Journal of Manufacturing Systems 82 (2025) 449–477

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

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

Review

Comprehensive evaluations of condition monitoring-based technologies in industrial maintenance: A systematic review

Mehdi Dadfarnia [a,b,*], Michael E. Sharp [a], Jeffrey W. Herrmann [c]

[a] National Institute of Standards and Technology, Communications Technology Laboratory, 100 Bureau Dr, Gaithersburg, MD 20899, USA
[b] University of Maryland, College Park, Department of Mechanical Engineering, 4298 Campus Dr, College Park, MD 20742, USA
[c] Catholic University of America, Department of Mechanical Engineering, 620 Michigan Ave NE, Washington DC 20064, USA

ARTICLE INFO

Keywords:
Condition monitoring
Condition-based maintenance
Evaluation methods
Prognostics and health management
Systematic review

ABSTRACT

Condition monitoring involves detecting, diagnosing, or predicting faults or failures in industrial equipment. Given advances in the underlying artificial intelligence solutions and internet of things-based technologies, condition monitoring has the potential to improve industrial maintenance processes rapidly. Adopting condition monitoring-based technologies requires evaluating their engineering and financial benefits to determine whether the investment is justified. An increasing number of studies describe procedures to evaluate condition monitoring-based maintenance, but the literature lacks a review of these evaluation studies to identify research opportunities and best practices. This systematic review aims to report and analyze the evaluation methods for using condition monitoring-based technologies in industrial maintenance. This review identified 465 relevant peer-reviewed studies between 2001 and 2023, from which 42 articles met the eligibility criteria. For each article, this paper analyzed facets of the evaluation process related to the study's characterizations of the industrial application, condition monitoring, maintenance deployment, evaluation techniques, performance measures, and economic analysis. Collectively, these results yield several insights. Few condition monitoring evaluation studies exist for manufacturing systems, unlike the domains of energy systems and transportation modes. Also, many studies lack details about condition monitoring and maintenance models. Additionally, the evaluation techniques across most studies can improve with combinations of analytical frameworks, simulation, and expanded sensitivity analysis. Lastly, the reviewed studies are difficult to directly compare due to heterogeneity in economic analysis, performance measures, and uncertainty analysis — indicating an opportunity for future research to structure comprehensive reporting items to enhance the comparability of domain-specific condition monitoring-based maintenance evaluations. Based on the literature review and analyses, this review suggests specific recommendations for future condition monitoring evaluation and opportunities for further research.

1. Introduction

Emerging advancements in artificial intelligence (AI) and the internet of things (IoT) create new opportunities for improving operations and maintenance in manufacturing as well as broader industrial settings, where enterprises use IoT to sense measurands and exploit data with analytics and AI methods [1]. These AI and IoT advancements have been rapidly embedded into condition monitoring-based technologies, providing users with potential improvements to the maintenance of industrial equipment by way of more options for better-performing anomaly detection, fault diagnosis, and failure prediction [2,3]. However, the impact of AI and IoT technologies on productivity growth has been low [4,5], and their adoption has especially been scant for small- and medium-sized manufacturing enterprises [6].

Maintenance-related uses of condition monitoring face a unique challenge to harness the power of AI and IoT technologies, as maintenance is typically under-prioritized by industrial firms, and novel technological interventions for maintenance are often hindered by factors such as the lack of technological expertise [6] and managerial commitment [7]. Crucially, many manufacturing enterprises perceive the value of maintenance to be invisible, and facilitating the approval of maintenance-related investments requires proof of profits, financial

* Corresponding author at: National Institute of Standards and Technology, Communications Technology Laboratory, 100 Bureau Dr, Gaithersburg, MD 20899, USA.

E-mail addresses: mehdi.dadfarnia@nist.gov (M. Dadfarnia), michael.sharp@nist.gov (M.E. Sharp), herrmannj@cua.edu (J.W. Herrmann).

https://doi.org/10.1016/j.jmsy.2025.06.015
Received 5 April 2025; Received in revised form 23 May 2025; Accepted 10 June 2025
0278-6125/Published by Elsevier Ltd on behalf of The Society of Manufacturing Engineers.

accountability, and strict cost justification [8,9]. The growing opportunities for using condition monitoring-based technologies in manufacturing maintenance, and more broadly, industrial maintenance, have culminated in increased studies that evaluate their engineering and financial benefits and an interest in methods or best practices for conducting those evaluations.

Many terms refer to condition monitoring-based technologies and related maintenance concepts. Condition monitoring systems (CMSs) are collections of software and hardware devices that monitor the state of one or more pieces of industrial equipment and aim to identify and detect any damage or abnormalities [10]. By using techniques such as oil analysis, vibration monitoring, thermography, and acoustic emission [11], CMSs can help operators and maintenance personnel maintain equipment and avoid unexpected downtime and costly repairs.

Though this review commonly uses CMSs to refer to condition monitoring-based technologies, concepts in industrial maintenance that are closely related to CMSs include statistical process control (SPC) [12], non-destructive testing (NDT) [13], and structural health monitoring (SHM) [11]. Like CMSs, SPC, NDT, and SHM refer to technologies that use measuring devices such as sensors to characterize an industrial artifact's (e.g., materials, components, machinery) health state or condition. These technologies provide equipment condition information to maintenance policies. CMS monitoring applications typically focus on (but are not limited to) rotating or reciprocating machinery, where the primary vibration source is from the equipment itself. These machinery can often be found in manufacturing and power plants [14]. SHM is analogous to CMS, but applied to civil, mechanical, and aerospace infrastructures. SPC is process-based and typically uses control charts to identify deviations in a process. NDT-based tools assess the condition of an equipment's material or structure without altering the equipment.

Other key concepts related to CMSs include reliability-centered maintenance (RCM) [15], condition-based maintenance (CBM) [16], predictive maintenance (PdM) [17], and prognostics and health management (PHM) [18]. CBM and PdM are maintenance policies that use CMS-type monitoring technologies, where PdM is an extension of CBM that focuses on equipment condition prognostics. RCM provides a decision support framework for implementing CBM or PdM. RCM helps identify the equipment operating context, failure modes associated with equipment operations, causes and effects of the failure modes, and the planning of maintenance tasks that address the failure modes. PHM approaches entail equipment lifecycle management that incorporates CBM or PdM maintenance policies with condition monitoring and maintenance decision support [11]. Fig. 1 summarizes the relationships between the CMS-adjacent monitoring and maintenance terms.

Several studies have reviewed the uses of CMSs and related technologies in industrial applications [19–22], and especially manufacturing systems [23–26]. However, using condition monitoring tools and technologies does not guarantee benefits to the industrial application [3]. Estimating their performance and their impact on equipment availability, productivity, and maintenance costs is needed to justify investments in condition monitoring technologies and CMS [27,28]. Once equipped, continuous or routine demonstration of benefits from condition-based maintenance may also be required, as monitoring performance may falter due to concept drift or hidden contexts [29].

Many studies review the performance measures that evaluations use to assess condition monitoring [30–33] or maintenance [34,35]. Though helpful, they do not provide methods or techniques to model the industrial application or to evaluate condition monitoring benefits using the performance measures. Moreover, studies that evaluate CMSs and related technologies tend to focus on the algorithmic performance of the AI tools that underlie a CMS [36–38]. These papers do not incorporate any economic justification for using condition monitoring-based technologies. Although studies on evaluating the impact of condition monitoring tools in terms of engineering and financial benefits exist [27,39,40], to the best of the authors' knowledge, no systematic review has analyzed these papers to uncover the commonalities and differences of these evaluation studies.

Developing comprehensive CMS evaluation methods requires insights from both from models of CMS used in industrial applications and techniques used to assess its engineering and financial benefits. This study aims to address gaps in the existing literature and derive insights about models and techniques from a survey of studies on CMS-related evaluation methods. This review systematically surveyed studies that evaluate condition monitoring with consideration of industrial equipment operations, maintenance decision-making, performance measures, and economic justifications for their use.

This paper reports survey results and synthesizes opportunities for future evaluation studies, giving special attention to the subset of manufacturing-related condition monitoring evaluation studies. The reporting concentrates on aspects of condition monitoring and maintenance modeling details that have been overshadowed by these evaluation studies. Considerable attention is paid to each evaluation study's analytical framework, simulation technique, and sensitivity analysis. Difficulties are noted in conducting a comparative analysis between the studies due to the heterogeneity of their reported results. Although this literature review uses three independent reviewers to ensure the validity of extracted data, future research may address this difficulty by comprehensively reporting results to enhance the comparability of evaluation studies.

The remainder of this paper is organized as follows. Section 2 describes the steps taken in this systematic literature review and the data items that were extracted from the reviewed papers. Section 3 has subsections for each model entity in the conceptual model for condition monitoring-enabled maintenance evaluation, where each entity comprises a set of data items. Each subsection describes the results for each model entity and its comprising data items, discusses the implications of the results, and suggests specific recommendations for future condition monitoring evaluation research, especially regarding manufacturing applications. Finally, Section 4 summarizes the paper's findings and makes some concluding remarks on the current state of evaluation research, with an emphasis on manufacturing maintenance applications.

2. Methodology

This systematic literature review used the methods outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) [41].

2.1. Eligibility criteria

The aim of this paper is to select and review articles that discuss methods for evaluating the impact of integrating CMSs or CMS-related technologies with the maintenance of an industrial application. Notably, this literature survey screened and assessed the eligibility of articles based on four criteria:

- Industrial Application
- Performance Measures
- Monitoring Feedback
- Evaluation Approach

The first criterion *Industrial Application* ensures that the selected articles focus on the impacts of a monitoring system on human-made industrial equipment. This criterion focuses on the industrial setting aspect of this literature review goal. The equipment could be a system or component part of an encompassing system. Here, a system is defined as any material artifact being operated on by an organization to fulfill some meaningful purpose [42]. This criterion excludes articles

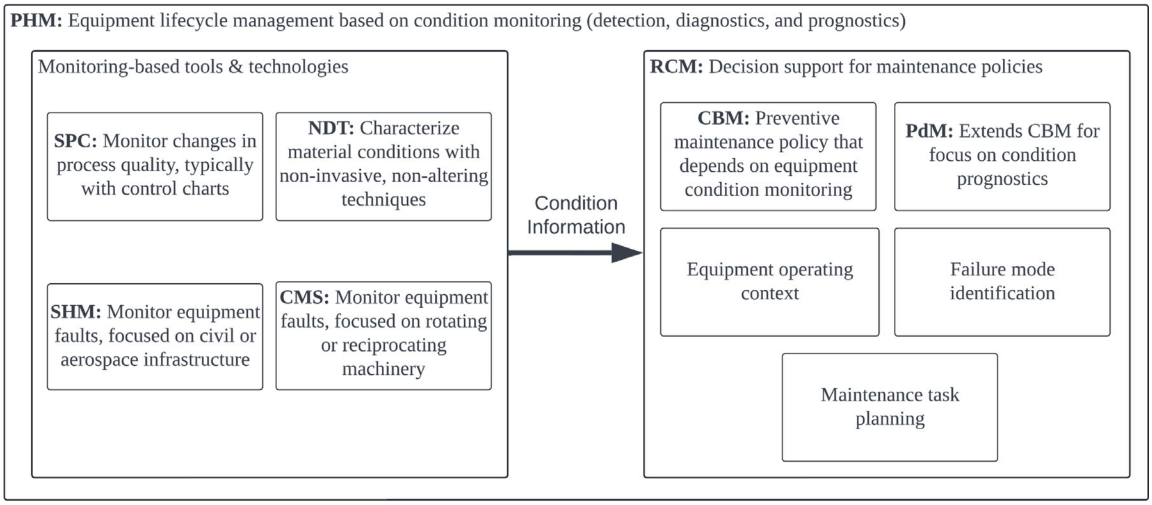

Fig. 1. Visualization of CMS-related concepts.

not applying CMSs to a human-made industrial system or component, as in [43]. Using a condition monitoring system in industrial settings, as opposed to non-industrial settings, requires the monitoring tool to meet strict engineering requirements and to produce explicit benefits to operations and maintenance.

The second criterion *Performance Measures* gives preference for articles that evaluate the impact of CMSs with metrics that go beyond including algorithm-level performance measures such as precision and recall [44]. This criterion focuses on the impact evaluation aspect of the literature review goal. The algorithm-level metrics evaluate a monitoring system's ability to detect, diagnose, or prognosticate faults or failures, and they are essential to understand a CMS's effectiveness. Because these metrics alone do not capture the impact of a CMS on equipment, this criterion excludes papers that focus only on such metrics, such as [36,45]. Changes to a monitoring system's algorithm-level performance can lead to changes in industrial equipment operations and maintenance performance. Focus on a monitoring system's algorithm-level metrics without considering the impact that a monitoring system has on an industrial equipment's performance does not capture the impact that a monitoring system's algorithm-level performance has on a monitoring system's benefits to equipment operations. Instead, this systematic review looks for studies measuring CMS impacts on equipment operations, maintenance, and business value. These key performance indicators are wide-ranging but include system or component productivity [34], availability [46], reliability [47], and financial cost–benefit metrics [28].

The third criterion *Monitoring Feedback* necessitates that the selected articles incorporate maintenance decision-making processes initiated or triggered due to feedback from monitoring data, acknowledging the relationship between maintenance and monitoring [48]. This criterion focuses on the literature review goal of integrating monitoring systems with maintenance. This criterion helps ensure that the selected papers for the literature review examine a monitoring system as part of a maintenance strategy and avoid impact evaluations from hypothetical CMS use without describing the mechanism by which the monitoring system interacts with maintenance (as in [49,50]). This criterion excludes articles that assume a CMS could reduce a certain fraction of faults or failures (as in [51]), assume a monitoring system can prognosticate failures with a pre-specified prognostic distance (such as [52]), or assume a hypothetical CMS would impact system or component failure rates (such as [53,54]).

The third criterion also excludes papers identifying the equipment (system or component) that could benefit from monitoring and maintenance [55–57]. Although this is a relevant precursor step to evaluating the impact of CMS-related technologies, plenty of reliability-centered maintenance (RCM) frameworks help decide the configuration and application of appropriate tasks for equipment risk mitigation [58].

Lastly, the fourth eligibility criterion *Evaluation Approach* ensures that the evaluation methods described in the selected articles present a structured procedure that can be repeatable for evaluating the impact of a CMS on equipment systems or components that are at least similar to the selected article's system or component. The evaluated CMS impact should also allow for comparing maintenance policies that use a CMS-type tool and policies not based on condition monitoring. This criterion focuses on the evaluation methods aspect of the literature review goal. This criterion excludes papers that show little detail for their evaluation methods (such as [59] or [60]) and papers that are specific to evaluating monitoring impact on the paper's particular, described equipment application and are difficult to generalize to other equipment (such as in [61]).

2.2. Information sources

Information sources were searched for using the electronic database Scopus because it has strong coverage of engineering sciences [62]. This search was conducted in Scopus on January 10, 2024. The search only included articles published within the time frame of January 2001 to December 2023.

CMS-related technologies have been around for decades before this selected time frame. To the best of the authors' knowledge, the earliest articles that cover the profitability and cost-effectiveness of condition monitoring technologies can be found as early as 35 years ago [63,64]. However, their evaluation of condition monitoring usage sticks to general guidelines and puts little emphasis on quantifying their risks and benefits.

Instead, the rationale for this coverage of dates for the literature search is that they coincide with a period, starting in the early 2000s [16,23], in which there was an explosion of research and development for harnessing (condition monitoring) data for managing equipment with PHM approaches. The increased interest in PHM research overlaps with advances in computing power and data collection, actualizing inspirations from prognostics in the medical field [65].

In addition to using the Scopus database, additional articles were identified from sources that reviewed challenges with integrating condition monitoring-like technologies in industrial settings [3,27,28,40,66,67].

2.3. Search strategy

For the literature search using the Scopus database, the focus was on search queries for four concepts across article titles, abstracts, and keywords:

- Equipment (system or component) monitoring, represented by: (“health monitoring” OR “condition monitoring” OR *destructive OR “statistical process control” OR prognosis OR diagnosis OR diagnostics OR prognostics)
- Maintenance-related decision-making, represented by: (“prognostics and health management” OR “condition-based maintenance” OR “condition based maintenance” OR “predictive maintenance” OR “preventive maintenance”)
- Performance measures, represented by: (evaluat* OR metric* OR measure* OR indicator* OR performance)
- Benefit analysis, represented by: (“risk analysis” OR “cost-benefit” OR “cost benefit” OR “cost-effective” OR “cost effective” OR “life-cycle cost” OR “life cycle cost” OR “lifecycle cost” OR “value of monitoring” OR “value of information”OR “present value” OR “rate of return” OR “discount rate” OR “hurdle rate” OR “cash flow” OR “payback period” OR “real options” OR “profitability index” OR “investment analysis” OR “return on investment”)

Furthermore, exclusionary criteria were imposed on the search strategy to maintain that the obtained articles are in English and are final publications (none should be articles in press). Conference reviews and editorials were also excluded.

2.4. Study selection process

Fig. 2 summarizes the study selection process of this literature review. The process initially identifies records through the search strategy in Section. Abstract and full-text assessments screen the records until a final set of studies are selected for this review.

All three reviewers (MD, JH, and MS) independently screened the abstracts of articles using a codebook that described the four eligibility criteria in Section 2.1. The codebook consists of a checklist of eligibility criteria for each screened abstract, along with reviewer notes about the quality or rigor of each abstract. If all three reviewers agreed that an article met all four criteria, they selected it for the full-text assessment stage of the screening process. If each of the three reviewers agreed that the article did not meet at least one criterion, they discarded the article. For the rest of the articles, the reviewers discussed the conflicting codes in detail and only selected articles for full-text assessment if only one reviewer determined that the article lacked only one criterion. Throughout the abstract screening step, discussions were held between the three reviewers to best maintain consistency with the criteria in the codebook.

For full-text assessment, two reviewers (MD and JH) independently screened each of the remaining articles against the four eligibility criteria using a codebook structured similarly to the one used for abstract screening. As with abstract screening, articles were selected for literature review if both reviewers agreed that the article met all four eligibility criteria, and articles were discarded if both reviewers agreed that the article was lacking of at least one criterion. Conflicting articles were sent to the third reviewer (MS), who made the final decision on whether an article met all four criteria.

Furthermore, at the full-text assessment stage, further exclusion criteria were imposed on the study selection process by excluding literature reviews (3 instances, e.g., [68]), papers that did not have a full-text available (5 instances), and records from sources that were not peer-reviewed (2 instances, a trade publication and an extended abstract for a presentation). In particular, the three literature reviews were excluded because they did not survey articles that discussed evaluation approaches (the fourth eligibility criterion). This assessment stage also excluded papers that overlapped the same style of evaluation methods with more recent publications from the same set of authors (such as [69] and [70]). In these 19 instances, the more recent publication references the earlier studies.

2.5. Data collection process

For the articles resulting from the study selection process for literature review synthesis, the reviewers developed a data collection codebook for the data items described in Section 2.6. After pilot testing and refining the data collection process on five randomly selected articles, one reviewer (MD) performed the initial data extraction for these articles, and the two other reviewers (JH and MS) verified and discussed the extracted data to agree upon the final data points.

2.6. Data items

Evaluations of condition monitoring-enabled maintenance policies typically require a model of the equipment’s operations and maintenance policy, as well as a procedure that studies the impact of the condition monitoring tool in the aforementioned model. After obtaining the resulting articles from the study selection process (Section 2.4), the reviewers (MD, JH, and MS) developed a conceptual model for condition monitoring-enabled maintenance evaluation that reflected both parts required for evaluations. The reviewers organized a preliminary set of data items aligned with the entities in the conceptual model. The purpose of defining data items is to identify the types of data extracted in the data collection process (Section 2.5).

The reviewers iteratively refined and updated the conceptual model and its relevant data items throughout the data collection process, resulting in Fig. 3. The solid arrows between entities and their data items in Fig. 3 represent a transfer of information or an action. The dashed lines between the evaluation entities represent supporting methods for deriving target information.

The first part of conducting evaluations is labeled as *condition monitoring-enabled maintenance application*. The entities comprising the model representing the equipment and its maintenance is derived from fundamental development steps relevant to building condition-based or predictive maintenance models [3,71]. These entities, each comprising their respective data items, are the target industrial equipment, condition monitoring data and algorithms, and maintenance deployment implementation.

Likewise, the second part of conducting evaluations is labeled as *condition monitoring-enabled maintenance evaluation procedure*. The entities that comprise the procedure to evaluate condition monitoring impact is derived from fundamental development goals relevant to evaluating technologies for industrial applications [34,72,73]. These entities, each comprising their respective data items, are performance measures, economic analysis, and the evaluation techniques used to obtain those measures and economics.

2.6.1. Industrial application

The industrial application follows from the first criterion in Section 2.1, where *industrial equipment* represents the human-made industrial system or component that a study integrates with condition monitoring-based maintenance.

