Lastly, especially for computer vision applications, these cases can arise if the model cannot predict the appearance of a relevant object in a given context [195]. Regardless of their exact definition, the consideration of corner cases is essential to properly define system boundaries and to reliably detect outliers in the data. Table 9 shows a systematic overview of the different corner cases, their characteristics, and examples for an automated driving application.

9. Advisory generation

With faults identified and RULs predicted, the true benefit of such an AHM system can only be realized through an effective post-prognostic decision making. This process entails the utilization of prognostic information to evaluate possible actions for aspects such as maintenance planning and execution, supply chain management, mission planning, and mission allocation. However, due to the sheer number of possible decision pathways, the complexity of information that needs to be processed typically exceeds the cognitive capacity of human decision-makers. Therefore, it requires an automated system that uses all available information across the various data sources to optimize key parameters such as life cycle costs, mission success rate, or turnaround times while allowing AHM users to collaborate in the decision-making process. The corresponding work flow for the development of such an effective advisory generation process is shown in Fig. 9 [197,198].

O5.1: Define displayed information for maintenance decision making

An essential prerequisite for the certifiable completion of an automated maintenance task is the availability of reliable maintenance data for the determination of appropriate intervention actions. Therefore, any information from the airborne AHM system displayed at the MCC will need to be exhaustive enough that all possible failure modes can clearly be identified and displayed accordingly. Additionally, the displaying unit must be compatible with other parts of the system while

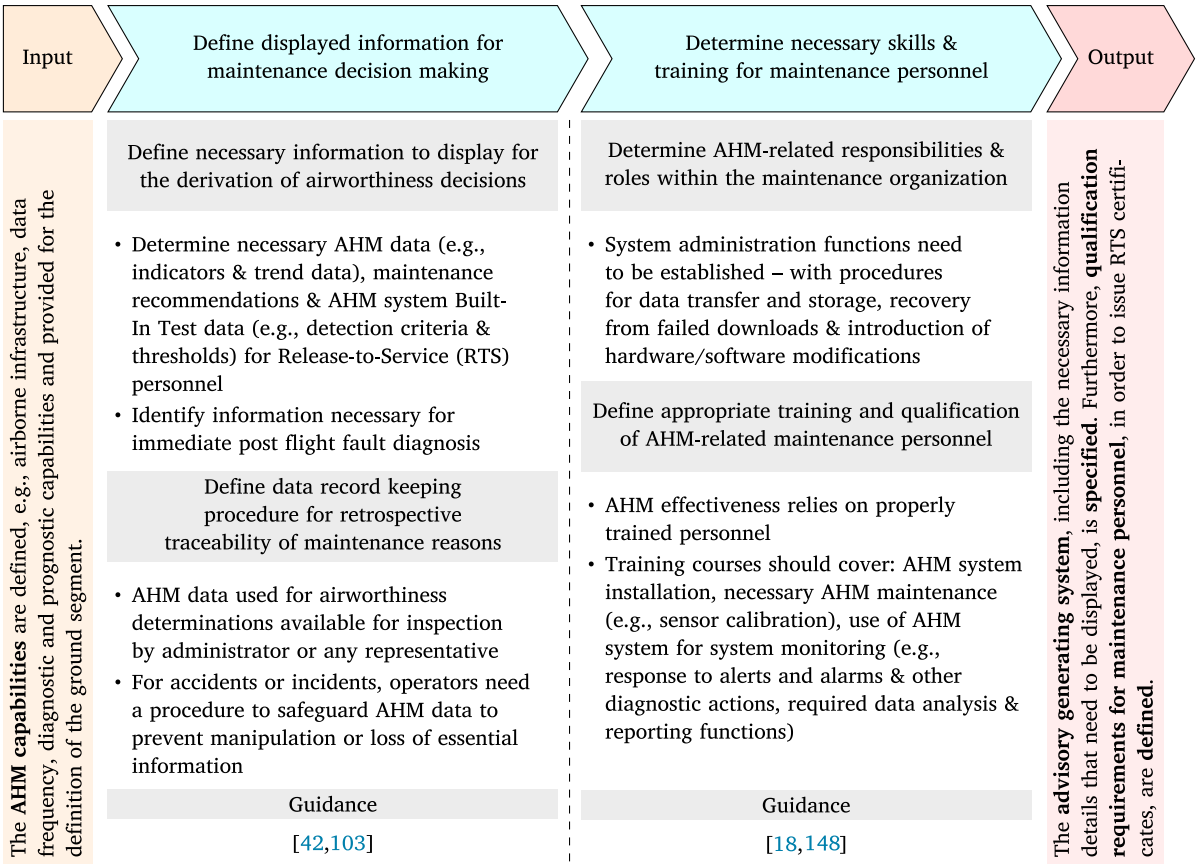

Fig. 9. Advisory generation.

providing a usable presentation [42,103]. In order to determine the formats for presenting and displaying data to an end-user, the work flow with the respective information requirements of each phase of condition monitoring needs to be assessed [103]. Furthermore, as emphasized by the FAA [42] on the example of ground based equipment for a HUMS application, the integrity and accuracy requirements for the ground infrastructure must be identical as for any other AHM system part.

With the installation of general-purpose COTS equipment in the ground based infrastructure, compliance with the integrity requirements may be difficult. As a remedy, the determination of compliance for COTS shall be based on equivalence. That is, similar to the qualification of the data analytics algorithms (see Section 8), any ground-based processing equipment that consists of commercial hardware and software must provide a satisfactory service history and an independent means of verification. Examples here are (a) physical inspections, (b) redundant processing by a second dissimilar computer with different COTS or (c) a combination of those. Once sufficient in-service experience has been gained and an appropriate performance has been demonstrated, the need for these independent verification means may be discontinued with regulatory approval [42].

O5.2: Determine necessary skills & training for maintenance personnel

In order to show compliance with the requirements for continued airworthiness, a qualified organization, e.g., in accordance with Annex II (Part-145) of Regulation (EU) 1312/2014 [50], needs to assure that the required maintenance – as laid out by the ICA – for an aircraft has been performed. Additionally, these maintenance tasks are only allowed to be accomplished by qualified personnel using methods, standards, and techniques as prescribed in the corresponding maintenance data [50, M.A.402]. This maintenance data contains, among others, (a) applicable procedures, standards, or information issued by a competent authority, (b) applicable airworthiness directives, and (c) ICA [50, M.A.401]. After completion of the maintenance tasks, an Release-to-Service (RTS) certificate needs to be issued [50, M.A.612].

Consequently, it becomes apparent that any maintenance recommendation based on automated AHM capabilities will need to be part of the maintenance data as well. Furthermore, in order to issue an RTS certificate, the corresponding certifying staff will need to demonstrate that they meet the requirements as described in Annex III (Part-66) of Regulation (EU) 1312/2014 [50, M.A.607]. However, the current qualification standards for certifying staff seem to be incompatible with requirements arising from automatically generated AHM advisories – since these standards are heavily aligned with practices of traditional manual inspections [50, Annex III - Appendix I].

