Vol. 4 No. 4 (2026)

RiESTech

JOURNAL RECENT IN ENGINEERING SCIENCE AND TECHNOLOGY

MBI

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Recent in Engineering Science and Technology (RiESTech)
PT MENCERDASKAN BANGSA INDONESIA (MBI)
Available online at: http://www.mbi-journals.com/index.php/riestech

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Recent in Engineering Science and Technology (RiESTech)

Vol. 4 No. 4, October 2026

FOCUS AND SCOPE

RIESTECH

Recent in Engineering Science and Technology **(RiESTech):** I*SSN: 2985-704X (online), ISSN: 2985-8321 (print)* a peer-reviewed quarterly engineering journal, publishes theoretical and experimental high-quality papers to promote engineering and technology's theory and practice. In addition to peer-reviewed original research papers, the Editorial Board welcomes original research reports, state-of-the-art reviews, and communications in the broadly defined field of recent engineering science and technology. **RiESTech** covers topics contributing to a better understanding of engineering, material science, computer science, environmental science, and their applications. **RiESTech** is concerned with scientific research on mechanical and civil engineering, Electrical/Electronics and Computer Engineering, and Metallurgical and Materials Engineering with specific analytical techniques and/or computational methods.

The frequency of RiESTech publications is four times a year namely in January, April, July, and October. The scope of RiESTech includes a wide spectrum of subjects namely:

Mechanical and Civil Engineering (Automotive Technologies; Construction Materials; Design and Manufacturing; Dynamics and Control; Energy Generation, Utilization, Conversion, and Storage; Fluid Mechanics and Hydraulics; Heat and Mass Transfer; Micro-Nano Sciences; Renewable and Sustainable Energy Technologies; Robotics and Mechatronics; Solid Mechanics and Structure; Thermal Sciences)

Electrical/Electronics and Computer Engineering (Instrumentation; Coding, Cryptography, and Information Protection; Communications, Networks, Mobile Computing, and Distributed Systems; Compilers and Operating Systems; Parallel Processing, and Dependability; Computer Vision and Robotics; Control Theory; Electromagnetic Waves, Microwave Techniques and Antennas; Embedded Systems; Integrated Circuits, VLSI Design, Testing, and CAD; Microelectromechanical Systems; Microelectronics, and Electronic Devices and Circuits; Power, Energy and Energy Conversion Systems; Signal, Image, and Speech Processing; Machine Learning and Data Science)

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Metallurgical and Materials Engineering (Advanced Materials Science; Ceramic and Inorganic Materials; Electronic-Magnetic Materials; Energy and Environment; Materials Characterization; Metallurgy Extractive; Polymers and Nanocomposites)

Environmental Science and Engineering (Waste Management, Climate Change, Zero Waste, Environmental Disaster Management, Circular Economy, Sustainable Development, Environmental Security, Environmental Management, Environmental Ecology, Conservation of Natural Resources And Environment, Environmental Impact Analysis, Planning and Environmental Administration, Environmental Health, Environmental Pollution, Environmental Accounting, and Environmental Information Systems)

Recent in Engineering Science and Technology (RiESTech)
PT MENCERDASKAN BANGSA INDONESIA (MBI)
Available online at: http://www.mbi-journals.com/index.php/riestech

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Recent in Engineering Science and Technology

(RiESTech)

Vol. 4 No. 4, October 2026

EDITOR TEAM

Editor in Chief

Prof. Iwan Susanto, Ph.D

Managing Editor

Prof. Dr. Ir. Kuncoro Diharjo S.T., M.T
Dr. Vika Rizkia

Editorial Board

Prof. Dr. Ir. Dwi Rahmalina MT, Universitas Pancasila, Indonesia
Prof. Ing-Song Yu, National Dong Hwa University, Taiwan
Prof. Chao-Yu Lee, National Formosa University, Taiwan
Prof. Ching-An Huang, Chang Gung University, Taiwan
Prof. Fabrice Gourbilleau, CIMAP CNRS/CEA/ENSICAEN/Université de Caen Normandie, France
Dr. Ir. Muhammad Amin, ST, MT, IPM, Universitas Samudra, Kota Langsa, Indonesia
Dr. Maykel Manawan, Universitas Pertahanan, Indonesia
Dr. Eng. Radon Dhelika, Universitas Indonesia
Dr. Ing. Haryanti Samekto, The University of Stuttgart, Germany (Alumni)
Dr. Ing. H. Agus Suhartono, BRIN, Indonesia
Yudhi Ariadi, Ph.D, Coventry University London, United Kingdom
Dien Taufan Lessy, S.ST, M.Sc Institute of Digital Signal Processing, Universiät Duisburg Essen
Noor Hidayati, S.T., M.S. Politeknik Negeri Jakarta, Indonesia

Peer-Reviewers

Prof. Dr. Tatun Hayatun Nufus, M.Si, Politeknik Negeri Jakarta, Indonesia
Dr. Rachmat Adhi Wibowo, M.Sc., AIT Austrian Institute of Technology Center for Energy Energy Conversion and Hydrogen, Giefinggasse 2, 1210 Vienna, Austria
Dhayanantha Prabu Jaihindh, Ph.D Academia Sinica, Institute of Atomic and Molecular Sciences, Taiwan
Dr. rer nat Eko Budiyanto, Max-Planck-Institut für Kohlenforschung, Germany
Sk Jahir Abbas, Ph.D, Shanghai Jiao Tong University School of Medicine, Shanghai, China
Wandi Wahyudi, Ph.D, Uppsala University, Sweden
Dr. Agus Budi Prasetyo, Pusat Riset Metalurgi, BRIN, Indonesia
Atul Verma, Ph,D., National Dong Hwa University, Shoufeng, Taiwan
Haolia Rahman, Ph.D, Politeknik Negeri Jakarta, Indonesia
Andy Tirta, S.T., M.Eng., Ph.D., Universitas Darma Persada, Indonesia
Dr. Vincent Irawan, Eindhoven University of Technology, Netherlands
Muhammad Hilmy Alfaruqi, S.T., M.Eng., Ph.D. Chonnam National University, South Korea
Tia Rahmiati, S.T., M.T, Politeknik Negeri Jakarta, Indonesia

Recent in Engineering Science and Technology (RiESTech)
PT MENCERDASKAN BANGSA INDONESIA (MBI)
Available online at: http://www.mbi-journals.com/index.php/riestech

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Layout and Typesetting:
Imam Sapto Nugroho, Universitas Indonesia (Alumni), Indonesia
Raihan Trinanda Agsya, Politeknik Negeri Jakarta, Indonesia

PUBLISHER
PT MENCERDASKAN BANGSA INDONESIA (MBI)
Address : 4th Floor Gedung STC Senayan Room 31-34, Jl. Asia Afrika Pintu IX, Jakarta 10270, Indonesia.

Recent in Engineering Science and Technology (RiESTech)
PT MENCERDASKAN BANGSA INDONESIA (MBI)
Available online at: http://www.mbi-journals.com/index.php/riestech

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Recent in Engineering Science and Technology (RiESTech)

Vol. 4 No. 4, October 2026

PREFACE

Journal RiESTech (e-ISSN: 2985-704X (online), p-ISSN: 2985-8321 (print); is a peer review journal published by PT Mencerdaskan Bangsa Indonesia. The RiESTech journal is published four times a year in January, April, July, and October. This journal provides direct open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge within the engineering field. This journal aims to provide a place for academics, researchers, and practitioners to publish original research articles or review articles. The scope of articles published in this journal relates to various topics in the field of outcomes of research activities.

The RiESTech journal publishes papers strictly following the RiESTech guidelines and templates for manuscript preparation. All submitted manuscripts will go through a double-blind peer review process. The paper is read by members of the editor (according to the area of specialization) and will be screened by the Managing Editor to meet the criteria required for RiESTech publication. Manuscripts will be sent to two reviewers based on their historical experience in reviewing manuscripts or based on their areas of specialization. RiESTech has review forms to keep the same item reviewed by two reviewers. Then the editorial board makes a decision on the comments or suggestions of the reviewers.

Reviewers provide an assessment of originality, clarity of presentation, contribution to the field/science. This journal publishes research articles, review articles/literature reviews, case reports and concept/policy articles, in all fields of Computer Science, Informatics Engineering, Multimedia, Arts. The article to be published is an original work and has never been published. Incoming articles will be reviewed by the reviewer team.

The Editorial Board will try to continue to improve the quality of the journal so that it can become an important reference in the development of engineering sciences. The greatest appreciation and gratitude to Mitra Bestari along with members of the Editorial Board and all parties involved in the publication of this journal. Complete writing instructions are displayed on the portal of this journal.

Regards,

Chief Editor

Recent in Engineering Science and Technology (RiESTech)
PT MENCERDASKAN BANGSA INDONESIA (MBI)
Available online at: http://www.mbi-journals.com/index.php/riestech

E-ISSN : 2985-704X
P-ISSN : 2985-8321

Recent in Engineering Science and Technology (RiESTech)

Vol. 4 No. 4, October 2026

Contents

Focus and Scope ii
Editor Team iii
Preface v
Contents vi

Articles

- ***A Risk-Based Integrity Management of CO_2 Pipelines in CCUS: Integration of Failure Mechanisms, Risk Assessment, and Inspection Strategies—A Review***
  Risdianto, Johny Wahyuadi Soedarsono, Yudha Pratesa
  185 - 208

- ***Preliminary LEAP-Based Scenario Analysis of Electricity Demand and Renewable Energy Development in South Sulawesi***
  La Ode Masri Ande Kolewora, Arham, Saiful. M, Ahmad Agus Setiawan, Muhammad Taufiq, Virtuous Ikmal Wicaksono
  209 – 223

- ***Auditing Physical Consistency in Engine Fault Detection Using Domain-Informed Monotonic XGBoost and Explainable Boosting Machine***
  Abdul Azis Abdillah
  224 - 244

- ***Analysis of the Effect of Air Inlet Filter Replacement on Gas Turbine Performance at Unit 1A of Bekasi Power Combined Cycle Power Plant***
  Gilang Fitra Ardiansyah, Arifia Eka Yuliana
  245 - 254

- ***The Analysis of the Development of an Electric Mini Excavator Based on Environmentally Friendly Technology***
  Fuad Zainuri, Rahmat Noval, M. Todaru
  255 - 262

RiESTech
Recent in Engineering Science and Technology

MBI

Article

Risk-Based Integrity Management of CO_2 Pipelines in CCUS: Integration of Failure Mechanisms, Risk Assessment, and Inspection Strategies—A Review

Risdianto [1,*], Johny Wahyuadi Soedarsono [1], Yudha Pratesa [1]

[1] Department of Metallurgical and Materials Engineering, Faculty of Engineering, Universitas Indonesia, Depok 16424, Indonesia
* Correspondence: arisdi434@gmail.com

Abstract: The rapid deployment of Carbon Capture, Utilization, and Storage (CCUS) systems has significantly increased the demand for safe and reliable CO_2 transportation infrastructure. Pipelines are widely recognized as the most economical and technically feasible option for large-scale CO_2 transport. however, the thermodynamic behavior of dense-phase CO_2, coupled with impurity interactions, rapid decompression, and environmentally assisted degradation, introduces integrity challenges that differ substantially from those encountered in conventional hydrocarbon pipelines. This review aims to integrate current knowledge on CO_2 pipeline failure mechanisms, probabilistic risk assessment approaches, and inspection strategies into a comprehensive Risk-Based Inspection (RBI) framework for CCUS applications. A structured literature review was conducted using peer-reviewed publications retrieved from the Scopus and Web of Science databases, complemented by internationally recognized engineering standards, including API 580, API 581, API 579-1/ASME FFS-1, and DNV recommended practices. The selected studies were systematically analyzed according to degradation mechanisms, probabilistic risk assessment methodologies, uncertainty treatment, and inspection and monitoring technologies. The review indicates that internal corrosion, impurity-induced degradation, and running fracture remain the dominant threats to CO_2 pipeline integrity, while probabilistic approaches, particularly Bayesian networks and reliability-based models, provide superior capability for representing operational uncertainty compared with conventional deterministic methods. Nevertheless, existing studies remain fragmented, with degradation assessment, probabilistic risk modeling, and inspection optimization frequently addressed as separate research topics. To overcome this limitation, this paper proposes a Dynamic Risk-Based Inspection (D-RBI) framework that integrates CO_2-specific degradation mechanisms, Bayesian risk updating, real-time monitoring, and adaptive inspection scheduling into a unified integrity management strategy. The proposed framework provides a conceptual foundation for developing adaptive, data-driven, and risk-informed integrity management systems capable of supporting the long-term safety and reliability of future CCUS pipeline networks.

Keywords: CO_2 pipeline; Carbon Capture, Utilization and Storage; Risk-Based Inspection; integrity management; Bayesian network; corrosion; fracture propagation.

Citation: Risdianto., Soedarsono, J. W., Pratesa, Y. (2026). Risk-Based Integrity Management of CO_2 Pipelines in CCUS: Integration of Failure Mechanisms, Risk Assessment, and Inspection Strategies—A Review. Recent in Engineering Science and Technology, 4(4), 185–208. Retrieved from https://www.mbi-jounals.com/index.php/riestech/article/view/167.

Academic Editor: Yudhi Ariadi

Received: 21 June 2026
Accepted: 27 July 2026
Published: 11 October 2026

Publisher's Note: MBI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2026 by the authors. Licensee MBI, Jakarta, Indonesia. This article is an open access article distributed under MBI license (https://mbi-journals.com/licenses/by/4.0/).

1. Introduction

Global efforts to achieve carbon neutrality and mitigate greenhouse gas emissions have accelerated the deployment of Carbon Capture, Utilization, and Storage (CCUS) technologies as one of the principal decarbonization strategies for hard-to-abate industrial sectors. The International Energy Agency (IEA) identifies CCUS as an essential technology for achieving global net-zero emission targets because it enables substantial reductions in carbon emissions from cement, steel, petrochemical, refining, and fossil-fuel power generation facilities [1].

A critical component of any CCUS system is the transportation network responsible for transferring captured carbon dioxide from emission sources to geological storage sites or utilization facilities. Among the available transportation alternatives, pipelines remain the most technically mature and economically attractive solution for continuous, large-capacity CO_2 transport over long distances [2]. To maximize transport efficiency, CO_2 pipelines are commonly operated under dense-phase or supercritical conditions, where fluid properties become highly sensitive to pressure, temperature, and impurity composition [3].

Unlike conventional natural gas transmission systems, captured CO_2 streams frequently contain impurities such as water (H_2O), hydrogen sulfide (H_2S), sulfur dioxide (SO_2), oxygen (O_2), and nitrogen oxides (NO_x). These impurities influence phase stability, electrochemical behavior, corrosion kinetics, fracture propagation, and decompression characteristics, thereby introducing degradation mechanisms that are fundamentally different from those encountered in hydrocarbon pipelines [4,5]. Internal corrosion caused by free water and acidic contaminants remains one of the most critical threats to pipeline integrity [6]. Furthermore, rapid decompression during accidental rupture can significantly accelerate crack propagation and reduce crack-arrest capability, increasing the likelihood of high-consequence failures [7]. Additional degradation mechanisms, including stress corrosion cracking (SCC), sulfide stress cracking (SSC), hydrogen-induced cracking (HIC), and impurity-assisted localized corrosion, have also been reported under specific operating conditions [8].

Risk-Based Inspection (RBI) has become a widely accepted methodology for optimizing inspection planning and maintenance strategies in the oil and gas industry. By integrating the Probability of Failure (PoF), Consequence of Failure (CoF), degradation mechanisms, and inspection effectiveness, RBI enables inspection resources to be allocated according to risk rather than fixed inspection intervals [9,10]. However, current RBI methodologies were primarily developed for conventional hydrocarbon facilities and do not explicitly consider the unique thermodynamic behavior, impurity effects, decompression characteristics, and environmentally assisted degradation mechanisms associated with dense-phase CO_2 transportation.

Although substantial progress has been made in understanding individual aspects of CO_2 pipeline integrity, existing research remains highly fragmented. Numerous studies investigate corrosion mechanisms, fracture mechanics, probabilistic reliability analysis, inspection technologies, or digital monitoring independently, with relatively limited efforts devoted to integrating these disciplines into a unified integrity management framework. This fragmentation has important practical consequences. In the absence of an integrated framework, degradation mechanisms may be evaluated independently from inspection planning, resulting in overly conservative inspection intervals, insufficient inspection coverage for emerging degradation mechanisms, inefficient allocation of maintenance resources, and an increased likelihood of undetected failures under variable operating conditions. Consequently, operators may either incur unnecessary inspection costs or underestimate evolving integrity risks, particularly in large-scale CCUS transportation networks.