Failure process represents how equipment is modeled to deteriorate into failure. Studies can model system or component failure and failure dependencies with many approaches. These approaches include failure distributions, linear aging models, system state models (e.g., Markov chains), and stochastic degradation processes [74,75]. This study focuses on whether studies represent equipment failure processes with predictions of time to failure, typically using a probability distribution of an equipment’s lifetime, or with component degradation models that

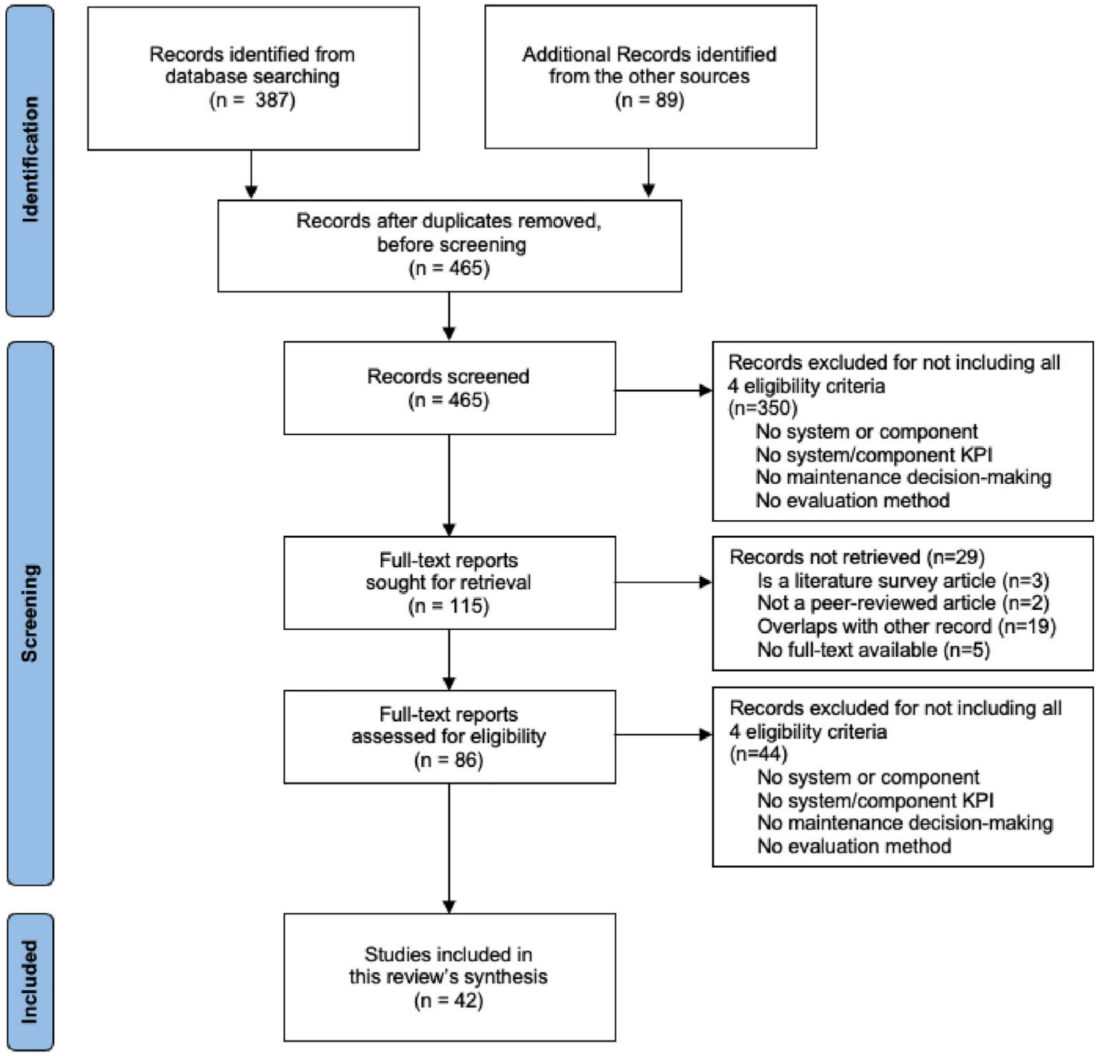

Fig. 2. PRISMA Search Flow Diagram.

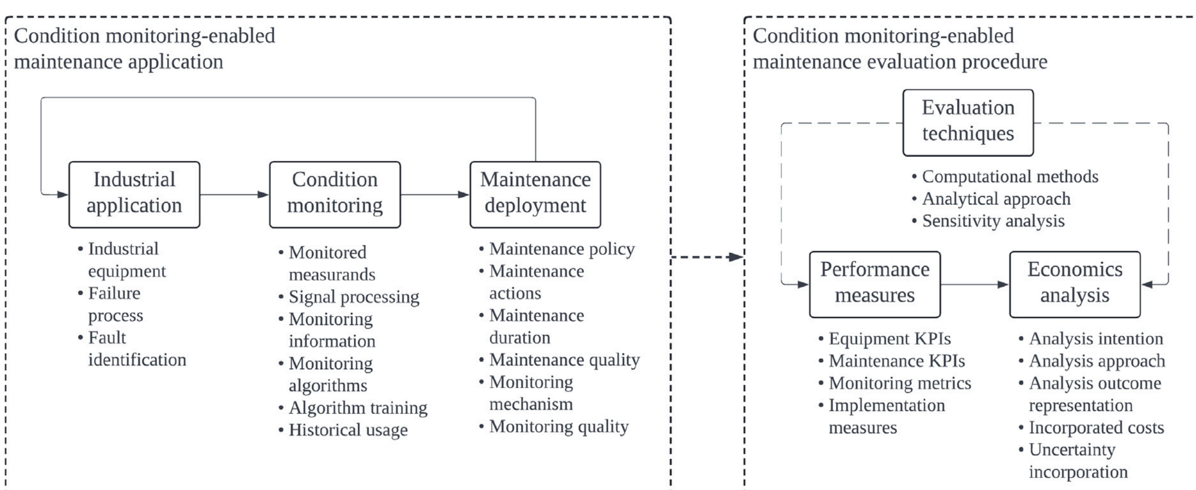

Fig. 3. Conceptual model for condition monitoring-enabled maintenance evaluation, and relevant data items.

go beyond observing the equipment's age and into improved failure prediction with estimating equipment condition [76].

Fault identification denotes the use of structured techniques to identify potential equipment faults and the risks associated with those faults. Risks from equipment faults comprise events in equipment operation, consequences of those events, uncertainties about each event or associated consequence, and background knowledge of the event and consequence [77]. In principle, a CMS should reduce the risk of equipment faults and failures [40]. However, it also introduces risks by missing faults or failures or creating false alarms and needing digital oversight or maintenance of the CMS itself [10]. Techniques for identifying faults whose risk could be decreased by CMS deployment include, but are not limited to, scenario analysis, the Ishikawa method, and Failure, Modes, Effects, and Criticality Analysis (FMECA) [78,79].

2.6.2. Condition monitoring

Condition monitoring refers to the process by which the CMS takes raw data inputs from the industrial application and transforms them into knowledge that may inform maintenance-related decision-making. This description is inspired by data communication and information flow guidelines for CMSs from the standard ISO-13374 [80].

Monitored measurands denotes the physical quantities whose data is collected from the equipment. These measurands include sensor signals such as temperature, vibration, voltage, acoustic emissions, cutting forces, and pressure [26,81–83]. A CMS collects these measurands using a measurement infrastructure ranging from a single sensor to an IoT infrastructure consisting of a connected network of sensors and actuators. In an industrial system embedded with IoT-based technologies, the collected measurands form a vital part of the system's IoT infrastructure [3].

Signal processing refers to techniques used to extract features from the incoming sensor data, often in the time-domain or frequency-domain. These techniques include Fourier analysis, wavelet analysis, moving averages, and principal component analysis [82,84]. Signal processing aims to extract the data's salient features and can also include simple statistical values such as kurtosis and root mean square (RMS) [85,86].

The *monitoring information* aspect describes the output knowledge that maintenance personnel seek to gain from integrating condition monitoring-based technologies into their decision-making processes. In regards to maintenance strategies, these outputs include part or machine conditions [87,88], condition-based risk assessments [18,89,90], anomaly detection, fault diagnosis, and failure prognosis [91]. Within the current literature, there is no shortage of methods to obtain various monitoring information outputs [22,85,92,93].

Monitoring algorithms use the extracted features from signal processing or sensor data as inputs for classifying or predicting the system information, which is denoted as *monitoring information* in this paper. Examples of these algorithms include Bayesian networks, support vector machines, neural networks, and other machine learning approximators [24,94–99].

Algorithm training denotes the adaptation of the monitoring algorithms' internal parameters to improve detection or prediction performance [42]. This training is typically done through hyperparameter optimization and search algorithms that build from the identified patterns in the data. Other training methods include system simulation, user feedback, heuristics, or closed-form physics-based equations.

Historical usage specifies whether a study indicates that the CMS used in its particular application has had prior use with other similar industrial equipment. Suppose such historical data exists and is sufficiently similar. In that case, it may be used directly for evaluating the condition monitoring-supported application or as a basis for training the monitoring algorithm [3,100–102]. This data also has applications as foundations for building or updating intelligent systems through transfer learning.

2.6.3. Maintenance deployment

Maintenance deployment refers to the decision-making logistics and risk-mitigating actions that maintenance personnel take to return or retain equipment to a state that meets their functional requirements. In the context of this paper, maintenance is a mechanism meant to decrease the probability of faults and failures of the industrial equipment, resulting in decreased operational risk and increased equipment reliability and associated cost–benefits [103].

The *maintenance policy* determines the type of actions and the timing of those actions that maintenance personnel plan to take when the equipment undergoes events requiring maintenance. Typically, policies can be classified based on equipment events such as equipment failure (failure-based or reactive maintenance), time-limits (time-based maintenance), equipment use-limits (usage-based maintenance), or condition state (condition-based maintenance or predictive maintenance) [74,104]; however policies can be classified into more nuanced categories [105]. Researchers typically consider time-based, usage-based, and condition-based maintenance as preventive maintenance.

Maintenance actions include corrective or preventive equipment repair and corrective or preventive equipment replacement. Policies prescribe corrective maintenance actions for failure equipment events but preventive maintenance actions when the equipment is operating [105].

The *maintenance duration* consideration is the equipment repair and replacement times. Though seemingly prosaic and often ignored, the duration times can significantly influence equipment and maintenance performance [74]. For the purposes of this paper, this duration is restricted to include a maintainer's time engaged with equipment repair. The travel and part procurement times are excluded due to extreme variability between facilities and jobs.

Maintenance quality is the degree to which a maintenance action can restore equipment functionality. Maintenance may induce perfect repair or replacement of equipment back to a new state (as good as new), minimal repair brings the equipment to the state right before maintenance was performed (as bad as old), imperfect repair improves the equipment's state to somewhere between perfect repair and minimal repair (better than old but worse than new), worse repair that sends the equipment to a worse or failed state (worse than old), and repairs where the equipment's deterioration rate alters due to the maintenance action rather than the equipment's state (hampered deterioration) [74,106].

Furthermore, the *monitoring mechanism* in maintenance deployment refers to the temporal means by which information about the equipment state alerts maintenance for any needed action. These mechanisms include continuous condition monitoring of equipment state, periodic inspections of the equipment, and aperiodic inspections (e.g., dynamically altering inspection times) [104].

Monitoring quality is the degree to which the monitoring feedback mechanism accurately depicts the equipment state. CMSs and inspections can produce imperfect equipment monitoring, resulting in false or misrepresentations of equipment state [106].

2.6.4. Evaluation techniques

Evaluation techniques are the supporting means and methods for calculating performance measures and assessing the risks and economics associated with integrating maintenance with CMSs.

Computational methods describe any techniques or tools used to evaluate condition-monitoring-enabled maintenance, and can include discrete-event simulation [107], agent-based simulations [108], Monte Carlo methods, or even back-of-the-envelope Fermi estimates.

Analytical formulation denotes whether a study mathematically conceptualizes some degree of the dynamics between equipment operations and maintenance decision-making (enabled by condition monitoring), and embeds this into CMS evaluation. This data item is meant to also consider whether the studies model these dynamics with Markov models, Bayesian decision networks, and dynamic risk assessment methods [79,89,109].

Sensitivity analysis refers to incorporating any study to understand the influences that evaluation model inputs have on model outputs [110]. The aim of including this data item is to understand the types of model inputs used in the studies.

2.6.5. Performance measures

Performance measures follow the third criterion in Section 2.1. They include the key performance indicators (KPIs) and metrics associated with either the industrial equipment, deployed maintenance activities, or condition monitoring.

Equipment KPIs capture the ability of the equipment (system or component) to perform its desired function. For example, productivity can be captured in manufacturing by production rate, throughput, and work-in-process [111]. These metrics can scale from capturing the performance of an entire system (the entire factory floor) or a constituent component (a single machine). Reliability and availability are two relevant KPIs [47] that many enterprises will target to improve when integrating condition monitoring information and maintenance [17,46, 112].

Maintenance KPIs capture the capability of all activities that are meant to restore, return, or retain equipment to a functional state (as per equipment requirements). Manufacturing maintenance KPIs can range across various enterprise maintenance objectives, from addressing personnel management to budgetary concerns [113]. Examples include the cost of maintenance actions, number of machine shutdowns, workforce utilization, corrective work costs, and percentage of maintenance work requiring rework [34]. Furthermore, KPIs directly relevant to condition monitoring can be divided into performance monitoring metrics and monitoring implementation measures.

Monitoring metrics represent the adoption of metrics that assess the degree by which the resulting monitoring information (detection, diagnostics, or prognostics) achieves its performance specification [33]. These metrics range from classification or detection metrics, such as accuracy and F-score [44,114], to metrics specific to regression or prognostics, such as mean absolute error, prognostic horizon, and Remaining Useful Life (RUL) prediction relative accuracy [31,115].

Implementation measures refers to qualities relevant to the application of condition monitoring, such as the redundancy and coverage of condition monitoring data [116], the interpretability of monitoring information outputs [117], or the monitoring output's ability to provide trustworthy information [118].

2.6.6. Economics analysis

Economics analysis involves calculating and understanding the monetary impacts of investments. Within the context of this paper, investments in condition monitoring technologies are meant to enable maintenance personnel to decrease equipment risks from faults and failures further. Understanding the impacts of investing in condition monitoring-enabled maintenance allows decision-makers to seek profitable investment decisions within the frame of acceptable equipment risk [73].

Analysis intention addresses the motives behind a study's modeling and evaluation of condition monitoring-enabled maintenance. An intended outcome of a study may be to analyze and evaluate CMS use to qualify the use of a new CMS-type tool or extract minimum requirements of a new CMS-type tool prior to integrating it into an equipment's maintenance practices [27]. Other reasons for CMS evaluation include justifying the continued use of a legacy CMS-type tool (especially in the context of an equipment's evolving environment and changing operation conditions [3]), comparing a legacy CMS tool with another CMS-type tool, or optimizing condition monitoring-enabled maintenance logistics [40].

Analysis approaches refers to the paradigms used to model and compare the cost evaluations of a condition monitoring-enabled maintenance application. Gauging the impact of maintenance on its ability to prevent or minimize adverse events is difficult, so studies often compare the use of condition monitoring against different maintenance policies, especially ones that do not use condition monitoring. Modeling approaches for enabling these comparisons include minimal total cost analysis, risk-based optimization, cost–benefit analysis, and cost-effectiveness analysis [73].

Analysis outcome representation describes the economic depictions of CMS economic evaluation outcomes. Costs in an evaluation study can be a monetary value comparison between maintenance policy cost outcomes. However, the costs may also be represented as return on investment (ROI), net present value (NPV), and internal rate of return (IRR) [119].

Incorporated costs indicate the calculated cost categories of an economic assessment, which range from costs related to the equipment itself, maintenance [72], and the monitoring system's implementation [33,52].

Uncertainty incorporation denotes whether the economic analysis accounted for uncertainties in their evaluation outcome that derive from uncertainties in the evaluation model inputs. These uncertainties can range from single-value probabilities of uncertainty to interval probabilities [120]. The goal is to distinguish between studies that quantify evaluation outcome uncertainties and studies that use point-value estimations of cost–benefits based on fixed input parameters [52]. Many parameters used to model and evaluate a condition monitoring-enabled maintenance application have associated uncertainty, from the performance of diagnostic or prognostic monitoring information that maintenance personnel receive [40,121] to uncertainty around internal equipment operations [117,122].

2.7. Limitations

This systematic literature review has several limitations. Notably, despite strict eligibility criteria, many of the data items extracted from the selected studies were heterogeneous and non-numeric, which voided quantitative synthesis and direct statistical comparisons of most data item measures. In addition to the heterogeneity of the extracted data items, the selected studies do not equally report similar levels of detail for each data item. Many of the studies in this review focus on a particular subset of data items while overlooking other items, further rendering the quantitative synthesis of extracted data item measures futile. This inconsistency in reported data items highlights the lack of reporting guidelines. To address limitations with respect to the inconsistency and heterogeneity of data items, this study verified collected data items with consultation between the three reviewers over the results from the codebook used in the study selection process.

Lastly, this literature review did not explicitly include study quality or bias criteria in the study selection or data collection process. Instead, this paper ensured study quality by focusing on refereed publications and carefully selected data items relevant to condition monitoring-based maintenance application and evaluation.

3. Results and discussion

Our search strategy described in Section 2.4 resulted in 465 records. After screening abstracts, 115 records remained. The reviewers agreed on the screening results for 74.2% of the articles during abstract screening. We excluded 29 of these 115 records during article retrieval, leaving 86 articles for full-text assessment. Of these 86 articles, 42 met the criteria to be eligible for literature review synthesis. During the full-text assessment, the reviewers agreed on the assessment of 82.6% of the articles. See Section 2.4 for conflict resolution processes for both abstract screening and full-text assessment.

This review selected 42 studies from an initial search of 465 records (Fig. 2). The types of documents in the 42 resulting records include 29 journal articles, 12 conference papers, and one book chapter. Fig. 4 shows the number of published records per year. Most of these records have been published since 2017 (23 studies).

Fig. 5 depicts the geographic distribution of the 42 resulting records. In this figure, node size indicates the number of publications with at least one author affiliated with an institution from the labeled country. The largest node, labeled "United States", represents 17 publications, followed by China with seven publications, Denmark and the United

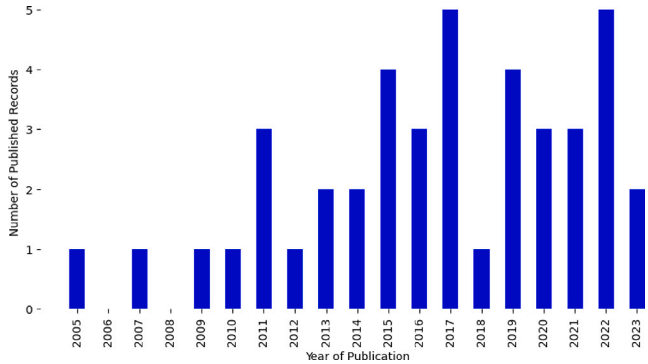

Fig. 4. Published records per year.

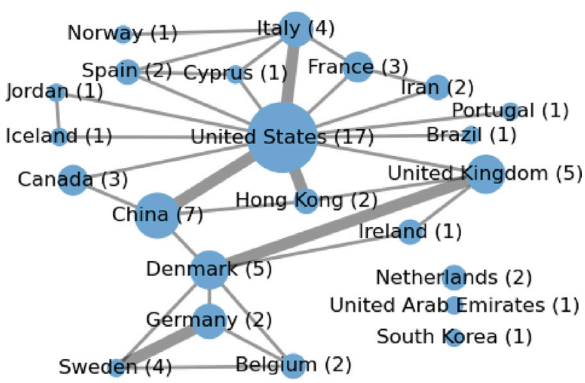

Fig. 5. Geographic distribution graph. Node size represents the number of publications from each country, specified by the number in each node label. Edge size represents the number of publications with authors from both countries, where thin lines represent one publication and thick lines represent two.

Kingdom with five publications each, and Sweden and Italy with four publications each. Edges mark publications with author affiliations at institutions from both countries at the end-nodes. A thin-lined edge indicates the presence of one publication with both country affiliations, while a thick-lined edge represents the presence of two publications. This graph indicates the leading countries in research on evaluating condition monitoring-based technologies, and the collaborations between the authors and institutions in these countries.

3.1. Industrial application

3.1.1. Results

Table 1 reviews the industrial equipment under focus in each of the selected studies, and Table 2 categorizes this equipment into broader engineering systems. Energy systems and transportation modes comprise most of the study applications, with wind turbines and aircraft forming the most common equipment in these two categories, respectively. More than one study examined production systems, civil structures (e.g., bridges), and generic equipment (including generic mechanical and electronic components).

Table 3 overviews each study's failure process models used to represent equipment deterioration. Time-to-failure estimates, such as the Weibull-based failure distributions in [123], tend to be the easiest way to model equipment failure behavior and comprise half of the studies (n = 21). An interesting example of time-to-failure estimations comes from the study by Kerres et al. [124], where some equipment components had an intermediary *time-to-deteriorate* distribution model to account for a third, *defected but operational* state between failure and operation states. In another case, Bakhshi and Sandborn [125] modify the equipment time-to-failure distributions to sample failure times concerning whether the equipment's reliability is affected by implementing new (condition monitoring) technology. Van Horenbeek et al. [126] incorporate other component damages beyond the monitored equipment associated with the equipment's sampled time-to-failure.

The selected studies exhibited various models to represent condition degradation measures and processes. For instance, Nielsen et al. [127] represented discrete categories of equipment deterioration with a discrete-time Markov model. However, the studies from Liang and Parlikad [128] and Meng et al. [129] used continuous-time stochastic processes to represent failure processes. A few studies used single-variable nonlinear deterioration models to represent generic condition indicators [130,131] and energy discharge [132]. Another subset of studies used decay or corrosion models as part of time-varying fragility curves [133–135]. Five papers make use of the Paris–Erdogan law to model crack fatigue mechanics [136–140], while two other studies also dealt with modeling crack deterioration [127,141].

A few papers also dealt with modeling multiple failure processes. For example, Yoon et al. [142] used time-varying stochastic models of multiple failure modes to represent health degradation. Both studies from Meng et al. [129] and Iannacone et al. [134] model both gradual deterioration and shock failures. Liang and Parlikad [128] modeled a continuous-time Markov chain to represent equipment conditions that can change state to a different condition, depending on a varying deterioration rate.

We note that five studies, rather than directly modeling a failure process, used datasets that represented a deteriorating equipment condition [143–146] or failure times [147]. Moreover, as for the *Both degradation and time-to-failure* category in Table 3, the study from Liu and Wang [148] had two use cases, one modeling failure process as a condition degradation and the other using a time-to-failure model.

In Table 4, we also list studies that described methods for identifying equipment faults and failure modes that may be suitable for condition monitoring. Fifteen of the reviewed studies evaluate CMS-type tools by first delving into the sources, causes, or drivers of equipment faults, as opposed to studies that evaluated the monitoring impact of these tools against pre-selected fault and failure modes.