While specific guidelines for minimum qualification requirements are very limited, RTCA [148, pp. 45–50] provides a generic list of qualification recommendations as part of an organizational ANSP/ISMS. Among others, these recommendations include

- Determination of necessary competencies for performing work affecting the ANSP
- Provision of training and documentation for each specific task to be performed
- Training for awareness of information security risks and their relation to aircraft safety
- Physical protection of digital and physical assets
- Restricted access to sensitive areas
- Establishment of recurrent training procedures to keep up-to-date with new technologies, system installations, and identified threats

Beyond these recommendations, the FAA [17] emphasizes the importance of properly trained personnel for an effective CBM program. In cases of airworthiness determinations with AHM data, e.g., verifying system functions without physical asset access, the responsible personnel shall be in possession of appropriate certification in the form of an 'Airframe and Powerplant' certificate. The EASA [18] extends that requirement by listing specific training course contents – resulting in the following aspects that should be covered:

- Installation of the HM system
- Line maintenance of the HM system (including necessary calibration)
- Usage of the HM system to monitor the vehicle (including data transfer, interface with data analysis, response to alerts and alarm processing, and fault-finding tasks)
- Necessary system administration functions (including recovery from failed downloads and the introduction of hardware/software modifications)
- Data analysis and reporting functions

10. Conclusion and recommendation

After reviewing the existing regulations and guidance material designed for aviation applications, we can conclude that – despite the various individual efforts – no conclusive framework to support the development of certifiable HM solutions exists. Furthermore, the level of ambiguity for the available information continuously increases for higher levels of the OSA CBM process framework. While the importance of E2E considerations for developed HM technologies is emphasized throughout all available literature, no common definition for the terminology seems to exist, let alone a corresponding workflow. However, practical examples from other applications (e.g., HUMS for rotorcrafts [18,42]) or industries (e.g., manufacturing industry [108]) offer useful insights that such a framework for AHM solutions can build upon. In addition, there are plenty of guidelines (e.g., SAE ARPs on the topic of IVHM) with technical details for the development of these solutions. An overview of the reviewed literature with their respective coverage of the OSA CBM cycle is shown in Table 10. Although the categorization is subjective in nature, we can see a strong emphasis of regulatory documents on the system design and requirement process with a lack of guidance for the development and evaluation of diagnostic and prognostic capabilities.

In summary, key take-aways from this review are:

1. The established OSA CBM framework offers a suitable workflow to support the development of certifiable AHM solutions. With its establishment in various application fields, the OSA CBM model comes with extensive guidance material to ensure a comprehensive consideration of E2E capabilities.

Table 10
Overview of existing guidance documents for the respective OSA CBM steps.

Document	OSA CBM Step						
	AHM System Design	Data Acquisition	Data Processing	State Detection	Diagnosis	Prognosis	Advisory Generation
Regulations & AMCs							
CS 29.1465 [18]	◕	◑	◕	◔	◔	○	◕
CS 25.1309 [43]	●[a]	○	○	○	○	○	○
AC 43-218 [17]	◔	○	◑	◔	○	○	◕
AC 29.2C - MG15 [42]	◑	◕	◑	◔	○	○	◔
AC 119-1A [144]	○	○	◔	○	○	○	◑
IP 180 [75] & IP 197 [76]	◑	◔	◔	◔	○	○	○
RTCA DO-160 [129]	○	●	○	○	○	○	○
RTCA DO-178 [130]	◕	○	●	○	○	○	○
RTCA DO-200 [121]	◕	○	●	○	○	○	○
RTCA DO-254 [199]	●[a]	◑	◔	○	○	○	○
RTCA DO-278 [141]	◕	○	●	○	○	○	○
RTCA DO-355 [148]	◔	○	●	○	○	○	◕
Standards & Recommended Practices							
ISO 25012 [125] & ISO 5259-2 [126] & ISO 9797 [152–154] & ISO 27001 [146]	○	○	●	○	○	○	○
ISO 17359 [108]	◑	◕	◔	◑[b]	◑[b]	◑[b]	◔[b]
ISO 13373 [128,139,200–203]	◕	●	◔	◕	◕	○	○
ISO 13381-1 [157]	◔	○	○	◔	◔	◕	○
ISO 13374-1 [103]	◔	◔	◔	◔	◔	◔	◑
SAE JA6268 [136]	◕	◔	◔	◔	◔	○	○
SAE ARP 6887 [72]	◑	○	○	◕	◑	◔	○
SAE ARP 5987 [74]	◑	◔	◔	◑	◑	○	○
SAE ARP 6883 [85]	●	◔	◔	◔	◔	◔	○
SAE ARP 6290 [109]	●	◑	◑	◔	○	○	○
SAE ARP 5783 [184]	○	○	○	●	●	◔	○
Reports & Academic Papers							
EASA [67]	●	◕	◔	◑	◔	◔	○
SAE AIR 7999 [183]	◔	○	○	◔	◕	◕	○
Saxena et al. [115]	◑	○	○	○	○	◕	○
Saxena et al. [120]	◕	◔	○	◔	◔	○	○
Saxena and Roemer [188]	○	○	○	◑	◕	◕	○

○ Not covered | ◔ Mentioned without further guidance | ◑ Highlighted with limited guidance | ◕ Extensively discussed with limited guidance | ● Full guidance and discussion

[a] Assuming an airborne AHM system is treated like conventional aircraft systems.

[b] References to other ISO standards is given.

2. The system design process for AHM technologies does not differ from the process of any other conventional airborne system. Thus, the existing ARPs that have been developed and iterated throughout the years for the traditional aircraft design can be seamlessly applied to support the development of conforming AHM solutions.
3. While there is an abundance of performance metrics for the evaluation of diagnostics and prognostics algorithm performance, research on how to translate High Level (HL) system specifications into actionable performance requirements is sparse.
4. Current maintenance practices and qualification requirements are strongly based on the assumption of legacy RCM approaches with repetitive manual functional checks and inspections. Therefore, a shift towards automated condition monitoring and remote airworthiness determination requires the definition of new or refined qualification requirements.

Furthermore, it has to be noted that by its design, this review is a purely theoretical work and merely outlines a potential approach for the development of certifiable AHM systems. Consequently, its practicality will need to be assessed through suitable use cases. Additionally, competent regulatory authorities will ultimately need to decide if an AHM solution developed in accordance to this framework will satisfy their requirements to justify issuance of maintenance credits. Therefore, this work can only contribute in showing possible pathways and foster further discussions within the regulatory boards.

Acronyms

AC Advisory Circular.

AHM Aircraft Health Management.

AI Artificial Intelligence.

AIR Aerospace Information Report.

AL Assurance Level.