Several review papers have discussed corrosion mechanisms, fracture control, or pipeline reliability individually. However, few studies have systematically integrate CO_2-specific degradation mechanisms, probabilistic risk assessment, uncertainty management, and inspection optimization into a single Risk-Based Inspection framework. Accordingly, the primary novelty of this review lies in proposing an integrated Dynamic Risk-Based Inspection (D-RBI) framework that combines CO_2-specific degradation mechanisms, Bayesian probabilistic updating, real-time monitoring, and adaptive inspection planning

within a unified integrity management concept. Unlike previous reviews, this study emphasizes the interaction between degradation assessment, uncertainty quantification, and inspection decision-making to support future digitalized CCUS infrastructure.

Therefore, this review aims to synthesize current knowledge on CO_2 pipeline failure mechanisms, probabilistic risk assessment methodologies, inspection technologies, and uncertainty treatment, and subsequently integrate these elements into a conceptual Dynamic Risk-Based Inspection framework suitable for future CCUS applications. The guiding research question addressed in this review is : *How can CO_2-specific degradation mechanisms, probabilistic risk assessment, uncertainty management, and inspection strategies be integrated to establish an adaptive Risk-Based Inspection framework for ensuring the long-term integrity of CO_2 transportation pipelines?*

Table 1. Comparative analysis of key literature and identified research gaps**.**

Author and Year	Main Paper Title	Core Methodology	Primary Contributions	Limitations / Gaps
Yarveisy et al. (2020) [11]	Dynamic risk assessment of corrosion in CO_2 transportation pipelines...	Probabilistic modeling & dynamic risk assessment	Established dynamic risk model integrating transient operations with probabilistic corrosion growth.	Limited to corrosion; not translated into tactile inspection scheduling or fracture control.
Khan et al. (2021) [12]	Risk-based pipeline integrity management: A road map for resilient pipelines	Conceptual framework & reliability analysis	Provided a roadmap for digital transformation of conventional RBI toward predictive analytics.	Generalized for conventional hydrocarbon; dense-phase CO_2 multi-impurity effects unaddressed.
Li et al. (2018) [5]	Corrosion behavior of carbon steel in dense-phase CO_2 pipelines...	Experimental laboratory investigation	Delivered electrochemical data on impurity effects (SO_2, H_2S, O_2, NOx) on pitting initiation.	Focused on lab-scale mechanisms without probabilistic risk modeling or field optimization.
Zhu and Wang (2023) [13]	Data-driven RBI framework for carbon transport infrastructure	Machine learning & data-driven RBI	Used operational data to forecast wall-thickness degradation and optimize inspection intervals.	Black-box statistical correlations were not sufficiently constrained by metallurgy or fracture limits.
Zheng et al. (2022) [14]	Integrity management of CCUS pipelines under uncertainty	Uncertainty quantification & simulation	Quantified epistemic and aleatory uncertainties under supply fluctuations from capture facilities.	Inspection optimization remained largely static rather than adaptively updated via sensors.

2. Review Methodology

This study was conducted as a structured literature review to critically evaluate recent developments in CO_2 pipeline integrity management for Carbon Capture, Utilization, and Storage (CCUS) systems. The review focused on integrating knowledge from degradation mechanisms, probabilistic risk assessment, inspection technologies, uncertainty analysis,

and integrity management frameworks to establish a comprehensive understanding of Risk-Based Inspection (RBI) implementation for CO_2 transportation infrastructure.

A systematic literature search was performed using the Scopus and Web of Science databases because of their extensive coverage of high-quality peer-reviewed engineering literature. The search strategy employed combinations of Boolean operators using keywords including "CO_2 pipeline", "CCUS", "carbon capture and storage", "pipeline integrity", "Risk-Based Inspection", "corrosion", "fracture propagation", "dense-phase CO_2", "Bayesian network", "probabilistic risk assessment", "inspection optimization", and "integrity management". In addition to journal articles, internationally recognized engineering standards and recommended practices, including API 580, API 581, API 579-1/ASME FFS-1, DNV-RP-F104, and DNV-ST-F101, were reviewed to provide industrial context and current engineering practice.

Publications were screened according to predefined inclusion criteria. Studies were retained when they discussed at least one of the following topics:

1. CO_2-specific degradation mechanisms, including corrosion, fracture, and environmentally assisted cracking;
2. Probabilistic, reliability-based, or uncertainty-based risk assessment methodologies;
3. Inspection, monitoring, or integrity management technologies applicable to CO_2 pipelines; or
4. Conceptual or practical Risk-Based Inspection (RBI) frameworks.

Studies unrelated to CO_2 transportation, duplicate publications, papers lacking sufficient technical information, and articles focusing exclusively on non-pipeline CCS components were excluded. Following the screening process, 78 publications were selected for detailed review and classified into four categories: numerical/simulation studies, experimental investigations, field applications, and conceptual or framework-oriented studies.

The selected publications were subsequently analyzed using a qualitative thematic synthesis approach. Each study was evaluated according to four analytical dimensions: (i) degradation mechanisms, (ii) probabilistic risk assessment methodology, (iii) inspection and monitoring technology, and (iv) integrity management strategy. Comparative analysis of these themes enabled the identification of current research trends, methodological limitations, and remaining knowledge gaps. The synthesized findings formed the scientific basis for developing the *Dynamic Risk-Based Inspection (D-RBI) framework* proposed in this review.

Although this review does not constitute a formal meta-analysis, the structured search strategy, predefined selection criteria, and thematic synthesis methodology provide a transparent and reproducible approach for identifying current advances and future research directions in CO_2 pipeline integrity management.

3. Results and Discussion

3.1. Overview of Selected Studies

The reviewed literature demonstrates a marked increase in research activity on CO_2 pipeline integrity following 2018, coinciding with the rapid global expansion of Carbon Capture, Utilization, and Storage (CCUS) initiatives and increasing international commitments toward carbon neutrality. This growing research interest reflects the recognition that CO_2 transportation infrastructure represents one of the most critical components determining the technical feasibility and long-term reliability of large-scale CCUS deployment. As investment in carbon transportation networks continues to increase, ensuring pipeline integrity has become an important multidisciplinary research topic

encompassing metallurgy, corrosion engineering, fracture mechanics, probabilistic reliability, structural health monitoring, and risk-informed decision making.

Analysis of the selected publications indicates that the majority of existing studies focus on individual technical aspects of pipeline integrity rather than on integrated integrity management. Corrosion mechanisms, fracture propagation, probabilistic reliability assessment, and monitoring technologies have all received considerable attention; however, relatively few publications attempt to establish systematic relationships among these disciplines within a unified Risk-Based Inspection (RBI) framework. Consequently, current knowledge remains compartmentalized, making it difficult to translate scientific findings into practical inspection and maintenance strategies for future CO_2 transportation systems.

The classification of the selected studies further illustrates this imbalance. As summarized in Table 2, numerical and simulation-based investigations account for approximately 41% of the reviewed publications, followed by experimental and laboratory investigations (29%), field studies (20%), and conceptual or framework-oriented studies (9%). The predominance of simulation-based research indicates that considerable effort has been devoted to modeling corrosion growth, fracture propagation, decompression behavior, and probabilistic reliability under controlled assumptions. Nevertheless, relatively limited validation has been conducted using operational pipeline data, particularly for dense-phase CO_2 transportation systems where long-term industrial experience remains scarce.

Table 2. Distribution of selected studies according to research category.

Category	Number	Percentage
Numerical/Simulation	32	41%
Experimental/Laboratory	23	29%
Field Study	16	20%
Conceptual/Framework	7	9%
Total	**78**	**100%**

The relatively small proportion of conceptual studies is particularly noteworthy because it indicates that, despite significant advances in individual research domains, comparatively little attention has been devoted to developing comprehensive integrity management methodologies specifically tailored to CO_2 transportation pipelines. Existing studies generally concentrate on solving isolated engineering problems, such as predicting corrosion rates, evaluating fracture resistance, or improving monitoring technologies, without adequately considering how these findings should collectively support inspection planning and operational decision-making. This observation directly supports the research gap identified in the Introduction. As CCUS infrastructure evolves toward large-scale interconnected pipeline networks, integrity management will increasingly require multidisciplinary approaches capable of integrating degradation assessment, uncertainty quantification, inspection effectiveness, and operational monitoring into a single decision-support framework. Therefore, the current literature suggests that future research should move beyond isolated degradation studies toward integrated integrity management methodologies that continuously incorporate operational information into adaptive inspection planning.

3.2. CO_2 Pipeline Operational Characteristics

The operational characteristics of CO_2 transportation pipelines differ fundamentally from those of conventional hydrocarbon transmission systems because of the unique thermodynamic behavior of carbon dioxide under dense-phase and supercritical operating conditions. To maximize transportation efficiency while minimizing compression costs, CO_2 is typically transported at pressures above its critical pressure (7.38 MPa) and temperatures near or above its critical temperature (31.1 °C), where relatively small variations in operating conditions may produce substantial changes in fluid density, compressibility, and phase stability [3]. Unlike natural gas pipelines, the physical behavior of dense-phase CO_2 is highly sensitive not only to pressure and temperature but also to the presence of impurities originating from different carbon capture processes. Depending on the capture technology and emission source, transported CO_2 streams may contain varying concentrations of water (H_2O), hydrogen sulfide (H_2S), sulfur dioxide (SO_2), oxygen (O_2), nitrogen oxides (NO_x), methane (CH_4), nitrogen (N_2), hydrogen (H_2), and other trace components. Even relatively small impurity concentrations may significantly alter phase equilibrium, decompression characteristics, and electrochemical behavior, thereby influencing both operational safety and long-term material degradation [4,5].

Among these impurities, water represents one of the most critical parameters controlling pipeline integrity. Under dry conditions, carbon steel generally exhibits excellent compatibility with dense-phase CO_2. However, when free water is present, dissolved CO_2 readily forms carbonic acid (H_2CO_3), initiating electrochemical reactions that may result in generalized corrosion, localized pitting, or under-deposit corrosion. The presence of acidic contaminants such as SO_2 and NO_x further decreases solution pH and accelerates corrosion kinetics, while H_2S may promote sulfide stress cracking and hydrogen-assisted degradation in susceptible pipeline steels [4–6]. Consequently, strict impurity control should be regarded not only as a process requirement but also as a fundamental component of integrity management for CO_2 transportation systems.

In addition to corrosion-related challenges, dense-phase CO_2 pipelines exhibit decompression behavior that differs significantly from conventional gas transmission pipelines. During accidental rupture, rapid depressurization may induce substantial Joule–Thomson cooling, phase transformation, and two-phase flow, resulting in decompression wave velocities that strongly influence crack propagation and crack-arrest capability [7]. These thermodynamic phenomena complicate fracture control because traditional design criteria developed for natural gas pipelines may underestimate fracture propagation risks under dense-phase CO_2 conditions.

The combined effects of thermodynamic behavior, impurity chemistry, and fracture mechanics create complex interactions among operating parameters, degradation mechanisms, and structural integrity. Variations in operating pressure, temperature, flow rate, or impurity concentration may simultaneously influence corrosion susceptibility, crack initiation, and fracture propagation, producing dynamic changes in pipeline risk throughout the operational life cycle. Therefore, integrity management for CO_2 transportation systems cannot rely solely on deterministic design margins or fixed inspection intervals but should incorporate continuous monitoring, probabilistic risk assessment, and adaptive inspection planning capable of responding to changing operating conditions.

These operational characteristics distinguish CO_2 transportation pipelines from conventional hydrocarbon systems and provide the technical justification for developing CO_2-specific Risk-Based Inspection methodologies. Rather than treating degradation mechanisms independently, future integrity management strategies should integrate thermodynamic behavior, material degradation, operational uncertainty, and inspection

effectiveness within a unified framework capable of supporting predictive and adaptive decision-making for large-scale CCUS infrastructure.

3.3. Failure Mechanisms in CO_2 Pipelines

The integrity of CO_2 transportation pipelines is governed by a complex interaction between environmental conditions, material properties, operating parameters, and fluid chemistry. Unlike conventional hydrocarbon pipelines, degradation in dense-phase CO_2 systems is influenced not only by mechanical loading but also by thermodynamic instability and impurity-induced electrochemical reactions. Consequently, pipeline failures generally originate from multiple interacting degradation mechanisms rather than a single isolated defect.

The reviewed literature indicates that corrosion-related degradation and fracture propagation represent the two dominant integrity threats for dense-phase CO_2 transportation pipelines. These mechanisms are closely interconnected because corrosion defects frequently serve as crack initiation sites, while fracture propagation is strongly influenced by material degradation, residual strength, and operating pressure. Therefore, understanding the interaction between corrosion and fracture is essential for developing effective Risk-Based Inspection (RBI) strategies capable of predicting failure probability under variable operating conditions.

Table 3. Synthesis of degradation mechanisms in CO_2 Pipelines

Degradation Mechanisms	Key Drivers / Contributing Factors	Main Effects on Pipeline Integrity	Typical Locations / Conditions	Indicators / Observable Signs	Key References
General CO_2 Corrosion	Presence of free water, carbonic acid formation, low pH, temperature, partial pressure of CO_2	Uniform wall thinning, loss of load-bearing capacity	Low points, water accumulation zones	Metal loss (wt%), corrosion rate increase	[4], [6], [18], [34]
Localised Corrosion (Pitting)	Chlorides, deposits, differential aeration cells, flow disturbances	Pit initiation and growth, potential perforation	Under deposits, welds, coating holidays	Pit depth, pit density, ILI anomaly	[4], [5], [6], [8], [34]
Erosion–Corrosion	High velocity, sand/solid particles, two-phase flow, impingement	Accelerated metal loss, grooving, coating damage	Bends, reducers, elbows, tees	Wall thinning pattern, groove formation	[6], [18], [29], [35]
Sulfide Stress Cracking (SSC)	H_2S presence, high strength steel, tensile stress, high pCO_2	Crack initiation and propagation, brittle failure	High strength pipe, HAZ, girth weld	Crack indications (ILI/UT), pressure/deformation drop	[6], [15], [18], [33]
Hydrogen-Induced Cracking (HIC)	Cathodic protection, high strength steel, moist CO_2 environment	Internal crack growth parallel to metal surface	High strength steel, coated/CP areas	Blistering, stepwise crack growth	[6], [15], [18], [33]
Decompression Fatigue (DCF)	Rapid depressurization, trapped gas, high stress, material toughness	Crack nucleation and propagation under decompression	Compressor stations, shutdown events	Axial crack, brittle fracture features	[15], [18], [33], [36]

Thermodynamic Phase Transition Effects	Pressure/temperature fluctuations, proximity to critical point, Joule–Thomson cooling	Two-phase flow, slugging, accelerated corrosion/erosion	Near critical region, valves, control points	Temperature drop, slug flow, corrosion spikes	[18], [19], [35], [37]
External Corrosion	Soil resistivity, MIC, coating damage, CP failure, water retention	External metal loss, coating disbondment	Buried sections, coating defects	CIS survey, coating holiday, current density change	[7], [18], [28]

3.3.1. Corrosion

Corrosion was identified as the most extensively investigated degradation mechanism in the reviewed literature, with approximately 87% of the selected publications discussing corrosion behavior under dense-phase CO_2 conditions. This predominance reflects the critical role of internal corrosion in determining the long-term structural integrity and service life of carbon steel pipelines used for CO_2 transportation. Unlike dry CO_2, which is generally considered non-corrosive to carbon steel, the presence of free water fundamentally changes the electrochemical environment. Once water condenses within the pipeline, dissolved CO_2 reacts to form carbonic acid (H_2CO_3), initiating electrochemical dissolution of iron according to well-established sweet corrosion mechanisms [6]. Depending on operating conditions, corrosion may occur as uniform wall thinning, localized pitting, mesa attack, or under-deposit corrosion, each of which contributes differently to structural degradation and remaining strength. The severity of corrosion is strongly influenced by impurity composition. Laboratory investigations have consistently demonstrated that impurities such as SO_2, NO_x, H_2S, and O_2 significantly modify corrosion kinetics and corrosion product stability [4,5]. Sulfur dioxide and nitrogen oxides reduce local solution pH, destabilize protective corrosion films, and promote localized corrosion, while hydrogen sulfide may facilitate hydrogen absorption into the steel substrate, increasing susceptibility to sulfide stress cracking (SSC) and hydrogen-induced cracking (HIC). Even trace concentrations of oxygen can accelerate cathodic reactions and increase general corrosion rates under moist conditions.

An important characteristic of dense-phase CO_2 transportation systems is that corrosion behavior is governed by the combined interaction between thermodynamic conditions and fluid chemistry rather than by individual parameters alone. Variations in operating pressure, temperature, impurity concentration, and water content may significantly alter phase equilibrium, resulting in intermittent condensation of aqueous phases that locally accelerate corrosion processes. Consequently, degradation rates cannot be represented accurately using deterministic corrosion allowances alone but require continuous monitoring and probabilistic assessment capable of accounting for operational variability. From an integrity management perspective, corrosion represents considerably more than a wall-thinning mechanism. Progressive metal loss reduces pipeline burst strength, decreases fracture resistance, and creates geometric discontinuities that significantly increase local stress concentration. Localized corrosion pits frequently act as preferential sites for crack initiation, particularly under cyclic loading or environmentally assisted degradation. Therefore, corrosion assessment should not be treated independently from fracture mechanics but rather as an integral component of structural reliability evaluation.