Six of these studies [124,141,149–152] list a set of fault and failure modes for which condition monitoring may be applicable. Halbert et al. [141] identified a failure mode for aircraft wings and the associated acceptable component risk level. Kerres et al. [124] cites failure modes of wind turbine components from a prior reliability-centered maintenance study and uses information about these failure modes to select the components for which condition monitoring may be suitable. The studies from May et al. [149], Puglia et al. [150], Turnbull & Carroll [151], and Vieira et al. [152] listed failure modes and associated failure rates of wind turbine subsystems that a CMS can monitor. Similarly to these six studies, a seventh study listed maintenance tasks and the associated impacts that condition monitoring could have to eliminate or reduce the necessity of those tasks [153].

Four studies used variants of failure mode and effects analysis (FMEA) techniques to identify faults and their associated risks to the industrial application [126,144,146,154]. These studies used FMEA-style techniques to record faults and failures that may occur, the mechanisms that can produce these faults and failures, maintenance activities to minimize or repair faults and failures, the likelihood and costs if the failures do occur, and whether condition monitoring tools can impact the failure modes.

Adams et al. [155] uses the risk filtering, ranking, and management (RFRM) method to list risks. However, the authors suggest that, within their study's framework, RFRM could be replaced by any other method for identifying condition monitoring targets that may impact manufacturing machine productivity, citing FMEAs as a possible alternative. Li et al. [133] uses hurricane and tornado scenario analysis to identify risks to power distribution systems. Liang & Parlikad [128] use Ishikawa diagrams, also known as fishbone diagrams, to visualize

Table 1
Industrial equipment in studies (N = 42).

Study	Equipment	Study	Equipment
Adams et al. [155]	Manufacturing machinery	Long et al. [136]	Welded joints on wind turbine monopiles (offshore)
Azadeh et al. [70]	Electricity delivery systems (power generation)	Mao et al. [132]	Battery units on distributed ground combat vehicles
Bakhshi and Sandborn [125]	Wind turbines in a wind farm	May et al. [149]	Wind turbine components in a wind farm (offshore)
Chang et al. [123]	Electrical lighting systems	Meng et al. [129]	Cloud-based software systems
Chen et al. [143]	Aircraft engines	Neves and Frangopol [131]	Bridge structures
Compare et al. [140]	Railroad car bogies	Nielsen et al. [127]	Wind turbine blades
Cot et al. [138]	Aircraft fuselage panels	Peng et al. [154]	Wind turbines
Erguido et al. [156]	Wind turbines in a wind farm	Puglia et al. [150]	Wind turbines
Florian et al. [157]	Manufacturing machine gearboxes	Rastegari [145]	Manufacturing machinery spindle units
Golmakani and Fattahipour [158]	Generic equipment	Reimann et al. [130]	Aircraft components
Halbert et al. [141]	Aircraft wings	Shamayleh et al. [146]	Medical equipment (immunoassay analyzer)
Hongsheng et al. [153]	Aircraft air conditioning systems	Tian et al. [159]	Wind turbine components in a wind farm
Iannacone et al. [134]	Bridge structures	Turnbull and Carroll [151]	Wind turbine components (offshore)
Kerres et al. [124]	Wind turbine components	Van Horenbeek et al. [126]	Wind turbine gearboxes (onshore)
Klerk et al. [135]	Flood defense structures	Vieira et al. [152]	Wind turbines (offshore)
Koochaki et al. [160]	Manufacturing systems (serial)	Wang and Pecht [161]	Aircraft electronics
Lei and Sandborn [162]	Wind turbine components	Wu et al. [163]	Electronics
Li et al. [133]	Electricity delivery systems (power distribution)	Yang and Letourneau [147]	Railroad cars
Liang and Parlikad [128]	Electricity delivery systems (power transformers)	Yoon et al. [142]	(a) Electro-hydrostatic actuators (b) Generic equipment
Liu and Wang [148]	(a) Battery systems (b) Generic mechanical equipment	Zhang et al. [137]	Welded joints on civil structures
Livera et al. [144]	Photovoltaic power plants	Zou et al. [139]	Welded joints on offshore marine structures

Table 2
Industrial equipment categories (N = 42).

Industry category	Count	Studies
Energy systems	18	[70,123–128,133,136,144,149–152,154,156,159,162]
Wind turbine-related	*13 of 18*	[124–127,136,149–152,154,156,159,162]
Transportation modes	9	[130,132,138,140,141,143,147,153,161]
Aircraft-related	*6 of 9*	[130,138,141,143,153,161]
Production systems	4	[145,155,157,160]
Generic equipment	4	[142,148,158,163]
Civil structures	4	[131,134,135,137]
Marine structures	1	[139]
Software	1	[129]
Medical devices	1	[146]

Table 3
Failure process models in studies (N = 42).

Failure process	Count	Studies
Time-to-failure estimates	21	[70,123–126,147,149–163]
Condition degradation measures	20	[127–146]
Both degradation and time-to-failure	1	[148]

Table 4
Studies that utilized methods to identify faults suitable for monitoring (n = 15).

Study	Fault identification method
Adams et al. [155]	Risk Filtering, Ranking, and Management (RFRM)
Halbert et al. [141]	Failure mode list
Hongsheng et al. [153]	Maintenance task list
Kerres et al. [124]	Failure mode list, citing a reliability-centered maintenance study
Klerk et al. [135]	Equipment fragility curves to identify risks with observed operational loads
Li et al. [133]	Scenario analysis
Liang & Parlikad [128]	Ishikawa diagram
Livera et al. [144]	Failure mode effects and criticality analysis (FMECA)
May et al. [149]	Failure mode list
Peng et al. [154]	Cost-based failure mode effects and criticality analysis (FMECA)
Puglia et al. [150]	Failure mode list
Shamayleh et al. [146]	Failure mode and effects analysis (FMEA)
Turnbull & Carroll [151]	Failure mode list
Van Horenbeek et al. [126]	Failure mode and effects analysis (FMEA)
Vieira et al. [152]	Failure mode list

the causes and effects of different power transformer subsystems that result in a failure mode. Lastly, Klerk et al. [135] uses fragility curves that characterize a flood defense dike's probability of failure against water levels. Rather than identifying specific failure mechanisms, this approach identifies risks associated with the equipment's observed operational load.

3.1.2. Discussion

The studies in this literature review address condition monitoring evaluation methods for various applications across many domains, including several manufacturing-related applications. However, energy systems and transportation modes, particularly wind turbines and aircraft, comprise many studies. Aircraft and wind turbines, particularly offshore ones, are costly engineering systems. They are not readily available for inspections and maintenance. Such circumstances lead researchers and practitioners to look into condition monitoring feasibility in lieu of costly inspections. We also note that, though many studies focused on a particular industrial application [150,153–155], the authors in some studies claimed that their condition monitoring evaluation methods apply to a broader spectrum of applications [127,132,134,140,142,147]. As such claims are difficult to prove, future research should examine the generalizability of these evaluation approaches.

Despite the potential of predicted maintenance and automated operations, only a few studies focus on manufacturing systems [145,155,157,160]. Of these studies, only Adams et al. [155] and Koochaki et al. [160] present case studies of condition monitoring for multi-stage manufacturing systems, where the production of goods results from a sequence of distinct manufacturing processes. The scope of the study from Koochaki et al. focuses only on two manufacturing processes.

The models used to represent equipment failure processes were split roughly into two categories: time-to-failure estimates and condition degradation. Most wind turbine-related studies (11 out of 13) modeled equipment failure with time-to-failure estimates, while 4 out of 6 aircraft-related studies used condition degradation processes. Rastegari [145] conducted the only manufacturing-related study (out of 4) that models failure with condition degradation measures. Interestingly, all five studies on civil or marine structures utilize condition degradation processes. Civil structures such as bridges have a longer lifecycle and more significant capital costs than smaller-scale electrical and mechanical components, which may be a factor in justifying the computationally intensive condition degradation models. Lastly, the lack of public datasets associated with failure process estimates in the selected studies further indicates the continued difficulty in obtaining or disclosing data.

Fault identification methods allow decision-makers to understand the risks associated with equipment faults and failure modes. They also allow the decision-maker to decipher the extent to which a CMS-type tool can monitor for such faults and failures, ultimately gauging the degree to which condition monitoring can decrease equipment operational risks. Over a third of the studies in this literature review (n = 15) described fault identification methods that pinpoint and compare uses of potential condition monitoring applications. Nearly half of these studies (n = 7) focus on wind turbine applications, meaning that over half of the thirteen studies related to wind turbines in our literature review described fault identification methods. Data about wind turbine components and their fault modes seem to be more accessible, as observed by the studies in our review that cite datasets for various failure modes, rates, and costs [124,126,149,151,152]. Instead, only one study related to manufacturing systems provided details on fault identification [155]. However, the authors argued that a key step in their framework is agnostic to the type of fault identification technique used to determine condition monitoring targets that impact manufacturing productivity.

Nearly one-half of the studies (n = 7) used lists of failure modes or maintenance tasks that can be affected by condition monitoring tools [124,141,149–153]. These lists are limited in fault identification as they do not determine the sources, causes, and risk drivers associated with equipment faults and failures. On the other hand, the studies that utilized the FMEA-style methods [126,144,146,154] or the RFRM method [155] presented a more structured approach to fault identification that also accounted for recording mechanisms or causes that result in faults and failures. Though these more structured methods provide input to maintenance that highlights failure modes best suited for condition monitoring, they are limited to identifying single failure modes or risks instead of combinations of (often dependent) failure modes.

The Ishikawa method from Liang & Parlikad [128] provides a structured analysis of causes and combinations of contributory factors that result in a failure mode. However, the diagram's separation of causal factors to failure modes could be misleading. Furthermore, scaling the diagrams to depict the causal factors of multiple failure modes may be infeasible. The scenario analysis from Li et al. [133], though helpful in determining failures and risks from hurricanes and tornadoes to power distribution equipment, can become biased towards high-risk scenarios or uncommon scenarios that do not reflect the equipment's normal operation and associated fault modes. The fragility curves from Klerk et al. [135] are limited to monitoring applications for which the obtained information depends on observed equipment operational loads, such as weight on a bridge or water levels for a dike structure.

3.2. Condition monitoring

3.2.1. Results

Tables 5 to 10 summarize the key aspects of condition monitoring in our reviewed studies. Table 5 shows that most studies (n = 25) discuss measurands intended to be taken as the sensory or IoT-based input by the condition monitoring tool in the study. Ten of these twenty-five studies come from the eighteen studies that evaluate the condition monitoring of energy systems, seven come from the nine transportation mode-related studies, three from the four civil structures-related studies, two from the four production system-related studies, and one study each related to medical devices, marine structures, and generic equipment. Furthermore, in ten studies, more than one physical quantity is meant to be observed via condition monitoring.

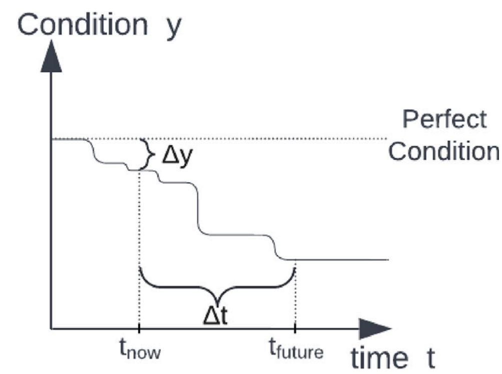

Fig. 6. In this oversimplified example of a single equipment condition indicator over time, Δy holds monitoring information that we categorize as *condition detection* and Δt contains information that we categorize as *condition prediction*.

The most common measurand is vibration. Eight of the studies include a vibration measure, half in applications related to the energy industry [124,126,149,156], two in production systems [145,157], one in medical devices [146], and one in generic equipment [142]. Other common measurands include wind velocity and damage size, with four studies. All mentions of wind velocity as a measurand come from energy system studies [125,127,136,144]. Three of the four studies that discuss damage size as a measurand come from studies on transportation modes [138,140,141], with the exception coming from a civil structure study [137].

Tables 6, 7, and 8 show the signal processing techniques, monitoring algorithms, and algorithm training routines used in studies that pertained to applications of energy systems, transportation modes, and the remaining industry categories listed in Table 2, respectively. Only eleven studies discussed any of these three data items in detail, of which five studies discuss all three [127,143,144,146,147]. Eight studies discuss signal processing techniques, and three share limited details [132,147,157]. Three studies discuss frequency domain analysis, and two use principal component analysis as part of a feature extraction and selection strategy.

A total of eight studies detail their use of monitoring algorithms. Four of these studies examined the use of machine learning algorithms strictly for forecasting remaining useful life [143,159] or future equipment condition [140,147]. Two studies focused on machine learning algorithms to classify faults, making use of support vector machines [146] and linear discriminant analysis [142], and one study applied a semi-supervised anomaly detection algorithm [127]. Livera et al. [144] implemented a machine learning pipeline that included equipment performance prediction, trend analysis, fault detection, and fault classification.

Another set of eight studies details algorithm training as part of the monitoring algorithm's model selection. However, six of these eight provide minimal details. Chen et al. [143] use the Adam optimization algorithm [164] to train their neural network. Two studies discussed hyperparameter tuning approaches: cross-validation [143] and receiver operating characteristic (ROC) curves [157].

All studies discussed the monitoring information provided by a CMS. We categorize the monitoring information in these studies in

Table 5
Monitored measurands in studies (n = 25).

Study	Industry	Measurand source	Study	Industry	Measurand source
Bakhshi and Sandborn [125]	ES	Wind velocity	Mao et al. [132]	TM	Voltage
Chen et al. [143]	TM	Temperature, pressure, rotational speed, mass flow rate, fuel flow rate, enthalpy	May et al. [149]	ES	Vibration, acoustic emissions, oil analysis, optical fiber analysis
Compare et al. [140]	TM	Damage size	Nielsen et al. [127]	ES	Acceleration, rotational speed, wind velocity, temperature, blade pitch
Cot et al. [138]	TM	Damage size	Rastegari [145]	OA	Vibration
Erguido et al. [156]	ES	Vibration, oil analysis	Shamayleh et al. [146]	OA	Vibration
Florian et al. [157]	OA	Vibration, electricity-related signals	Van Horenbeek et al. [126]	ES	Vibration, oil analysis
Halbert et al. [141]	TM	Damage size	Vieira et al. [152]	ES	Stress measures, acoustic emissions, acceleration, corrosion
Iannacone et al. [134]	OA	Material corrosion, stiffness	Wang and Pecht [161]	TM	Canaries: alerting devices that fail faster than equipment
Kerres et al. [124]	ES	Vibration	Yang and Letourneau [147]	TM	Wheel axle impacts
Klerk et al. [135]	OA	Pore pressure	Yoon et al. [142]	OA	Case study (a) Vibration Case study (b) Pressure, displacement, rotational speed
Liang and Parlikad [128]	ES	Dissolved gas analysis	Zhang et al. [137]	OA	Damage size
Livera et al. [144]	ES	Temperatures, wind velocity, electric current, voltage, power	Zou et al. [139]	OA	Magnetic particle inspection
Long et al. [136]	ES	Wind speed, stress range			

Notes: ES = Energy System. TM = Transportation Mode. OA = Other Application.

Table 9 as providing maintenance with information related to the equipment's current condition (*condition detection*) (n = 29) or information about a forecasted time (*condition prediction*) (n = 10) in which the equipment reaches a specific condition, such as failure. Three studies' monitoring information included condition monitoring and prediction [132,148,151]. The simple example in Fig. 6 shows the distinction between condition detection and prediction. The four studies on production systems and the four on civil structures used condition monitoring tools that relayed condition detection. Most studies on energy systems (n = 13) and roughly half on transportation modes (n = 5) also dealt with condition detection.

Lastly, four studies indicated a prior history of CMS use (see Table 10). These studies cited condition monitoring performance rates from prior studies to justify similar monitoring performance in evaluating the study's condition monitoring-supported application.

For example, Nielsen et al. [127] cite the monitoring performance of other wind turbines and assume that this performance will be transferable to the monitoring of the larger wind turbines in their study. The other three studies also dealt with condition monitoring applied to energy systems. Other studies have discussed using expert interviews, undisclosed data collection repositories, and Fermi estimates to describe their condition monitoring processes. However, they did not explicitly cite any results from historical usage of their condition monitoring tool with other industrial equipment.

3.2.2. Discussion

Our results reveal that condition monitoring evaluation studies focus on the type of monitoring information that a CMS provides (see Table 9). Monitoring information is the key output from a CMS. However, it relies on input data measurands and algorithms that process them to provide a CMS user with information about the equipment's condition. Yet only slightly over half of the studies (n = 25) discuss monitored measurands. Roughly a quarter of the studies (n = 11) go into detail about the processing techniques, monitoring algorithms, and algorithm training routines that underlie a CMS's ability to output monitoring information (see Tables 6 to 8).

Table 6
Signal processing (SP) techniques, monitoring algorithms (MA), and algorithm training (AT) routine - energy systems (n = 3).

Study	Class	Details
Livera et al. [144]	SP	Specified data pre-processing routine
	MA	Performance prediction (XGBoost); Fault detection; Trend analysis; AD (S-H-ESD); Fault classification (FBP)
	AT	Yes (unspecified)
Nielsen et al. [127]	SP	Frequency domain (Band-pass filtering); Feature selection(PCA)
	MA	Semi-supervised AD
	AT	Yes (unspecified)
Tian et al. [159]	SP	NR
	MA	RUL estimation (Artificial neural networks)
	AT	Yes (unspecified)

Notes: NR = not reported. XGBoost = Extreme Gradient Boosting. S-H-ESD = Seasonal Hybrid Extreme Studentized Deviates. AD = Anomaly Detection. PCA = Principal Component Analysis. FBP = Facebook Prophet algorithm. RUL = Remaining Useful Life.

All reviewed studies (N = 42) specify the condition monitoring information output from CMSs. Maintenance personnel use this monitoring information to decide whether anomalies are detected or repair times are predicted. However, most of these studies do not include details on the measurand data, signal processing techniques, and monitoring algorithms used to generate the output condition monitoring information. Instead, many studies resort to modeling characteristics of the monitoring information that determine the quality of that information, such as the accuracy or reliability of the information.

For example, the studies from Peng et al. [154] and Kerres et al. [124] assume a constant detection rate representing their monitoring system's information outputs. On the contrary, Adams et al. [155], the authors assume perfect machine condition information from the monitoring systems. The monitoring information modeled by May et al. [149] detects faulty equipment conditions either when faults appear or with a 6-month advanced warning, and each of these fault detection modes is associated with a fixed detection rate.

Table 7
Signal processing (SP) techniques, monitoring algorithms (MA), and algorithm training (AT) routine - transportation modes (n = 4).

Study	Class	Details
Chen et al. [143]	SP	Normalization (Z-score standardization); Clustering (FCM)
	MA	RUL estimation (Bi-LSTM neural networks)
	AT	Network training (Adam); HT (CV)
Compare et al. [140]	SP	NR
	MA	Prognostics (PF; Model-based interpolation)
	AT	NR
Mao et al. [132]	SP	Data pre-processing
	MA	NR
	AT	NR
Yang & Letourneau [147]	SP	Data pre-processing
	MA	Prognostics (Decision tree; Naive Bayes)
	AT	Yes (unspecified)

Notes: NR = not reported. RUL = Remaining Useful Life. Bi-LSTM = Bi-directional Long-Short Term Memory. FCM = Fuzzy C-Means algorithm. PF = Particle Filtering. HT = Hyperparameter Tuning. CV = Cross-Validation.

Table 8
Signal processing (SP) techniques, monitoring algorithms (MA), and algorithm training (AT) routine - remaining studies (n = 4).

Study	Class	Details
Florian et al. [157]	SP	Data pre-processing
	MA	NR
	AT	HT (ROC curves)
Rastegari [145]	SP	Frequency domain (FFT)
	MA	NR
	AT	NR
Shamayleh et al. [146]	SP	Frequency domain (FFT); Feature selection (PCA; correlations)
	MA	Fault classification (Linear SVM)
	AT	Yes (unspecified)
Yoon et al. [142]	SP	NR
	MA	Fault classification (LDA)
	AT	Yes (unspecified)

Notes: NR = not reported. FFT = Fast Fourier Transform. SVM = Support Vector Machine. ROC = Receiver Operating Characteristic. HT = Hyperparameter Tuning. LDA = Linear Discriminant Analysis. PCA = Principal Component Analysis.

The low degree to which these studies directly apply monitoring algorithms on input data to obtain condition monitoring information points to a missed opportunity in evaluating condition monitoring-enabled maintenance. Obtaining condition information using measurand data and monitoring algorithms, rather than approximating condition monitoring performance with detection rates, can help bind real operational and environmental equipment conditions to the validity of a condition monitoring evaluation study.

For example, researchers can examine the performance of a particular monitoring algorithm for detection or forecasting instead of varying operating conditions, gaining intuition on the limitations of that particular algorithm in its industrial application. Future research should examine the roles of training data size, signal processing, monitoring algorithms, and algorithm training on CMSs and their impacts on maintenance by contrasting different data preprocessing techniques, various monitoring algorithms, and numerous training algorithms. For instance, Yang and Letourneau [147] compared decision tree-based prognostics models and Naive Bayes-based prognostics models regarding saved maintenance costs.

3.3. Maintenance deployment

3.3.1. Results

Tables 11 to 13 focus on the maintenance policies considered in each of our reviewed studies and the actions prescribed by the policies. As our review focuses on condition monitoring technologies, all but one

Table 9
Monitoring information in studies (N = 42).

Information type	Count	Studies
Condition detection	29	[124–131,133–139,141,142,144–146,149,152–158,160]
Condition prediction	10	[70,123,140,143,147,150,159,161–163]
Both	3	[132,148,151]

Table 10
Studies that cited historical usage (n = 4).

Study	Industry	Historical usage
Bakhshi & Sandborn [125]	Energy system	Monitoring performance comes from prior studies
Liang & Parlikad [128]	Energy system	Monitoring performance comes from prior studies
Nielsen et al. [127]	Energy system	Monitoring performance comes from prior studies
Tian et al. [159]	Energy system	Monitoring performance comes from prior studies

study (n = 41) considered condition-based maintenance (CBM) in their evaluations. Golmakani and Fattahipour [158] analyze maintenance that monitors equipment conditions via periodic inspections, which falls under the purview of our time-based maintenance classification, briefly described in Section 2.6.3.