AMC Acceptable Means of Compliance

AMM Aircraft Maintenance Manual

AMOC Alternative Means of Compliance

ANSP Aircraft Network Security Program

AOG Aircraft on Ground

ARP Aerospace Recommended Practice

BIT Built-In Test

BVID Barely Visible Impact Damage

CAMO Continuing Airworthiness Management Organization

CBA Certification by Analysis

CBM Condition Based Maintenance

CM Condition Monitoring

CMR Certification Maintenance Requirement

ConOps Concept of Operations

COTS Commercial Off-The-Shelf

CS Certification Specification

DAH Design Approval Holder

DET Detailed Inspection

DMC Direct Maintenance Cost

DOC Direct Operating Cost

DPAL Data Process Assurance Level

DQR Data Quality Requirement

E2E End-to-End

EASA European Union Aviation Safety Agency

EOL End of Life

EWIS Electrical Wiring Interconnect System

FAA Federal Aviation Administration

FEC Failure Effect Category

FHA Functional Hazard Assessment

FMEA Failure Mode and Effects Analysis

FTA Fault Tree Analysis

FWS Flight Warning System

GM Guidance Material

GSE Ground Support Equipment

GSIS Ground Support Information Systems

GUM Guide to the Expression of Uncertainty in Measurement

GVI General Visual Inspection

HEART Human Error Assessment and Reduction Technique

HIRF High Intensity Radiated Fields

HL High Level

HM Health Management

HRC Health-ready Component

HUMS Health and Usage Monitoring System

IAHM Integrated Aircraft Health Management

ICA Instructions for Continued Airworthiness

IDG Integrated Drive Generator

IMRBPB International MRB Policy Board

IP Issue Paper

ISHM Integrated System Health Management

ISMS Information Security Management System

ISO International Organization for Standardization

IVHM Integrated Vehicle Health Management

LOT Level of Threat

MCC Maintenance Control Center

MEL Minimum Equipment List

ML Machine Learning

MP Maintenance Program

MRBR Maintenance Review Board Report

MSG-3 Maintenance Steering Group – 3rd Generation

NDT Nondestructive Testing

NSE No Safety Effect

OD Operational Domain

ODD Operational Design Domain

OSA CBM Open System Architecture for Condition Based Maintenance

PCA Principal Component Analysis

PHM Prognostics and Health Management

POD Probability of Detection

PSSA Preliminary System Safety Assessment

RCM Reliability Centered Maintenance

RTS Release-to-Service

RUL Remaining Useful Lifetime

SATAA Sense, Acquire, Transfer, Analyze, Act

SATCOM Satellite Communication

SB Service Bulletin

SDI Special Detailed Inspection

SHM Structural Health Monitoring

SSA System Safety Assessment

STC Supplemental Type Certificate

TBO Time Between Overhaul

TCH Type Certificate Holder

UQ Uncertainty Quantification

V&V Verification & Validation

VHF Very High Frequency

VHM Vibration Health Monitoring

CRediT authorship contribution statement

Robert Meissner: Writing – original draft, Visualization, Resources, Methodology, Investigation, Conceptualization. **Ahmad Ali Pohya:** Writing – review & editing, Supervision, Resources, Project administration. **Oliver Weiss:** Supervision, Methodology, Funding acquisition, Conceptualization. **David Piotrowski:** Writing – review & editing, Validation, Investigation. **Gerko Wende:** Supervision, Resources, Funding acquisition.

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.

Acknowledgments

The authors would like to thank Dr. Heidi M. Meissner and Felix Kranich (Airbus) for their assistance in reviewing this document.

Data availability

No data was used for the research described in the article.

References

[1] M. Hinsch, Industrial Aviation Management, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-662-54739-7, 2019, http://dx.doi.org/10.1007/978-3-662-54740-3.

[2] N. Hölzel, Ein Bewertungsansatz zur Analyse von Zustandsmanagementsystemen in Verkehrsflugzeugen unter Berücksichtigung neuer Instandhaltungskonzepte (Ph.D. thesis), DLR, 2019, http://dx.doi.org/10.15480/882.2543.

[3] IATA, Aircraft Operational Availability, first ed., International Air Transport Association, Montréal, ISBN: 978-92-9229-712-1, 2018, URL: https://www.iata.org/contentassets/bf8ca67c8bcd4358b3d004b0d6d0916f/aoa-1sted-2018.pdf.

[4] O. Weiss, Maintenance of tomorrow: The AHM path from airbus' perspective, 2018, URL: https://www.iata.org/contentassets/fafa409c883d41198aeb87628c848851/1115_oliver_weiss_iata-pao-conference—airbus-ahm-presentation-v2.pdf.

[5] R. Ahmad, S. Kamaruddin, An overview of time-based and condition-based maintenance in industrial application, Comput. Ind. Eng. 63 (1) (2012) 135–149, http://dx.doi.org/10.1016/j.cie.2012.02.002.

[6] R.K. Mobley, An Introduction to Predictive Maintenance, second ed., Butterworth-Heinemann, Amsterdam and New York, ISBN: 978-0-7506-7531-4, 2002.

[7] R. Meissner, A. Rahn, K. Wicke, Developing prescriptive maintenance strategies in the aviation industry based on a discrete-event simulation framework for post-prognostics decision making, Reliab. Eng. Syst. Saf. 214 (2021) http://dx.doi.org/10.1016/j.ress.2021.107812.

[8] R. Meissner, H. Meyer, K. Wicke, Concept and economic evaluation of prescriptive maintenance strategies for an automated condition monitoring system, Int. J. Progn. Heal. Manag. 12 (3) (2021) http://dx.doi.org/10.36001/IJPHM.2021.v12.i3.2911, URL: http://papers.phmsociety.org/index.php/ijphm/article/view/2911.

[9] P. Söderholm, Continuous improvements of complex technical systems: A theoretical quality management framework supported by requirements management and health management, Total. Qual. Manag. Bus. Excel. 15 (4) (2004) 511–525, http://dx.doi.org/10.1080/1478336042000183569.

[10] A.K. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mech. Syst. Signal Process. 20 (7) (2006) 1483–1510, http://dx.doi.org/10.1016/j.ymssp.2005.09.012.

[11] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Machinery health prognostics: A systematic review from data acquisition to RUL prediction, Mech. Syst. Signal Process. 104 (2018) 799–834, http://dx.doi.org/10.1016/j.ymssp.2017.11.016.

[12] E. Zio, Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice, Reliab. Eng. Syst. Saf. 218 (2022) 108119, http://dx.doi.org/10.1016/j.ress.2021.108119.

[13] I.K. Jennions (Ed.), Integrated Vehicle Health Management: The Technology, in: Integrated Vehicle Health Management Ser, SAE International, Warrendale, ISBN: 9780768079524, 2013.

[14] M. Fudge, T.R. Stagliano, S. Tsiao, FAA, Non-traditional flight safety systems and integrated vehicle health management systems [electronic resource]: Descriptions of proposed and existing systems and enabling technologies and verification methods: Final report, 2003, URL: https://nla.gov.au/nla.cat-vn4183524.

[15] M.J. Roemer, C.S. Byington, G.J. Kacprzynski, G. Vachtsevanos, An overview of selected prognostic technologies with application to engine health management, in: Volume 2: Aircraft Engine; Ceramics; Coal, Biomass and Alternative Fuels; Controls, Diagnostics and Instrumentation; Environmental and Regulatory Affairs, ASMEDC, ISBN: 0-7918-4237-1, 2006, pp. 707–715, http://dx.doi.org/10.1115/GT2006-90677.

[16] G.W. Vogl, B.A. Weiss, M.A. Donmez, Standards for prognostics and health management (PHM) techniques within manufacturing operations, in: Proceedings of the Annual Conference of the PHM Society 2014, 2014, http://dx.doi.org/10.36001/phmconf.2014.v6i1.2503.

[17] FAA, Operational authorization of integrated aircraft health management systems: AC 43-218, 2022, URL: https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_43-218_FAA_Web.pdf.

[18] EASA, Certification specifications, acceptable means of compliance and guidance material for large rotorcraft (amendment 11): CS-29, 2023, URL: https://www.easa.europa.eu/en/document-library/certification-specifications/cs-29-amendment-11.

[19] A. Arnaiz, S. Ferreiro, M. Buderath, New decision support system based on operational risk assessment to improve aircraft operability, Proc. Inst. Mech. Eng. Part O: J. Risk Reliab. 224 (3) (2010) 137–147, http://dx.doi.org/10.1243/1748006XJRR282.

[20] C.R. Farrar, K. Worden, An introduction to structural health monitoring, Philos. Trans. Ser. A, Math. Phys. Eng. Sci. 365 (1851) (2007) 303–315, http://dx.doi.org/10.1098/rsta.2006.1928.

[21] IATA, From aircraft health monitoring to aircraft health management: White paper on AHM – position updates, 2023, URL: https://www.iata.org/contentassets/bf8ca67c8bcd4358b3d004b0d6d0916f/ahm-wp-2nded-2023.pdf.

[22] S. Uckun, K. Goebel, P.J. Lucas, Standardizing research methods for prognostics, in: 2008 International Conference on Prognostics and Health Management, IEEE, ISBN: 978-1-4244-1935-7, 2008, pp. 1–10, http://dx.doi.org/10.1109/PHM.2008.4711437.

[23] D. Piotrowski, Implementation of SHM at Delta air lines, in: 12th International Symposium on NDT in Aerospace 2020, 2020, URL: https://www.ndt.net/?id=25638.

[24] S. Lang, AVIATAR – deep dive into prediction with AVIATAR with in-service examples from airlines, PHM Soc. Asia-Pacific Conf. 4 (1) (2023) http://dx.doi.org/10.36001/phmap.2023.v4i1.3776.

[25] D. Piotrowski, Health monitoring in civil aviation, in: 1st International Conference for CBM in Aerospace, 2022.