These observations suggest that upstream impurity control should be regarded as a primary integrity barrier rather than merely a process-quality requirement. Maintaining water content below the CO_2 solubility limit, controlling acidic impurities, and continuously monitoring corrosion activity represent essential preventive measures for minimizing degradation and extending pipeline service life.

3.3.2. Fracture and Cracking

Fracture propagation constitutes the second major integrity threat identified in the reviewed literature, with approximately 49% of the selected studies addressing ductile fracture, brittle fracture, or running fracture behavior. Compared with conventional natural gas pipelines, fracture control in dense-phase CO_2 transportation systems is considerably more challenging because decompression behavior strongly influences crack propagation characteristics following accidental rupture.

During pipeline failure, rapid depressurization generates a decompression wave accompanied by substantial Joule–Thomson cooling and phase transformation. Unlike methane pipelines, where decompression generally follows predictable gas expansion, dense-phase CO_2 undergoes complex thermodynamic changes involving liquid, gaseous, and solid phases. These phenomena reduce crack-arrest capability and may substantially increase the likelihood of unstable running fracture over long pipeline distances [7].

Fracture resistance depends on multiple interacting parameters, including operating pressure, pipe diameter, wall thickness, material toughness, fracture toughness, crack geometry, and decompression wave velocity [15]. Consequently, fracture assessment for CO_2 pipelines requires consideration of both structural mechanics and fluid thermodynamics. Conventional fracture-control methodologies developed for hydrocarbon pipelines may therefore underestimate fracture propagation risks when applied directly to dense-phase CO_2 transportation systems. In addition to rapid fracture propagation, environmentally assisted cracking mechanisms have received increasing attention in recent years. Approximately 31% of the reviewed studies discussed degradation mechanisms including Stress Corrosion Cracking (SCC), Sulfide Stress Cracking (SSC), Hydrogen-Induced Cracking (HIC), and Hydrogen-Assisted Cracking (HAC). These mechanisms generally develop through the combined effects of tensile stress, susceptible microstructures, and aggressive chemical environments. Their occurrence is particularly significant in high-strength low-alloy steels exposed to impurity-containing CO_2 streams.

Corrosion and fracture should not be considered independent degradation mechanisms because they frequently interact throughout the pipeline life cycle. Corrosion defects create localized stress concentration, while fracture processes accelerate once crack initiation has occurred at corrosion sites. Similarly, hydrogen generated during corrosion reactions may diffuse into the steel matrix, reducing ductility and fracture toughness and promoting environmentally assisted cracking. This interaction significantly increases failure probability compared with isolated corrosion or fracture mechanisms.

The literature therefore suggests that effective integrity management should evaluate fracture susceptibility using integrated approaches that simultaneously consider corrosion growth, material degradation, operational loading, and fracture mechanics. Such multidisciplinary assessment is particularly important for Risk-Based Inspection because inspection priorities should reflect not only defect size but also the potential interaction between degradation mechanisms that may substantially increase the Probability of Failure (PoF).

Table 4. Frequency of failure mechanisms discussed in the reviewed literature.

Failure Mechanism	Number of Studies	Percentage*
CO_2 (sweet) corrosion	54	87%
Impurity-induced corrosion	34	44%
SCC / cracking	24	31%
Fracture (ductile/brittle)	38	49%
Running fracture	12	15%
Other degradation mechanisms	16	21%

**Note: Percentage is based on reviewed studies discussing the corresponding mechanisms.*

3.4. Risk Assessment Approaches

Risk assessment forms the foundation of modern pipeline integrity management because it provides a systematic framework for evaluating the likelihood and consequences of pipeline failure under uncertain operating conditions. In conventional pipeline systems, risk assessment has traditionally been based on deterministic engineering approaches that employ conservative safety factors, predefined inspection intervals, and fixed acceptance criteria. Although these approaches have demonstrated satisfactory performance in hydrocarbon transportation systems, they often fail to adequately represent the complex degradation behavior and operational uncertainty associated with dense-phase CO_2 pipelines.

Table 5. Methods for risk Assesment of CO_2 Pipelines

Risk Assessment Method	Purpose / Application	Key Inputs	Key Outputs	Advantages	Limitations	Key References
Deterministic Assessment (Code-Based)	Check compliance with design codes and standards	Material properties, design pressure, minimum wall thickness, operating conditions	Safety factor, utilization ratio	Simple, straightforward, widely accepted	Does not account for uncertainty and variability	[9], [10], [18]
Reliability-Based Assessment (RBA)	Quantify probability of failure from structural reliability	Limit state functions, statistics of material/loads, degradation data	Probability of Failure (PoF), reliability index (β)	Considers variability and uncertainty	Requires statistical data and modeling expertise	[10], [16], [17], [24]
Risk-Based Inspection (RBI)	Prioritize inspection based on risk (PoF × CoF)	PoF, Consequence of Failure (CoF), inspection data	Risk ranking, inspection priority, inspection interval	Optimizes resources, industry standardized (API 580/581)	Static decision between inspection cycles	[9], [10], [12]
Bayesian Network (BN)	Model causal relationships and update risk dynamically	Operational data, degradation model,	Updated PoF (posterior), uncertainty reduction	Captures dependencies, updates with new data	Requires sufficient data and expert knowledge	[11], [14], [16], [24]

		inspection evidence, expert judgment (prior probabilities)				
Monte Carlo Simulation (MCS)	Quantify uncertainty through random sampling	Probability distributions of loads, material properties, corrosion rate, model parameters	PoF distribution, confidence interval	Flexible, handles complex problems	Computationally intensive	[16], [17], [24]
Digital Twin-Based Assessment	Real-time prediction using virtual pipeline model	Real-time sensor data, physics-based models, historical data	Real-time risk, RUL, what-if analysis	Real-time monitoring and scenario analysis	High data and model requirements	[12], [18], [42]

The reviewed literature indicates a clear transition from deterministic methodologies toward probabilistic and reliability-based risk assessment approaches. Approximately half of the selected studies employ probabilistic techniques to quantify the uncertainty associated with corrosion growth, fracture initiation, inspection accuracy, material properties, and operational variability. This transition reflects the increasing recognition that pipeline degradation is inherently stochastic and cannot be represented accurately by single deterministic values. Among the available probabilistic approaches, *Monte Carlo Simulation (MCS)* remains one of the most widely adopted techniques for propagating uncertainty through degradation models. MCS enables statistical evaluation of multiple uncertain variables simultaneously, including corrosion rate, wall-thickness loss, material strength, operating pressure, and inspection uncertainty. Consequently, Monte Carlo methods provide probability distributions for failure events rather than single conservative estimates, thereby supporting more realistic integrity assessment.

Reliability-based approaches have also gained considerable attention because they explicitly quantify structural safety using probabilistic limit-state functions and reliability indices. Instead of relying solely on safety margins, reliability analysis evaluates the probability that structural resistance becomes lower than operational loading during service. Such methodologies provide a rational basis for estimating Probability of Failure (PoF) while accounting for uncertainties associated with material properties, manufacturing tolerances, degradation mechanisms, and inspection effectiveness.

Recent studies increasingly utilize *Bayesian Networks (BNs)* to represent complex dependency relationships among degradation mechanisms, operational conditions, and inspection outcomes. Unlike conventional probabilistic models, Bayesian Networks allow causal relationships among multiple variables to be represented explicitly within a probabilistic graphical model. Parent nodes typically represent operational parameters, including pressure, temperature, water content, impurity concentration, and material properties, whereas intermediate nodes describe degradation processes such as corrosion rate, crack initiation, hydrogen embrittlement, and fracture susceptibility. Leaf nodes subsequently estimate the Probability of Failure associated with different integrity limit states. One of the principal advantages of Bayesian Networks is their capability to update failure probability continuously as new operational evidence becomes available. Inspection findings, corrosion monitoring data, fiber-optic sensing measurements, or operational transients may be incorporated through Bayesian updating, allowing posterior probabilities to evolve dynamically throughout the operational life of the pipeline. This adaptive capability makes Bayesian Networks particularly attractive for supporting Risk-

Based Inspection (RBI) in emerging CCUS infrastructure, where operating conditions and impurity compositions may vary considerably over time.

The establishment of Bayesian Network models generally involves two complementary processes. First, prior probabilities are derived from historical inspection records, published reliability data, laboratory investigations, or expert judgement where field observations remain limited. Second, Conditional Probability Tables (CPTs) are constructed using a combination of statistical analysis, physics-based degradation models, experimental observations, and engineering expertise to quantify probabilistic relationships between parent and child nodes. As additional monitoring information becomes available, Bayesian inference updates posterior probabilities according to Bayes' theorem, thereby reducing epistemic uncertainty and improving prediction accuracy throughout pipeline operation. Despite these significant advances, the reviewed literature indicates that most probabilistic studies remain focused primarily on failure prediction rather than inspection optimization. Risk models frequently terminate at Probability of Failure estimation without explicitly translating updated risk information into inspection intervals, maintenance prioritization, or resource allocation. Consequently, substantial opportunities remain for integrating probabilistic risk assessment with adaptive inspection planning, thereby enabling risk-informed decision-making that continuously reflects actual pipeline operating conditions.

Overall, the literature demonstrates that probabilistic methodologies provide a substantially more realistic representation of CO_2 pipeline integrity than deterministic approaches because they explicitly account for uncertainty, operational variability, and continuous learning. These characteristics provide the theoretical foundation for the Dynamic Risk-Based Inspection (D-RBI) framework proposed in this study, where Bayesian updating continuously links operational monitoring with inspection planning.

3.5. Inspection and Monitoring Strategies

Inspection and monitoring constitute essential components of pipeline integrity management because they provide the operational information required to identify degradation before structural failure occurs. Within the Risk-Based Inspection (RBI) philosophy, inspection should not merely detect existing defects but should also reduce uncertainty associated with degradation mechanisms and improve confidence in subsequent risk assessment. Consequently, inspection effectiveness directly influences the reliability of Probability of Failure estimation and the optimization of maintenance strategies.

Table 6. Inspection Technologies and strategies for CO_2 Pipelines

Inspection Technology / Strategy	Type	Detectable Features	Best Application / Location	Advantages	Limitations	Key References
Inline Inspection (ILI)	In-line (Smart PIG)	Metal loss, pitting, cracks, deformation, geometry	Long-distance pipelines, mainline	High coverage, accurate sizing, minimal downtime	High cost, requires launcher/receiver	[9], [10], [39], [40]
Corrosion Monitoring Probes	Online Monitoring	Corrosion rate (uniform/local), environmental parameters	High risk locations, corrosion loops	Real-time data, early warning	Limited to local area	[18], [28], [41]

Distributed Fiber Optic Sensing (DFOS)	Continuous Monitoring	Temperature, strain, vibration, acoustic signals	Long-distance, remote pipelines	Distributed, real-time, long range	High installation complexity and cost	[12], [18], [29]
Direct Assessment (DA)	In-situ	Wall thickness, crack sizing, mechanical properties	High risk spots, anomaly verification	High accuracy	Requires shutdown, labor intensive	[9], [10], [30]
External Corrosion Survey (CIS)	External Inspection	Coating condition, corrosion rate, CP effectiveness	Buried pipelines	Detects external corrosion, coating defects	Cannot detect internal damage	[7], [28], [31]
Aerial / Satellite Monitoring	Remote Monitoring	Right-of-way, ground movement, third-party interference	Long ROW, remote areas	Large area coverage, low cost per km	Indirect method, lower resolution for defects	[12], [31], [32]
Inspection Strategy (Adaptive Planning)	Management Strategy	Risk-based interval, data fusion, adaptive planning	All pipeline systems	Optimizes resources, adapts to changing risk	Requires dynamic data and advanced analytics	[9], [10], [12], [24]

The reviewed literature indicates that Inline Inspection (ILI) remains the primary inspection technology for long-distance CO_2 transportation pipelines. Modern intelligent pigging systems employ various sensing principles, including Magnetic Flux Leakage (MFL), Ultrasonic Testing (UT), and Electromagnetic Acoustic Transducer (EMAT) technologies to detect internal corrosion, wall-thickness reduction, geometric deformation, longitudinal cracking, and manufacturing defects. Among these techniques, MFL is widely applied for metal-loss assessment, whereas UT generally provides superior sizing accuracy for corrosion defects and crack characterization. EMAT technology has attracted increasing interest because of its capability to detect stress corrosion cracking and other crack-like defects without requiring liquid couplants.

Although periodic Inline Inspection remains indispensable, conventional inspection programs are generally based on predetermined inspection intervals that may not adequately reflect changing operating conditions. Dense-phase CO_2 transportation systems are characterized by continuously varying impurity concentrations, pressure fluctuations, and thermodynamic conditions that may significantly alter degradation rates between scheduled inspections. Consequently, fixed inspection intervals may either result in unnecessary inspection costs or allow critical degradation to develop undetected.

To overcome these limitations, recent studies increasingly emphasize continuous structural health monitoring through permanently installed sensing systems. Distributed Fiber Optic Sensing technologies, including Distributed Temperature Sensing (DTS) and Distributed Acoustic Sensing (DAS), provide continuous monitoring over extensive pipeline distances by detecting thermal anomalies, acoustic emissions, mechanical strain, and leak-related disturbances. These technologies enable rapid identification of abnormal operating conditions that may indicate corrosion activity, leakage, or fracture initiation before catastrophic failure occurs.

Additional monitoring technologies, including Linear Polarization Resistance (LPR) probes, Electrical Resistance (ER) sensors, corrosion coupons, acoustic emission monitoring, pressure transient analysis, and wireless sensor networks, have further enhanced the capability for real-time degradation assessment. These monitoring systems provide continuous information regarding corrosion activity, environmental conditions,

and structural response, thereby significantly reducing uncertainty associated with degradation prediction.

Recent advances in digitalization have introduced machine learning and artificial intelligence techniques into pipeline monitoring systems. Data-driven algorithms have demonstrated promising capability for detecting anomalies, forecasting corrosion growth, identifying leakage patterns, and optimizing maintenance scheduling using large operational datasets. Nevertheless, purely data-driven approaches remain limited because statistical correlations alone cannot adequately represent the underlying physical mechanisms governing corrosion, fracture propagation, and thermodynamic behavior. Consequently, increasing attention has shifted toward physics-informed machine learning, which combines physical degradation models with data-driven learning to improve both prediction accuracy and interpretability. From an RBI perspective, the primary objective of inspection extends beyond defect detection. Inspection should function as a mechanism for reducing uncertainty within probabilistic risk models. Each inspection campaign provides new evidence regarding pipeline condition, allowing degradation models and Bayesian Network probabilities to be updated using actual field observations. Consequently, inspection effectiveness should be evaluated according to its ability to reduce uncertainty and improve risk prediction rather than solely according to defect detection capability.

The reviewed literature further indicates that current inspection optimization for CO_2 pipelines remains relatively immature. Most inspection strategies are still developed independently of probabilistic risk assessment and operational monitoring, limiting their ability to respond dynamically to evolving degradation conditions. Future integrity management systems should therefore integrate inspection technologies, continuous monitoring, Bayesian updating, and predictive analytics within a unified decision-support platform capable of continuously adjusting inspection priorities according to changing operational risk.

Such adaptive inspection strategies represent a significant evolution beyond conventional RBI practices and constitute one of the principal components of the Dynamic Risk-Based Inspection (D-RBI) framework proposed in this review. By continuously incorporating monitoring evidence into probabilistic risk assessment, inspection intervals can be optimized according to actual degradation conditions, thereby improving safety while reducing unnecessary inspection and maintenance expenditures.

3.6. Uncertainty Treatment

Uncertainty is one of the fundamental challenges in the integrity management of CO_2 transportation pipelines because degradation mechanisms are governed by complex interactions among operating conditions, material characteristics, environmental exposure, and impurity composition. Compared with conventional hydrocarbon pipelines, dense-phase CO_2 transportation systems are associated with greater uncertainty due to limited long-term operational experience, evolving CCUS infrastructure, and variability in captured CO_2 stream quality. Consequently, uncertainty influences not only degradation prediction but also inspection planning, maintenance optimization, and long-term risk-informed decision making. Zheng et al. [14] emphasized that uncertainty remains one of the principal obstacles to establishing reliable integrity management strategies for future CCUS pipeline networks, while Aven [16] highlighted that uncertainty should be considered an integral component of risk rather than merely a statistical parameter.

From a risk analysis perspective, uncertainty can generally be classified into *aleatory uncertainty* and *epistemic uncertainty*. Aleatory uncertainty represents the inherent variability of physical processes and cannot be eliminated because it originates from natural randomness, such as fluctuations in operating pressure, temperature, flow rate, impurity concentration, and corrosion growth. In contrast, epistemic uncertainty arises

from incomplete knowledge regarding degradation mechanisms, material behavior, inspection effectiveness, and predictive models. Unlike aleatory uncertainty, epistemic uncertainty can be progressively reduced through additional inspection data, laboratory investigations, operational experience, and improved analytical models. The distinction between these two uncertainty types has become a fundamental concept in modern reliability engineering and probabilistic risk assessment because each requires different treatment within integrity management frameworks [16,31,38].