Usage-based maintenance (UBM) was considered in only a minority of transportation mode-related studies (n = 3). Twenty-four of the reviewed studies described time-based maintenance (TBM), comprising a slight majority of studies related to energy system applications (n = 12) and studies related to the remaining industrial applications (n = 10). Twenty-two studies described a maintenance policy that is strictly failure-based or reactive maintenance (RM), comprising slightly more than half of the energy system-related studies (n = 10) and the studies about the remaining industrial applications (n = 8).

Across Tables 11 and 12, we also observe five studies that presented a maintenance policy that combined condition monitoring-enabled maintenance with time-based inspections and maintenance [126,128,144,156,161]. Similarly, one study combined CBM with UBM [138]. Generally, models of TBM, UBM, and CBM imply that any equipment failures not prevented via inspections or monitoring will be mended with actions associated with reactive maintenance.

Regarding maintenance actions that a maintenance policy prescribes, we focused on two broad categories: equipment replacements and repair. Our expectation in framing these two broad categories is that replacing equipment or some component that comprises the equipment implies that the replacement is new and in perfect operating condition, while repair actions are not constrained to resulting in perfect equipment operations. Half of the studies in the review specified repair actions (n = 21), comprising a slight majority of studies related to energy system applications (n = 10) and studies related to the remaining industrial applications (n = 19). A more significant number of studies specified replacement actions (n = 26), comprising a majority of energy system-related studies (n = 13) and about half of the transportation mode-related studies (n = 5).

Two studies used replace and repair interchangeably [123,142]. In four studies, the maintenance actions were either not specified at all [152,153] or they were conceived as generic maintenance activities [134,156]. Two studies did not define repair or replace actions but divided maintenance actions into lower-cost predictive or preventive work and higher-cost corrective work [143,162]. Three studies specified maintenance actions beyond the repair and replace dichotomy [126,131,151]. Neves and Frangopol [131] compare several different types of repair actions, each with its implications in terms of costs, maintenance duration, and impact on equipment condition and safety. Turnbull and Carroll [151] define different types of failure

Table 11
Maintenance policies and actions considered in the energy system-related studies (n = 18).

Study	Policies					Actions		
	RM	TBM	UBM	CBM	Other	Repair	Replace	Other
Azadeh et al. [70]	Yes	Yes	No	Yes		Yes	No	
Bakhshi & Sandborn [125]	Yes	No	No	Yes		Yes	Yes	Repair via CMS-induced component reliability improvement
Chang et al. [123]	Yes	No	No	Yes		Yes	Yes	Both used interchangeably
Erguido et al. [156]	No	Yes	No	Yes	CBM added to TBM policy	No	No	Generic action
Kerres et al. [124]	Yes	Yes	No	Yes		No	Yes	
Lei & Sandborn [162]	Yes	No	No	Yes		No	No	Actions: high-cost corrective, or low-cost predictive work
Li et al. [133]	Yes	Yes	No	Yes		No	Yes	
Liang & Parlikad [128]	No	Yes	No	Yes	CBM added to TBM policy	Yes	Yes	
Livera et al. [144]	No	Yes	No	Yes	Single policy integrating both	Yes	Yes	Additional *clean* action for particular fault mode
Long et al. [136]	Yes	No	No	Yes		Yes	No	
May et al. [149]	No	Yes	No	Yes		No	Yes	
Nielsen et al. [127]	No	Yes	No	Yes		Yes	Yes	
Peng et al. [154]	Yes	No	No	Yes		Yes	Yes	
Puglia et al. [150]	Yes	Yes	No	Yes		Yes	Yes	
Tian et al. [159]	No	Yes	No	Yes		No	Yes	
Turnbull & Carroll [151]	Yes	No	No	Yes		Yes	Yes	Classified into minor or major forms of repair or replace
Van Horenbeek et al. [126]	No	Yes	No	Yes	CBM added to TBM policy	Yes	Yes	Another *Replace secondarily damaged components* action
Vieira et al. [152]	No	Yes	No	Yes		NR	NR	Actions not specified
Count	10	12	0	18		10	13	

*Notes: RM = Reactive Maintenance. TBM = Time-Based Maintenance. UBM = Usage-Based Maintenance. CBM = Condition-Based Maintenance. NR = Not Reported.

Table 12
Maintenance policies and actions considered in the transportation mode-related studies (n = 9).

Study	Policies					Actions		
	RM	TBM	UBM	CBM	Other	Repair	Replace	Other
Chen et al. [143]	No	No	No	Yes		No	No	Actions: high-cost corrective, or low-cost preventive work
Compare et al. [140]	No	No	No	Yes		No	Yes	
Cot et al. [138]	No	No	Yes	Yes	CBM added to UBM policy	No	Yes	
Halbert et al. [141]	No	No	Yes	Yes		Yes	No	
Hongsheng et al. [153]	Yes	Yes	No	Yes		NR	NR	Actions not specified
Mao et al. [132]	No	No	No	Yes		No	Yes	
Reimann et al. [130]	Yes	No	Yes	Yes		Yes	No	
Wang & Pecht [161]	Yes	Yes	No	Yes	Both standalone CBM and CBM added to TBM	No	Yes	
Yang & Letourneau [147]	Yes	No	No	Yes		No	Yes	
Count	4	2	3	9		2	5	

*Notes: RM = Reactive Maintenance. TBM = Time-Based Maintenance. UBM = Usage-Based Maintenance. CBM = Condition-Based Maintenance. NR = Not Reported.

Table 13
Maintenance policies and actions considered in the remaining studies (n = 15).

Study	Policies					Actions		
	RM	TBM	UBM	CBM	Other	Repair	Replace	Other
Adams et al. [155]	Yes	Yes	No	Yes		Yes	Yes	
Florian et al. [157]	Yes	Yes	No	Yes		No	Yes	
Golmakani & Fattahipour [158]	No	Yes	No	No		No	Yes	
Iannacone et al. [134]	No	Yes	No	Yes		No	No	Generic action
Klerk et al. [135]	Yes	No	No	Yes		Yes	No	Repair is done via reinforcement
Koochaki et al. [160]	No	Yes	No	Yes		No	Yes	
Liu & Wang [148]	No	Yes	No	Yes		No	Yes	
Meng et al. [129]	No	Yes	No	Yes		Yes	No	
Neves & Frangopol [131]	Yes	Yes	No	Yes		Yes	Yes	Several types of repair
Rastegari [145]	No	No	No	Yes		No	Yes	
Shamayleh et al. [146]	Yes	No	No	Yes		Yes	No	
Wu et al. [163]	Yes	No	No	Yes		Yes	No	
Yoon et al. [142]	No	No	No	Yes		Yes	Yes	Both used interchangeably
Zhang et al. [137]	Yes	Yes	No	Yes		Yes	No	
Zou et al. [139]	Yes	Yes	No	Yes		Yes	No	
Count	8	10	0	14		9	8	

*Notes: RM = Reactive Maintenance. TBM = Time-Based Maintenance. UBM = Usage-Based Maintenance. CBM = Condition-Based Maintenance.

modes and rates for various equipment components, which each failure mode corresponding to a major or minor repair or replacement. Van Horenbeek et al. [126] define a secondary category of equipment component replacements for severe component failures that led to other component failures.

From Tables 11 to 13, we also observed that only a minority of studies included at least two distinct types of maintenance actions (n = 12), regardless of whether the studies made distinctions between forms of repair and replace or distinctions between corrective and non-corrective work. Surprisingly, we found different types of maintenance actions in only one of the studies related to transport modes [143], one of the studies related to the production systems [155], and one of the studies related to civil structures [131]. Half of the energy-related studies modeled distinct types of maintenance actions.

Tables 14 to 16 portray the modeling choices in the reviewed studies for two key characteristics: maintenance duration and quality, relevant to executing maintenance actions. These two maintenance model characteristics determine the time it takes to execute a maintenance action and the end-result of the equipment's condition once the maintenance action has concluded. Rather than extract the maintenance duration's temporal values or the action's quality impact measures, our focus is to frame the mechanisms used to model maintenance duration and the classifications of maintenance action quality impacts.

As shown in Tables 14 to 16, studies focusing on energy system-related applications are more likely to model and disclose maintenance duration and maintenance action quality. The availability of this maintenance information for energy system-related applications follows the trend that more data is available to disclose for this industrial sector. We also speculate that maintenance duration is a key concern for hard-to-access energy systems, such as offshore wind turbines. Of the eighteen energy system-related studies in Table 14, a plurality included both maintenance duration and quality (n = 7), with fewer studies modeling only maintenance duration (n = 4) or only maintenance quality (n = 3). Of the nine transportation mode-related studies in Table 15, only Wang and Pecht [161] incorporates details about both maintenance duration and quality (n = 1). At the same time, studies modeling only maintenance duration (n = 2) or only maintenance quality (n = 2) are also uncommon. Among the fifteen remaining studies (Table 16), only Koochaki et al. [160] includes details of both duration and quality in their manufacturing system's maintenance model (n = 1). Some of these fifteen studies include modeling details for maintenance quality without maintenance duration (n = 5), but including modeling details without maintenance duration without maintenance quality is uncommon (n = 1). The six studies that model maintenance quality are from several industrial equipment applications: civil structures (n = 2), generic equipment (n = 2), marine structures (n = 1), and production systems (n = 1).

The studies in our review used several different representations of maintenance duration. The most common representation, appearing in nine works [123,124,127,128,146,150,151,153,160], took the form of specifying a fixed maintenance duration parameter, such as a constant-value time-to-repair. Similarly, three studies used fixed parameters to represent equipment downtime [125,144,162]. Equipment downtime often aggregates maintenance duration with other parameters such as maintenance lead time, the duration between identifying a maintenance need and commencing maintenance action, or logistical spare parts waiting time. Two studies, instead specify a fixed parameter, represent maintenance duration with unspecified variables [126,132]. Azadeh et al. [70] takes a different approach to modeling maintenance duration by sampling time-to-repair values from an exponential distribution.

The maintenance action quality models used in the reviewed papers can fall into several classes. Sixteen of our reviewed studies model actions that result in equipment with operating conditions that are as good as new. Nine of these sixteen models exclusively model the replace actions to result in as good as new equipment quality [124,125,127,128,131,140,148,160,161], as opposed to four models that exclusively model repair actions to result in as good as new [70,136,137,139]. The remaining three studies from these sixteen indicate that either repair or replacement can result in as good as new quality [126,142,150]. Nielsen et al. [127] is the sole study that models some of its repair actions as minimally impactful actions that result in equipment conditions on par with the equipment state immediately preceding maintenance, which we dub as bad as old. Another four studies model maintenance actions to result in improved equipment conditions that are not quite as good as new, which we denote as maintenance action quality that is better than old but worse than new [130,131,156,159]. Among these four studies, Reimann et al. and Neves and Frangopol model repair actions to be better than old but worse than new, but Tian et al. models replace actions as such. Lastly, three studies model their repair actions' impact as hampered deterioration, where an equipment's deterioration rate decreases due to maintenance rather than any change in the equipment's operating condition [125,128,131].

Our review also considered monitoring information characteristics, shown in Tables 17 to 19. We specifically extract data from the reviewed papers about the modeled monitoring mechanism, or frequency in which maintenance personnel retrieve equipment condition information, and the quality of that retrieved monitoring information. The monitoring mechanism and the monitoring quality characterize the interface between condition monitoring and maintenance deployment. Maintenance personnel use feedback from the monitoring system about the equipment's conditions to inform decisions regarding actions the personnel must take to support equipment operations. The success of these maintenance actions, actions informed by monitored conditions, relies on the timing of incoming monitoring information and the quality of that information. The timing of monitoring information retrieval can affect maintenance personnel logistics for initiating maintenance actions, while the quality of the monitoring information can determine whether unnecessary maintenance actions are taken due to false monitoring alarms or equipment conditions requiring maintenance attention are missed due to missed alarms. Of the forty-two studies, roughly three-quarters include some degree of information regarding both monitoring mechanisms and quality (n = 32). The remainder of the studies solely include information about either the monitoring mechanism (n = 6) or monitoring quality (n = 4).

Our survey results in Tables 17 to 19 exemplify several monitoring mechanisms that deliver retrieved information about the equipment state to maintenance for any necessary action. Most reviewed studies indicate that information about equipment state is communicated to maintenance via continuous condition monitoring (n = 28). Per industrial equipment application, these studies form the majority of energy system-related studies (n = 11), transportation mode-related studies (n = 6), and studies pertaining to the remaining industrial applications (n = 11). Half of the production system-related studies from Table 19 modeled the monitoring mechanism with continuous monitoring as well [155,160]. Several of these studies combine continuous monitoring with inspections [126–128,134,136] or, in the case of Compare et al. [140], periodic RUL estimations.

Of the studies that do not model monitoring mechanisms with continuous monitoring, the mechanisms are highly varied. A subset of studies (n = 7) modeled periodic monitoring or estimation updates of equipment conditions [125,132,133,141,157,159,162]. The study from Bakhshi and Sandborn uniquely presented the continual monitoring of equipment that comprises a system, in their case wind turbines that comprise a wind farm, which results in the periodic monitoring of each constituent equipment. Another subset of studies (n = 3) modeled non-periodic monitoring or inspections for their maintenance scheme [124,139,158]. Golmakani and Fattahipour investigate age-based inspection schemes in their study [158], and Kerres et al. samples CMS monitoring times from a time-to-detect distribution [124].

In addition to monitoring mechanism considerations, most of the studies in our review explicitly model the quality of monitoring information that maintenance personnel receive (n = 36). Per industrial equipment application, these studies form the majority of energy system-related studies (n = 14), all of the transportation mode-related studies (n = 9), and nearly all of the studies about the remaining industrial applications (n = 13). We classify the modeled quality of monitoring information as either perfect, which is always accurate, or imperfect, which may provide misleading information about equipment conditions to maintenance personnel.

As expected for a review of papers evaluating monitoring system performance, most studies model monitoring systems that may be prone to communicating inaccurate, imperfect information (n = 28). None of the energy system-related studies in Table 17 modeled perfect monitoring information, though Puglia et al. [150] compared CMSs with perfect and imperfect RUL predictions. Very few of the transportation mode-related studies in Table 18 (n = 2) or the remaining industrial application studies in Table 19 (n = 4) modeled perfect-quality monitoring information. However, these do include manufacturing system-related studies from Adams et al. [155] and Koochaki et al. [160].

The reviewed studies modeled various sources of imperfections in monitoring information. Many of these studies allowed for monitoring information to allow for missed alarms (false negatives) or false alarms (false positives) when detecting equipment conditions (n = 22). Half of these studies come from the energy system-related applications [123,124,126–128,136,149,150,152,154,156], five of the studies come from the transportation mode-related applications [132,140,141,147,161], and five come from the remaining industrial applications [137,139,142,146,157] (including the production system-related study from Florian et al.). Among these studies, Compare et al. [140], Zhang et al. [137], Liu and Wang [148], and Yoon et al. [142] incorporate time-varying rates of false alarms or missed alarms. The inclusion of time-varying rates of false or missed alarms requires more model complexity, but they more accurately represent the fluctuations of monitoring performance during equipment operation.

In addition to allowing for missed or false alarms in detection, the studies from Azadeh et al. [70], Lei and Sandborn [162], Tian et al. [159], Chen et al. [143], and Mao et al. [132] modeled imperfect monitoring quality with RUL estimate uncertainties or prognostic errors (n = 5). Lastly, studies from Erguido et al. [156], Van Horenbeek et al. [126], Yang and Letourneau [147], and Wu et al. [163] modeled monitoring information that may be communicated to maintenance personnel at points of time where they may be less valuable (n = 4), such as a late prognostic alarm that does not allow ample time for maintenance to prevent an impending equipment failure.

3.3.2. Discussion

Concerning the maintenance policies modeled by condition monitoring evaluation studies, our review indicates that thirty-six of these forty-two studies model RM, TBM, or UBM policies alongside their CBM models. Modeling alternative policies to CBM is a point of comparison that can elucidate the impacts of condition monitoring in a maintenance policy. Surprisingly, only three studies modeled UBM policies [130,138,141], though this may be an artifact of the focus on aircraft applications in these three studies. Also, we found that although five studies integrated TBM policy-type inspections alongside the condition monitoring-triggered actions of CBM policy models [126,128,144,156,161], only Wang and Pecht [161] models both a CBM policy with and without incorporating TBM-style policies.

The implemented policy prescribes maintenance actions to address equipment maintenance needs. Although no standard definitions exist to delineate between different types of maintenance actions, our review found references to repair or replace actions in thirty-six studies. Another two studies referred to maintenance actions as corrective, preventive, or predictive work [143,162]. However, out of these thirty-eight studies, only twelve make distinctions for at least two types of maintenance actions. Furthermore, the studies from Neves and Frangopol [131] and Turnbull and Carroll [151] go beyond two types of maintenance actions and instead classify various types of repair and replacement actions. Van Horenbeek et al. [126] consider actions for secondarily damaged equipment. Integrating distinct maintenance actions indicates that the evaluation study considers the need for different maintenance actions to respond to various equipment failure modes and the distinct implications ascribed to executing each action (such as costs, maintenance duration, and degree of equipment condition improvement).

We note that while the majority of twelve studies that incorporate at least two types of maintenance actions pertain to energy system-related applications [125–128,144,150,151,154,162], the remaining three relate to manufacturing systems [155], transportation modes [143], and civil structures [131]. This imbalance towards energy system-related studies supports our finding that condition monitoring evaluations have been much more prevalent in this industrial sector and provide much more detail regarding models used for evaluations. In particular, evaluations of condition monitoring in manufacturing system applications should consider a larger degree of maintenance actions to reflect the diversity of maintenance work orders on the production floor.

Regarding details about the maintenance action characteristics, two important characteristics are the duration of a maintenance action

Table 14
Maintenance model characteristics in the energy system-related studies (n = 18).

Study	Maintenance duration	Maintenance quality
Azadeh et al. [70]	Sampled from probability distribution	AGAN
Bakhshi & Sandborn [125]	Fixed maintenance downtime for each component	Replace: AGAN Repair: HD
Chang et al. [123]	Fixed duration for time-to-repair or -replace	NR
Erguido et al. [156]	NR	BTO-WTN
Kerres et al. [124]	Fixed duration for inspections and maintenance types	AGAN
Lei & Sandborn [162]	Fixed downtime for corrective work, zero downtime for predictive work	NR
Li et al. [133]	NR	NR
Liang & Parlikad [128]	Fixed duration per maintenance type	Replace: AGAN Repair: HD
Livera et al. [144]	Fixed downtime per fault mode type	NR
Long et al. [136]	NR	AGAN
May et al. [149]	NR	NR
Nielsen et al. [127]	Fixed durations defined per maintenance type	Replace: AGAN Repair: Bernoulli distribution for each action to result in AGAN or ABAO
Peng et al. [154]	NR	NR
Puglia et al. [150]	Fixed durations defined per maintenance type	AGAN
Tian et al. [159]	NR, but authors model a fixed maintenance lead time	NR for wind turbine components, BTO-WTN for wind turbines
Turnbull & Carroll [151]	Fixed repair times per component and maintenance type	NR
Van Horenbeek et al. [126]	Time-to-repair variables per fault mode type, and a logistics waiting time parameter	AGAN
Vieira et al. [152]	NR	NR

Notes: AGAN = As Good As New. HD = Hampered Deterioration. BTO-WTN = Better Than Old but Worse Than New. ABAO = As Bad As Old.

Table 15
Maintenance model characteristics in the transportation mode-related studies (n = 9).

Study	Maintenance duration	Maintenance quality
Chen et al. [143]	NR	NR
Compare et al. [140]	NR	AGAN
Cot et al. [138]	NR	NR
Halbert et al. [141]	NR	NR
Hongsheng et al. [153]	Fixed durations per maintenance type	NR
Mao et al. [132]	Negligible duration for preventive work, and a maintenance-time variable for corrective work	NR
Reimann et al. [130]	NR	BTO-WTN
Wang & Pecht [161]	Negligible maintenance duration	AGAN
Yang & Letourneau [147]	NR	NR

Notes: NR = Not Reported. AGAN = As Good As New. BTO-WTN = Better Than Old but Worse Than New.

Table 16
Maintenance model characteristics in the remaining studies (n = 15).

Study	Maintenance duration	Maintenance quality
Adams et al. [155]	NR	NR
Florian et al. [157]	NR	NR
Golmakani & Fattahipour [158]	NR	NR
Iannacone et al. [134]	NR	NR
Klerk et al. [135]	NR	NR
Koochaki et al. [160]	Fixed duration for time-to-repair	AGAN
Liu & Wang [148]	NR	AGAN
Meng et al. [129]	NR	NR
Neves & Frangopol [131]	NR	AGAN, HD, or BTO-WTN, depending on maintenance action
Rastegari [145]	NR	NR
Shamayleh et al. [146]	Fixed duration for time-to-repair	NR
Wu et al. [163]	NR	NR
Yoon et al. [142]	NR	AGAN
Zhang et al. [137]	NR	AGAN
Zou et al. [139]	NR	AGAN

Notes: NR = Not Reported. AGAN = As Good As New. HD = Hampered Deterioration. BTO-WTN = Better Than Old but Worse Than New.

and the end result of an equipment's condition due to the action. Fixed, constant-value parameters proved to be the most common representation of maintenance duration in our reviewed studies. Azadeh et al. [70] modeled a more realistic maintenance duration representation by sampling a time-to-repair value from a probability distribution.

Furthermore, with respect to modeling accurate representations of maintenance practices, studies from Tian et al. [159] and Van Horenbeek et al. [126] included maintenance duration-related model parameters to represent maintenance lead time and logistical waiting time, respectively. Maintenance action models that result in as-good-as-new equipment conditions were the most common maintenance action quality models in our reviewed studies. Very few studies consider maintenance actions that result in equipment conditions that are as-bad-as-old, better-than-old-but-worse-than-new, or hampered in their deterioration rates. Unfortunately, even fewer studies consider more than one type of equipment condition end-result that signifies maintenance action quality [125,127,128,131]. Modeling the diversity in maintenance effects on equipment condition is an opportunity to more accurately characterize maintenance actions, especially when there are several types of possible actions in the maintenance model, and each impacts the equipment.