[26] J. Groenenboom, The changing MRO landscape, 2019, URL: https://www.iata.org/contentassets/81005748740046de878439e6c54f2355/d1-1100-1130-the-changing-mro-landscape_icf.pdf.

[27] A. Kählert, S. Giljohann, U. Klingauf, Cost-benefit analysis and specification of component-level PHM systems in aircraft, Univers. J. Mech. Eng. 4 (4) (2016) 88–98, http://dx.doi.org/10.13189/ujme.2016.040403.

[28] A. Hess, G. Calvello, P. Frith, Challenges, issues, and lessons learned chasing the "Big P". Real predictive prognostics. Part 1, in: D.A. Williamson (Ed.), 2005 IEEE Aerospace Conference Proceedings, IEEE, Piscataway, N.J., ISBN: 0-7803-8870-4, 2005, pp. 3610–3619, http://dx.doi.org/10.1109/AERO.2005.1559666.

[29] A. Hess, G. Calvello, P. Frith, S.J. Engel, D. Hoitsma, Challenges, issues, and lessons learned chasing the "Big P": Real predictive prognostics part 2, in: 2006 IEEE Aerospace Conference, IEEE, ISBN: 0-7803-9545-X, 2006, pp. 1–19, http://dx.doi.org/10.1109/AERO.2006.1656124.

[30] J.P. Sprong, X. Jiang, H. Polinder, A deployment of prognostics to optimize aircraft maintenance - A literature review, Annu. Conf. the PHM Soc. 11 (1) (2019) http://dx.doi.org/10.36001/phmconf.2019.v11i1.776, URL: http://phmpapers.org/index.php/phmconf/article/view/776.

[31] P.A. Sandborn, C. Wilkinson, A maintenance planning and business case development model for the application of prognostics and health management (PHM) to electronic systems, Microelectron. Reliab. 47 (12) (2007) 1889–1901, http://dx.doi.org/10.1016/j.microrel.2007.02.016.

[32] K.R. Wheeler, T. Kurtoglu, S.D. Poll, A survey of health management user objectives related to diagnostic and prognostic metrics, in: 29th Computers and Information in Engineering Conference, in: Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, ASME, New York, N.Y., ISBN: 978-0-7918-4899-9, 2010, pp. 1287–1298, http://dx.doi.org/10.1115/DETC2009-87073.

[33] M.J. Ashby, R.J. Byer, An approach for conducting a cost benefit analysis of aircraft engine prognostics and health management functions, in: 2002 IEEE Aerospace Conference Proceedings, IEEE, Piscataway, NJ, ISBN: 0-7803-7231-X, 2002, 6–2847–6–2856, http://dx.doi.org/10.1109/AERO.2002.1036124.

[34] L.R. Rodrigues, T. Yoneyama, C.L. Nascimento, How aircraft operators can benefit from PHM techniques, in: 2012 IEEE Aerospace Conference, IEEE, ISBN: 978-1-4577-0556-4, 2012, pp. 1–8, http://dx.doi.org/10.1109/AERO.2012.6187376.

[35] B.P. Leao, K.T. Fitzgibbon, L.C. Puttini, G.P.B. de Melo, Cost-benefit analysis methodology for PHM applied to legacy commercial aircraft, in: 2008 IEEE Aerospace Conference, IEEE, ISBN: 978-1-4244-1487-1, 2008, pp. 1–13, http://dx.doi.org/10.1109/AERO.2008.4526599.

[36] A.G. Starr, A structured approach to the selection of condition based maintenance, in: Fifth International Conference on FACTORY 2000 - the Technology Exploitation Process, IEEE, ISBN: 0 85296 682 2, 1997, pp. 131–138, http://dx.doi.org/10.1049/cp:19970134.

[37] M. Haarman, P. de Klerk, P. Decaigny, M. Mulders, C. Vassiliadis, H. Sijtsema, I. Gallo, Predictive maintenance 4.0: Beyond the hype, 2018, URL: https://www.pwc.nl/nl/assets/documents/pwc-predictive-maintenance-beyond-the-hype-40.pdf.

[38] C. Teubert, A.A. Pohya, G. Gorospe, An analysis of barriers preventing the widespread adoption of predictive and prescriptive maintenance in aviation, 2023, URL: https://ntrs.nasa.gov/api/citations/20230000841/downloads/Teubert_An%20Analysis%20of%20Barriers%20Preventing%20the%20Widespread%20Adoption%20of%20Predictive%20and%20Prescriptive%20Maintenance%20in%20Aviation%20(1).pdf.

[39] W.J.C. Verhagen, B.F. Santos, F. Freeman, P. van Kessel, D. Zarouchas, T. Loutas, R.C.K. Yeun, I. Heiets, Condition-based maintenance in aviation: Challenges and opportunities, Aerosp. 10 (9) (2023) 762, http://dx.doi.org/10.3390/aerospace10090762.

[40] G. Vachtsevanos, R. Rajamani, Criticality of prognostics in the operations of autonomous aircraft, SAE Int. J. Aerosp. 16 (3) (2023) http://dx.doi.org/10.4271/01-16-03-0022.

[41] K. Goebel, R. Rajamani, Policy, regulations and standards in prognostics and health management, Int. J. Progn. Heal. Manag. 12 (1) (2021) http://dx.doi.org/10.36001/ijphm.2021.v12i1.2908.

[42] FAA, Certification of transport category rotorcraft: AC 29-2C, 1999, URL: https://www.faa.gov/documentlibrary/media/advisory_circular/ac_29-2c.pdf.

[43] EASA, Certification specification for large aeroplanes (amendment 28): CS-25, 2023, URL: https://www.easa.europa.eu/en/document-library/certification-specifications/cs-25-amendment-28.

[44] R. Rajamani, Unsettled issues concerning integrated vehicle health management systems and maintenance credits, 2020, http://dx.doi.org/10.4271/EPR2020006, URL: https://www.sae.org/publications/technical-papers/content/epr2020006/.

[45] International MRB Policy Board, International MRB/MTB process standard (IMPS), 2022, URL: https://www.easa.europa.eu/en/domains/aircraft-products/international-maintenance-review-board-policy-board-IMRBPB.

[46] Department of Defense, MIL-HDBK-516C: Airworthiness certification criteria, 2014, URL: http://everyspec.com/MIL-HDBK/MIL-HDBK-0500-0599/MIL-HDBK-516C_52120/.

[47] F. de Florio, Airworthiness: An Introduction to Aircraft Certification and Operations, third ed., Butterworth-Heinemann, Oxford, ISBN: 978-0081008881, 2016.

[48] European Union, Regulation (EU) 2018/1139 of the European parliament and of the council, 2018, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2018:212:FULL.

[49] European Union, Commission regulation (EU) no 748/2012, 2012, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2012:224:FULL.

[50] European Union, Commission regulation (EU) no 1321/2014, 2014, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2014:362:FULL.

[51] Airlines for America, Operator / manufacturer scheduled maintenance: MSG-3 volume 1 - fixed wing aircraft, 2022.

[52] Airlines for America, Operator / manufacturer scheduled maintenance: MSG-3 volume 2 - rotorcraft, 2022.

[53] F.S. Nowlan, H.F. Heap, Reliability-Centered Maintenance, United Airlines, San Francisco, 1978, URL: https://apps.dtic.mil/dtic/tr/fulltext/u2/a066579.pdf.

[54] B.S. Dhillon, Engineering Maintenance: A Modern Approach, CRC Press, Boca Raton, FL, ISBN: 1-58716-142-7, 2002.

[55] A. Ahmadi, P. Söderholm, U. Kumar, On aircraft scheduled maintenance program development, J. Qual. Maint. Eng. 16 (3) (2010) 229–255, http://dx.doi.org/10.1108/13552511011072899.

[56] B.S. Blanchard, J. Blyler, System Engineering Management, Fifth edition, in: Wiley series in systems engineering and management, John Wiley & Sons, Inc, Hoboken, New Jersey, ISBN: 9781119047827, 2016.