A major source of uncertainty in CO_2 transportation systems is the variability of the transported fluid itself. The composition of captured CO_2 depends strongly on the carbon capture technology and the industrial source, including coal-fired power plants, cement production, steel manufacturing, and petrochemical facilities. As a result, concentrations of impurities such as H_2O, H_2S, SO_2, O_2, and NO_x may vary significantly among projects and even during normal operation. These compositional variations directly affect corrosion kinetics, phase equilibrium, fracture behavior, and thermodynamic stability, leading to substantial uncertainty in degradation prediction. Previous investigations have demonstrated that even trace concentrations of impurities may significantly modify electrochemical reactions and localized corrosion behavior in dense-phase CO_2 environments [4,5,14,34]. Among the various impurities, free water is considered the most influential parameter because it controls the initiation of electrochemical corrosion. Maintaining water content below the CO_2 solubility limit prevents the formation of an aqueous phase and minimizes carbonic acid corrosion. Likewise, acidic impurities such as SO_2 and NO_x reduce local pH and accelerate localized corrosion, whereas H_2S increases susceptibility to sulfide stress cracking (SSC) and hydrogen-induced cracking (HIC) in high-strength pipeline steels. Consequently, impurity specification should be regarded not only as a process requirement but also as a primary integrity control measure for minimizing uncertainty associated with degradation mechanisms [4–6,8].

Additional uncertainty arises from the prediction models used to estimate degradation and structural reliability. Corrosion growth models developed from laboratory experiments frequently differ from field observations because laboratory conditions cannot fully reproduce transient operating conditions, impurity interactions, localized flow characteristics, or environmental fluctuations encountered in service. Similarly, fracture assessment depends on uncertain variables such as fracture toughness, defect geometry, residual stress, and decompression behavior, all of which may vary during pipeline operation. These uncertainties directly affect predictions of Remaining Useful Life (RUL) and Probability of Failure (PoF), indicating that deterministic degradation models alone are insufficient for supporting adaptive integrity management [6,15,16,18].

Inspection activities play a crucial role in reducing epistemic uncertainty because they provide direct evidence of the actual structural condition of the pipeline. Information obtained from inline inspection (ILI), corrosion monitoring systems, non-destructive examination, and field observations continuously improves the understanding of degradation behavior and enables predictive models to be refined. Within the Risk-Based Inspection (RBI) philosophy, inspection should therefore be viewed not only as a defect detection activity but also as a mechanism for uncertainty reduction that increases confidence in subsequent risk assessment and maintenance planning [9,10,12].

Recent developments in probabilistic integrity management have increasingly adopted *Bayesian inference* to manage uncertainty throughout the operational life cycle of pipeline systems. Bayesian methods provide a rigorous mathematical framework for combining prior knowledge derived from historical inspection records, laboratory experiments, published reliability data, and expert judgement with newly acquired operational evidence. In practical applications, prior probabilities are updated as new monitoring information

becomes available, producing posterior probability distributions that continuously improve the estimation of degradation states and Probability of Failure. Han et al. [24] demonstrated that Bayesian Networks are particularly effective for representing causal relationships among operational variables, degradation mechanisms, and failure events, while Yarveisy et al. [11] successfully incorporated Bayesian updating into dynamic corrosion risk assessment for CO_2 transportation pipelines.

Despite these advances, the reviewed literature indicates that uncertainty treatment remains insufficiently integrated with inspection optimization. Most published studies focus on uncertainty quantification or probabilistic failure prediction independently, whereas relatively few explicitly translate uncertainty reduction into adaptive inspection intervals, maintenance prioritization, or inspection resource allocation. Consequently, many existing integrity management systems remain reactive rather than predictive, limiting their capability to respond efficiently to evolving degradation conditions in large-scale CCUS infrastructure [12–14]. Based on the perspective of Risk-Based Inspection, uncertainty should therefore be regarded as a dynamic parameter that evolves continuously throughout the service life of the pipeline. Effective integrity management requires the integration of uncertainty quantification, real-time monitoring, probabilistic risk assessment, and adaptive inspection planning within a unified decision-support framework. Such an approach enables inspection intervals and maintenance priorities to be continuously updated according to the latest operational evidence, thereby reducing both technical uncertainty and unnecessary inspection expenditure while enhancing pipeline safety and reliability [9,11,14,24].

Overall, the literature demonstrates that uncertainty treatment should become a central component of future CO_2 pipeline integrity management rather than a supplementary consideration. Integrating uncertainty quantification with Bayesian updating, continuous monitoring, and Dynamic Risk-Based Inspection (D-RBI) provides a scientifically robust foundation for predictive maintenance and supports the long-term integrity, safety, and economic performance of future CCUS transportation networks.

3.7. Dynamic Risk Based Inspection (D-RBI) Framework for CO_2 Pipelines

The comprehensive review presented in the preceding sections reveals that current integrity management practices for CO_2 transportation pipelines remain largely fragmented. Corrosion studies primarily focus on degradation mechanisms, fracture research emphasizes structural failure behaviour, while probabilistic risk assessment, inspection planning, and uncertainty analysis are generally investigated independently. Although each discipline contributes significantly to pipeline integrity management, limited integration among these approaches restricts the capability of existing Risk-Based Inspection (RBI) systems to respond effectively to continuously evolving operational conditions [9,10,14].

Conventional RBI, as defined in API 580 and API 581, evaluates inspection priorities based on the combination of Probability of Failure (PoF) and Consequence of Failure (CoF) using periodically updated inspection information [9,10]. This methodology has demonstrated considerable success for conventional oil and gas pipelines; however, its application to dense-phase CO_2 transportation systems is challenged by operational variability, impurity-induced degradation, thermodynamic phase transitions, and limited long-term operational experience. DNV-RP-F104 also emphasizes that CO_2 pipeline integrity requires consideration of operating conditions, fluid composition, and material degradation mechanisms that differ substantially from conventional hydrocarbon pipelines [18].

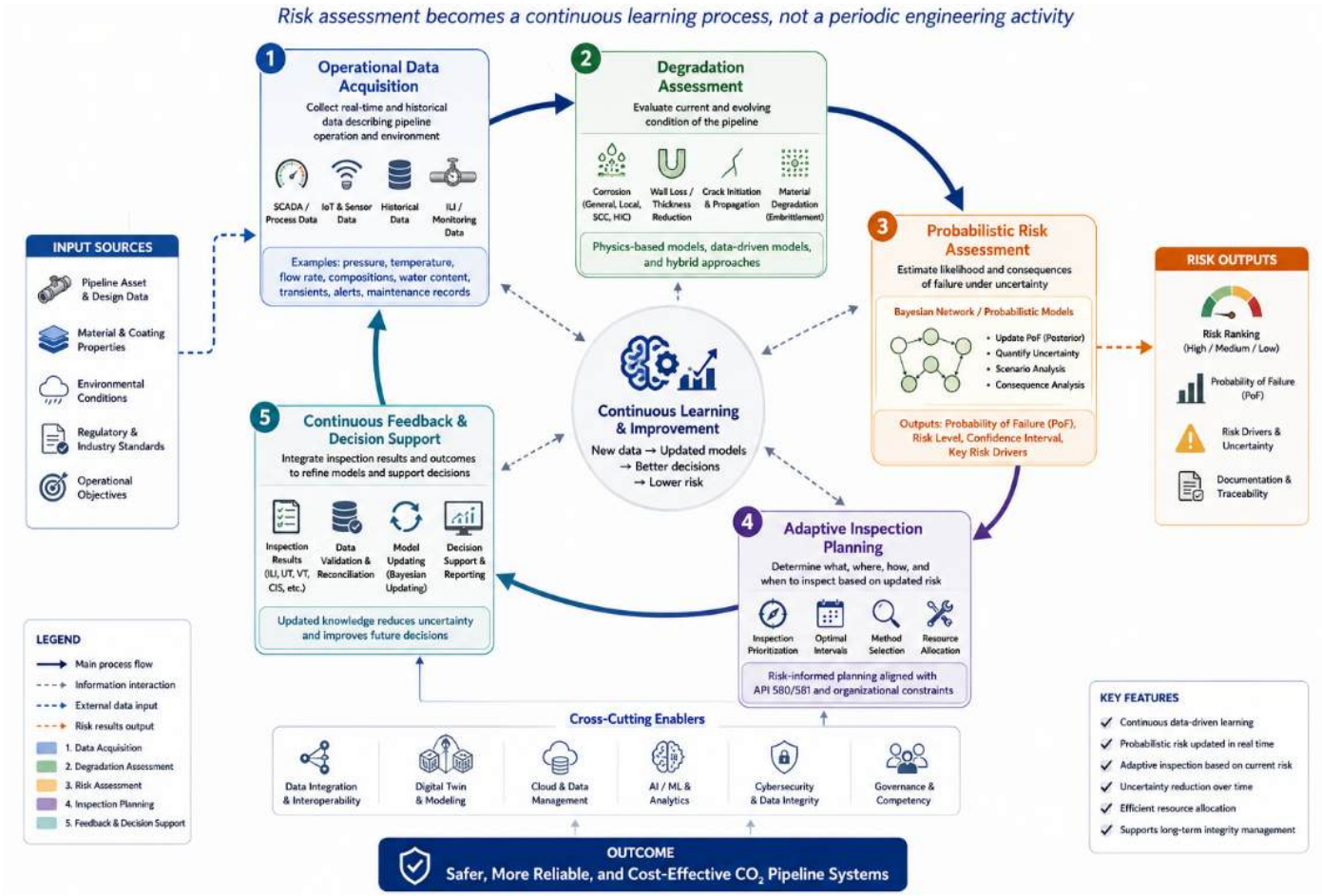


Figure 1. Dynamic Risk-Based Inspection (D-RBI) framework

Based on the reviewed literature, this study proposes an integrated *Dynamic Risk-Based Inspection (D-RBI)* framework in which risk assessment becomes a continuous learning process rather than a periodic engineering activity. As illustrated in Figure 1, the framework integrates five interconnected components: (i) operational data acquisition, (ii) degradation assessment, (iii) probabilistic risk assessment, (iv) adaptive inspection planning, and (v) continuous feedback and decision support. Unlike conventional RBI, which updates inspection plans only after scheduled inspections, D-RBI continuously incorporates operational information obtained from Supervisory Control and Data Acquisition (SCADA), Inline Inspection (ILI), corrosion monitoring systems, Distributed Fiber Optic Sensing (DFOS), and other structural health monitoring technologies [9,12,18].

Within the proposed framework, degradation assessment simultaneously considers corrosion growth, fracture mechanics, material degradation, and structural reliability. Nešić [6] demonstrated that corrosion progressively reduces pipeline structural resistance, whereas Zargarzadeh et al. [15] reported that fracture propagation depends strongly on decompression behaviour, material toughness, and defect geometry. Therefore, corrosion and fracture should not be treated as independent failure mechanisms but as interacting degradation processes that collectively determine structural reliability throughout pipeline operation.

The integration of degradation assessment with probabilistic risk evaluation enables inspection priorities to evolve dynamically according to actual operating conditions rather than predefined inspection intervals. Consequently, pipeline segments experiencing accelerated degradation, increased operational uncertainty, or elevated failure probability

receive higher inspection priority, whereas low-risk sections may safely operate with extended inspection intervals. Such adaptive inspection planning improves maintenance efficiency while maintaining acceptable levels of structural safety and operational reliability [9,10,12]. Overall, the proposed D-RBI framework establishes a conceptual foundation for transitioning from conventional static integrity management toward intelligent, adaptive integrity systems capable of supporting the long-term deployment of CCUS transportation infrastructure.

3.8 Bayesian Network-Based Dynamic Risk Updating

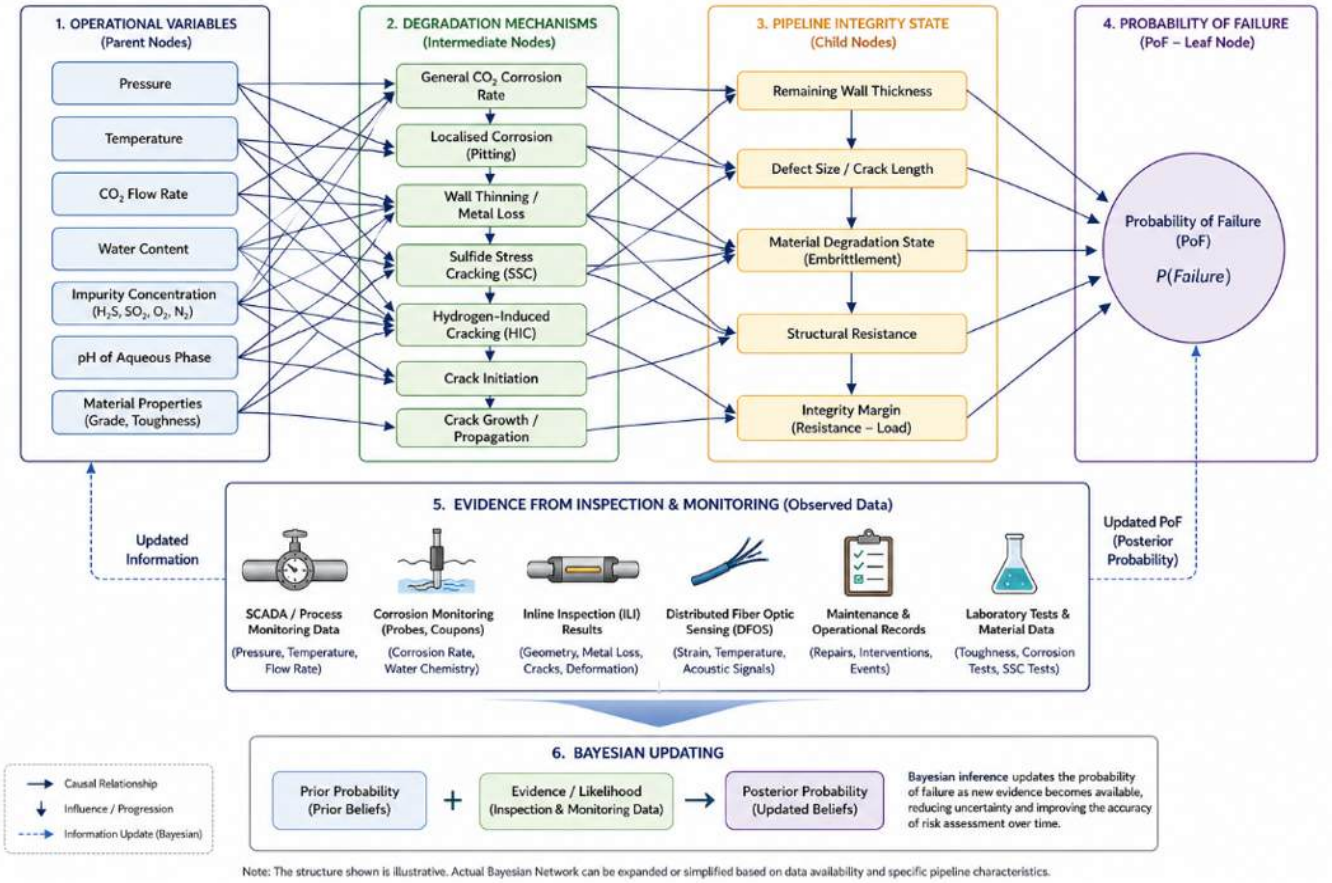


Figure 2. Conceptual Bayesian Network for Dynamic Risk Updating in CO_2 Pipeline Integrity Management.

An essential component of the proposed Dynamic Risk-Based Inspection framework is the application of Bayesian Networks (BNs) for continuously updating pipeline risk throughout the operational life cycle. Compared with conventional deterministic approaches, Bayesian Networks provide a probabilistic graphical framework capable of representing causal relationships among operational conditions, degradation mechanisms, inspection observations, and structural failure events under uncertain environments [11,24].

In Bayesian Networks, operational variables such as pressure, temperature, flow rate, water content, impurity concentration, and material properties are represented as parent nodes, while degradation processes—including corrosion rate, crack initiation, hydrogen-assisted cracking, and fracture susceptibility—are represented as intermediate nodes. The resulting structural condition determines the child node, which estimates the Probability of Failure (PoF). Relationships among these variables are quantified using Conditional

Probability Tables (CPTs) constructed from historical inspection records, laboratory investigations, published reliability data, physics-based degradation models, and expert judgement [11,16,24].