Lastly, emphasizing the lack of detail for the maintenance model characteristics in the survey, our results also reveal that only nine studies specify both characteristics in their maintenance models. Overall,

Table 17
Monitoring information characteristics used by maintenance models in the energy system-related studies (n = 18).

Study	Monitoring mechanism	Monitoring quality
Azadeh et al. [70]	CM	Imperfect (prognostics errors)
Bakhshi & Sandborn [125]	Periodic monitoring of a subset of equipment	NR
Chang et al. [123]	CM	Imperfect (FN)
Erguido et al. [156]	CM	Imperfect (FN, and possible time-inefficient/late detections)
Kerres et al. [124]	Monitoring times sampled from time-to-detect distribution	Imperfect monitoring (FN) Perfect inspections
Lei & Sandborn [162]	Periodic estimations of RUL	Imperfect (uncertainty in RUL estimates)
Li et al. [133]	Intervalled monitoring	NR
Liang & Parlikad [128]	CM, and periodic inspections	Imperfect (FN)
Livera et al. [144]	CM	NR
Long et al. [136]	CM, and inspections	Imperfect monitoring (uncertainties in data) Imperfect inspections (FN)
May et al. [149]	NR	Imperfect (FN and FP)
Nielsen et al. [127]	CM, and inspections	Imperfect (FN and FP)
Peng et al. [154]	CM	Imperfect (FN)
Puglia et al. [150]	NR	Compares perfect and imperfect (FN)
Tian et al. [159]	Periodic estimation of RUL prediction	Imperfect (RUL prediction error)
Turnbull & Carroll [151]	CM	NR
Van Horenbeek et al. [126]	CM and inspections	Imperfect (FP and FN, and possible time-inefficient/late detections)
Vieira et al. [152]	CM	Imperfect (FN)

Notes: NR = Not Reported. RUL = Remaining Useful Life. FN = False Negatives. FP = False Positives. CM = Continuous Monitoring.

opportunities exist to further examine the integration of maintenance duration and maintenance action quality details into maintenance models. This examination can help ascertain their impact on condition monitoring evaluation studies and the degree of modeling required for accurate evaluation results. The existing condition monitoring evaluation studies for manufacturing applications are significantly lacking in considering these maintenance model characteristics, as only Koochaki et al. [160] incorporate details regarding maintenance action duration or quality in their study.

Our review also offers insights into the modeled monitoring information that interfaces with maintenance personnel and informs their actions. The first of the two monitoring information characteristics that we focused on is the monitoring mechanism, where we found that most studies in our review assume continuous information retrieval of equipment conditions. Regarding the monitoring mechanism results, we identify an opportunity to examine further the impact of the frequency and timing of condition monitoring information retrieval on maintenance personnel, equipment operations, and business value. For instance, the points in time that maintenance personnel receive monitoring information may be periodic, as specified by Florian et al. [157], or non-periodic, akin to the age-based inspection schemes discussed by Golmakani and Fattahipour [158]. We note that high frequencies of receiving monitoring information mimic continuous monitoring. Another aspect of examining monitoring frequency and timing are the constraints in a condition monitoring application. For example, Bakhshi and Sandborn [125] modeled the monitoring limitations of applying a single CMS to numerous wind turbines in a wind farm. Including such constraints provides more factual representations in impact evaluation.

Table 18
Monitoring information characteristics used by maintenance models in the transportation mode-related studies (n = 9).

Study	Monitoring mechanism	Monitoring quality
Chen et al. [143]	CM	Imperfect (incorrect RUL interval predictions)
Compare et al. [140]	CM with periodic RUL estimations	Imperfect (time-varying FP and FN)
Cot et al. [138]	CM	Perfect
Halbert et al. [141]	Periodic, constant-interval inspections or monitoring	Imperfect (FP)
Hongsheng et al. [153]	NR	Imperfect (FP and FN)
Mao et al. [132]	Fixed intervals for detection and RUL estimates	Imperfect (possible inaccurate detections and RUL estimates)
Reimann et al. [130]	CM	Perfect
Wang & Pecht [161]	CM	Imperfect (FN)
Yang & Letourneau [147]	CM	Imperfect (FP and time-inefficient detections that are too early or too late)

Notes: RUL = Remaining Useful Life. FN = False Negatives. FP = False Positives. CM = Continuous Monitoring.

Our review focused on the second monitoring information characteristic, which is the quality of the monitored data communicated to maintenance personnel. Unsurprising for a review of condition monitoring evaluation studies, we found that many studies modeled CMSs that conveyed imperfect, inaccurate monitoring information regarding fault or failure detection, diagnosis, or prognosis.

However, many studies asserted that their monitoring system reported inaccurate information at a fixed rate, such as a constant value representing a CMS's detection rate. Interestingly, we observe a recent trend of integrating time-varying rates of false alarms or missed alarms associated with detection or prognosis, found in the studies by Compare et al. [140], Zhang et al. [137], Liu and Wang [148], and Yoon et al. [142]. Modeling inaccuracies in condition monitoring information with time-varying estimates of monitoring performance helps reflect changing operating environments for the equipment and the CMS. We note that modeling uncertainty into measurand data [136] or detection thresholds [131] can also help model dynamically changing imperfect monitoring information that better reflects real-world scenarios.

Only four studies considered the timeliness of monitoring information [126,147,156,163]. Future evaluation studies could consider how the timeliness of monitoring information, relative to maintenance lead time, affects maintenance effectiveness.

Regarding both monitoring information characteristics, this literature review recommends that future condition monitoring evaluation research examine the relationship between the frequency of monitoring information retrieval and the quality of monitoring information. Exploring any trade-offs between monitoring frequency and monitoring quality can elucidate a better understanding of the benefits and risks of using a CMS.

3.4. Evaluation techniques

3.4.1. Results

Tables 21 to 23 present the evaluation techniques used across all of the studies in this literature survey. The aim of reviewing the evaluation techniques in these studies is to understand the analytical frameworks, computational methods, and sensitivity analyses used to evaluate condition monitoring applications.

This review reveals that slightly over half of the selected studies include analytical formulations that capture the dynamics of condition monitoring and maintenance (n = 24). The review focused on studies that included some degree of maintenance decision-making, equipment operations, and condition monitoring information in their condition monitoring-based maintenance evaluations. Interestingly, only three studies compared the solutions of analytical formulations of condition-based maintenance decision-making to approximate solutions from numerical methods [70,138,161].

This review also reveals a wide range of computational methods in nearly all surveyed papers (n = 41). For each study, this review identified the relevant techniques or tools that the authors described and used to compute evaluations for condition monitoring-enabled maintenance. This review revealed that each of the surveyed studies provides varying details about their computational methods, and several complementary computational techniques were used in numerous studies.

The varying levels of details regarding computational methods that are present in the surveyed studies makes it challenging to delineate strict classifications for this data item. However, to organize the surveyed computational methods, Table 20 introduces a broad classification, which is by no means exclusive, of the tools and techniques that

Table 19
Monitoring information characteristics used by maintenance models in the remaining studies (n = 15).

Study	Monitoring mechanism	Monitoring quality
Adams et al. [155]	CM	Perfect
Florian et al. [157]	Periodic monitoring (scoring frequency)	Imperfect (FP and FN)
Golmakani & Fattahipour [158]	Non-periodic, age-based inspections	Perfect
Iannacone et al. [134]	CM, and inspections	Imperfect (monitoring errors)
Klerk et al. [135]	CM	NR
Koochaki et al. [160]	CM	Perfect
Liu & Wang [148]	CM	Imperfect Case 1: Detection from a fixed normal distribution Case 2: RUL prediction from an improving, time-varying distribution
Meng et al. [129]	CM	Perfect
Neves & Frangopol [131]	NR	Imperfect (probabilistic detection thresholds to represent measurement uncertainties)
Rastegari [145]	CM	NR
Shamayleh et al. [146]	CM	Imperfect (FP and FN)
Wu et al. [163]	CM	Imperfect (time-inefficient predictions that are too early or too late)
Yoon et al. [142]	CM	Imperfect (time-varying FP and FN)
Zhang et al. [137]	CM	Imperfect (time-varying FP and FN)
Zou et al. [139]	Non-periodic monitoring	Imperfect (FN)

Notes: FN = False Negatives. FP = False Positives. CM = Continuous Monitoring. RUL = Remaining Useful Life.

Table 20
Major classifications for the tools and techniques that comprise the computational methods.

	First principles-based	Data-driven
System Representations	-Models that characterize the system	-Datasets of system observations
System Analyses	-Numerical methods -Simulations -Monte Carlo methods	-Machine learning-based detections or predictions

comprise this data item. Using the definition of a system as described in Section 2.1 as any material artifact being operated on to fulfill a purpose, this classification separates the computational methods as either pertaining to the construction of system representations or execution of system analyses.

Data-driven system representations typically consist of datasets of observations from the system. These system observations can come in various formats, including quantified, numeric sensor signals or human-generated textual information. Data-driven system analysis techniques include machine learning algorithms that make use of these datasets. These algorithms learn from the datasets to make predictions or inferences regarding the system.

Several studies in this literature review conducted evaluations with computational methods that used data-driven system representations and system analyses. These studies trained and applied machine learning algorithms to given datasets (n = 4). These studies obtained insight into the ability of those algorithms to detect conditions or forecast faults present in the dataset [143,144,146,147]. The datasets in these studies contain a history of equipment operation via time-series sensor signals.

First principles-based system representations typically use models derived from a system's underlying physics, engineering, and heuristic behaviors [165]. Characterizations of particular interest are whether a system is time-invariant, nonlinear, discrete-state, and event-driven, as these indicate that the system can be modeled as a discrete-event system [166]. This class of system representations are used extensively for modeling production systems, transportation systems, and health care systems. Related to discrete-event systems are agent-based model representations, which can be considered a special case of discrete-event systems where agents are entities that interact with other agents and their environment [167].

Also of interest are the output variables of a discrete-event system. If the output variables are random variables, then probabilistic models are used to characterize system behaviors. These probabilistic models include Markov models and Bayesian decision models. System behaviors, regardless of being probabilistic, may also be characterized with general-purpose systems engineering models.

First principles-based system analysis techniques mainly include numerical methods that compute approximate solutions to mathematical problems. Numerical methods are used to execute computations for simulations, which capture the dynamics of the system model's evolution over time. Discrete-event simulation in particular refers to the study of the dynamics of discrete-event systems. In addition to simulations, Monte Carlo methods refer to a class of numerical methods that use random numbers to estimate the value of an unknown stochastic or deterministic quantity.

The use of nondescript numerical methods (n = 11) and nondescript simulations (n = 6) frequently appeared in our review. However, discrete-event simulations also frequently appeared as a computational tool (n = 9). Roughly half of the studies that utilized discrete-event simulations came from the energy system-related studies in Table 21 (n = 4). Two manufacturing-related evaluation studies also utilized discrete-event simulations [155,160], while the other two manufacturing studies did not specify many details about their computational methods. Moreover, only one study computed results with agent-based simulations [156].

Monte Carlo methods were used in many studies across our survey (n = 15), with roughly half of these studies about energy system-related evaluation studies (n = 8) and a third of them to evaluation studies that focus on the remaining industrial applications (n = 5). Monte Carlo methods are often used with (stochastic) discrete-event simulation models [125,142,151,163] and probabilistic graphical models such as Markov models or Bayesian models [127,134,139,140].

Several studies represented the underlying stochasticity of their system with Bayesian decision models, particularly to understand maintenance decision-making behavior and monitoring information value

Table 21
Analytical Formulation (AF), Computations Methods (CM), and Sensitivity Analyses (SA) associated with the evaluation techniques in the energy system-related studies (n = 18).

Study	AF	CM	SA
Azadeh et al. [70]	Yes	DES with MM	-Equipment fail-time distribution parameter -Prognostics error level
Bakhshi & Sandborn [125]	No	DES, MCM	-Equipment fail-time distribution parameter
Chang et al. [123]	No	DES	-Equipment fail-time distribution parameters
Erguido et al. [156]	Yes	ABS	No
Kerres et al. [124]	No	MCM	-Electricity costs -Equipment fail-time distribution parameter -Wait time between maintenance alert and maintenance start
Lei & Sandborn [162]	Yes	MCM	No
Li et al. [133]	No	MCM	No
Liang & Parlikad [128]	Yes	Numerical methods with MM	-Repair costs -Downtime costs
Livera et al. [144]	No	MLAD	-Maintenance frequency -Energy yield -Electricity costs
Long et al. [136]	Yes	Simulation with BDM	-Maintenance costs -Cost discounting rate
May et al. [149]	Yes	Simulation with MM	-Increased sensors -CMS detection rates
Nielsen et al. [127]	Yes	MCM and BDM	-Maintenance frequency -Detection threshold -Maintenance costs -Maintenance quality
Peng et al. [154]	No	Fuzzy arithmetics	No
Puglia et al. [150]	No	Numerical methods	-Cost discounting rate
Tian et al. [159]	Yes	Simulation	-Detection threshold -Maintenance frequency
Turnbull & Carroll [151]	No	DES, MCM	-Component failure rate parameters
Van Horenbeek et al. [126]	No	MCM	-Monitoring performance parameters
Vieira et al. [152]	Yes	MCM	-Equipment service life -Maintenance frequency -Monitoring performance parameters

Notes: DES = Discrete-Event Simulation. MCM = Monte Carlo Method. ABS = Agent-Based Simulation. MLAD = Machine learning Applied to Dataset. BDM = Bayesian Decision Model. MM = Markov Model.

(n = 7). The majority of evaluation studies on civil or marine structure applications construct Bayesian decision models [134,135,137,139,148], while these models do appear in two energy system-related studies. Another set of studies use Markov models (n = 4), primarily to capture the dynamics of equipment operational state, though in the case of Compare et al. the Markov model aims to represent not the equipment state itself but the outcome of a monitoring or inspection check of the equipment's state [140]. Most of these papers appear in energy system-related studies [70,128,149]. Finally, Mao et al. [132] focused on using the Systems Modeling Language (SysML) to represent PHM applications and maintenance behaviors.

Lastly, with respect to computational methods, Peng et al. [154] uniquely used fuzzy numbers to quantify the risk of wind turbine component faults and failures, both with and without a CMS. With fuzzy arithmetic, this study's evaluation quantifies the uncertainty in failure mode costs.

As for sensitivity analyses conducted in the selected studies, this paper identified evaluation model inputs that each study uses to observe influences on their respective model outputs. Most studies conducted some sensitivity analysis on their input model parameters (n = 30),

Table 22
Analytical Formulation (AF), Computations Methods (CM), and Sensitivity Analyses (SA) associated with the evaluation techniques in the transportation mode-related studies (n = 9).

Study	AF	CM	SA
Chen et al. [143]	Yes	MLAD	-Maintenance costs
Compare et al. [140]	Yes	Numerical methods, MM, and MCM	-Monitoring performance parameters
Cot et al. [138]	Yes	Simulation	-Detection threshold
Halbert et al. [141]	No	Numerical methods	
Hongsheng et al. [153]	Yes	MCM	-Monitoring performance parameters
Mao et al. [132]	No	Simulation, SysML	No
Reimann et al. [130]	No	Simulation	No
Wang & Pecht [161]	Yes	DES	-Wait time between maintenance alert and maintenance start -Prognostic distance
Yang & Letourneau [147]	No	MLAD	No

Notes: MLAD = Machine Learning Applied to Dataset. MCM = Monte Carlo Method. MM = Markov Model. SysML = Systems Modeling Language. DES = Discrete-Event Simulation.

Table 23
Analytical Formulation (AF), Computations Methods (CM), and Sensitivity Analyses (SA) associated with the evaluation techniques in the remaining studies (n = 15).

Study	AF	CM	SA
Adams et al. [155]	No	DES	No
Florian et al. [157]	Yes	Numerical methods	-Maintenance costs -Detection frequency -Monitoring performance parameter -Rate of detecting false positive upon inspection
Golmakani & Fattahipour [158]	Yes	Numerical methods	-Inspection costs
Iannacone et al. [134]	Yes	Numerical methods, MCM, BDM	No
Klerk et al. [135]	Yes	Numerical methods, BDM	-Equipment operational load/stress level
Koochaki et al. [160]	No	DES and plant modeling	-Equipment failure rate
Liu & Wang [148]	Yes	Numerical methods, simulation, BDM	-Maintenance costs -Equipment fail-time distribution parameter -Monitoring performance parameter
Meng et al. [129]	Yes	Numerical methods	-Maintenance frequency -Maintenance costs -Equipment degradation parameter
Neves & Frangopol [131]	Yes	MCM	No
Rastegari [145]	No	NR	No
Shamayleh et al. [146]	No	MLAD	No
Wu et al. [163]	No	DES, MCM	-Number of monitored components -Monitoring costs -Duration of monitoring -Prognostic distance
Yoon et al. [142]	Yes	DES, MCM	-Maintenance costs -Monitoring performance parameters
Zhang et al. [137]	Yes	Numerical methods with BDM	-Monitoring performance degradation parameters -Duration of monitoring
Zou et al. [139]	Yes	MCM with BDM	-Maintenance costs

Notes: DES = Discrete-Event Simulation. MCM = Monte Carlo Method. MLAD = Machine Learning Applied to Dataset. BDM = Bayesian Decision Model.

including two of the four evaluation studies related to manufacturing systems [157,160]. Roughly half of the studies conducting sensitivity analysis do so concerning at least two input model parameters (n = 16), including nine energy system-related evaluation studies. Only one of the nine transportation mode-related evaluation studies conducts sensitivity analysis on at least two parameters [161], as well as only one of the four manufacturing-related studies [157]. Although Florian et al. [157] is the sole manufacturing-related study that conducts sensitivity analysis, the study considers four parameters relevant to condition monitoring performance and maintenance costs.

Several parameters that undergo sensitivity analysis relate to equipment deterioration or lifetime. These parameters include equipment failure-time distribution variables, equipment or component failure rates, service life, load or stress levels, and equipment degradation parameters. Of the studies that consider equipment deterioration or lifetime (n = 10), six belong to energy system-related evaluation studies, and none belong to transportation mode-related studies.

Quite a few parameters used for sensitivity analysis pertain to maintenance deployment and execution. These parameters encompass maintenance lead times and wait times, maintenance or inspection costs, downtime costs, maintenance frequency, and maintenance quality. These parameters appeared in a large portion of the studies (n = 15), seven of them pertaining to energy system-related applications and two to transportation mode-related applications.

Many sensitivity analysis parameters relate to condition monitoring performance and operation. These parameters include monitoring performance parameters (including prognostic error and detection rates), monitoring performance degradation, detection threshold, prognostic distance, sensor coverage or number of monitored components, monitoring duration, monitoring frequency, and monitoring costs. Numerous studies conducted a sensitivity analysis on these parameters (n = 15), six belonging to energy system-related evaluations and four to transportation mode-related evaluations.

Lastly, the remaining parameters considered for sensitivity analysis deal with electricity costs, energy yield, and cost discounting rates (n = 4). These four studies are energy system-related.

3.4.2. Discussion

Prior works proclaim that analytical methods or simulations drive condition monitoring-based maintenance evaluation techniques and remark on the need for more analytical approaches. Bakhshi and Sandborn [125] note that most studies calculate maintenance costs for wind turbines through analytical methods or simulations. Wang and Pecht [161] propose an analytical method for the cost–benefit analysis of canary-based PHMs, but compare analytical results to simulations and note the strength of simulators for complex, analytically-insolvable maintenance models. In Compare et al. [3], the authors assert that many studies typically use simulation-based methods, rather than any general analytical approach, to evaluate predictive maintenance cost–benefits. In another work, Compare et al. [140] discuss a few evaluation studies that present an analytical approach but not one that fully captures the dynamics of condition-based maintenance and its PHM algorithms. As a result of the discussions in these prior works, this section attempts to understand better the analytical formulations underpinning condition monitoring evaluation studies.

Analysis of the results found that delineating the degree to which evaluation studies develop an analytical framework of condition-based maintenance dynamics is subjective. The survey results show that twenty-four reviewed studies go beyond aggregating maintenance costs, including parameters or variables that capture maintenance decision-making, equipment operations, and condition monitoring information. However, only three studies derive analytical solutions. In these cases, deriving analytical solutions helps verify approximated solutions from numerical methods. Verification with analytical solutions is often infeasible with the complexity of condition-based maintenance policies [70], though when possible, evaluation studies should implement such verification tests even if they encompass comparisons to a subset of numerical results.

Concerning the computational methods used within the surveyed studies, the review reveals that the selected studies most frequently use Monte Carlo methods (n = 15). Maintenance models typically use stochastic processes to represent equipment conditions, maintenance outcomes, and other logistics. Monte Carlo simulations allow researchers to predict those stochastic processes via sampling experiments. Monte Carlo simulation has general applicability for many applications [168]. As expected, many modeling paradigms, such as the discrete-event simulation model in the study by Yoon et al. [142], use Monte Carlo sampling to conduct experiments.

A review of the computational methods used in the selected studies also points to issues related to equipment and sensor data availability. The availability of equipment degradation datasets enables researchers to test monitoring algorithms on their ability to detect faults or predict failures [143,144,146,147]. However, without discounting the value these datasets can bring to studying CMSs and their algorithms, they have limitations. For one, the equipment operation scenarios they represent may not capture the full breadth of scenarios the equipment may encounter. Furthermore, the datasets tend to not include the decision-making processes that occur between condition monitoring information retrieval and maintenance execution.

Discrete-event modeling and simulation appeared in several of our studies (n = 9), including two manufacturing related evaluations [155,160]. Although these two manufacturing-related studies form too small a sample size to make any definitive conclusions, they were the only two of this review's selected studies that provided details about their computation methods. Discrete-event models represent a system's state at each point in time by capturing the entities that make up a system and the interactions of those entities that cause changes to the system state at discrete points in time. Given these capabilities, discrete-event models can model the equipment operations and maintenance decision-making dynamics, allowing researchers to explore extreme or unexpected scenarios [169]. The agent-based simulations utilized by Erguido et al. [156] also point to an opportunity for conducting condition monitoring-based maintenance evaluations with discrete-event simulations. Adopting agent-based modeling in discrete-event simulations allows researchers to model the behavior of autonomous agents in a network, where a set of rules governs each agent's behavior. The resulting simulation can provide insights into emergent behaviors in the system. Integrating agent-based modeling of maintenance decision-making can allow discrete-event simulations to more accurately represent maintenance.