[57] P. Wang, Civil Aircraft Electrical Power System Safety Assessment: Issues and Practices, Elsevier Science, Saint Louis, ISBN: 978-0-08-100721-1, 2017.

[58] M. Rausand, Reliability centered maintenance, Reliab. Eng. Syst. Saf. 60 (1998) 121–132.

[59] U. Papa, S. Ponte, Preliminary design of an unmanned aircraft system for aircraft general visual inspection, Electron. 7 (12) (2018) 435, http://dx.doi.org/10.3390/electronics7120435.

[60] C. Che, H. Wang, Q. Fu, X. Ni, Combining multiple deep learning algorithms for prognostic and health management of aircraft, Aerosp. Sci. Technol. 94 (2019) 105423, http://dx.doi.org/10.1016/j.ast.2019.105423.

[61] T.B. Sheridan, Telerobotics, Automation, and Human Supervisory Control, MIT Press, Cambridge Mass., ISBN: 0-262-19316-7, 1992.

[62] L. Save, B. Feuerberg, Designing human-automation interaction: A new level of automation taxonomy, in: D. de Waard, K. Brookhuis, F. Dehais, C. Weikert, S. Röttger, D. Manzey, S. Biede, F. Reuzeau, P. Terrier (Eds.), Human Factors: A View from an Integrative Perspective, ISBN: 978-0-945289-44-9, 2012, pp. 43–56, URL: https://www.hfes-europe.org/wp-content/uploads/2014/06/Save.pdf.

[63] T.B. Sheridan, W.L. Verplank, Human and computer control of undersea teleoperators, 1978, URL: https://apps.dtic.mil/sti/pdfs/ADA057655.pdf.

[64] J. Nuñez, A. Colomer, S. Bonelli, J.-P. Imbert, G. Granger, Harvis: D1.1 state of the art review, 2019, URL: https://www.harvis-project.eu/wp-content/uploads/2020/03/HARVIS-D1.1_SoA_v3.1.pdf.

[65] R.W. Proud, J.J. Hart, R.B. Mrozinski, Methods for determining the level of autonomy to design into a human spaceflight vehicle: A function specific approach, in: Performance Metrics for Intelligent Systems Workshop, 2003, URL: https://ntrs.nasa.gov/api/citations/20100017272/downloads/20100017272.pdf.

[66] R. Parasuraman, T.B. Sheridan, C.D. Wickens, A model for types and levels of human interaction with automation, IEEE Trans. Syst. Man Cybern. Part A Syst. Humans : A Publ. the IEEE Syst. Man Cybern. Soc. 30 (3) (2000) 286–297, http://dx.doi.org/10.1109/3468.844354.

[67] EASA, EASA concept paper: Guidance for level 1 & 2 machine learning applications, 2024, URL: https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-concept-paper-issue-2.

[68] DIN, 13306:2010: Maintenance - maintenance terminology, 2010.

[69] Department of Defense, Condition based maintenance plus DoD guidebook, 2008, URL: https://www.dau.edu/guidebooks/Shared%20Documents/Condition%20Based%20Maintenance%20Plus%20(CBM+)%20Guidebook.pdf.

[70] M.G. Thurston, An open standard for web-based condition-based maintenance systems, in: 2001 IEEE Autotestcon Proceedings. IEEE Systems Readiness Technology Conference. (Cat. No.01CH37237), IEEE, ISBN: 0-7803-7094-5, 2001, pp. 401–415, http://dx.doi.org/10.1109/AUTEST.2001.949021.

[71] S. Khan, Towards MRO 4.0: Challenges for digitalization and mapping emerging technologies, 2023, http://dx.doi.org/10.4271/EPR2023007.

[72] SAE, ARP 6887: Validation and verification of integrated vehicle health management systems, 2024, http://dx.doi.org/10.4271/ARP6887, URL: https://www.sae.org/standards/content/arp6887.

[73] SAE, AIR 6334: A guide to extending times between overhaul for rotorcraft power train transmissions using monitoring data, 2020, http://dx.doi.org/10.4271/AIR6334, URL: https://www.sae.org/standards/content/air6334/.

[74] SAE, ARP 5987: A process for utilizing aerospace propulsion health management systems for maintenance credit, 2018, http://dx.doi.org/10.4271/ARP5987A, URL: https://www.sae.org/standards/content/arp5987a/.

[75] EASA, IP 180 aircraft health monitoring (AHM) integration in MSG-3, 2018, URL: https://www.easa.europa.eu/download/imrbpb/IP%20180%20-%20AHM%20integration%20in%20MSG-3.pdf.

[76] EASA, IP 197 amendment to IP180 to clarify system features to be certified by type certification staff, 2021, URL: https://www.easa.europa.eu/sites/default/files/dfu/ip_197_-_amendment_to_ip180.pdf.

[77] SAE, (WIP) ARP 7122: Utilizing Integrated Vehicle Health Management Systems for Airworthiness Credit. (not published). URL: https://www.sae.org/standards/content/arp7122/.

[78] FAA, Installation of structural monitoring systems comparative vacuum monitor (CVM) sensors: STC ST04103ny, 2022.

[79] D. Piotrowski, Navigating civil aviation regulations for health monitoring, in: 1st International Conference for CBM in Aerospace, 2022.

[80] J.R. Linn, D. Roach, Moving SHM into routine use - progress in CVM applications, 2019, URL: https://a4andtforum.com/wp-content/uploads/2019/09/180930-John-Linn-Moving-SHM-into-Routine-Use.pdf.

[81] SAE, (WIP) AIR 8621: Perspectives on Qualification of SHM Systems for Detection of Damage in Structure. (not published). URL: https://www.sae.org/standards/content/air8621/.

[82] W. Doeland, Certification by analysis, or: Modelling & simulation, 2019.

[83] E. Lewis, W. Doeland, EASA contribution to DLR workshop: Certification by analysis in HYTAZER, 2024.

[84] NASA, A guide for aircraft certification by analysis, 2021, URL: https://ntrs.nasa.gov/api/citations/20210015404/downloads/NASA-CR-20210015404%20updated.pdf.

[85] SAE, ARP 6883: Guidelines for writing IVHM requirements for aerospace systems, 2019, http://dx.doi.org/10.4271/ARP6883, URL: https://www.sae.org/standards/content/arp6883/.

[86] R. Plant, R. Gamble, Methodologies for the development of knowledge-based systems, 1982–2002, Knowl. Eng. Rev. 18 (1) (2003) 47–81, http://dx.doi.org/10.1017/S026988890300064X.

[87] L.E. Schwer, An overview of the ASME V&V-10 guide for verification and validation in computational solid mechanics, in: 20th International Conference on Structural Mechanics in Reactor Technology, SMiRT 20, 2009, URL: https://repository.lib.ncsu.edu/server/api/core/bitstreams/a45e16f4-5cce-44f4-b2bd-fa760019c8aa/content.

[88] I. Babuska, J. Oden, Verification and validation in computational engineering and science: Basic concepts, Comput. Methods Appl. Mech. Eng. 193 (36–38) (2004) 4057–4066, http://dx.doi.org/10.1016/j.cma.2004.03.002.

[89] NASA, NASA handbook for models and simulation: An implementation guide for NASA-STD-7009a, 2019, URL: https://standards.nasa.gov/sites/default/files/standards/NASA/A/0/nasa-hdbk-7009a.pdf.

[90] IEEE, IEEE standard for software verification and validation, 1998.

[91] O. Schwabe, E. Shehab, J. Erkoyuncu, Uncertainty quantification metrics for whole product life cycle cost estimates in aerospace innovation, Prog. Aerosp. Sci. 77 (2015) 1–24, http://dx.doi.org/10.1016/j.paerosci.2015.06.002.

[92] JCGM, Evaluation of measurement data - guide to the expression of uncertainty in measurement, 2008, http://dx.doi.org/10.59161/JCGM100-2008E.