One of the principal advantages of Bayesian Networks lies in their capability to incorporate newly acquired operational evidence through Bayesian updating. Inspection findings, corrosion monitoring results, sensor measurements, pressure transients, and maintenance records are continuously integrated into the probabilistic model, allowing posterior probabilities to replace initial prior estimates. Consequently, epistemic uncertainty decreases progressively as additional operational information becomes available, resulting in increasingly reliable predictions of pipeline degradation and structural integrity [11,16]. Recent studies have demonstrated the effectiveness of Bayesian Networks in supporting pipeline integrity management. Yarveisy et al. [11] successfully integrated corrosion prediction with probabilistic inspection planning using Bayesian inference, while Han et al. [24] showed that Bayesian updating significantly improves failure prediction accuracy by incorporating continuously evolving inspection evidence. Zheng et al. [14] further emphasized that probabilistic learning approaches represent one of the most promising methodologies for supporting integrity management within emerging CCUS infrastructure, where operational uncertainty remains relatively high.

The integration of Bayesian Networks within the D-RBI framework therefore enables continuous updating of both Probability of Failure and inspection priorities. Instead of relying on periodic engineering judgement alone, inspection decisions become data-driven and adaptive, allowing maintenance resources to be directed toward pipeline segments exhibiting the highest updated risk. Such continuous probabilistic learning distinguishes Dynamic RBI from conventional static RBI methodologies and represents a significant advancement in intelligent pipeline integrity management.

3.9. Practical Implications and Future Research Directions

The findings of this review provide several practical implications for pipeline operators, regulators, and researchers involved in the development of Carbon Capture, Utilization, and Storage (CCUS) infrastructure. As commercial deployment of CO_2 transportation networks expands worldwide, maintaining pipeline integrity will become increasingly important for ensuring operational safety, environmental protection, regulatory compliance, and public acceptance. Existing Risk-Based Inspection methodologies described in API 580 and API 581 remain highly relevant; however, they require adaptation to address the operational uncertainty and degradation mechanisms unique to dense-phase CO_2 transportation systems [9,10,18]. From an industrial perspective, the proposed Dynamic Risk-Based Inspection framework offers a systematic strategy for integrating operational monitoring, degradation assessment, Bayesian risk updating, and adaptive inspection planning into a unified asset integrity management system. Continuous evaluation of operational data enables pipeline operators to optimize inspection intervals, improve maintenance prioritization, and reduce unnecessary inspection costs while maintaining acceptable levels of structural reliability [12,14].

The framework also supports ongoing digital transformation within pipeline asset management. Recent developments in SCADA systems, Distributed Fiber Optic Sensing (DFOS), Internet of Things (IoT) technologies, cloud-based asset management platforms, and Digital Twin applications provide unprecedented opportunities for continuously monitoring pipeline condition throughout its operational life cycle. When integrated with Bayesian inference and probabilistic degradation models, these technologies enable near real-time integrity assessment and predictive maintenance, significantly reducing epistemic uncertainty while improving operational decision-making [11,18,24]. Despite these promising developments, several research gaps remain. Long-term field data

describing corrosion evolution, fracture behaviour, and impurity effects under dense-phase CO_2 conditions are still limited, resulting in significant uncertainty within current degradation models [6,14,18]. Furthermore, most existing probabilistic studies focus primarily on estimating Probability of Failure without explicitly integrating inspection optimization, maintenance scheduling, and operational decision support into a unified framework. Consequently, future research should emphasize validation of Dynamic RBI using operational datasets obtained from commercial CO_2 pipeline systems.

Future investigations should also explore the integration of *Digital Twin technology, physics-informed machine learning, and Bayesian inference* within adaptive integrity management systems. Physics-informed machine learning has the potential to improve degradation prediction while preserving engineering interpretability, whereas Digital Twin technology enables continuous synchronization between virtual pipeline models and field observations. Combined with Bayesian updating, these emerging technologies are expected to provide the foundation for next-generation intelligent integrity management systems capable of supporting autonomous inspection planning and predictive maintenance under uncertain operating environments [12,16,24].

Overall, this review demonstrates that future CO_2 pipeline integrity management should evolve from conventional static inspection practices toward integrated Dynamic Risk-Based Inspection systems that combine engineering mechanics, corrosion science, probabilistic reliability, uncertainty quantification, digital monitoring, and intelligent decision support. Such multidisciplinary integration provides a robust scientific foundation for improving pipeline safety, operational reliability, inspection efficiency, and lifecycle cost performance while supporting the sustainable deployment of global CCUS infrastructure [9,11,14,18,24].

4. Conclusions

This review article critically examined the current state of integrity management for CO_2 transportation pipelines within Carbon Capture, Utilization, and Storage (CCUS) systems by integrating recent advances in degradation mechanisms, probabilistic risk assessment, inspection technologies, uncertainty treatment, and intelligent integrity management. Unlike conventional reviews that address these topics separately, this study synthesized the interrelationships among corrosion, fracture propagation, risk assessment, and inspection planning to establish a comprehensive understanding of CO_2 pipeline integrity.

The literature indicates that corrosion and fracture remain the dominant degradation mechanisms affecting the structural integrity of dense-phase CO_2 pipelines. Their occurrence is strongly influenced by operational parameters, impurity composition, thermodynamic conditions, and material characteristics. These interacting degradation processes increase the complexity of integrity assessment and demonstrate that deterministic approaches based solely on fixed inspection intervals are insufficient for representing the continuously evolving condition of CO_2 transportation systems [6,9,18]. Recent developments in probabilistic risk assessment have significantly improved the capability to quantify uncertainty and predict pipeline failure. Bayesian Networks, reliability-based assessment, and probabilistic degradation models enable continuous updating of the Probability of Failure (PoF) using operational monitoring and inspection evidence, thereby reducing epistemic uncertainty and supporting more reliable decision-making [11,14,16,24]. However, the review also reveals that most existing studies focus primarily on individual degradation mechanisms or probabilistic modelling without fully integrating inspection optimization and real-time operational monitoring.

To address this research gap, this review proposed a ***Dynamic Risk-Based Inspection (D-RBI)*** framework that integrates operational data acquisition, degradation assessment,

Bayesian risk updating, adaptive inspection planning, and continuous feedback within a unified integrity management cycle. The proposed framework extends conventional Risk-Based Inspection by transforming inspection planning from a periodic activity into a continuous learning process that dynamically responds to changes in operating conditions, degradation behaviour, and inspection findings. Such integration has the potential to improve pipeline safety, optimize maintenance resources, reduce lifecycle costs, and enhance the reliability of future CCUS transportation infrastructure. Despite these advances, several research challenges remain. Long-term field data for dense-phase CO_2 transportation are still limited, and further validation of probabilistic degradation models using operational datasets is required. Future studies should therefore focus on integrating Bayesian inference, Digital Twin technology, physics-informed machine learning, and continuous structural health monitoring to establish intelligent integrity management systems capable of supporting predictive maintenance under uncertain operating environments.

Overall, this review demonstrates that future CO_2 pipeline integrity management should evolve beyond conventional static Risk-Based Inspection toward adaptive, data-driven, and uncertainty-aware integrity management frameworks. The proposed Dynamic Risk-Based Inspection concept provides a practical and scientifically robust foundation for supporting the safe, reliable, and economically sustainable deployment of large-scale CCUS pipeline networks.

Supplementary Materials: Not applicable.

Author Contributions: Conceptualization, Risdianto, J.W.S. and Y.P.; methodology, Risdianto; formal analysis, Risdianto; investigation, Risdianto; writing—original draft preparation, Risdianto; writing—review and editing, J.W.S. and Y.P.; supervision, J.W.S. and Y.P. All authors have read and agreed to the published version of the manuscript.

Funding: This study received no specific funding from government, commercial, or non-profit organizations.

Acknowledgments: The authors acknowledge the Department of Metallurgical and Materials Engineering, Universitas Indonesia, for academic support during preparation of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. International Energy Agency. CCUS in Clean Energy Transitions; IEA: Paris, France, 2022.
2. Knoope, M.; Ramírez, A.; Faaij, A.; van der Ham, L. A state-of-the-art review of techno-economic models predicting the costs of CO_2 pipeline transport. International Journal of Greenhouse Gas Control 2013, 16, 241–270.
3. Zhang, Z.; Mahgerefteh, H.; Brown, S. Modelling of full-bore rupture of dense-phase CO_2 pipelines. Process Safety and Environmental Protection 2006, 84, 529–539.
4. Dugstad, A.; Halseid, M.; Morland, B.; Clausen, S. Corrosion of transport pipelines for CO_2. Energy Procedia 2011, 4, 3063–3070.
5. Li, X.; Wang, S.; Yan, J.; Zhang, L. Corrosion behavior of carbon steel in dense-phase CO_2 transport pipelines containing impurities. Corrosion Science 2018, 146, 1–12.
6. Nešić, S. Key issues related to modelling of internal corrosion of oil and gas pipelines—A review. Corrosion Science 2007, 49, 4308–4338.
7. Cosham, A.; Eiber, R. Fracture control in carbon dioxide pipelines: A review. International Journal of Greenhouse Gas Control 2008, 2, 533–551.

8. Choi, Y.-S.; Nešić, S.; Young, D. Effect of impurities on the corrosion behavior of CO_2 transmission pipeline steel in supercritical CO_2 environments. Environmental Science & Technology 2013, 47, 296–302.
9. API 580. Risk-Based Inspection; American Petroleum Institute: Washington, DC, USA, 2016.
10. API 581. Risk-Based Inspection Technology; American Petroleum Institute: Washington, DC, USA, 2016.
11. Yarveisy, R.; Khan, F.; Abbassi, R.; Garaniya, V. Dynamic risk assessment of corrosion in CO_2 transportation pipelines using probabilistic modeling. Process Safety and Environmental Protection 2020, 138, 127–139.
12. Khan, F.; Yarveisy, R.; Abbassi, R.; Garaniya, V. Risk-based pipeline integrity management: A road map for resilient pipelines. Journal of Pipeline Science and Engineering 2021, 1, 100001.
13. Zhu, X.; Wang, H. Data-driven RBI framework for carbon transport infrastructure. Process Safety and Environmental Protection 2023, 172, 113–126.
14. Zheng, J.; Liu, Y.; Xu, J. Integrity management of CCUS pipelines under uncertainty. International Journal of Greenhouse Gas Control 2022, 118, 103683.
15. Zargarzadeh, P.; Chahardehi, A.; Brennan, F. Development of a failure assessment diagram based method for engineering criticality assessment of CO_2 transportation pipelines. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering 2013, 227, 275–286.
16. Aven, T. Risk Analysis; Wiley: Hoboken, NJ, USA, 2015.
17. API 579-1/ASME FFS-1. Fitness-For-Service; American Petroleum Institute and ASME: Washington, DC, USA, 2021.
18. DNV-RP-F104. Design and Operation of Carbon Dioxide Pipelines; Det Norske Veritas: Høvik, Norway, 2021.
19. DNVGL-ST-F101. Submarine Pipeline Systems; DNV: Høvik, Norway, 2021.
20. Dugstad, A.; Morland, B.; Clausen, S. Internal corrosion control in CO2 transport pipelines. Energy Procedia 2013, 37, 2877–2884.
21. Eiber, R.; Cosham, A.; Maxey, W. Fracture propagation control for CO_2 pipelines. Proceedings of IPC Conference 2010, IPC2010-31352, 1–10.
22. Fu, B.; Cheng, Y. F.; Luo, J. L. Corrosion of pipelines in supercritical CO_2 environments. Corrosion Reviews 2014, 32(3-4), 109–128.
23. Guo, Q.; Wang, Y.; Sun, W. Machine learning-based pipeline integrity prediction under uncertain conditions. Reliability Engineering & System Safety 2022, 220, 108292.
24. Han, J.; Lee, K.; Kim, W. Bayesian network applications in pipeline integrity management. Process Safety and Environmental Protection 2020, 141, 89–101.
25. Han, Z.; Weng, W.; Huang, G. Dynamic risk assessment methodologies for process systems. Journal of Loss Prevention in the Process Industries 2019, 57, 17–29.
26. Heddle, G.; Herzog, H.; Klett, M. The economics of CO_2 storage. MIT Laboratory for Energy and the Environment Report 2003, MIT-LFEE-2003-002, 1–120.
27. Herzog, H. Scaling up carbon dioxide capture and storage: From megatons to gigatons. Energy Economics 2011, 33(4), 597–604.
28. IEC 61511. Functional Safety - Safety Instrumented Systems for the Process Industry Sector; International Electrotechnical Commission: Geneva, Switzerland, 2016.

29. ISO 31000. Risk Management - Guidelines; International Organization for Standardization: Geneva, Switzerland, 2018.
30. Kiefner, J. F.; Maxey, W. A. Running fracture behavior in high-pressure pipelines. Journal of Pressure Vessel Technology 2008, 130(4), 041701.
31. Klever, F. J.; Stewart, G. Reliability-based pipeline design and assessment. Marine Structures 2016, 49, 1–20.
32. Koornneef, J.; van Keulen, T.; Faaij, A.; Turkenburg, W. Life cycle assessment of CCS systems. International Journal of Greenhouse Gas Control 2008, 2(4), 448–467.
33. Leis, B. N. Fracture mechanics for CO2 pipelines. Proceedings of IPC Conference 2012, IPC2012-90243, 1–12.
34. Li, H.; Jakobsen, J.; Wilhelmsen, Ø.; Yan, J. Impurity effects in CO_2 transport for CCS. Energy Procedia 2011, 4, 3060–3067.
35. Line, O.; Rostrup, S. Material selection for CO_2 pipelines. Corrosion Engineering, Science and Technology 2015, 50(7), 489–498.
36. Liu, M.; Shuai, J.; Wang, J.; Xu, K. Quantitative risk assessment for pipeline leakage accidents. Process Safety and Environmental Protection 2019, 123, 232–241.
37. Mahgerefteh, H.; Brown, S.; Denton, G. Modeling decompression and fracture propagation in CO_2 pipelines. Chemical Engineering Science 2012, 74, 200–210.
38. Melchers, R. E. Structural Reliability Analysis and Prediction; Wiley: Chichester, UK, 2001.
39. Mohitpour, M.; Golshan, H.; Murray, A. Pipeline Design and Construction: A Practical Approach; ASME Press: New York, NY, USA, 2007.
40. Nešić, S.; Sun, W. Corrosion in oil and gas production systems. Corrosion 2010, 66(5), 055001.
41. Nessim, M.; Stephens, M.; Zhou, W. Pipeline risk management methodologies. Proceedings of IPC Conference 2009, IPC2009-72144, 1–9.
42. Ovidi, O.; Davidson, J. Pipeline digital twin technologies for integrity management. Journal of Natural Gas Science and Engineering 2020, 81, 103434.
43. Papavinasam, S. Corrosion Control in the Oil and Gas Industry; Elsevier: Amsterdam, Netherlands, 2014.
44. Park, J.; Kim, D.; Lee, S. AI-assisted corrosion prediction in CO_2 pipelines. Corrosion Engineering, Science and Technology 2021, 56(6), 523–534.
45. Race, J. M.; Seevam, P. N.; Downie, M. J. Challenges for the design of CO_2 pipelines. Pipeline & Gas Journal 2007, 234(12), 44–51.
46. Svensson, R.; Odenberger, M.; Filipsson, S. Transportation systems for CO_2. Energy Policy 2004, 32(11), 1301–1310.
47. Seevam, P. N.; Race, J. M.; Downie, M. J. Transporting the captured carbon dioxide. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy 2008, 222(5), 453–465.
48. Stephens, M. J. Risk-based inspection and integrity management. Proceedings of ASME Pressure Vessels and Piping Conference 2006, PVP2006-93108, 1–8.
49. Veritas, D. N. Recommended Practice for Corroded Pipelines; DNV-RP-F101: Høvik, Norway, 2010.
50. Wilkowski, G.; Maxey, W.; Eiber, R. Fracture propagation studies in gas transmission pipelines. Battelle Memorial Institute Report 2000, BMI-2194, 1–85.

51. Zhou, W.; Huang, G. Risk assessment methodologies for pipelines: A review. Journal of Loss Prevention in the Process Industries 2012, 25(3), 507–516.
52. Zhang, L.; Wang, Y.; Cheng, Y. F. Corrosion mechanisms in dense-phase CO_2 transportation pipelines. Corrosion Science 2017, 123, 36–48.
53. Zhang, Z.; Mahgerefteh, H.; Brown, S. Thermodynamic modeling of CO_2 decompression. Chemical Engineering Science 2007, 62(21), 5898–5908.
54. Zhao, Y.; Teixeira, A. P. Reliability analysis and maintenance optimization for pipelines. Reliability Engineering & System Safety 2018, 169, 434–443.
55. Choi, Y.-S.; Nešić, S. Corrosion behavior in supercritical CO_2 systems. Corrosion 2011, 67(9), 095001.
56. Chandel, S.; Saini, R. P. Reliability-centered maintenance and risk-based inspection approaches in energy pipelines. Journal of Loss Prevention in the Process Industries 2021, 69, 104380.
57. Brennan, F. P.; Zargarzadeh, P. Structural integrity assessment of CO_2 pipelines. Engineering Fracture Mechanics 2015, 148, 1–15.
58. Biezma, M. V.; Agudo, D.; San Cristóbal, J. R. Probabilistic modeling for pipeline integrity management. Engineering Failure Analysis 2020, 115, 104617.