A few reviewed papers use Bayesian decision models to understand information value in the context of maintenance decision-making (n = 7). These studies construct decision trees for Bayesian decision-making to analyze the value of information (VOI) that can come from monitoring or inspections. Since, at each time step, maintenance personnel have many options for decisions regarding the acquisition of information or maintenance actions, and the outcome of each of these decisions can result in many subsequent decisions, the maintenance decision tree for each time step can become complex. Evaluating the outcomes of each decision tree at each time step can quickly become computationally expensive [3,127]. The bulk of studies that use Bayesian decision models to consider in their computational methods relate to civil or marine structures [134,135,137,139,148], where the longer operational timescales of months or years make it possible to scale-down on the size of decision trees.

Markov models also appeared in several studies (n = 4), representing equipment operational state. These models are appropriate when discrete states can describe equipment operation or degradation and possess simple analytical representations [3]. However, definitions of equipment states and transition state probabilities require careful model development, as demonstrated in the study by Compare et al. [140], and the scale of states and transitions can rapidly grow when considering the impact of condition-based maintenance on equipment states, as demonstrated in the study by Liang and Parlikan [128]. The effort required for model development and the scalability of states and transitions can make Markov model simulations computationally expensive.

A unique approach for obtaining simulation results appeared in the study from Mao et al. [132], where the authors used SysML as a modeling framework to simulate and evaluate PHM systems. SysML allows researchers to model structural and behavioral representations of equipment operations and maintenance, and among its strengths is the ability to trace and manage requirements for condition monitoring, maintenance deployment, and the equipment itself.

Regarding sensitivity analysis, our review reveals that many studies considered at least two input model parameters (n = 16). We could classify many of these parameters as related to equipment deterioration and lifetime, maintenance deployment and execution, or monitoring performance and operation. From this classification, we can observe the types of model parameters considered for analyzing each equipment, maintenance, or monitoring.

Furthermore, we can obtain insights regarding any selection process or rationale for the parameters under sensitivity analysis. We find that several papers studied parameters related to both maintenance and condition monitoring [127,142,157,159,161], including the manufacturing-related study from Florian et al. Fewer sensitivity analyses look into parameters related to the equipment, and either maintenance or condition monitoring [69,124,129]. However, two studies cover parameters related to each of equipment, maintenance, and monitoring [148,152]. Although conducting a sensitivity analysis of several parameters related to just one of these three classes can be valuable [163], this trend suggests a gap and an opportunity for evaluation studies to consider sensitivity analysis of parameters from equipment, maintenance, and monitoring.

3.5. Performance measures

3.5.1. Results

Tables 24 to 26 show the performance measures used in our reviewed evaluation studies, divided into each table concerning whether the study focused on energy system applications, transportation mode applications, or other industrial applications. Most studies use more than one performance measure (n = 32). We extracted the performance measures from each evaluation study and categorized them according to whether they are associated with equipment key performance indicators (KPIs), maintenance KPIs, or monitoring metrics. No study in our review used any implementation measure as discussed in Section 2.6.5.

Equipment KPIs appear in less than half of our reviewed studies (n = 18). The most commonly reported equipment-level KPIs are availability (n = 7) and energy production (n = 6). Of the studies that relate to energy systems and use equipment KPIs (n = 10), six report energy production measures, four report availability, and Van Horenbeek et al. [126] reports downtime (closely related to availability). In addition to energy production, the studies from Bakhshi and Sandborn [125], and Livera et al. [144] also account for reliability and availability measures, respectively. Few studies that relate to transportation modes ascribe performance measures to their equipment (n = 2), and both report equipment availability measures.

Of the equipment KPIs in the remaining studies related to industrial applications in Table 26 (n = 6), three studies report equipment reliability or risk measures. Another three studies report KPIs related to their manufacturing applications. Adams et al. [155] considers overall equipment effectiveness and production costs. Koochaki et al. [160] presents measures for availability and production line efficiency. Rastegari [145] reports production speed. Finally, we note that measures of availability, downtime, and reliability are closely linked with maintenance, but we categorize them as equipment KPIs as they convey information about equipment operation and status.

All of our reviewed studies use maintenance-related KPIs to measure some degree of performance in their evaluations (n = 42). A performance measure that we denote as maintenance action costs appears in most of our reviewed studies (n = 40).

The two studies that did not focus on reporting maintenance action costs were from studies related to transportation modes: Cot et al. [138] provided estimates of the number of maintenance actions, and Mao et al. [132] measured the ratio of necessary, true-alarmed preventative repairs to all repairs. The energy system-related study from Livera et al. [144] reported the number of maintenance actions in addition to maintenance costs, while both Azadeh et al. [70] and Chang et al. [123] included a measure for mean times to execute a maintenance action. Of the manufacturing-related evaluations, Adams et al. [155] also used maintenance labor hours as an evaluation metric.

Monitoring metrics are used in a slight majority of our reviewed studies to evaluate performance (n = 23), and most of these studies include more than one monitoring metric (n = 15). Ten of the eighteen energy system-related studies account for monitoring metrics, as do six of nine transportation mode-related studies and seven of the fifteen studies that pertain to the remaining industrial applications. Florian et al. [157] is the sole manufacturing-related study that harbors monitoring metrics, specifying true positive rate (TPR, also known as sensitivity), false positive rate (FPR), the area under the receiver operating characteristic curve (AUROC), and the F1-score.

The most commonly reported monitoring metrics are relevant to measuring classification or detection performance measures rather than measuring the performance of regression or prognostics. These commonly reported metrics are true positive counts or TPRs (n = 14), false positive counts or FPRs (n = 8), and accuracy (n = 5). Other monitoring metrics include prognostic distance (n = 3), RUL utilization (n = 3), false negative counts or false negative rates (n = 3), prediction error (n = 2), and true negative counts or true negative rates (n = 1). Furthermore, detection efficiency appeared in two studies, which Van Horenbeek et al. [126] uses to gauge how soon the CMS detects equipment degradation before its failure. Compare et al. [140] introduces a cumulative first false positive metric, a time-varying metric representing the equipment survival probability up to an inspection or monitoring check, considering that it may have failed at or before prior checks. Both studies from Yoon et al. [142] and Compare et al. [140] consider time-varying monitoring metrics to best represent monitoring performance. Though the inclusion of time-varying metrics requires more modeling complexity, but they provide a better representation of monitoring performance as it may change throughout equipment operation due to factors such as environmental disturbances or changes in equipment operations.

We note that several studies use performance measures from each category of equipment KPIs, maintenance KPIs, and monitoring metrics (n = 10). Six of these studies relate to energy systems [123,124,126,128,144,151], two deal with transportation modes [132,141], one corresponds to marine structures [139], and another relates to generic equipment [142]. On the other end of the spectrum, a few studies only use performance measures that we consider to be classified as maintenance KPIs without considering equipment KPIs or monitoring metrics (n = 11). Four of these studies relate to energy systems [70,133,136,154], three correspond to transportation modes [130,138,153], two deal with civil structures [131,134], one relates to software [129], and another to generic equipment [158].

Table 24
Performance measures for equipment, maintenance, and monitoring in the energy system-related studies (n = 18).

Study	Equipment KPIs	Maintenance KPIs	Monitoring metrics
Azadeh et al. [70]	No	MAC, MTTF	No
Bakhshi & Sandborn [125]	EP, Reliability Improvement	MAC	No
Chang et al. [123]	Availability	MAC, MTTR	Prognostic distance
Erguido et al. [156]	No	MAC	TPR, Detection efficiency
Kerres et al. [124]	Availability	MAC	TPR
Lei & Sandborn [162]	EP	MAC	No
Li et al. [133]	No	MAC	No
Liang & Parlikad [128]	Availability	MAC	TPR
Livera et al. [144]	EP, Availability	Number of actions, MAC	Accuracy
Long et al. [136]	No	MAC	No
May et al. [149]	No	MAC	FPR, TPR Ratio of TPR made 6 months in advance
Nielsen et al. [127]	No	MAC	FPR, TPR
Peng et al. [154]	No	MAC	No
Puglia et al. [150]	EP	MAC	No
Tian et al. [159]	No	MAC	Prediction error
Turnbull & Carroll [151]	EP	MAC	TPR, RUL utilization
Van Horenbeek et al. [126]	Downtime	MAC	TPR, Detection efficiency
Vieira et al. [152]	EP	MAC	No

Notes: MAC = Maintenance Action Costs. MTTF = Mean Time To Failure. EP = Energy Production. MTTR = Mean Time To Repair. TPR = True Positive Rate. FPR = False Positive Rate. RUL = Remaining Useful Life.

3.5.2. Discussion

Performance measures provide insights into the impact of CMS-enabled maintenance policies on industrial equipment and their operations. These performance measures often inform the economics models, discussed in more detail in Section 3.6, that capture the costs and benefits of CMS integration. Our review categorizes the sheer number of performance measures we identified in the studies relating to the industrial equipment, the implemented maintenance policy, or condition monitoring algorithmic performance. Our review notes that the variety of industrial applications, condition monitoring tools, and maintenance deployment schemes undermines any comparison between performance measure results and studies. Instead, we observe trends in the types of performance measures that the studies utilized, and we focus on gaps and opportunities that may be suitable for evaluating condition-enabled maintenance in manufacturing applications.

We can categorize the equipment KPIs in our reviewed studies into KPIs related to availability versus domain-specific KPIs. The former KPIs are intimately linked with maintenance, as maintenance is traditionally expected to improve an industrial equipment's availability and reliability [34]. Domain-specific KPIs provide vital information about a CMS's impact on equipment functionalities. However, we observed that only evaluation studies of CMS applications to energy systems and manufacturing systems use a limited number of domain-specific KPIs to characterize equipment productivity and efficiency. We also note that the adoption of equipment KPIs is relatively low (n = 18), and even lower for the adoption of domain-specific equipment KPIs (n = 9). Using equipment-level KPIs can help gauge the overall impact of both using a CMS and improving a CMS's performance concerning equipment performance. Although we find it promising that our small

Table 25
Performance measures for equipment, maintenance, and monitoring in the transportation mode-related studies (n = 9).

Study	Equipment KPIs	Maintenance KPIs	Monitoring metrics
Chen et al. [143]	No	MAC	RUL accuracy
Compare et al. [140]	No	MAC	False negatives, False positives, Cumulative first false positives
Cot et al. [138]	No	Number of actions	No
Halbert et al. [141]	Availability	MAC	TPR, FPR
Hongsheng et al. [153]	No	MAC	No
Mao et al. [132]	Availability	Ratio of true-alarm preventive repairs to all repairs	TPR, RUL accuracy, RUL utilization
Reimann et al. [130]	No	MAC	No
Wang & Pecht [161]	No	MAC	Prognostic distance
Yang & Letourneau [147]	No	MAC	False positives, False negatives, True positives, RUL utilization

Notes: MAC = Maintenance Action Costs. RUL = Remaining Useful Life. TPR = True Positive Rate. FPR = False Positive Rate.

Table 26
Performance measures for equipment, maintenance, and monitoring in the remaining studies (n = 15).

Study	Equipment KPIs	Maintenance KPIs	Monitoring metrics
Adams et al. [155]	OEE, PC	MAC, LH	No
Florian et al. [157]	No	MAC	TPR, FPR, AUROC, F1-score
Golmakani & Fattahipour [158]	No	MAC	No
Iannacone et al. [134]	No	MAC	No
Klerk et al. [135]	Risk level	MAC	No
Koochaki et al. [160]	Availability, Production line efficiency	MAC	No
Liu & Wang [148]	No	MAC	Prediction error, Number of predictions
Meng et al. [129]	No	MAC	No
Neves & Frangopol [131]	No	MAC	No
Rastegari [145]	Production speed	MAC	No
Shamayleh et al. [146]	No	Inventory costs, MAC	Accuracy, TPR, FPR, TNR
Wu et al. [163]	No	MAC	Prognostic distance, Accuracy
Yoon et al. [142]	Reliability index	MAC	Time-varying FPR and FNR
Zhang et al. [137]	No	MAC	TPR, FPR
Zou et al. [139]	Reliability index	MAC	TPR

Notes: MAC = Maintenance Action Costs. OEE = Overall Equipment Effectiveness. PC = Production Costs. LH = Labor Hours. TPR = True Positive Rate. FPR = False Positive Rate. AUROC = Area Under the Receiver Operating Characteristic curve. TNR = True Negative Rate. FNR = False Negative Rate.

sample size of manufacturing system-related studies presents results with equipment performance measures, an opportunity exists to expand on the various manufacturing performance measures used in condition monitoring evaluation studies.

Our review reveals that aggregating costs from maintenance related actions is the most common measure of maintenance performance. We do note that maintenance action costs can take many forms, such as an aggregation of maintenance task costs in the study by Puglia et al. [150] or maintenance costs per unit time in the study by Chen et al. [143]. They may include labor costs, as in the study by Adams et al. [155], or distinguish between maintenance costs due to failure versus preventative actions, as in the study by Tian et al. [159]. The actions quantified in maintenance costs can become industrial application-specific. For example, costs of maintenance actions can be affected by dispatch logistics for maintaining wind farms [156] or the available personnel and types of facilities in which aircraft undergo maintenance [130]. Though the aggregation of maintenance costs is beneficial in that they can seamlessly be used in economic evaluations of CMS impact, evaluation studies can benefit from considering maintenance KPIs such as maintenance personnel efficiency or utilization, maintenance work order turnover, the number of each type of maintenance action, counts of rework, and ratios between costs of particular maintenance actions to total maintenance costs [34].

Concerning the monitoring metrics adopted in the reviewed studies, we find that several studies use accuracy or true negative rates (TNR), which can be misleading for CMSs. The industrial applications in our review skew towards operating without faults or failures, and likewise, the CMS data inputs will typically represent normal equipment operations, creating a data imbalance that leads to misleading accuracy or TNR scores. Three studies also solely rely on true positive rates (TPR), which reveal only aspects of monitoring performance. Six studies pair TPR scores with false positive rates (FPR), providing more context to monitoring performance. However, given the imbalance of CMS data inputs, precision would be a more apt pairing with TPR. Precision focuses on the accuracy of positive detections or predictions, avoiding the misrepresentation of false positives when the data is imbalanced. To our surprise, none of the studies used precision as a metric. However, the manufacturing-related study Florian et al. [157] relied on F1-scores, a harmonic balance between TPR and precision. We recommend that future evaluations report precision, TPR, F1-scores, and the area under the precision–recall curve (AUPRC). Another opportunity comes from investigating time-varying versions of these metrics, as in Yoon et al. [142] and Compare et al. [140], as they may provide more context to monitoring performance.

We also found that, among monitoring metrics, only a few reviewed studies used metrics that primarily focus on the performance of prognostics or regression, such as RUL utilization, detection efficiency, prognostic distance, and mean absolute error. Other relevant metrics we suggest for measuring the performance of fault forecasts and prognostics include the $\alpha - \lambda$ performance and convergence [31].

As none of our evaluation studies used implementation measures, considering facets related to the trustworthiness, prediction capability, or interpretability of monitoring information outputs can be an avenue for further research. This category of implementation measures can support the selection of condition monitoring tools.

3.6. Economics analysis

3.6.1. Results

Tables 27 to 32 highlight key aspects of the economics analysis efforts that went into evaluating condition monitoring-enabled maintenance in our reviewed studies. We interpret economics analysis in these evaluation studies to include approaches and techniques to understand the financial impact of CMS-related investments.

Table 27 aligns each of the studies in our review with the underlying motives or intentions behind evaluating the condition monitoring-enabled maintenance policy's economic impact. We note that several papers in our review conduct research to address more than one intention. Most studies assessed the benefits of integrating a CMS-like

Table 27
Analysis and evaluation intention in the studies (N = 42).

Intention	Total	Studies by industrial category		
	Count	Energy	Transport	Other
Assess CMS Integration	35	[70,123–127,133, 136,144,149–152, 154,156,159]	[130,132,138,141, 143,147,153,161]	[134,135,137,139, 142,146,148,155, 157,160,163]
Compare CMSs	11	[127,128,136]	[132,140,141,143, 147]	[134,137,157]
Optimize Maintenance Policy	9	[127,128,144,159, 162]	[130,161]	[139,158]
Assess Continued Use of CMS	1			[145]

tool with their equipment and maintenance practices (n = 35), distributed across studies on energy system-related applications (n = 16), transportation mode-related applications (n = 8), and the remaining industrial applications (n = 11). In these thirty-five studies, assessing the integration of CMS-type tools entails either estimating a condition monitoring-enabled maintenance policy's ability to reduce costs and improve equipment operations, or extracting minimum design requirements of a CMS-type tool for a particular application.

Several studies in Table 27 focused on comparing various condition monitoring tools or algorithms concerning their impact on financial benefits (n = 11). Transportation mode-related studies focused more on comparing condition monitoring tools (n = 5), and only a few of the eighteen energy system-related studies focused on this type of analysis (n = 3). Several studies also tried to optimize condition monitoring-enabled maintenance policies by finding a set of condition monitoring or maintenance model parameters that maximize financial benefits (n = 9). Most of these studies are related to energy system applications (n = 5). Regarding manufacturing-related studies, the majority take the perspective of assessing condition monitoring integration (n = 3), though Florian et al. [157] additionally conducts comparisons of condition monitoring tools. Uniquely, Rastegari [145] evaluates a machine's continued use of a vibration monitoring system.

Table 28 highlights the various analysis approaches used to understand condition monitoring investments and economic benefits. Many evaluation studies used an approach that we dub minimal total cost analysis (n = 14), where the study aggregates maintenance, fault or failure, and operations costs for each maintenance policy. Often, the study will compare the aggregated costs across several policies to identify the one with the least cost or consider whether a policy fits a predetermined budget.

A majority of the studies conduct a cost–benefit analysis (n = 23). With cost–benefit analyses, the evaluation calculates the cost–benefits of each maintenance policy implementation with a baseline policy. Then, if a maintenance policy uses CMS-like tools, the study calculates the net benefit of that policy by subtracting the CMS investment costs from the cost–benefits. These studies often consider the time-value of the net benefits across a certain period of time.

A few studies derived the economic benefits of CMS impacts with a cost-effectiveness analysis approach (n = 6). In this approach, studies evaluate the difference between two different maintenance implementation costs, and compare this difference to a difference between the resulting equipment KPIs of each implemented maintenance policy. Liang & Parlikad [128] and Mao et al. [132] used both cost–benefit and cost-effectiveness analyses. Mao et al. [132] compares the cost–benefits of a prognostics-driven, condition-based maintenance policy against a baseline policy, before comparing the cost–benefits of each maintenance policy in terms of equipment availability and maintenance quality metrics. Liang & Parlikad [128] similarly calculates the cost–benefits of two CMS-enabled maintenance policies against a baseline, time-based maintenance policy, then compares the cost–benefits to how each policy impacts an equipment availability metric.

We also note the diversity of approaches among the manufacturing-related studies. Rastegari [145] and Koochaki et al. [160] conducted evaluations of economic benefits with cost-effectiveness analysis, Adams et al. [155] used cost–benefit analysis, and Florian et al. [157] used minimal total cost analysis.

Table 29 classifies the studies based on their approach for representing and communicating the economic costs and benefits of any analysis approaches. We found seven classes of economic analysis outcome representations. Most studies presented their economic findings as an aggregation of costs and benefits for each implemented maintenance policy (n = 24). These aggregated costs typically come from models of equipment operations, maintenance actions, and the technical value provided by a specific condition monitoring technology to decrease faults and failures [33]. Representations of aggregated costs appear in nearly half of energy system-related studies (n = 8) and transportation mode-related studies (n = 7), as well as three of the four manufacturing-related studies [145,157,160]. Several studies also present financial benefits in terms of the return on investment (n = 6), which is a ratio that indicates the profitability of an investment in condition monitoring tools. Three energy system-related studies use return on investment [123,125,156], as opposed to a single transportation mode study [132] and a single manufacturing system study [155].

Net present value (n = 15), payback period (n = 3), and break-even points (n = 2) are three other representations of financial benefits that we found in our review. All three of these representations incorporate the time value of money. Slightly over half of studies that use net present value are related to energy systems (n = 8), as opposed to a single transportation mode-related study that uses net present value. A few studies represent analysis outcomes with value of information (n = 7) and real options valuation (n = 1). With real options valuation, the study from Lei & Sandborn [162] examines the opportunities to invest in options for performing maintenance actions in the future, given RUL predictions.

A third of the studies used more than one representation of financial benefits (n = 14). For example, Shamayleh et al. [146] used both payback period and net present value (NPV) representations, Liu & Wang [148] used break-even points and value of information (VOI), and Bakhsi & Sandborn [125] provided both NPV and return on investment outcomes.

Tables 30 to 32 present a breakdown of costs incorporated into each study's economic analysis and evaluation. Costs derived from equipment operation and performance appear in many studies (n = 18), mostly from energy system-related applications (n = 12). Costs related to maintenance are incorporated in nearly all of the reviewed papers (n = 41). Costs that are relevant to the monitoring system's implementation, from costs of hardware to financial impacts of incorrect or misleading fault detections and forecasts, appear in a slight majority of studies (n = 24), distributed across the majority of energy system-related studies (n = 12) and roughly half of the transportation mode-related studies (n = 4) and studies related to other industrial applications (n = 8).

Tables 30 to 32 also show whether each study considers any measure of uncertainty in their evaluation outcomes. Only one-third of the reviewed studies incorporated uncertainty into their outcomes (n = 14), most coming from energy system-related applications (n = 8) and none

Table 28
Analysis approaches in the studies (N = 42).