[93] W.E. Walker, P. Harremoës, J. Rotmans, J.P. van der Sluijs, M. van Asselt, P. Janssen, M.P. Krayer von Krauss, Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support, Integr. Assess. 4 (1) (2003) 5–17, http://dx.doi.org/10.1076/iaij.4.1.5.16466.

[94] D. Marikyan, S. Papagiannidis, G. Stewart, Technology acceptance research: Meta-analysis, J. Inf. Sci. (2023) http://dx.doi.org/10.1177/01655515231191177.

[95] Y. Hwang, Investigating enterprise systems adoption: uncertainty avoidance, intrinsic motivation, and the technology acceptance model, Eur. J. Inf. Syst. 14 (2) (2005) 150–161, http://dx.doi.org/10.1057/palgrave.ejis.3000532.

[96] S. Sankararaman, Significance, interpretation, and quantification of uncertainty in prognostics and remaining useful life prediction, Mech. Syst. Signal Process. 52–53 (2015) 228–247, http://dx.doi.org/10.1016/j.ymssp.2014.05.029.

[97] W.L. Oberkampf, C.J. Roy, Verification and Validation in Scientific Computing, Cambridge University Press, ISBN: 9780521113601, 2013, http://dx.doi.org/10.1017/cbo9780511760396.

[98] International Organization for Standardization, ISO 31000: Risk management - guidelines, 2018.

[99] B. Thacker, Uncertainty quantification and validation assessment, 2016, URL: https://faa.niar.wichita.edu/Portals/0/Uncertainty%20Quantification%20and%20Validation%20ASsessment%20-%20Thacker-FAA_Workshop.pdf.

[100] EASA, Notification of a proposal to issue a certification memorandum: Modelling & simulation – CS-25 structural certification specifications, 2020, URL: https://www.easa.europa.eu/en/proposed-cm-s-014-modelling-simulation-consultation#group-downloads.

[101] B. de Jonge, P.A. Scarf, A review on maintenance optimization, Eur. J. Oper. Res. 285 (3) (2020) 805–824, http://dx.doi.org/10.1016/j.ejor.2019.09.047.

[102] T. Wilmering (Ed.), Integrated Vehicle Health Management: Systems of Systems Integration, SAE International, Warrendale, Pennsylvania, ISBN: 9780768084290, 2017.

[103] International Organization for Standardization, ISO 13374-1: Condition monitoring and diagnostics of machines - data processing, communication and presentation: Part 1: General guidelines, 2003.

[104] IEEE, IEEE standard framework for prognostics and health management of electronic systems, 2017, http://dx.doi.org/10.1109/IEEESTD.2017.8227036.

[105] T. Sreenuch, A. Tsourdos, I.K. Jennions, Distributed embedded condition monitoring systems based on OSA-cbm standard, Comput. Stand. Interfaces 35 (2) (2013) 238–246, http://dx.doi.org/10.1016/j.csi.2012.10.002.

[106] B.D. Larder, M.W. Davis, HUMS condition based maintenance credit validation, in: 63rd American Helicopter Society International Annual Forum, Vertical Flight Society, ISBN: 9781617829307, 2007.

[107] International Organization for Standardization, ISO 15288: Systems and software engineering - system life cycle processes, 2023.

[108] International Organization for Standardization, ISO 17359: Condition monitoring and diagnostics of machines – general guidelines, 2018.

[109] SAE, ARP 6290: Guidelines for the development of architectures for integrated vehicle health management systems, 2023, http://dx.doi.org/10.4271/ARP6290, URL: https://www.sae.org/standards/content/arp6290.

[110] M. Bengtsson, Condition Based Maintenance Systems: An Investigation of Technical Constituents and Organizational Aspects (Ph.D. thesis), Mälardalen University, Västerås, 2004.

[111] M. Lebold, K. Reichard, D. Boylan, Utilizing DCOM in an open system architecture framework for machinery monitoring and diagnostics, in: 2003 IEEE Aerospace Conference Proceedings (Cat. No.03TH8652), IEEE, ISBN: 0-7803-7651-X, 2003, pp. 1227–1236, http://dx.doi.org/10.1109/AERO.2003.1235237.

[112] EASA, Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation, 2023, URL: https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20.

[113] High-Level Expert Group on Artificial Intelligence, Ethics guidelines for trustworthy AI, 2019.

[114] SAE, ARP 5580: Recommended failure modes and effects analysis (FMEA) practices for non-automobile applications, 2020, http://dx.doi.org/10.4271/ARP5580, URL: https://www.sae.org/standards/content/arp5580/.

[115] A. Saxena, I. Roychoudhury, J. Celaya, B. Saha, S. Saha, K. Goebel, Requirements flowdown for prognostics and health management, in: Infotech@Aerospace 2012, American Institute of Aeronautics and Astronautics, Reston, Virigina, ISBN: 978-1-60086-939-6, 2012, http://dx.doi.org/10.2514/6.2012-2554.

[116] H. Bello, D. Geißler, L. Ray, S. Müller-Divéky, P. Müller, S. Kittrell, M. Liu, B. Zhou, P. Lukowicz, Towards certifiable AI in aviation: Landscape, challenges, and opportunities, 2024, http://dx.doi.org/10.48550/arXiv.2409.08666.

[117] European Union, Commission implementing regulation (EU) 2022/1426, 2022, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2022:221:FULL.

[118] SAE, J 3016: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles, 2021, URL: https://www.sae.org/standards/content/j3016/.

[119] SAE, ARP 6275: Determination of costs and benefits from implementing an integrated vehicle health management system, 2021, http://dx.doi.org/10.4271/ARP6275A, URL: https://www.sae.org/standards/content/arp6275a/.