RiESTech
Recent in Engineering Science and Technology

MBI

Article

Preliminary LEAP-Based Scenario Analysis of Electricity Demand and Renewable Energy Development in South Sulawesi

Arham [1], La Ode Masri Ande Kolewora [1,*], Saiful M [1], Ahmad Agus Setiawan [2], Muhammad Taufiq [1], and Virtuous Ikmal Wicaksono [1]

[1] Master's Program in Systems Engineering, Universitas Gadjah Mada, Indonesia
[2] Department of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada, Indonesia
* Correspondence: la.ode.masri.ande@mail.ugm.ac.id

Abstract: South Sulawesi possesses abundant renewable energy potential, particularly from hydro, wind, solar, and biomass resources. However, rapid population growth and increasing energy consumption across sectors create challenges in ensuring sustainable energy supply and planning. This study aims to project the mid-term renewable energy supply and demand in South Sulawesi using the Long-range Energy Alternatives Planning (LEAP) model. By simulating various scenarios—including the business-as-usual (BAU) and renewable policy intervention scenarios—this research analyzes the region's future energy trajectory up to 2035. The findings reveal that South Sulawesi's average population growth of 1.028% (approximately 142,000 people per year) and electricity demand growth of 7.94% annually significantly increase regional energy needs. Without intervention, fossil fuels will continue to dominate the energy mix, but policy-driven scenarios demonstrate that renewable energy contributions could exceed 45% of the total supply. The study provides valuable insights for regional energy planning and supports policy formulation toward achieving a sustainable, low-carbon energy system in South Sulawesi.

Keywords: LEAP, renewable energy, energy forecasting, South Sulawesi, sustainable development.

Citation: Arham., Kolewora, L. M. A., Saiful, M., Setiawan, A. A., Taufiq, M., Wicaksono, V. I. (2026). Preliminary LEAP-Based Scenario Analysis of Electricity Demand and Renewable Energy Development in South Sulawesi. Recent in Engineering Science and Technology, 4(4), 209–223. Retrieved from https://www.mbi-jounals.com/index.php/riestech/article/view/165.

Academic Editor: Yudhi Ariadi

Received: 8 June 2026
Accepted: 16 October 2026
Published: 31 October 2026

Publisher's Note: MBI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2026 by the authors. Licensee MBI, Jakarta, Indonesia. This article is an open access article distributed under MBI license (https://mbi-journals.com/licenses/by/4.0/).

1. Introduction

The increasing population and economic growth in South Sulawesi have led to a continuous rise in energy demand across all sectors. Based on data from the *Badan Pusat Statistik Sulawesi Selatan* (South Sulawesi Statistics Agency), the province's population has grown at an average rate of 1.028% per year, equivalent to approximately 142,000 additional people annually. This demographic growth contributes directly to higher household and industrial energy consumption, accelerating the region's overall energy requirements.

Electricity consumption in South Sulawesi has shown a steady upward trend over the past decade. The household sector remains the largest consumer, followed by business and public sectors, while industrial consumption—although smaller—continues to increase steadily. According to the RUPTL 2025–2034, electricity sales in South Sulawesi grew at an average annual rate of 7.94% between 2015 and 2024, increasing from 4,480 GWh in 2015 to 9,225 GWh in 2024. These figures represent electricity sales across the household, public, business, and industrial sectors, reflecting the region's growing electricity demand [1].

Despite these increases in demand, South Sulawesi holds substantial renewable energy potential that remains underutilized. The province's *Regional Energy General Plan (RUED-P Sulsel 2015–2050)* identifies various renewable sources: hydroelectric power plants (2,946.8 MW), wind power (4,193 MWe), solar PV (7,588 MWp), micro-hydro

(7,622.9 MW), and geothermal resources (800 MW). However, only a small portion of these potentials has been developed to date. The full realization of these resources could play a vital role in achieving regional and national renewable energy targets.

Given this context, developing a long-term and data-driven energy planning model is essential. The *Long-range Energy Alternatives Planning* (LEAP) system provides a comprehensive analytical framework for evaluating future energy scenarios, considering population growth, economic development, technology adoption, and policy measures. By applying the LEAP model, this study aims to project the mid-term supply and demand of renewable energy in South Sulawesi and assess the implications of various policy scenarios [2], [3].

Ultimately, the results are expected to assist policymakers, local governments, and stakeholders in designing effective strategies to accelerate the energy transition, reduce dependence on fossil fuels, and achieve a more sustainable and resilient energy system for the province[4], [5].

1. Materials and Experiment Methods

2.1 Power Sector Structure

The power sector in South Sulawesi plays a strategic role in supporting the island's economic development and industrial growth. It forms part of the interconnected Sulawesi Power System, managed primarily by PT PLN (Persero), Indonesia's state-owned electricity company. The system operates as part of the Sulbagsel (South, Southeast, and West Sulawesi) grid, ensuring electricity transmission and distribution across provinces through high-voltage networks and substations. Electricity generation in South Sulawesi is characterized by a diverse mix of energy sources, with thermal power plants particularly those using coal and diesel still dominating the generation capacity. However, the province has made significant progress in integrating renewable energy sources into its power mix, driven by national energy transition targets and regional initiatives. As of recent data, hydropower (PLTA) and wind power (PLTB) represent the largest contributors among renewables, supported by several operational plants in Sidrap, Jeneponto, and Tana Toraja.

South Sulawesi's electricity system consists of generation, transmission, and distribution subsystems. The generation segment includes both PLN-owned and Independent Power Producer (IPP) facilities, comprising hydropower plants, coal-fired power plants, diesel generators, and renewable installations such as wind and solar farms. The transmission network, operated under PLN's regional units, connects major load centers including Makassar, Parepare, Palopo, and Bone, forming the backbone of electricity delivery across the province. The distribution network serves urban and rural areas, supporting residential, business, and industrial customers with varying reliability and capacity.

According to the RUPTL 2025–2034, the South Sulawesi power system is projected to experience continued demand growth of approximately 7–8% per year, primarily driven by population increase, economic activity, and industrial expansion. The household sector remains the largest electricity consumer, accounting for over 50% of total demand, followed by the business and public sectors. Meanwhile, the industrial sector though smaller in share is expected to grow steadily as the province develops new manufacturing and processing zones.

To address the growing demand and reduce dependence on fossil fuels, the government and PLN are promoting the expansion of renewable energy generation, particularly through the development of additional hydropower, solar photovoltaic

(PLTS), and wind power projects. These initiatives are supported by the provincial Regional Energy General Plan (RUED-P Sulsel), which emphasizes diversification, energy security, and environmental sustainability. In summary, the power sector structure of South Sulawesi reflects a transitional system—moving from fossil fuel dependency toward a more balanced, renewable-based energy mix. Continued investment in grid infrastructure, renewable integration, and energy efficiency will be essential to ensure sustainable and reliable electricity supply for the province's long-term development[6].

2.2 Analysis of Electricity Demand

Electricity demand in South Sulawesi has shown a consistent upward trend over the period 2015–2024. The total number of customers increased from 4.48 million in 2015 to an estimated 9.23 million in 2024, representing an average annual growth rate of approximately 8%. This steady rise reflects both population growth and economic expansion across various consumer sectors.

The household sector remains the largest consumer group, growing from 2.32 million customers in 2015 to 3.61 million in 2024. This growth indicates the continuous electrification of residential areas and an increase in household electricity usage driven by lifestyle changes and appliance adoption[7].

The business sector also shows a positive trend, increasing from 0.93 million customers in 2015 to 1.33 million in 2024, highlighting the expansion of small and medium enterprises in the region. Meanwhile, the public sector experienced moderate growth, from 0.41 million to 0.67 million customers, in line with the development of public facilities and government infrastructure[8], [9].

The most notable increase occurred in the industrial sector, which rose significantly from 0.83 million customers in 2015 to 3.62 million in 2024. This sharp growth, especially between 2020 and 2023, suggests a surge in industrial activity and investment, possibly linked to regional development programs and the establishment of industrial zones in South Sulawesi.

Annual growth in total electricity demand fluctuated between 0.5% and 19.1%, with the highest spike in 2022, reflecting post-pandemic economic recovery and industrial resurgence. Overall, the data indicate robust demand growth across all sectors, reinforcing the need for strategic planning in electricity generation and distribution to sustain economic development in South Sulawesi.

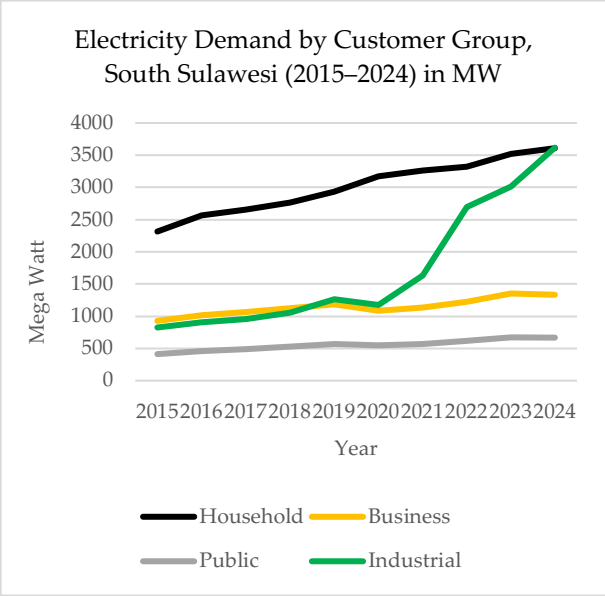


Figure 1. Electricity Demand by Year

2. 3 Analysis of Electricity Supply

In 2025, the total electricity production in South Sulawesi is projected to reach 9,798.53 MWh, generated from a mix of fossil fuels and renewable energy sources. The data indicate that coal remains the dominant source of power generation, contributing approximately 5,864.52 MWh, or about 60% of the total production. This heavy reliance on coal reflects the region's continued dependence on conventional energy sources to meet growing electricity demand.

Among renewable sources, hydropower plays a significant role, producing around 1,498.26 MWh (15% of total output). Smaller-scale renewables such as minihydro (370.51 MWh) and microhydro (4.31 MWh) complement the larger hydro facilities, reflecting ongoing efforts to utilize local water resources for power generation[10].

Wind energy contributes 465.39 MWh, showing potential for further development given South Sulawesi's favorable wind conditions, particularly in coastal and highland areas. In contrast, solar power remains minimal, producing only 2.38 MWh, suggesting that solar energy infrastructure is still in its early stages.

On the fossil fuel side, gas and steam plants produce 1,260.39 MWh, while diesel and gas-only generation contribute 250.83 MWh and 81.93 MWh, respectively. The relatively small share of oil-based generation indicates a gradual transition toward cleaner and more efficient technologies.

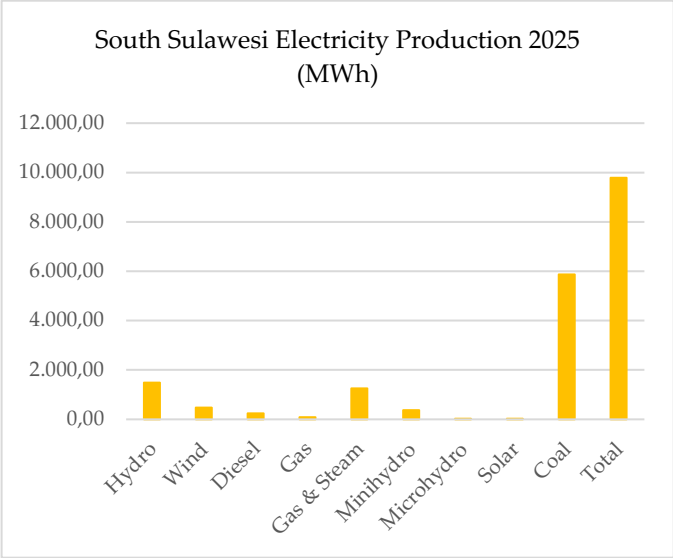


Figure 2. Electricity Production in South Sulawesi

2. 4 Methodology

2.4.1 South Sulawesi's LEAP Model Framework

The LEAP (Long-range Energy Alternatives Planning) model framework developed for South Sulawesi provides a systematic approach to analyze and forecast the province's future energy demand, supply, and emissions under different development scenarios. This framework integrates both bottom-up demand modeling and top-down supply analysis, allowing for a comprehensive assessment of the region's energy system over time[11], [12].

The framework begins by structuring the demand side according to key consumer sectors—household, business, public, and industrial—based on historical electricity consumption data from 2015 to 2024. Each sector's growth trend is modeled using socioeconomic indicators such as population growth, regional GDP, and industrial

expansion rates. These factors serve as the primary drivers of electricity demand in the projection period up to 2040 (or as defined in the study).

On the supply side, the LEAP framework incorporates data on electricity generation by energy source, including coal, hydro, wind, diesel, gas, steam, minihydro, microhydro, and solar, as observed in 2025. These inputs form the baseline scenario, representing the existing energy mix and technological structure. The framework then allows for the development of alternative scenarios—such as renewable energy expansion, energy efficiency improvement, or decarbonization pathways—to evaluate their long-term impacts on production, cost, and emissions[13], [14]. The model structure follows LEAP's hierarchical logic:

1. Current Accounts – representing the base year (2025) energy situation of South Sulawesi.
2. Demand Modules – projecting energy use by sector and subsector.
3. Transformation Modules – modeling electricity generation and fuel conversion processes.
4. Resources and Emissions Modules – estimating fuel inputs, resource availability, and environmental outputs (CO_2 and other emissions).

Through this framework, the LEAP model serves as a decision-support tool for policymakers and planners in South Sulawesi, enabling them to explore different energy development strategies and assess their implications for energy security, sustainability, and emission reduction.

Table 1. Economic Growth Development

Year	Quarter	Growth (%)
2021	Q1	-0.21
2021	Q2	7.70
2021	Q3	3.24
2021	Q4	7.89
2022	Q1	4.31
2022	Q2	5.22
2022	Q3	5.69
2022	Q4	5.12
2023	Q1	5.30
2023	Q2	5.00
2023	Q3	4.04
2023	Q4	3.82
2024	Q1	4.82
2024	Q2	4.98
2024	Q3	5.08
2024	Q4	5.18
2025	Q1	5.78

The LEAP (Long-range Energy Alternatives Planning) model for South Sulawesi is developed based on a series of key assumptions that define the structure and behavior of both the energy demand and supply systems over the modeling period. These assumptions are essential for ensuring that the model reflects realistic socioeconomic, technological, and policy conditions within the province[15].

On the demand side, the model assumes that future electricity consumption will be driven by the growth of four major consumer sectors: household, business, public, and industrial. The projection of each sector is influenced by macroeconomic and demographic factors such as population growth, gross regional domestic product

(GRDP), and sectoral economic expansion. Economic growth is assumed to remain stable in the medium term, averaging around 5–6 percent annually, based on recent regional trends. Population growth is projected to continue moderately, leading to increased household electricity demand. Energy intensity (energy use per unit of economic output) is assumed to gradually decline due to efficiency improvements and technological adoption[16].

On the supply side, the baseline scenario assumes that the existing energy mix observed in 2025 dominated by coal and supported by hydro, wind, and other renewables will serve as the reference structure. The capacity of renewable energy sources such as hydro, minihydro, wind, and solar is expected to expand gradually in line with national energy transition policies and regional renewable potential. Meanwhile, fossil fuel-based generation (coal, diesel, and gas) is assumed to maintain a significant share in the short term but gradually decrease under alternative or policy-driven scenarios[17].

Additional assumptions include the use of current technology efficiencies, plant lifetimes, and capacity factors based on available data from PLN and regional planning documents. The model also assumes that no major disruptions—such as extreme policy shifts, large-scale energy crises, or drastic demographic changes—will occur during the simulation period. Carbon emission factors are based on standard IPCC values, and fuel price assumptions follow national energy planning parameters (RUEN or RUPTL data).

Overall, these assumptions provide the analytical foundation for simulating different development pathways—such as business-as-usual, renewable energy expansion, or low-carbon scenarios—allowing the LEAP model to assess their impacts on energy demand, supply structure, and emission outcomes in South Sulawesi.

While the LEAP (Long-range Energy Alternatives Planning) model provides a flexible and comprehensive framework for analyzing South Sulawesi's energy system, several limitations should be acknowledged to ensure a balanced interpretation of the results.

First, the model's accuracy is highly dependent on the quality and availability of input data. In the case of South Sulawesi, historical energy consumption and production data may contain inconsistencies or incomplete records, particularly for smaller renewable energy systems such as microhydro and solar. Limited access to up-to-date regional statistics can introduce uncertainty into demand projections and energy balance estimations.