Approach	Total count	Studies by industrial category		
		Energy	Transport	Other
MTCA	14	[70,127,144, 159]	[130,140, 143,153,161]	[129,139, 142,157,158]
CBA	23	[123–126, 128,133,136, 149,150,152, 154,156,162]	[132,141, 147]	[134,135, 137,146,148, 155,163]
CEA	6	[128,151]	[132]	[131,145, 160]

Notes: MTCA = Minimal Total Cost Analysis. CBA = Cost-Benefit Analysis. CEA = Cost-Effectiveness Analysis.

from transportation mode-related applications. Adams et al. [155] is the only manufacturing-related study considering any uncertainty in their results. Overall, only three energy system-related studies incorporate uncertainties in their economic analysis evaluation outcomes while also incorporating costs from all three categories of equipment, maintenance, and condition monitoring [124–126]. The manufacturing-related study from Adams et al. [155] is the only other study that comprehensively incorporates all three categories of costs along with measures of uncertainty.

3.6.2. Discussion

Our goal with categorizing the reviewed papers into each of Tables 27 to 32 is to understand the economic analysis methods used to calculate and communicate the costs and benefits of installing and using a CMS. Table 27 depicts why a study intends to evaluate condition monitoring tools and calculate their economic benefits. On the other hand, Tables 28 and 29 discern how each study analyzes and communicates the evaluated economic benefits. Tables 30 through 32 enumerate the source of costs that are incorporated into economic benefits analysis and whether these benefits account for uncertainties in evaluation outcomes that propagate from uncertainties in inputs in the evaluation study.

The evaluation study's underlying intention or motivation provides the overarching context for which the study presents economic benefits. Most studies assess whether a maintenance policy designed with condition monitoring tools can reduce maintenance costs and improve equipment operations. Fewer studies delve into a comparative analysis of condition monitoring algorithms and tools, optimization of condition-based maintenance policy parameters, or assessments that justify the continued use of condition monitoring tools in light of potentially changing environmental conditions.

A possible reason for the lack of such studies is that these studies already assume a CMS-integrated, condition-based maintenance policy. On the other hand, focus on assessing CMS integration tends to consider whether a CMS-type tool is beneficial to the industrial application. Nine out of eleven studies compare condition monitoring tools and assess the integration of condition monitoring tools, including the manufacturing-related study Florian et al. [157]. Likewise, six out of nine studies optimize maintenance policy parameters and assess condition monitoring integration. Our framing highlights opportunities for future studies to investigate maintenance optimization and condition monitoring algorithm comparisons beyond assessing the feasibility of designing industrial applications that use CMSs.

Regarding the economic analysis approaches, our results reveal that one-third (n = 14) of the reviewed studies do not go beyond identifying the maintenance policy, which results in minimal total costs. Studies conducting cost–benefit analysis (CBA) or cost-effectiveness analysis (CEA) provide more nuance in identifying financial costs and benefits by directly comparing condition-based maintenance policies to baseline strategies and explicitly accounting for the investment costs

Table 29
Analysis outcome representations in the studies (N = 42).

Representation	Total count	Studies by industrial category		
		Energy	Transport	Other
AC	24	[70,126–128, 151,154,156, 159]	[130,132,140, 143,147,153, 161]	[129,134,137, 139,142,145, 157,158,160]
ROI	6	[123,125,156]	[132]	[155,163]
NPV	15	[124–126,133, 144,149,150, 152]	[141]	[131,135,137, 139,142,146]
PP	3		[132]	[146,163]
B/E	2	[123]		[148]
VOI	7	[127,136]		[134,135,137, 139,148]
ROV	1	[162]		

Notes: AC = Aggregated Costs Only. ROI = Return On Investment. NPV = Net Present Value. B/E = Break-Even point. PP = Payback Period. VOI = Value Of Information. ROV = Real Options Value.

Table 30
Incorporated costs and uncertainties in the energy system-related studies (n = 18).

Study	Incorporated costs			Uncertainty incorporation
	EO	MA	CMS	
Azadeh et al. [70]	No	Yes	No	Yes
Bakhshi & Sandborn [125]	Yes	Yes	Yes	Yes
Chang et al. [123]	Yes	Yes	Yes	No
Erguido et al. [156]	No	Yes	Yes	No
Kerres et al. [124]	Yes	Yes	Yes	Yes
Lei & Sandborn [162]	Yes	Yes	No	Yes
Li et al. [133]	Yes	Yes	Yes	No
Liang & Parlikad [128]	Yes	Yes	Yes	No
Livera et al. [144]	Yes	Yes	No	No
Long et al. [136]	No	Yes	Yes	No
May et al. [149]	Yes	Yes	Yes	No
Nielsen et al. [127]	No	Yes	No	Yes
Peng et al. [154]	No	Yes	Yes	Yes
Puglia et al. [150]	Yes	Yes	Yes	No
Tian et al. [159]	No	Yes	No	No
Turnbull & Carroll [151]	Yes	Yes	No	Yes
Van Horenbeek et al. [126]	Yes	Yes	Yes	Yes
Vieira et al. [152]	Yes	Yes	Yes	No
Count	12	18	12	8

Notes: EO = Equipment Operations-related costs. MA = Maintenance Activity-related costs. CMS = Condition Monitoring System-related costs.

for condition monitoring. Two studies even combine elements of both CBA and CEA [128,132]. Even with the financial quantification and incorporation of equipment or maintenance performance indicators into a CBA, an additional, supporting CEA can effectively demonstrate the value the evaluation placed on those performance indicators.

Concerning analysis outcome representations, we found that many studies report the economic impacts of each maintenance policy as an aggregation of costs and benefits, and most studies do not consider the time value of money. Net present value, one of the most common financial metrics for investment analysis in the manufacturing sector [119], only appears in fifteen studies. Unfortunately, other representations that indicate the time value of money, payback period, and break-even points were sparingly used. On a more promising note, we found that many studies used multiple representations to demonstrate economic analysis outcomes [125,146,148]. Several analysis representations provide more context for analysis outcomes, and evaluations should include indicators for the time value of money.

Several papers represent cost outcomes of maintenance-related decision-making with VOI and real options valuation (ROV). As we have mentioned in Section 3.4.2, though very promising, estimations of VOI may require complex, computationally expensive decision tree analysis. As for ROV, the single study that calculated expected real

Table 31
Incorporated costs and uncertainties in the transportation mode-related studies (n = 9).

Study	Incorporated costs			Uncertainty
	EO	MA	CMS	incorporation
Chen et al. [143]	No	Yes	No	No
Compare et al. [140]	No	Yes	Yes	No
Cot et al. [138]	NR	NR	NR	NR
Halbert et al. [141]	No	Yes	Yes	No
Hongsheng et al. [153]	Yes	Yes	No	No
Mao et al. [132]	Yes	Yes	Yes	No
Reimann et al. [130]	No	Yes	No	No
Wang & Pecht [161]	Yes	Yes	Yes	No
Yang & Letourneau [147]	No	Yes	No	No
Count	3	8	4	0

Notes: EO = Equipment Operations-related costs. MA = Maintenance Activity-related costs. CMS = Condition Monitoring System-related costs. NR = Not Reported.

Table 32
Incorporated costs and uncertainties in the remaining studies (n = 15).

Study	Incorporated costs			Uncertainty
	EO	MA	CMS	incorporation
Adams et al. [155]	Yes	Yes	Yes	Yes
Florian et al. [157]	No	Yes	Yes	No
Golmakani & Fattahipour [158]	No	Yes	No	No
Iannacone et al. [134]	Yes	Yes	No	Yes
Klerk et al. [135]	No	Yes	Yes	Yes
Koochaki et al. [160]	No	Yes	No	No
Liu & Wang [148]	No	Yes	Yes	No
Meng et al. [129]	No	Yes	No	NR
Neves & Frangopol [131]	No	Yes	No	Yes
Rastegari [145]	Yes	Yes	No	No
Shamayleh et al. [146]	No	Yes	Yes	No
Wu et al. [163]	No	Yes	Yes	No
Yoon et al. [142]	No	Yes	No	Yes
Zhang et al. [137]	No	Yes	Yes	Yes
Zou et al. [139]	No	Yes	Yes	No
Count	3	15	8	6

Notes: EO = Equipment Operations-related costs. MA = Maintenance Activity-related costs. CMS = Condition Monitoring System-related costs. NR = Not Reported.

options values [162] limited their analysis to a single maintenance action event on a single equipment component.

We also note that we did not observe any strong relationships between analysis representations in Table 29, the analysis approaches in Table 28, or the intent of analysis in Table 27. These facets of economic analysis may be considered independently when designing an evaluation study.

As for costs incorporated into the economic analysis, unfortunately, we only found that twelve studies incorporate costs from the three cost categories of equipment operations-related costs, maintenance activity-related costs, and CMS-related costs. Furthermore, Adams et al. [155] is the only manufacturing-related study that incorporated costs from all three cost categories.

Maintenance-related costs appeared in almost every study, as they directly come from the maintenance action costs presented as maintenance KPIs in Section 3.5.2. Ideally, evaluations would also quantify and incorporate costs from equipment operations and performance, condition monitoring investment, and performance-related costs.

Unfortunately, accounting for uncertainty in evaluation outcomes was also limited in our review, with the lack of uncertainty analysis in any of the transportation mode-related studies being surprising. The low level of uncertainty incorporation and cost inclusion from the three aforementioned categories points to a missed opportunity in condition monitoring evaluation methods.

4. Conclusion

This systematic review of peer-reviewed research articles from January 2001 to December 2023 discussed methods for evaluating the impact of integrating CMS-related technologies on the maintenance processes of industrial applications. Through specified eligibility criteria and study selection process, this review selected articles that evaluate the impact of CMS-related technologies that go beyond algorithm-level condition monitoring performance measures and include measures of impact on equipment performance, maintenance operations, and business value.

This paper introduced a conceptual model for condition monitoring-enabled maintenance evaluation, organizing it into industrial applications, condition monitoring, maintenance deployment, evaluation techniques, performance measures, and economic analysis. This model was used to guide this paper's data collection process. Data items were collected from each selected study, which provided insights into the state of research in condition monitoring evaluation and helped us identify opportunities for future work. Though the wide range of industrial applications in this review included transportation modes, energy systems, and civil structures, this paper focused on opportunities for evaluating condition monitoring tools and technologies in the context of manufacturing applications.

The review of the selected studies faces limitations due to widely varying levels of detail reported by each study and the heterogeneity and inconsistency of the extracted study data items, which creates obstacles to drawing in-depth comparisons and analyses. The heterogeneity and inconsistently reported data items highlight the need for reporting guidelines to help standardize future evaluation studies.

Nonetheless, the results yield some interesting insights into the shortcomings of the existing evaluation methods and opportunities for future research.

In Section 3.1, the results reveal a need for more condition monitoring evaluation studies in manufacturing system applications, as most evaluation studies focus on energy systems or transportation modes. Furthermore, evaluation research could benefit from more studies that disclose details about their equipment failure process datasets and any fault identification methods used to identify these equipment failure processes.

In Section 3.2, the results indicate that the evaluation studies aptly disclose the type of monitoring information a CMS provides, whether that information is about detecting an equipment's condition or forecasting its faults or failure. However, future studies should describe details about their input training data, processing techniques, monitoring algorithms, and algorithmic training routines that underlie a CMS's ability to output monitoring information.

In Section 3.3, the results reveal that the selected studies generally consider various maintenance policies compared to condition-based maintenance. However, future evaluation research should include more detailed maintenance action models to reflect different maintenance needs during equipment operation, especially in manufacturing-related studies with many types of maintenance work orders. This section also noted that future studies should examine the relationship between the frequency of monitoring information retrieved by maintenance and the quality of monitoring information.

In Section 3.4, the results found that analytical frameworks can underpin evaluation studies, but deriving analytical solutions were often infeasible. Computational methods in this review included Monte Carlo methods and discrete-event simulations. Future research should consider both but disclose details about their implementations. Bayesian decision models and Markov models were also frequently used to model maintenance decision-making, though future evaluations should use either deliberately given their computational drawbacks. This section also brought forth a research need for studies to incorporate a diverse, larger set of parameters for sensitivity analysis.

In Section 3.5, trends were observed in the results with respect to the types of performance measures that the selected studies utilized. Few studies measured performance indicators related to industrial equipment, such as manufacturing production rate. Maintenance-related performance indicators were mainly limited to aggregated costs from maintenance actions. Commonly reported metrics for monitoring performance included accuracy or true negative rates, which can be misleading for CMSs, with little to no focus on precision, F1-scores, time-varying metrics, or prognostics-related metrics. Future research should address these shortcomings in reported performance measures.

Lastly, in Section 3.6, the results indicate that there are several opportunities in future research with the economic analysis aspect of these evaluations that can help enhance the interpretability and comparability of results, such as clarifying the intent of the analysis, taking a CBA and CEA approach to economic evaluation, including the time value of money, and quantifying uncertainties in analysis outcomes. Analysis and communication of the economic aspects of the evaluations are especially important for justifying manufacturing maintenance expenditures.

This review of forty-two studies demonstrates the pervasive use of condition monitoring tools and technologies across many industrial applications. This paper suggests specific recommendations for future condition monitoring evaluation research for manufacturing and industrial applications. Based on survey results, this paper posits that these suggestions can improve methods used to evaluate CMSs and validate their benefits in industrial settings.

CRediT authorship contribution statement

Mehdi Dadfarnia: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Michael E. Sharp:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Jeffrey W. Herrmann:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization.

Disclaimer

The views and opinions expressed herein do not necessarily state or reflect those of the National Institute of Standards and Technology (NIST). Certain commercial entities, equipment, or materials may be identified in this document to illustrate a point or concept. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

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.

References

[1] Moeuf A, Lamouri S, Pellerin R, Tamayo-Giraldo S, Tobon-Valencia E, Eburdy R. Identification of critical success factors, risks and opportunities of industry 4.0 in SMEs. Int J Prod Res 2020;58(5):1384–400.

[2] Zio E. Prognostics and health management of industrial equipment. Diagn Progn Eng Syst: Methods Tech 2013;333–56.

[3] Compare M, Baraldi P, Zio E. Challenges to IoT-enabled predictive maintenance for industry 4.0. IEEE Internet Things J 2019;7(5):4585–97.

[4] Lee J, Davari H, Singh J, Pandhare V. Industrial artificial intelligence for industry 4.0-based manufacturing systems. Manuf Lett 2018;18:20–3.

[5] Furman J, Holdren JP, Munoz C, Smith M, Zients J. Artificial intelligence, automation, and the economy. Executive office of the United StatesPresident; 2016, p. 8.

[6] Hansen EB, Bøgh S. Artificial intelligence and internet of things in small and medium-sized enterprises: A survey. J Manuf Syst 2021;58:362–72.

[7] Lundgren C, Bokrantz J, Skoogh A. Hindering factors in smart maintenance implementation. In: SPS2022. IOS Press; 2022, p. 629–37.

[8] Bokrantz J, Skoogh A, Berlin C, Wuest T, Stahre J. Smart maintenance: an empirically grounded conceptualization. Int J Prod Econ 2020;223:107534.

[9] Bokrantz J, Skoogh A, Berlin C, Wuest T, Stahre J. Smart maintenance: a research agenda for industrial maintenance management. Int J Prod Econ 2020;224:107547.

[10] Sharp M, Dadfarnia M, Sprock T, Thomas D. Procedural guide for system-level impact evaluation of industrial artificial intelligence-driven technologies: Application to risk-based investment analysis for condition monitoring systems in manufacturing. J Manuf Sci Eng 2022;144(7):071008.

[11] Tinga T, Loendersloot R. Aligning PHM, SHM and CBM by understanding the physical system failure behaviour. In: PHM society European conference, vol. 2, 2014.

[12] Woodall WH, Montgomery DC. Research issues and ideas in statistical process control. J Qual Technol 1999;31(4):376–86.

[13] Gholizadeh S. A review of non-destructive testing methods of composite materials. Procedia Struct Integr 2016;1:50–7.

[14] Farrar CR, Worden K. An introduction to structural health monitoring. Philos Trans R Soc A: Math Phys Eng Sci 2007;365(1851):303–15.

[15] Society of Automotive Engineers. Potential Failure Mode and Effects Analysis in Design (Design FMEA), Potential Failure Mode and Effects Analysis in Manufacturing and Assembly Processes (Process FMEA). SAE J1739: 2009.

[16] Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. Mech Syst Signal Process 2014;42(1–2):314–34.

[17] Zio E, Compare M. Evaluating maintenance policies by quantitative modeling and analysis. Reliab Eng Syst Saf 2013;109:53–65.

[18] Moradi R, Groth KM. Modernizing risk assessment: A systematic integration of PRA and PHM techniques. Reliab Eng Syst Saf 2020;204:107194.

[19] Lee J, Jin C, Liu Z, Davari Ardakani H. Introduction to data-driven methodologies for prognostics and health management. Probab Progn Heal Manag Energy Syst 2017;9–32.

[20] Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. Comput Chem Eng 2003;27(3):293–311.

[21] Yucesan YA, Dourado A, Viana FA. A survey of modeling for prognosis and health management of industrial equipment. Adv Eng Inform 2021;50:101404.

[22] Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK. Applications of machine learning to machine fault diagnosis: A review and roadmap. Mech Syst Signal Process 2020;138:106587.

[23] Xia T, Dong Y, Xiao L, Du S, Pan E, Xi L. Recent advances in prognostics and health management for advanced manufacturing paradigms. Reliab Eng Syst Saf 2018;178:255–68.

[24] Ademujimi TT, Brundage MP, Prabhu VV. A review of current machine learning techniques used in manufacturing diagnosis. In: Advances in production management systems. the path to intelligent, collaborative and sustainable manufacturing: IFIP WG 5.7 international conference, APMS 2017, hamburg, Germany, September 3-7, 2017, proceedings, Part I. Springer; 2017, p. 407–15.

[25] Zhou Y, Xue W. Review of tool condition monitoring methods in milling processes. Int J Adv Manuf Technol 2018;96:2509–23.

[26] Serin G, Sener B, Ozbayoglu AM, Unver HO. Review of tool condition monitoring in machining and opportunities for deep learning. Int J Adv Manuf Technol 2020;109(3):953–74.

[27] Zio E. Some challenges and opportunities in reliability engineering. IEEE Trans Reliab 2016;65(4):1769–82.

[28] Sandborn P, Wilkinson C, Sharon KL, Jazouli T, Bakhshi R. PHM cost and return on investment. In: Prognostics and health management of electronics: Fundamentals, machine learning, and the internet of things. Wiley Online Library; 2018, p. 221–60.

[29] Bayram F, Ahmed BS, Kassler A. From concept drift to model degradation: An overview on performance-aware drift detectors. Knowl-Based Syst 2022;245:108632.

[30] Leao BP, Yoneyama T, Rocha GC, Fitzgibbon KT. Prognostics performance metrics and their relation to requirements, design, verification and cost-benefit. In: 2008 international conference on prognostics and health management. IEEE; 2008, p. 1–8.

[31] Saxena A, Celaya J, Saha B, Saha S, Goebel K. Metrics for offline evaluation of prognostic performance. Int J Progn Heal Manag 2010;1(1):4–23.

[32] Sharp ME. Simple metrics for evaluating and conveying prognostic model performance to users with varied backgrounds. In: Annual conference of the PHM society, vol. 5, 2013.

[33] Dzakowic JE, Valentine GS. Advanced Techniques for the verification and validation of prognostics & health management capabilities. Machinery Failure Prevention Technologies (MFPT 60); 2007, p. 1–11.

[34] Lundgren C, Bokrantz J, Skoogh A. Performance indicators for measuring the effects of Smart Maintenance. Int J Prod Perform Manag 2021;70(6):1291–316.

[35] Juuso EK, Lahdelma S. Intelligent performance measures for condition-based maintenance. J Qual Maint Eng 2013;19(3):278–94.