[120] A. Saxena, I. Roychoudhury, K. Goebel, W. Lin, Towards requirements in systems engineering for aerospace IVHM design, in: AIAA Infotech@Aerospace (I@a) Conference, American Institute of Aeronautics and Astronautics, Reston, Virginia, 2013, http://dx.doi.org/10.2514/6.2013-4659.
[121] RTCA, DO-200: Standards for processing aeronautical data, 2024.
[122] D.A. Tobon-Mejia, K. Medjaher, N. Zerhouni, The ISO 13381-1 standard's failure prognostics process through an example, in: 2010 Prognostics and System Health Management Conference, PHM-2010 Macau, IEEE, Piscataway, NJ, ISBN: 978-1-4244-4756-5, 2010, pp. 1–12, http://dx.doi.org/10.1109/PHM.2010.5413482.
[123] G.W. Vogl, B.A. Weiss, M.A. Donmez, Standards related to prognostics and health management (PHM) for manufacturing, 2014, http://dx.doi.org/10.6028/NIST.IR.8012.
[124] MLEAP Consortium, EASA research – machine learning application approval (MLEAP) final report, 2024, URL: https://www.easa.europa.eu/sites/default/files/dfu/mleap-d4-public-report-issue01.pdf.
[125] International Organization for Standardization, ISO 25012: Software engineering - software product quality requirements and evaluation - data quality model, 2008.
[126] International Organization for Standardization, ISO 5259-2: Artificial intelligence - data quality for analytics and machine learning (ML): Part 2: Data quality measures, 2024.
[127] N.A. Emran, Data completeness measures, in: A. Abraham, A.K. Muda, Y.-H. Choo (Eds.), Pattern Analysis, Intelligent Security and the Internet of Things, in: Advances in Intelligent Systems and Computing, vol. 355, Springer International Publishing, Cham, ISBN: 978-3-319-17397-9, 2015, pp. 117–130, http://dx.doi.org/10.1007/978-3-319-17398-6{_}11.
[128] International Organization for Standardization, ISO 13373-1: Condition monitoring and diagnostics of machines - vibration condition monitoring: Part 1: General procedures, 2002.
[129] RTCA, DO-160: Environmental conditions and test procedures for airborne equipment, 2010.
[130] RTCA, DO-178: Software considerations in airborne systems and equipment certification, 2011.
[131] SAE, ARP 6461: Guidelines for implementation of structural health monitoring on fixed wing aircraft, 2021.
[132] SAE, (WIP) ARP 6835: Propulsion System Monitoring for Continued Airworthiness. (not published). URL: https://www.sae.org/standards/content/arp6835/.
[133] A.A. Pohya, K. Wicke, T. Kilian, Introducing variance-based global sensitivity analysis for uncertainty enabled operational and economic aircraft technology assessment, Aerosp. Sci. Technol. 122 (2022) 107441, http://dx.doi.org/10.1016/j.ast.2022.107441.
[134] D. Gonzalez-Jimenez, J. Del-Olmo, J. Poza, F. Garramiola, P. Madina, Data-driven fault diagnosis for electric drives: A review, Sensors (Basel, Switzerland) 21 (12) (2021) http://dx.doi.org/10.3390/s21124024.
[135] J. SEKERA, A. NOVÁK, The future of data communication in aviation 4.0 environment, INCAS BULLETIN 13 (3) (2021) 165–178, http://dx.doi.org/10.13111/2066-8201.2021.13.3.14.
[136] SAE, JA 6268: Design & run-time information exchange for health-ready components, 2023, http://dx.doi.org/10.4271/JA6268{_}202303, URL: https://www.sae.org/standards/content/ja6268_202303/.
[137] Civil Aviation Authority, CAP 753: Helicopter vibration health monitoring (VHM): Guidance material for operators utilising VHM in rotor and rotor drive systems of helicopters, 2018, URL: https://www.caa.co.uk/publication/download/13249.
[138] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, From data mining to knowledge discovery in databases, AI Mag. 17 (3) (1996) 37–54, http://dx.doi.org/10.1609/aimag.v17i3.1230, URL: https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1230.
[139] International Organization for Standardization, ISO 13373-2: Condition monitoring and diagnostics of machines - vibration condition monitoring: Part 2: Processing, analysis and presentation of vibration data, 2016.
[140] C. Cremona, J. Santos, Structural health monitoring as a big-data problem, Struct. Eng. Int. 28 (3) (2018) 243–254, http://dx.doi.org/10.1080/10168664.2018.1461536.
[141] RTCA, DO-278: Software integrity assurance considerations for communication, navigation, surveillance and air traffic management (CNS/ATM) systems, 2011.
[142] European Union, Commission delegated regulation (EU) 2022/1645, 2022, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2022:248:FULL.
[143] European Union, Commission implementing regulation (EU) 2023/203, 2023, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2023:031:FULL.
[144] FAA, Operational authorization of aircraft network security program: AC 119-1A, 2023, URL: https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_119-1A.pdf.
[145] EASA, Executive director decision 2023/009/R, 2023, URL: https://www.easa.europa.eu/en/document-library/agency-decisions/ed-decision-2023009r.
[146] International Organization for Standardization, ISO 27001: Information security, cybersecurity and privacy protection – information security management systems – requirements, 2023.
[147] RTCA, DO-326: Airworthiness security process specification, 2014.
[148] RTCA, DO-355: Information security guidance for continued airworthiness, 2020.
[149] RTCA, DO-356: Airworthiness security methods and considerations, 2018.
[150] C. Torens, Safety versus security in aviation, comparing DO-178C with security standards, in: AIAA Scitech 2020 Forum, American Institute of Aeronautics and Astronautics, Reston, Virginia, ISBN: 978-1-62410-595-1, 2020, http://dx.doi.org/10.2514/6.2020-0242.
[151] SAE, ARP 4754: Guidelines for development of civil aircraft and systems, 2023, http://dx.doi.org/10.4271/ARP4754B, URL: https://www.sae.org/standards/content/arp4754b.
[152] International Organization for Standardization, ISO 9797-1: Information technology - security techniques - message authentication codes: Part 1: Mechanisms using a Block Cipher 2011.
[153] International Organization for Standardization, ISO 9797-2: Information technology - security techniques - message authentication codes: Part 2: Mechanisms using a dedicated hash-function, 2021.
[154] International Organization for Standardization, ISO 9797-3: Information technology - security techniques - message authentication codes: Part 3: Mechanisms using a universal hash-function, 2011.
[155] CAAS, Aircraft network security program (ANSP): AC 121-7-2, 2018, URL: https://www.caas.gov.sg/docs/default-source/pdf/ac121-7-2(rev-0)-aircraft-network-security-progamme-(ansp).pdf.
[156] Department of Defense, JSSG-2001b: Air vehicle, 2004, URL: http://everyspec.com/USAF/USAF-General/JSSG-2001B_25156/.
[157] International Organization for Standardization, ISO 13381-1: Condition monitoring and diagnostics of machines - prognostics: Part 1: General guidelines, 2015.
[158] K.B. Marais, M.R. Robichaud, Analysis of trends in aviation maintenance risk: An empirical approach, Reliab. Eng. Syst. Saf. 106 (2012) 104–118, http://dx.doi.org/10.1016/j.ress.2012.06.003.
[159] Airbus, Maintenance briefing notes: Human performance. Error management, 2007.
[160] B.S. Dhillon, Human Reliability, Error, and Human Factors in Engineering Maintenance: With Reference to Aviation and Power Generation, CRC Press, Boca Raton, ISBN: 978-1-4398-0383-7, 2009.
[161] A. Hobbs, A. Williamson, Aircraft maintenance safety survey - results, 2000, URL: https://www.atsb.gov.au/sites/default/files/media/30080/sir199706_002.pdf.
[162] K.A. Latorella, P.V. Prabhu, A review of human error in aviation maintenance and inspection, Int. J. Ind. Ergon. 26 (2) (2000) 133–161, http://dx.doi.org/10.1016/S0169-8141(99)00063-3.
[163] W. Rankin, R. Hibit, J. Allen, R. Sargent, Development and evaluation of the maintenance error decision aid (MEDA) process, Int. J. Ind. Ergon. 26 (2) (2000) 261–276, http://dx.doi.org/10.1016/S0169-8141(99)00070-0.
[164] Civil Aviation Authority, CAP 718: Human factors in aircraft maintenance and inspection, 2002, URL: https://www.caa.co.uk/publication/download/12248.
[165] S.M. Goldman, E.R. Fiedler, R.E. King, General aviation maintenance-related accidents: A review of ten years of NTSB data, 2002, URL: https://rosap.ntl.bts.gov/view/dot/57936.
[166] B.S. Dhillon, Human Reliability and Error in Transportation Systems, in: Springer series in reliability engineering, Springer, London, ISBN: 978-1-84628-811-1, 2007.
[167] FAA, Aviation maintenance technician handbook - general, 2023, URL: https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/amtg_handbook.pdf.
[168] A. Hobbs, An overview of human factors in aviation maintenance, 2008, URL: https://www.sec.gov/comments/s7-07-18/s70718-4171034-172090.pdf.
[169] N. McDonald, S. Corrigan, C. Daly, S. Cromie, Safety management systems and safety culture in aircraft maintenance organisations, Saf. Sci. 34 (1–3) (2000) 151–176, http://dx.doi.org/10.1016/S0925-7535(00)00011-4.
[170] Civil Aviation Authority, CAP 716: Aviation maintenance human factors (EASA / JAR145 approved organisations), 2003, URL: https://www.caa.co.uk/publication/download/12242.
[171] J.L. Reynolds, C.G. Drury, J. Sharit, F. Cerny, The effects of different forms of space restriction on inspection performance, Proc. Hum. Factors Ergon. Soc. Annual Meet. 38 (10) (1994) 631–635, http://dx.doi.org/10.1177/154193129403801019.
[172] J.C. Williams, A data-based method for assessing and reducing human error to improve operational performance, in: Conference Record for 1988 IEEE Fourth Conference on Human Factors and Power Plants, IEEE, 1988, pp. 436–450, http://dx.doi.org/10.1109/HFPP.1988.27540.
[173] C. Fualdes, Composites at airbus: Damage tolerance methodology, in: FAA (Ed.), Composite Damage Tolerance and Maintenance Workshop, 2006.
[174] L.T. Ostrom, C.A. Wilhelmsen, Developing risk models for aviation maintenance and inspection, Int. J. Aviat. Psychol. 18 (1) (2008) 30–42, http://dx.doi.org/10.1080/10508410701749407.