Second, LEAP operates primarily as a deterministic model, meaning that it does not automatically account for uncertainties in future socioeconomic or technological developments. Variables such as economic growth, population change, fuel prices, and technology costs are treated as fixed assumptions within each scenario. Consequently, sudden economic shocks, policy changes, or external factors (for example, global energy price fluctuations) cannot be dynamically captured without manual scenario adjustments.

Third, the spatial dimension of energy planning is not explicitly represented in LEAP. The model focuses on aggregate regional data and does not simulate geographic variations in energy supply or distribution infrastructure within South Sulawesi. As a result, it cannot fully evaluate localized grid constraints, transmission losses, or site-specific renewable resource availability.

Fourth, LEAP's environmental module, while useful for estimating greenhouse gas emissions, relies on default emission factors that may not perfectly reflect local fuel characteristics or technology performance. This can lead to small discrepancies between modeled and actual emission outcomes.

Finally, the model does not perform economic optimization or cost-benefit analysis automatically. It estimates energy requirements and emissions based on user-defined pathways rather than identifying the least-cost or most efficient solutions. This limits its capability to determine the optimal investment portfolio without integration with other analytical tools, such as MARKAL or TIMES[18], [19].

In summary, while LEAP offers valuable insights for scenario analysis and long-term energy planning, its results should be interpreted with caution. The model's outputs are most effective when used as guidelines for policy discussion and comparative scenario evaluation, rather than as precise forecasts or investment recommendations.

1.4.1 Scenario Development

In this study, two main scenarios were developed within the LEAP model to evaluate the future electricity outlook in South Sulawesi: the RUPTL Scenario and the National Target (RUEN) Scenario. The RUPTL Scenario (Rencana Usaha Penyediaan Tenaga Listrik) reflects the official electricity supply plan of PLN, emphasizing a realistic projection based on current infrastructure, investment plans, and committed generation projects. It assumes gradual growth in electricity demand consistent with regional economic development, with the energy mix still dominated by coal and hydropower, while renewable penetration increases moderately following PLN's 2021–2030 planning targets[20], [21].

The National Target (RUEN) Scenario (Rencana Umum Energi Nasional) represents a more ambitious pathway aligned with Indonesia's national energy policy, which aims to achieve 23% renewable energy share by 2025 and 31% by 2050. This scenario emphasizes stronger integration of renewables such as solar, wind, and hydro, along with energy efficiency improvements and gradual reduction of fossil fuel dependency[22].

2. Results and Discussion

3.1. Demand Projection

The projected final energy demand in South Sulawesi shows a consistent upward trend from 2024 to 2034. Total electricity consumption increases from approximately 9 TWh in 2024 to around 12.5 TWh by 2034, reflecting steady economic growth and rising energy needs across all sectors.

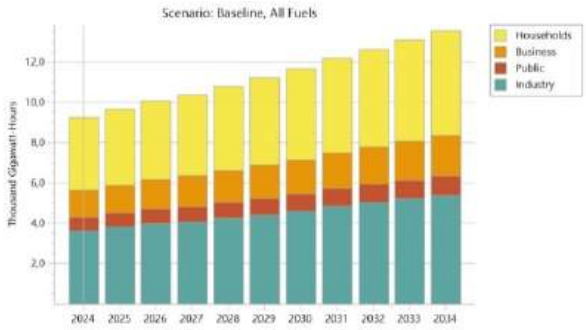


Figure 3. Demand Projections in South Sulawesi

Among the four main sectors households, business, public, and industry the industrial sector remains the largest energy consumer throughout the projection period. Industrial electricity demand grows significantly, indicating continued expansion in manufacturing and processing activities within the province[23], [24]. The household sector also contributes a substantial share of total demand, rising steadily as electrification rates improve and living standards increase. The business sector shows moderate growth,

corresponding to the development of commercial areas and small enterprises. Meanwhile, the public sector which includes government buildings and public services—maintains a smaller but gradually increasing share[25], [26]. Overall, the projection suggests that total electricity demand will continue to grow at an average annual rate of around 3–4%, aligning with South Sulawesi's expected economic development and population growth. This rising demand underscores the importance of expanding generation capacity, promoting energy efficiency, and accelerating renewable energy integration to ensure a reliable and sustainable energy supply in the coming decade.

3.2 Electricity Generations Projections

3.2.1 Comparison Based on Installaed Capacity

The comparison between the RUPTL Scenario and the RUEN Target Scenario highlights different pathways for electricity generation capacity development in South Sulawesi from 2024 to 2034. Under the RUPTL Scenario, total installed capacity shows steady growth from around 3.3 GW in 2024 to approximately 5.1 GW in 2034. The generation mix remains dominated by coal-fired power plants (PLTU) and large hydropower (PLTA), while renewable sources such as mini-hydro (PLTM), solar (PLTS), and wind (PLTB) increase only moderately. This scenario reflects a more conservative expansion approach, focusing on reliability and gradual integration of renewables in alignment with current PLN planning documents.

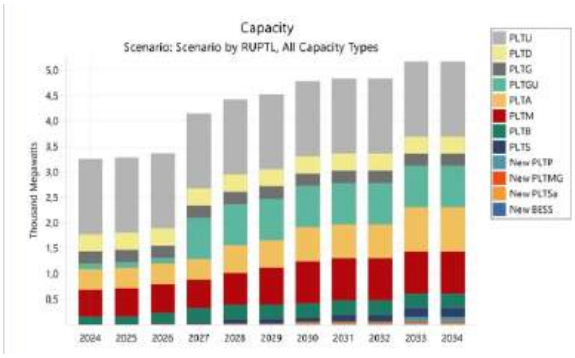


Figure 4. RUPTL Scenario

In contrast, the RUEN Target Scenario (National Energy Policy Plan) projects a stronger commitment to renewable energy integration. Although total installed capacity reaches a similar level of around 5 GW, the composition differs significantly. The share of renewable energy plants including solar (PLTS), geothermal (PLTP), and mini-hydro (PLTM) expands considerably, while dependence on coal (PLTU) gradually decreases. This reflects national efforts to achieve the renewable energy mix target of 23% by 2025 and beyond, in line with Indonesia's clean energy transition goals[27].

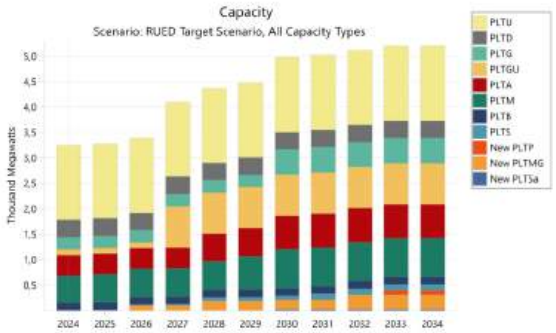


Figure 5. RUED Target Scenario

3.2.2 Comparison Based on Output

Under the RUPTL scenario, the generation mix remains dominated by coal-fired power plants (PLTU), as represented by the large grey portion of each bar in the chart. Despite moderate growth in renewable energy sources such as hydropower (PLTA), mini-hydro (PLTM), biomass (PLTSa), and solar power (PLTS), coal continues to contribute the largest share of electricity generation throughout the projection period[28], [29].

Diesel (PLTD) and gas turbine (PLTG) plants also maintain a noticeable presence, indicating limited diversification of the generation portfolio. This scenario represents a business-as-usual pathway, where energy security and capacity expansion are prioritized over the rapid integration of renewables. Consequently, the RUPTL pathway implies continued reliance on fossil fuels, leading to higher carbon emissions and slower progress toward Indonesia's clean energy targets[30], [31].

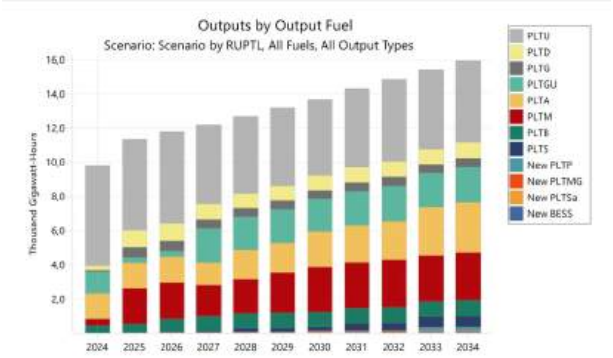


Figure 6. RUPTL Scenario Based on Output

In contrast, the RUED target scenario envisions a significant transformation of the energy mix toward renewable sources. While the total electricity generation output follows a similar upward trend, the contribution of PLTU (coal), shown in yellow, gradually declines relative to other technologies[32], [33].

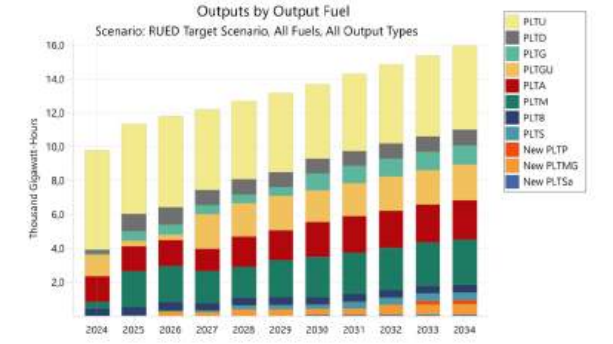


Figure 7. RUED Scenario Based on Output

Renewable energy sources such as PLTA, PLTM, PLTB, and PLTS experience substantial growth, complemented by the introduction of new plants including New PLTP (geothermal) and New PLTS (solar). This indicates a strategic shift toward low-carbon and sustainable energy systems, in line with Indonesia's national commitment to increase the share of renewable energy in its primary energy mix and reduce greenhouse gas emissions[34], [35].

3.3 Renewable Energy Projections

Projected capacity of renewable energy power plants from 2024 to 2034, measured in megawatts (MW). The capacity includes several renewable technologies, namely hydropower (PLTA), mini-hydro (PLTM), biomass (PLTSa), solar power (PLTS), geothermal power (New PLTP), and new solar installations (New PLTS).

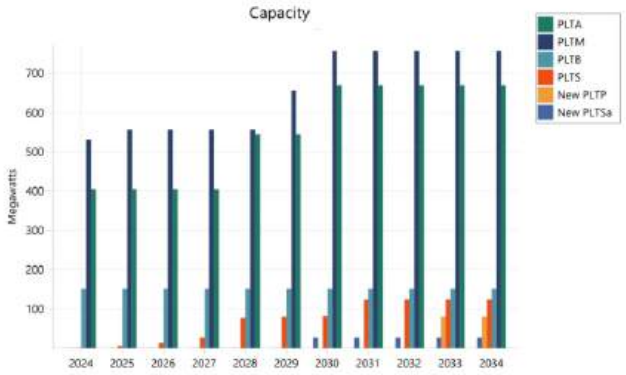


Figure 8. Renewable Energy Projection

Throughout the projection period, the total installed renewable capacity exhibits a steady upward trend. The most dominant contributors to renewable generation capacity are PLTM and PLTA. PLTM maintains the largest share from the beginning of the period, starting at around 530 MW in 2024 and increasing to approximately 750 MW by 2034. PLTA also shows significant growth, rising from about 400 MW in 2024 to more than 600 MW by 2034.

Biomass (PLTSa) and solar (PLTS) capacities gradually expand over time, indicating continuous development and diversification of renewable sources. PLTS capacity grows noticeably starting in 2026 and continues to rise steadily through 2034. In addition, the appearance of new geothermal (New PLTP) and solar (New PLTS) projects from around 2027 onwards marks an important milestone in the renewable expansion plan[2]. Overall, the figure reflects a strong commitment to renewable energy development. The increasing capacity across multiple renewable technologies suggests that future energy policies prioritize sustainability, diversification, and carbon emission reduction. The integration of new geothermal and solar power plants also indicates an effort to enhance system reliability while supporting Indonesia's national renewable energy targets for 2030 and beyond[36], [37].

3.4 Supply and Demand Balance

The balance of energy supply and demand across various energy sources indicates that the total primary energy supply reaches 36,207 MWh, mainly sourced from imports of solid fuels (13,348 MWh) and natural gas (9,118 MWh), along with domestic generation from hydropower (6,503 MWh) and other renewables (5,132 MWh). During the transformation phase, these primary energy sources are converted into secondary energy, predominantly electricity. Negative values in the transformation section for solid fuels, natural gas, hydropower, and renewables indicate their use in electricity generation, which produces 15,953 MWh of electricity. After accounting for transformation and distribution losses, the total transformed energy is −22,649 MWh. [38].

	Solid Fuels	Natural Gas	Hydropower	Renewables	Electricity	Oil Products	Total
Production	-	-	6,503	5,132	-	-	11,635
Imports	13,348	9,118	-	-	0,000	2,108	24,574
Exports	-	-	-	-	-0,003	-	-0,003
Total Primary Supply	13,348	9,118	6,503	5,132	-0,003	2,108	36,207
Electrical Generation	-13,348	-9,118	-6,503	-5,132	15,953	-2,108	-20,257
Transformation and Distribution	-	-	-	-	-2,392	-	-2,392
Total Transformation	-13,348	-9,118	-6,503	-5,132	13,560	-2,108	-22,649
Households	-	-	-	-	5,215	-	5,215
Business	-	-	-	-	2,001	-	2,001
Public	-	-	-	-	0,958	-	0,958
Industry	-	-	-	-	5,383	-	5,383
Total Final Energy Demand	-	-	-	-	13,557	-	13,557
Unmet Requirements	-	-	-	-	-	-	-

Figure 9. Supply and Demand Balance

Final energy demand totals 13,557 MWh, distributed among households (5,215 MWh), industry (5,383 MWh), business (2,001 MWh), and the public sector (0.958 MWh). These figures represent final energy consumption across different sectors[39], [40]. Overall, the energy system demonstrates equilibrium between total supply and final demand, with no unmet requirements observed, indicating that the available energy resources and conversion capacity are sufficient to satisfy the projected consumption needs.

3.5 Renewable Cost Data

The data compares key technical and economic parameters of various power plant types, including fossil-based and renewable technologies, for the years 2018 and 2040. Parameters include capital cost, operation and maintenance (O&M) cost, efficiency, lifetime, fuel cost, and maximum capacity factor.

Power plant type	Capital costs (USD/kW) in 2018/2040	Fixed O&M costs (USD/kW/year) in 2018/2040	Variable O&M costs (USD/MWh) in 2018/2040	Efficiency	Lifetime (years)	Fuel costs (USD/GJ)	Maximum capacity factor
Coal	1400/1340 [A,B]	41.2/39.3 [A,B]	0.12/0.12 [A,B]	37% [A,B]	30 [A]	2.4 [G]	80% [A,B]
Open cycle gas turbine (OCGT)	770/700 [A,B]	23/22.5 [A,B]	0 [A,B]	33% [A,B]	25 [A]	6.2 [E]	95% [A,B]
Closed cycle gas turbine (CCGT)	750/680 [A,B]	23/22.5 [A,B]	0.13/0.13 [A,B]	56% [A,B]	25 [A]	6.2 [E]	95% [A,B]
Large hydro	2000/2000 [A,B,C,D]	38/34.7 [A,B]	0.65/0.6 [A,B]	–	50 [A]	–	36% [H]
Small hydro	2600/2600 [A]	53/48.8 [A]	0.5/0.46 [A]	–	50 [A]	–	50% [C]
Geothermal	3700/3050 [A,B,C,D]	18/15.9 [A,B]	0.25/0.22 [A,B]	–	30 [A]	–	80% [A,B]
Biomass	1700/1500 [A,B,C,D]	48/40.8 [A,B]	3/2.59 [A,B]	31% [A]	25 [A]	1.3 [F]	80% [A,B]
Solar PV	1190/BAS: 610 RE:340 [C,D]	11.9/BAS: 6.1 RE: 3.4 [C,D]	0 [A,B]	–	25 [A]	–	17% [K]
Wind	1500/BAS: 1210 RE: 1020 [A,B,C,D]	30/26.5 [A,B]	0 [A,B]	–	27 [A]	–	25% [I,J]
Diesel	800/800 [A,B]	8/8 [A,B]	6.4/6.4 [A,B]	45% [A]	25 [A]	15.9 [C]	80% [A,B]
Pumped hydro storage	860/860 [A]	8/8 [A]	1.3/1.3 [A]	80% [A]	35 [A]	–	100%
Battery storage	1570/730 [D]	39.3/18.2 [D]	0 [D]	85% [D]	15 [D]	–	100%

Figure 10. Cost Analysis

Coal power plants have a capital cost of 1,400 USD/kW in 2018, projected to decline slightly to 1,340 USD/kW by 2040. They maintain relatively low O&M costs but exhibit moderate efficiency at 37% and a high fuel cost of 2.4 USD/GJ, with an operational lifetime of 30 years. Gas-based plants, including open-cycle (OCGT) and closed-cycle (CCGT) turbines, show lower capital costs (700–770 USD/kW) but have higher fuel costs at 6.2 USD/GJ. CCGT systems achieve higher efficiency (56%) compared to OCGT (33%).