[36] Wu D, Jennings C, Terpenny J, Gao RX, Kumara S. A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. J Manuf Sci Eng 2017;139(7):071018.
[37] Gouarir A, Martínez-Arellano G, Terrazas G, Benardos P, Ratchev S. In-process tool wear prediction system based on machine learning techniques and force analysis. Procedia CIRP 2018;77:501–4.
[38] Dadfarnia M, Sharp M, Sprock T. Understanding and evaluating naive diagnostics algorithms applicable in multistage manufacturing from a risk management perspective. In: International manufacturing science and engineering conference, vol. 84263, American Society of Mechanical Engineers; 2020, V002T07A042.
[39] Dadfarnia M, Sharp ME, Drozdov S, Herrmann JW. A simulation-based approach to assess condition monitoring-enabled maintenance in manufacturing. In: 2023 7th international conference on system reliability and safety. IEEE; 2023, p. 413–22.
[40] Saxena A, Roychoudhury I, Celaya J, Saha S, Saha B, Goebel K. Requirements specification for prognostics performance-an overview. Aiaa infotech@ aerospace 2010; 2010, p. 3398.
[41] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj 2021;372.
[42] Leukel J, González J, Riekert M. Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review. J Manuf Syst 2021;61:87–96.
[43] Wang GC, Gross K. Telemetry parameter synthesis system to support machine learning tuning and validation. In: 2018 international conference on computational science and computational intelligence. IEEE; 2018, p. 941–6.
[44] Tharwat A. Classification assessment methods. Appl Comput Inform 2021;17(1):168–92.
[45] Oviedo EHS, Travé-Massuyès L, Subias A, Alonso C, Pavlov M. Feature extraction and health status prediction in PV systems. Adv Eng Inform 2022;53:101696.
[46] Compare M, Bellani L, Zio E. Availability model of a PHM-equipped component. IEEE Trans Reliab 2017;66(2):487–501.
[47] Saleh JH, Marais K. Reliability: How much is it worth? Beyond its estimation or prediction, the (net) present value of reliability. Reliab Eng Syst Saf 2006;91(6):665–73.
[48] Dadfarnia M, Sharp M. Key elements to contextualize ai-driven condition monitoring systems towards their risk-based evaluation. In: 2022 5th international conference on artificial intelligence for industries. IEEE; 2022, p. 38–41.
[49] Cortez J, Keller K, Poblete J. Establishing an approach to systematically improve the effectiveness of fleet support. In: 2008 IEEE AUTOTESTCON. IEEE; 2008, p. 169–75.
[50] Fan I-S, Verma R, Jennions I, Dubicka A. A business analysis methodology for integrated equipment health management in high value assets. In: IET and IAM asset management conference 2011. IET; 2011, p. 1–8.
[51] Williams R, Crabtree C, Hogg S. Quantifying the economic benefits of wind turbine condition monitoring. In: Turbo expo: power for land, sea, and air, vol. 45660, American Society of Mechanical Engineers; 2014, V03BT46A022.
[52] Feldman K, Jazouli T, Sandborn PA. A methodology for determining the return on investment associated with prognostics and health management. IEEE Trans Reliab 2009;58(2):305–16.
[53] Leão BP, Fitzgibbon KT, Puttini LC, de Melo GP. Cost-benefit analysis methodology for PHM applied to legacy commercial aircraft. In: 2008 IEEE aerospace conference. IEEE; 2008, p. 1–13.
[54] Hoyle C, Mehr A, Tumer I, Chen W. On quantifying cost-benefit of ISHM in aerospace systems. In: 2007 IEEE aerospace conference. IEEE; 2007, p. 1–7.
[55] Brundage MP, Sexton T, Hodkiewicz M, Morris KC, Arinez J, Ameri F, et al. Where do we start? Guidance for technology implementation in maintenance management for manufacturing. In: International manufacturing science and engineering conference, vol. 58745, American Society of Mechanical Engineers; 2019, V001T02A016.
[56] Netto AC, Melani AHdA, Michalski MAdC, Murad CA, Belinello MM, de Souza GF. Defining operator driven reliability inspection routes through reliability, risk analysis and quality control techniques. In: 2021 annual reliability and maintainability symposium. IEEE; 2021, p. 1–7.
[57] Milje R. Engineering methodology for selecting condition based maintenance. Norway: University of Stavanger; 2011.
[58] JA1011 S. Evaluation criteria for reliability-centered maintenance (RCM) processes. Soc Automot Eng 1999.
[59] Zornio P, Boudreaux M. Case study: How digital transformation paved the way for one refinery's predictive maintenance strategy. In: Offshore technology conference. OTC; 2019, D021S016R007.
[60] Schlake BW, Barkan CP, Edwards JR. Impact of automated inspection technology on unit train performance. In: Joint rail conference, vol. 49071, 2010, p. 281–8.
[61] Kato T, Suzuoki Y. Robustness assessment of life-cycle-management of CV cables based on degradation diagnosis. In: 2007 IEEE international conference on solid dielectrics. IEEE; 2007, p. 192–5.
[62] Mongeon P, Paul-Hus A. The journal coverage of web of science and scopus: a comparative analysis. Scientometrics 2016;106:213–28.
[63] Shrieve P. Implementing a cost effective machinery condition monitoring program. Profitab Cond Monit 1993;3–10.
[64] Nicholls C. Cost-effective condition monitoring. In: COMADEM 89 international: proceedings of the first international congress on condition monitoring and diagnostic engineering management. Springer; 1989, p. 335–47.
[65] Abu-Hanna A, Lucas PJ. Prognostic models in medicine. Methods Inf Med 2001;40(01):1–5.
[66] Vogl GW, Weiss BA, Helu M. A review of diagnostic and prognostic capabilities and best practices for manufacturing. J Intell Manuf 2019;30:79–95.
[67] Zhang W-H, Lu D-G, Qin J, Thöns S, Faber MH. Value of information analysis in civil and infrastructure engineering: a review. J Infrastruct Preserv Resil 2021;2(1):16.
[68] Brijder R, Hagen CH, Cortés A, Irizar A, Thibbotuwa UC, Helsen S, et al. Review of corrosion monitoring and prognostics in offshore wind turbine structures: Current status and feasible approaches. Front Energy Res 2022;10:991343.
[69] Azadeh A, Asadzadeh SM, Seif J. An integrated simulation-analysis of variance methodology for effective analysis of CBM alternatives. Int J Comput Integr Manuf 2014;27(7):624–37.
[70] Azadeh A, Asadzadeh S, Salehi N, Firoozi M. Condition-based maintenance effectiveness for series–parallel power generation system—A combined Markovian simulation model. Reliab Eng Syst Saf 2015;142:357–68.
[71] Bousdekis A, Magoutas B, Apostolou D, Mentzas G. A proactive decision making framework for condition-based maintenance. Ind Manag Data Syst 2015;115(7):1225–50.
[72] Thomas D, Weiss B. Maintenance costs and advanced maintenance techniques in manufacturing machinery: Survey and analysis. Int J Progn Heal Manag 2021;12(1).
[73] Chen C, Reniers G, Khakzad N, Yang M. Operational safety economics: Foundations, current approaches and paths for future research. Saf Sci 2021;141:105326.
[74] Van Horenbeek A, Pintelon L, Muchiri P. Maintenance optimization models and criteria. Int J Syst Assur Eng Manag 2010;1:189–200.
[75] Zeng Z, Barros A, Coit D. Dependent failure behavior modeling for risk and reliability: A systematic and critical literature review. Reliab Eng Syst Saf 2023;109515.
[76] Arts J, Boute RN, Loeys S, van Staden HE. Fifty years of maintenance optimization: Reflections and perspectives. European J Oper Res 2024.
[77] Aven T. On how to define, understand and describe risk. Reliab Eng Syst Saf 2010;95(6):623–31.
[78] The International Organization for Standardization and The International Electrotechnical Commission. Risk management — Risk assessment techniques. ISO.IEC 31010:2019.
[79] Chemweno P, Pintelon L, Muchiri PN, Van Horenbeek A. Risk assessment methodologies in maintenance decision making: A review of dependability modelling approaches. Reliab Eng Syst Saf 2018;173:64–77.
[80] The International Organization for Standardization. Condition monitoring and diagnostics of machines — Data processing, communication and presentation — Part 1: General guidelines. ISO 13374-1:2003.
[81] Oh H, Azarian MH, Cheng S, Pecht MG. Sensor systems for PHM. In: Prognostics and health management of electronics: Fundamentals, machine learning, and the internet of things. Wiley Online Library; 2018, p. 39–60.
[82] Mohanraj T, Shankar S, Rajasekar R, Sakthivel N, Pramanik A. Tool condition monitoring techniques in milling process—A review. J Mater Res Technol 2020;9(1):1032–42.
[83] Kuntoğlu M, Aslan A, Pimenov DY, Usca ÜA, Salur E, Gupta MK, et al. A review of indirect tool condition monitoring systems and decision-making methods in turning: Critical analysis and trends. Sensors 2020;21(1):108.
[84] Jaros R, Byrtus R, Dohnal J, Danys L, Baros J, Koziorek J, et al. Advanced signal processing methods for condition monitoring. Arch Comput Methods Eng 2023;30(3):1553–77.
[85] Lei Y, Li N, Guo L, Li N, Yan T, Lin J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mech Syst Signal Process 2018;104:799–834.
[86] Zhu J, Nostrand T, Spiegel C, Morton B. Survey of condition indicators for condition monitoring systems. In: Annual conference of the PHM society, vol. 6, 2014.
[87] Djeziri MA, Benmoussa S, Zio E. Review on health indices extraction and trend modeling for remaining useful life estimation. In: Artificial intelligence techniques for a scalable energy transition: Advanced methods, digital technologies, decision support tools, and applications. Springer; 2020, p. 183–223.
[88] Atamuradov V, Medjaher K, Camci F, Zerhouni N, Dersin P, Lamoureux B. Machine health indicator construction framework for failure diagnostics and prognostics. J Signal Process Syst 2020;92:591–609.
[89] Zio E. The future of risk assessment. Reliab Eng Syst Saf 2018;177:176–90.
[90] Xing J, Zeng Z, Zio E. A framework for dynamic risk assessment with condition monitoring data and inspection data. Reliab Eng Syst Saf 2019;191:106552.
[91] Pashami S, Nowaczyk S, Fan Y, Jakubowski J, Paiva N, Davari N, et al. Explainable predictive maintenance. 2023, arXiv preprint arXiv:2306.05120.

[92] Gao Z, Cecati C, Ding SX. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. IEEE Trans Ind Electron 2015;62(6):3757–67.
[93] Gao Z, Cecati C, Ding SX, et al. A survey of fault diagnosis and fault-tolerant techniques—Part II: Fault diagnosis with knowledge-based and hybrid/active approaches. IEEE Trans Ind Electron 2015.
[94] Wang J, Ma Y, Zhang L, Gao RX, Wu D. Deep learning for smart manufacturing: Methods and applications. J Manuf Syst 2018;48:144–56.
[95] Diez-Olivan A, Del Ser J, Galar D, Sierra B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0. Inf Fusion 2019;50:92–111.
[96] Liu R, Yang B, Zio E, Chen X. Artificial intelligence for fault diagnosis of rotating machinery: A review. Mech Syst Signal Process 2018;108:33–47.
[97] Carvalho TP, Soares FA, Vita R, Francisco RdP, Basto JP, Alcalá SG. A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 2019;137:106024.
[98] Xu Z, Saleh JH. Machine learning for reliability engineering and safety applications: Review of current status and future opportunities. Reliab Eng Syst Saf 2021;211:107530.
[99] Ward T, Jenab K, Ortega-Moody J, Staub S. A comprehensive review of machine learning techniques for condition-based maintenance. Int J Progn Heal Manag 2024;15(2).
[100] Moradi R, Groth KM. On the application of transfer learning in prognostics and health management. 2020, arXiv preprint arXiv:2007.01965.
[101] Kouw WM, Loog M. A review of domain adaptation without target labels. IEEE Trans Pattern Anal Mach Intell 2019;43(3):766–85.
[102] Denkena B, Klemme H, Stiehl TH. Failure sensitivity and similarity of process signals among multiple machine tools. Procedia CIRP 2023;120:922–7.
[103] Marais KB, Saleh JH. Beyond its cost, the value of maintenance: an analytical framework for capturing its net present value. Reliab Eng Syst Saf 2009;94(2):644–57.
[104] De Jonge B, Scarf PA. A review on maintenance optimization. European J Oper Res 2020;285(3):805–24.
[105] Wang H. A survey of maintenance policies of deteriorating systems. European J Oper Res 2002;139(3):469–89.
[106] Alaswad S, Xiang Y. A review on condition-based maintenance optimization models for stochastically deteriorating system. Reliab Eng Syst Saf 2017;157:54–63.
[107] Alrabghi A, Tiwari A. A novel approach for modelling complex maintenance systems using discrete event simulation. Reliab Eng Syst Saf 2016;154:160–70.
[108] Kaegi M, Mock R, Kröger W. Analyzing maintenance strategies by agent-based simulations: A feasibility study. Reliab Eng Syst Saf 2009;94(9):1416–21.
[109] Maidana RG, Parhizkar T, Gomola A, Utne IB, Mosleh A. Supervised dynamic probabilistic risk assessment: Review and comparison of methods. Reliab Eng Syst Saf 2023;230:108889.
[110] Saltelli A, Ratto M, Tarantola S, Campolongo F, Commission E, et al. Sensitivity analysis practices: Strategies for model-based inference. Reliab Eng Syst Saf 2006;91(10–11):1109–25.
[111] Li J, Meerkov SM, Zhang L. Production systems engineering: main results and recommendations for management. Int J Prod Res 2013;51(23–24):7209–34.
[112] Compare M, Bellani L, Zio E. Reliability model of a component equipped with PHM capabilities. Reliab Eng Syst Saf 2017;168:4–11.
[113] Van Horenbeek A, Pintelon L. Development of a maintenance performance measurement framework—using the analytic network process (ANP) for maintenance performance indicator selection. Omega 2014;42(1):33–46.
[114] Ruff L, Kauffmann JR, Vandermeulen RA, Montavon G, Samek W, Kloft M, et al. A unifying review of deep and shallow anomaly detection. Proc IEEE 2021;109(5):756–95.
[115] Saxena A, Celaya J, Balaban E, Goebel K, Saha B, Saha S, et al. Metrics for evaluating performance of prognostic techniques. In: 2008 international conference on prognostics and health management. IEEE; 2008, p. 1–17.
[116] Lewis AD, Groth KM. Metrics for evaluating the performance of complex engineering system health monitoring models. Reliab Eng Syst Saf 2022;223:108473.
[117] Zio E. Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice. Reliab Eng Syst Saf 2022;218:108119.
[118] Zeng Z, Di Maio F, Zio E, Kang R. A hierarchical decision-making framework for the assessment of the prediction capability of prognostic methods. Proc Inst Mech Eng Part O: J Risk Reliab 2017;231(1):36–52.
[119] Thomas D. Investment analysis methods: a practitioner's guide to understanding the basic principles for investment decisions in manufacturing. In: NIST Advanced manufacturing series, (200-5). National Institute of Standards and Technology; 2017.
[120] Aven T, Zio E. Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliab Eng Syst Saf 2011;96(1):64–74.
[121] Nemani V, Biggio L, Huan X, Hu Z, Fink O, Tran A, et al. Uncertainty quantification in machine learning for engineering design and health prognostics: A tutorial. Mech Syst Signal Process 2023;205:110796.
[122] Papananias M, McLeay TE, Mahfouf M, Kadirkamanathan V. A Bayesian framework to estimate part quality and associated uncertainties in multistage manufacturing. Comput Ind 2019;105:35–47.
[123] Chang M-H, Sandborn P, Pecht M, Yung WK, Wang W. A return on investment analysis of applying health monitoring to LED lighting systems. Microelectron Reliab 2015;55(3–4):527–37.
[124] Kerres B, Fischer K, Madlener R. Economic evaluation of maintenance strategies for wind turbines: a stochastic analysis. IET Renew Power Gener 2015;9(7):766–74.
[125] Bakhshi R, Sandborn PA. A return on investment model for the implementation of new technologies on wind turbines. IEEE Trans Sustain Energy 2017;9(1):284–92.
[126] Van Horenbeek A, Van Ostaeyen J, Duflou JR, Pintelon L. Quantifying the added value of an imperfectly performing condition monitoring system—Application to a wind turbine gearbox. Reliab Eng Syst Saf 2013;111:45–57.
[127] Nielsen JS, Tcherniak D, Ulriksen MD. A case study on risk-based maintenance of wind turbine blades with structural health monitoring. Struct Infrastruct Eng 2021;17(3):302–18.
[128] Liang Z, Parlikad A. A Markovian model for power transformer maintenance. Int J Electr Power Energy Syst 2018;99:175–82.
[129] Meng H, Zhang X, Zhu L, Wang L, Yang Z. Optimizing software rejuvenation policy based on CDM for cloud system. In: 2017 12th IEEE conference on industrial electronics and applications. IEEE; 2017, p. 1850–4.
[130] Reimann J, Kacprzynski G, Cabral D, Marini R. Using condition based maintenance to improve the profitability of performance based logistic contracts. In: Annual conference of the PHM society, vol. 1, 2009.
[131] Neves LC, Frangopol DM. Condition, safety and cost profiles for deteriorating structures with emphasis on bridges. Reliab Eng Syst Saf 2005;89(2):185–98.
[132] Mao K, Zhu Y, Chen Z, Tao X, Xue Q, Wu H, et al. A visual model-based evaluation framework of cloud-based prognostics and health management. In: 2017 IEEE international conference on smart cloud. IEEE; 2017, p. 33–40.
[133] Li Y, Salman AM, Braik A, Bjarnadóttir S, Salarieh B. Risk-based management of electric power distribution systems subjected to hurricane and tornado hazards. In: Engineering for extremes: decision-making in an uncertain world. Springer; 2021, p. 143–66.
[134] Iannacone L, Francesco Giordano P, Gardoni P, Pina Limongelli M. Quantifying the value of information from inspecting and monitoring engineering systems subject to gradual and shock deterioration. Struct Heal Monit 2022;21(1):72–89.
[135] Klerk WJ, Schweckendiek T, den Heijer F, Kok M. Value of information of structural health monitoring in asset management of flood defences. Infrastructures 2019;4(3):56.
[136] Long L, Mai QA, Morato PG, Sørensen JD, Thöns S. Information value-based optimization of structural and environmental monitoring for offshore wind turbines support structures. Renew Energy 2020;159:1036–46.
[137] Zhang W-H, Qin J, Lu D-G, Liu M, Faber MH. Quantification of the value of condition monitoring system with time-varying monitoring performance in the context of risk-based inspection. Reliab Eng Syst Saf 2023;231:108993.
[138] Cot LD, Wang Y, Bès C, Gogu C. Scheduled and SHM structural airframe maintenance applications using a new probabilistic model. In: EWSHM-7th European workshop on structural health monitoring. 2014.
[139] Zou G, Banisoleiman K, González A, Faber MH. Probabilistic investigations into the value of information: a comparison of condition-based and time-based maintenance strategies. Ocean Eng 2019;188:106181.
[140] Compare M, Antonello F, Pinciroli L, Zio E. A general model for life-cycle cost analysis of condition-based maintenance enabled by PHM capabilities. Reliab Eng Syst Saf 2022;224:108499.
[141] Halbert K, Fitzwater L, Davis C, Ihn J-b, Jones D, Mcfarland J, et al. Cost/benefit analysis for system-level integration of non-deterministic analysis and maintenance technology. In: 53rd AIAA/aSME/ASCE/AHS/aSC structures, structural dynamics and materials conference 20th AIAA/aSME/AHS adaptive structures conference 14th AIAA. 2012, p. 1437.
[142] Yoon JT, Youn BD, Yoo M, Kim Y, Kim S. Life-cycle maintenance cost analysis framework considering time-dependent false and missed alarms for fault diagnosis. Reliab Eng Syst Saf 2019;184:181–92.
[143] Chen C, Shi J, Lu N, Zhu ZH, Jiang B. Data-driven predictive maintenance strategy considering the uncertainty in remaining useful life prediction. Neurocomputing 2022;494:79–88.
[144] Livera A, Theristis M, Micheli L, Fernandez EF, Stein JS, et al. Operation and maintenance decision support system for photovoltaic systems. IEEE Access 2022;10:42481–96.
[145] Rastegari A. Vibration analysis of machine tool spindle units. In: Asset intelligence through integration and interoperability and contemporary vibration engineering technologies: proceedings of the 12th world congress on engineering asset management and the 13th international conference on vibration engineering and technology of machinery. Springer; 2019, p. 511–22.
[146] Shamayleh A, Awad M, Farhat J. IoT based predictive maintenance management of medical equipment. J Med Syst 2020;44(4):72.
[147] Yang C, Létourneau S. Model evaluation for prognostics: Estimating cost saving for the end users. In: Sixth international conference on machine learning and applications. IEEE; 2007, p. 304–9.

[148] Liu X, Wang P. Valuation of continuous monitoring systems for engineering system design in recurrent maintenance decision scenarios. J Mech Des 2022;144(9):091702.

[149] May A, McMillan D, Thöns S. Economic analysis of condition monitoring systems for offshore wind turbine sub-systems. IET Renew Power Gener 2015;9(8):900–7.

[150] Puglia G, Bangalore P, Tjernberg LB. Cost efficient maintenance strategies for wind power systems using LCC. In: 2014 international conference on probabilistic methods applied to power systems. IEEE; 2014, p. 1–6.

[151] Turnbull A, Carroll J. Cost benefit of implementing advanced monitoring and predictive maintenance strategies for offshore wind farms. Energies 2021;14(16):4922.

[152] Vieira M, Snyder B, Henriques E, White C, Reis L. Economic viability of implementing structural health monitoring systems on the support structures of bottom-fixed offshore wind. Energies 2023;16(13):4885.

[153] Hongsheng Y, Hongfu Z, Jianzhong S, Zhi L. Cost effectiveness evaluation model for civil aircraft maintenance based on prognostics and health management. In: 2017 international conference on sensing, diagnostics, prognostics, and control. IEEE; 2017, p. 161–7.

[154] Peng Y, Asgarpoor S, Qiao W, Foruzan E. Fuzzy cost-based FMECA for wind turbines considering condition monitoring systems. In: 2016 North American power symposium. IEEE; 2016, p. 1–6.

[155] Adams S, Malinowski M, Heddy G, Choo B, Beling PA. The WEAR methodology for prognostics and health management implementation in manufacturing. J Manuf Syst 2017;45:82–96.

[156] Erguido A, Castellano E, Gómez JF, Márquez AC. Optimization of maintenance strategies and ROI analysis of CMS through RAM-lcc analysis. a wind energy sector case study.. In: Proceedings of the 8th European workshop on structural health monitoring. 2016.

[157] Florian E, Sgarbossa F, Zennaro I. Machine learning-based predictive maintenance: A cost-oriented model for implementation. Int J Prod Econ 2021;236:108114.

[158] Reza Golmakani H, Fattahipour F. Age-based inspection scheme for condition-based maintenance. J Qual Maint Eng 2011;17(1):93–110.

[159] Tian Z, Jin T, Wu B, Ding F. Condition based maintenance optimization for wind power generation systems under continuous monitoring. Renew Energy 2011;36(5):1502–9.

[160] Koochaki J, Bokhorst J, Wortmann H, Klingenberg W. Evaluating condition based maintenance effectiveness for two processes in series. J Qual Maint Eng 2011;17(4):398–414.

[161] Wang W, Pecht M. Economic analysis of canary-based prognostics and health management. IEEE Trans Ind Electron 2010;58(7):3077–89.

[162] Lei X, Sandborn PA. PHM-based wind turbine maintenance optimization using real options. Int J Progn Heal Manag 2016;7(1):1–14.

[163] Wu B, Jia X, Xia L, Lei C, Wang Y. Cost-benefit analysis of the application of prognostics and health management technology. In: 2013 international conference on quality, reliability, risk, maintenance, and safety engineering. IEEE; 2013, p. 1923–7.

[164] Kingma DP, Adam JLB. A method for stochastic optimization. In: int. conf. learn. representations, vol. 6, 2015.

[165] Bradley W, Kim J, Kilwein Z, Blakely L, Eydenberg M, Jalvin J, et al. Perspectives on the integration between first-principles and data-driven modeling. Comput Chem Eng 2022;166:107898.

[166] Cassandras CG. Discrete-event systems. In: Handbook of networked and embedded control systems. Springer; 2005, p. 71–89.

[167] Law AM. In: Simulation modeling and analysis, vol. 5, Mcgraw-hill New York; 2015.

[168] Janssen H. Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence. Reliab Eng Syst Saf 2013;109:123–32.

[169] Turati P, Pedroni N, Zio E. An adaptive simulation framework for the exploration of extreme and unexpected events in dynamic engineered systems. Risk Anal 2017;37(1):147–59.