[175] F.W. Spencer, Visual inspection research project report on benchmark inspections, 1996, URL: https://apps.dtic.mil/sti/pdfs/ADA321199.pdf.
[176] D. Erhart, L.T. Ostrom, C.A. Wilhelmsen, Visual detectability of dents on a composite aircraft inspection specimen: An initial study, Int. J. Appl. Aviat. Stud. 4 (2) (2004) 111–122, URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2e6025d2aa89e6a56934a7694fa7e7adec5f26ed#page=111.
[177] L.G. Williams, M.S. Borow, The effect of rate and direction of display movement upon visual search, Hum. Factors 5 (1963) 139–146, http://dx.doi.org/10.1177/001872086300500204.
[178] C.G. Drury, J. Watson, Good practices in visual inspection, 2002, URL: http://www.dviaviation.com/files/45146949.pdf.
[179] C.G. Drury, The effect of speed of working on industrial inspection accuracy, Appl. Ergon. 4 (1) (1973) 2–7, http://dx.doi.org/10.1016/0003-6870(73)90002-1.
[180] T.P. Carvalho, F.A.A.M.N. Soares, R. Vita, R.P. Da Francisco, J.P. Basto, S.G.S. Alcalá, A systematic literature review of machine learning methods applied to predictive maintenance, Comput. Ind. Eng. 137 (2019) 106024, http://dx.doi.org/10.1016/j.cie.2019.106024.
[181] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, D.J. Inman, A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications, Mech. Syst. Signal Process. 147 (2021) 107077, http://dx.doi.org/10.1016/j.ymssp.2020.107077.
[182] K. Ranasinghe, R. Sabatini, A. Gardi, S. Bijjahalli, R. Kapoor, T. Fahey, K. Thangavel, Advances in integrated system health management for mission-essential and safety-critical aerospace applications, Prog. Aerosp. Sci. 128 (2022) 100758, http://dx.doi.org/10.1016/j.paerosci.2021.100758.
[183] SAE, AIR 7999: Diagnostic and prognostic metrics for aerospace propulsion health management systems, 2020, http://dx.doi.org/10.4271/AIR7999, URL: https://www.sae.org/standards/content/air7999/.
[184] SAE, ARP 5783: Health and usage monitoring metrics monitoring the monitor, 2018, http://dx.doi.org/10.4271/ARP5783, URL: https://www.sae.org/standards/content/arp5783/.
[185] A. Saxena, J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, M. Schwabacher, Metrics for evaluating performance of prognostic techniques, in: 2008 International Conference on Prognostics and Health Management, IEEE, ISBN: 978-1-4244-1935-7, 2008, pp. 1–17, http://dx.doi.org/10.1109/PHM.2008.4711436.
[186] A. Saxena, J. Celaya, B. Saha, S. Saha, K. Goebel, On applying the prognostic performance metrics, in: Annual Conference of the Prognostics and Health Management Society, 2009, URL: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20100023445.pdf.
[187] A. Saxena, J. Celaya, B. Saha, S. Saha, K. Goebel, Metrics for offline evaluation of prognostic performance, Int. J. Progn. Heal. Manag. 1 (1) (2010) URL: http://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2010/ijPHM_10_001.pdf.
[188] A. Saxena, M. Roemer, IVHM assessment metrics, in: I.K. Jennions (Ed.), Integrated Vehicle Health Management: The Technology, in: Integrated Vehicle Health Management Ser, SAE International, Warrendale, ISBN: 9780768079524, 2013, pp. 107–127.
[189] IEEE, Trial-use standard for testability and diagnosability characteristics and metrics, 2005, http://dx.doi.org/10.1109/IEEESTD.2005.95926.
[190] T. Kurtoglu, O.J. Mengshoel, S. Poll, A framework for systematic benchmarking of monitoring and diagnostic systems, in: 2008 International Conference on Prognostics and Health Management, IEEE, ISBN: 978-1-4244-1935-7, 2008, pp. 1–13, http://dx.doi.org/10.1109/PHM.2008.4711454.
[191] R.F. Orsagh, M.J. Roemer, C.J. Savage, M. Lebold, Development of performance and effectiveness metrics for gas turbine diagnostic technologies, in: 2002 IEEE Aerospace Conference Proceedings, IEEE, Piscataway, NJ, ISBN: 0-7803-7231-X, 2002, 6–2825–6–2834, http://dx.doi.org/10.1109/AERO.2002.1036121.
[192] J. Cohen, A coefficient of agreement for nominal scales, Educ. Psychol. Meas. 20 (1) (1960) 37–46, http://dx.doi.org/10.1177/001316446002000104.
[193] J. Breitenstein, J.-A. Termohlen, D. Lipinski, T. Fingscheidt, Systematization of corner cases for visual perception in automated driving, in: 2020 IEEE Intelligent Vehicles Symposium, IV, IEEE, ISBN: 978-1-7281-6673-5, 2020, pp. 1257–1264, http://dx.doi.org/10.1109/IV47402.2020.9304789.
[194] F. Heidecker, J. Breitenstein, K. Rosch, J. Lohdefink, M. Bieshaar, C. Stiller, T. Fingscheidt, B. Sick, An application-driven conceptualization of corner cases for perception in highly automated driving, in: 2021 IEEE Intelligent Vehicles Symposium, IV, IEEE, ISBN: 978-1-7281-5394-0, 2021, pp. 644–651, http://dx.doi.org/10.1109/IV48863.2021.9575933.
[195] J.-A. Bolte, A. Bar, D. Lipinski, T. Fingscheidt, Towards corner case detection for autonomous driving, in: 2019 IEEE Intelligent Vehicles Symposium, IV, IEEE, ISBN: 978-1-7281-0560-4, 2019, pp. 438–445, http://dx.doi.org/10.1109/IVS.2019.8813817.
[196] T. Ouyang, V.S. Marco, Y. Isobe, H. Asoh, Y. Oiwa, Y. Seo, Corner case data description and detection, in: 2021 IEEE/ACM 1st Workshop on AI Engineering - Software Engineering for AI, WAIN, IEEE, ISBN: 978-1-6654-4470-5, 2021, pp. 19–26, http://dx.doi.org/10.1109/WAIN52551.2021.00009.
[197] K. Goebel, What comes after prognostic, 2019.
[198] N. Iyer, K. Goebel, P. Bonissone, Framework for post-prognostic decision support, in: 2006 IEEE Aerospace Conference, IEEE, ISBN: 0-7803-9545-X, 2006, pp. 1–10, http://dx.doi.org/10.1109/AERO.2006.1656108.
[199] RTCA, DO-254: Design assurance guidance for airborne electronic hardware, 2000.
[200] International Organization for Standardization, ISO 13373-3: Condition monitoring and diagnostics of machines - vibration condition monitoring: Part 3: Guidelines for vibration diagnosis, 2015.
[201] International Organization for Standardization, ISO 13373-4: Condition monitoring and diagnostics of machines - vibration condition monitoring: Part 4: Diagnostic techniques for gas and steam turbines with fluid-film bearings, 2022.
[202] International Organization for Standardization, ISO 13373-5: Condition monitoring and diagnostics of machines - vibration condition monitoring: Part 5: Diagnostic techniques for fans and blowers, 2021.
[203] International Organization for Standardization, ISO 13373-9: Condition monitoring and diagnostics of machines - vibration condition monitoring: Part 9: Diagnostic techniques for electric motors, 2018.