Hydropower systems, both large and small, demonstrate long operational lifetimes (up to 50 years) and low variable O&M costs, though their capacity factors vary between 36% and 50%. Geothermal and biomass plants have relatively high capital costs (3,700 and 1,700 USD/kW, respectively) but offer stable output and moderate fuel costs.

Solar PV and wind technologies show substantial cost reductions by 2040, with solar capital costs decreasing from 1,190 to around 610 USD/kW and wind from 1,500 to approximately 1,020 USD/kW. However, both technologies exhibit lower capacity factors—17% for solar and 25% for wind—due to their intermittent nature. Diesel generation remains relatively costly in terms of fuel (15.9 USD/GJ) and less efficient (45%), making it less favorable for large-scale or long-term applications. In contrast, energy

storage systems, such as pumped hydro and battery storage, present increasingly competitive options for supporting grid reliability. Pumped hydro storage maintains a long lifespan (35 years) and full availability (100% capacity factor), while battery storage achieves high efficiency (85%) and continues to experience rapid cost declines, from 1,570 USD/kW in 2018 to 730 USD/kW by 2040.

3. Conclusion

This study provides a mid-term projection of South Sulawesi's renewable energy supply and demand using the Long-range Energy Alternatives Planning (LEAP) model. The simulation results reveal a steady increase in electricity demand, rising from approximately 9 TWh in 2024 to 12.5 TWh by 2034, driven by population growth, industrial expansion, and economic development. The industrial and household sectors remain the primary contributors to the overall energy demand throughout the projection period.

The comparison between the RUPTL and RUEN scenarios highlights two distinct development pathways. The RUPTL scenario reflects a business-as-usual approach, maintaining high dependence on coal-fired power plants, while renewable energy integration progresses gradually. In contrast, the RUEN scenario demonstrates a more ambitious trajectory toward renewable energy expansion, particularly in hydropower, mini-hydro, solar, biomass, and geothermal generation. Under this policy-driven scenario, the renewable energy share could exceed 45% of total electricity production by 2034.

The renewable energy capacity projection indicates continuous growth across multiple technologies. Hydropower and mini-hydro remain the dominant contributors, followed by the gradual development of solar and geothermal plants. This diversification reflects a growing policy commitment to sustainability, energy security, and carbon emission reduction. The supply-demand balance analysis confirms that South Sulawesi's projected energy system is capable of meeting future electricity needs without major deficits, provided that planned capacity expansions are implemented.

From an economic perspective, renewable technologies show significant cost reduction potential by 2040, especially for solar PV, wind, and battery storage, making them increasingly competitive with fossil-based power plants. These findings underscore the importance of accelerating renewable energy investment, enhancing grid infrastructure, and integrating storage technologies to support a reliable and low-carbon energy transition.

In conclusion, the results emphasize the necessity of aligning regional energy planning with national policy targets to ensure sustainable development in South Sulawesi. Strengthening institutional coordination, promoting technological innovation, and expanding renewable deployment will be essential strategies to achieve a balanced, resilient, and environmentally sustainable power system by 2035.

Acknowledgment: The authors would like to express their gratitude to the Department of Systems Engineering, Faculty of Engineering, Universitas Gadjah Mada, for providing academic guidance and research support throughout this study. All authors are final-year master's students in the Systems Engineering Program, specializing in Renewable Energy Systems. This research was conducted as part of the academic requirements for completing the Master of Engineering degree and reflects the authors' shared commitment to advancing sustainable energy development in Indonesia.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] PT PLN (Persero), Rencana Usaha Penyediaan Tenaga Listrik (RUPTL) 2025–2034. Jakarta, Indonesia: PT PLN (Persero), 2025.

[2] V. Cattaneo et al., "Forecasting silver demand and supply by 2030: Impact of silver-intensive photovoltaic cells and sectoral competition," Resour. Conserv. Recycl., vol. 224, no. June 2025, p. 108562, 2026, doi: 10.1016/j.resconrec.2025.108562.

[3] H. Choi, J. Kim, S. Lee, M. S. Lee, and J. Kim, "Reinforcement learning-based framework for integrated green methanol supply–demand management with demand forecasting under renewable energy uncertainty," Energy Convers. Manag., vol. 346, no. June, p. 120511, 2025, doi: 10.1016/j.enconman.2025.120511.

[4] M. Aldarraji, B. Vega-Márquez, B. Pontes, B. Mahmood, and J. C. Riquelme, "Addressing energy challenges in Iraq: Forecasting power supply and demand using artificial intelligence models," Heliyon, vol. 10, no. 4, p. e25821, 2024, doi: 10.1016/j.heliyon.2024.e25821.

[5] Q. V. Dinh, V. N. Dinh, and P. G. Leahy, "A diffusion model approach to forecast multi-sector demand growth for green hydrogen generated from offshore wind power," Renew. Sustain. Energy Rev., vol. 223, no. July, p. 116038, 2025, doi: 10.1016/j.rser.2025.116038.

[6] W. Zhou, J. Chang, W. Zuo, and F. Wang, "A seasonal grey model for forecasting energy imports demand from information differences perspective," Appl. Math. Model., vol. 140, no. August 2024, p. 115907, 2025, doi: 10.1016/j.apm.2024.115907.

[7] S. Ahmed, R. K. Chakrabortty, and D. L. Essam, "A multi-source multi-layer-based transfer learning approach for forecasting customer demands of newly launched products," Intell. Syst. with Appl., vol. 27, no. October 2024, p. 200548, 2025, doi: 10.1016/j.iswa.2025.200548.

[8] H. J. Lee, K. Kameda, S. Manzhos, and M. Ihara, "A novel encoding method for high-dimensional categorical data for electricity demand forecasting in distributed energy systems," Appl. Energy, vol. 392, no. November 2024, p. 125989, 2025, doi: 10.1016/j.apenergy.2025.125989.

[9] S. Smyl, B. N. Oreshkin, P. Pełka, and G. Dudek, "Any-quantile probabilistic forecasting of short-term electricity demand: Fusing uncertainties from diverse sources," Inf. Fusion, vol. 127, no. July 2025, p. 103637, 2026, doi: 10.1016/j.inffus.2025.103637.

[10] S. H. Barati, "A system dynamics approach for leveraging blockchain technology to enhance demand forecasting in supply chain management," Supply Chain Anal., vol. 10, no. March, p. 100115, 2025, doi: 10.1016/j.sca.2025.100115.

[1] M. H. Ghodusinejad, H. Yousefi, and M. Sotoodeh, "Demand forecasting and PV/EV-integrated energy system modeling for achieving a sustainable island," Sustain. Futur., vol. 10, no. September 2024, p. 100857, 2025, doi: 10.1016/j.sftr.2025.100857.

[12] B. Rinchi, S. Al-Dahidi, M. E. B. Abdalla, O. Ayadi, and A. Al-Akhras, "Economic and environmental analysis of forecasting errors in air conditioning energy demand under fixed and time-of-use tariffs: A case study," Case Stud. Therm. Eng., vol. 72, no. March, p. 106409, 2025, doi: 10.1016/j.csite.2025.106409.

[13] S. Zhang, B. Yu, Z. Zhou, J. Tan, X. C. Yuan, and Y. M. Wei, "Copper demand forecast for the transport sector under China's Carbon neutrality target," Resour. Conserv. Recycl., vol. 220, no. October 2024, p. 108346, 2025, doi: 10.1016/j.resconrec.2025.108346.

[14] N. Sinha and C. Lucheroni, "Demand and supply curve forecasting using a monotonic autoencoder for short-term day-ahead electricity market bid curves," Appl. Energy, vol. 397, no. June, p. 126262, 2025, doi: 10.1016/j.apenergy.2025.126262.

[15] L. Guangxiao, Y. Yilu, L. Zongjie, H. Zhaohui, and L. Changyun, "Electricity demand forecasting and power supply planning under carbon neutral targets," Energy Reports, vol. 13, no. November 2024, pp. 2740–2751, 2025, doi: 10.1016/j.egyr.2025.02.015.

[16] A. Carrillo-Galvez et al., "Electricity demand forecasting in green ports: Modelling and future research directions," Transp. Policy, vol. 171, no. July, pp. 1012–1024, 2025, doi: 10.1016/j.tranpol.2025.07.013.

[17] K. Zhou, Y. Chu, and R. Hu, "Energy supply-demand interaction model integrating uncertainty forecasting and peer-to-peer energy trading," Energy, vol. 285, no. October, p. 129436, 2023, doi: 10.1016/j.energy.2023.129436.

[18] S. Y. Kim, S. K. Heo, K. J. Nam, T. Y. Woo, and C. K. Yoo, "Flexible renewable energy planning based on multi-step forecasting of interregional electricity supply and demand: Graph-enhanced AI approach," Energy, vol. 282, no. February, p. 128858, 2023, doi: 10.1016/j.energy.2023.128858.

[19] S. Mateo-Barcos, D. Ribó-Pérez, J. Rodríguez-García, and M. Alcázar-Ortega, "Forecasting electricity demand of municipalities through artificial neural networks and metered supply

point classification," Energy Reports, vol. 11, no. January, pp. 3533–3549, 2024, doi: 10.1016/j.egyr.2024.03.023.

[20] A. T. Mustafa and O. S. A.-D. Al-Yozbaky, "Forecasting Energy Demand and Generation Using Time Series Models: A Comparative Analysis of Classical, Grey, Fuzzy, and Intelligent Approaches," Franklin Open, p. 100350, 2025, doi: 10.1016/j.fraope.2025.100350.

[21] G. Rajarathnam et al., "Forecasting energy demand, supply, and carbon footprint for greenfield planning in the context of net zero systems: Case study on Western Sydney Aerotropolis in New South Wales, Australia," Smart Energy, vol. 12, no. January, p. 100117, 2023, doi: 10.1016/j.segy.2023.100117.

[22] S. El Bakali, H. Ouadi, F. Giri, S. Gheouany, N. Mounir, and I. Jrhilifa, "Forecasting Occupants' Presence and Photovoltaic Power using an Enhanced Stacking Algorithm for Demand/Supply Energy Management," IFAC-PapersOnLine, vol. 58, no. 13, pp. 290–295, 2024, doi: 10.1016/j.ifacol.2024.07.497.

[23] I. Sokolova et al., "From abundant resource to critical commodity: Forecasting manganese supply and assessing its sustainability," Sustain. Mater. Technol., vol. 44, no. September 2024, p. e01349, 2025, doi: 10.1016/j.susmat.2025.e01349.

[24] J. B. S. O. De Andrade Guerra, L. Dutra, N. B. C. Schwinden, and S. F. De Andrade, "Future scenarios and trends in energy generation in Brazil: Supply and demand and mitigation forecasts," J. Clean. Prod., vol. 103, pp. 197–210, 2015, doi: 10.1016/j.jclepro.2014.09.082.

[25] A. Moradzadeh, L. Ben-Brahim, A. Arefi, A. Oshnoei, and S. M. Muyeen, "Hybrid deep learning models for water demand forecasting in greenhouses: Exploring the energy Nexus in Urban agriculture," Energy Nexus, vol. 20, no. September, p. 100546, 2025, doi: 10.1016/j.nexus.2025.100546.

[26] L. Massidda and M. Marrocu, "Hybrid forecasting of demand flexibility: A top-down approach for thermostatically controlled loads," Energy AI, vol. 20, no. October 2024, p. 100487, 2025, doi: 10.1016/j.egyai.2025.100487.

[27] A. Roozbeh Nia, A. Awasthi, and N. Bhuiyan, "Industry 4.0 and demand forecasting of the energy supply chain: A literature review," Comput. Ind. Eng., vol. 154, no. October 2020, p. 107128, 2021, doi: 10.1016/j.cie.2021.107128.

[28] J. Guo, Q. Meng, B. Du, and H. Sun, "Probability forecasting for multivariate urban water demand using temporal convolutional network based on quantile regression and Parzen window," Eng. Appl. Artif. Intell., vol. 162, no. PB, p. 112528, 2025, doi: 10.1016/j.engappai.2025.112528.

[29] K. C. Singh, S. Baskaran, and P. Marimuthu, "Reinforced learning for demand side management of smart microgrid based forecasted hybrid renewable energy scenarios," Comput. Electr. Eng., vol. 123, no. PB, p. 110127, 2025, doi: 10.1016/j.compeleceng.2025.110127.

[30] B. Vijay Kumar, "Intelligent demand-side energy management based IoT using multi-attention fusion residual convolutional neural network for day-ahead energy forecasting," Energy, vol. 333, no. June, p. 136929, 2025, doi: 10.1016/j.energy.2025.136929.

[31] S. A. Alsamraee and S. Khanna, "Long-term electricity demand forecasting of a university campus based on advanced deep learning artificial neural network algorithms," Results Eng., vol. 27, no. March, p. 106683, 2025, doi: 10.1016/j.rineng.2025.106683.

[32] A. Patney and R. Patel, "Optimizing Electricity Demand Forecasting with Ensemble Learning: A Dynamic Factor Perspective," Procedia Comput. Sci., vol. 258, pp. 1454–1465, 2025, doi: 10.1016/j.procs.2025.04.378.

[33] D. Matkovic, T. M. Pilski, and T. Capuder, "Participation of electric vehicle charging station aggregators in the day-ahead energy market using demand forecasting and uncertainty-based pricing," Energy, vol. 328, no. March, p. 136299, 2025, doi: 10.1016/j.energy.2025.136299.

[34] H. Kırkgöz and O. Kurt, "Modeling bitcoin network energy demand: Price-adjusted hybrid deep learning approach to complex time series forecasting," Chaos, Solitons and Fractals, vol. 200, no. July, 2025, doi: 10.1016/j.chaos.2025.117075.

[35] A. Xu et al., "Multi-Scale Spatio-Temporal Graph Neural Network for Enhanced Water Demand Forecasting," Water Res., vol. 288, no. PB, p. 124711, 2025, doi: 10.1016/j.watres.2025.124711.

[36] A. H. Asfaw, G. Teklu, A. Birhan, Y. Eshetu, E. Regasa, and T. Wondimu, "Mathematical modeling of Ethiopia's energy demand by sectors and energy types, with forecasts for the next 30 years," Heliyon, vol. 10, no. 22, p. e40185, 2024, doi: 10.1016/j.heliyon.2024.e40185.

[37] J. Wang, J. Fan, S. Wu, and A. Farnoosh, "Medium- and long-term hydrogen demand forecast in China: A multi-sector and multi-region analysis," Energy, vol. 335, no. August, p. 137992, 2025, doi: 10.1016/j.energy.2025.137992.

[38] F. Hussain, M. Hasanuzzaman, and N. A. Rahim, "Multivariate machine learning algorithms for energy demand forecasting and load behavior analysis," Energy Convers. Manag. X, vol. 26, no. October 2024, p. 100903, 2025, doi: 10.1016/j.ecmx.2025.100903.

[39] X. Zhou et al., "Noise-robust and sector-aware representation learning for natural gas demand forecasting," Expert Syst. Appl., vol. 299, no. PA, p. 129964, 2026, doi: 10.1016/j.eswa.2025.129964.

[40] B. Palaniyappan and S. K. Ramu, "Optimized LSTM-based electric power consumption forecasting for dynamic electricity pricing in demand response scheme of smart grid," Results Eng., vol. 25, no. November 2024, p. 104356, 2025, doi: 10.1016/j.rineng.2025.104356.

RiESTech
Recent in Engineering Science and Technology

MBI

Article

Auditing Physical Consistency in Engine Fault Detection Using Domain-Informed Monotonic XGBoost and Explainable Boosting Machine: A Case Study on a Public Marine Diesel Engine Dataset

Abdul Azis Abdillah [1,*]

[1] Department of Mechanical Engineering, School of Engineering, University of Birmingham, United Kingdom
* Correspondence: axa2072@student.bham.ac.uk

Abstract: Machine-learning models for engine fault detection are usually judged by accuracy, although a model can score well while learning relationships that contradict engine physics. This study proposes a physical-consistency audit that combines domain-informed monotonic constraints in XGBoost with the glass-box Explainable Boosting Machine (EBM). Three XGBoost configurations (unconstrained, domain-constrained and reverse-constrained) are compared, and their difference is summarised by a physical-consistency gap (PCG) with a confidence interval and a decision rule. The audit is applied to a public marine diesel engine dataset (IEEE DataPort; 19,533 records, six sensors). Its label is documented by the dataset creator (1 = no failure, 0 = failure), and the records are a synthetic extension of 174 shipboard observations. Under a leakage-safe protocol, the best of six classifiers reached a modest test ROC-AUC of 0.708. Engineering-consistent constraints had no measurable cost (test ΔAUC +0.0001), whereas reversed constraints reduced ROC-AUC by 0.015, giving PCG = −0.0154 (95% CI −0.0208 to −0.0101). SHAP and EBM agreed that fault risk falls with oil pressure and rises with oil temperature. Semi-synthetic labels, 30 null replicates and a positive control on the AI4I 2020 dataset showed that the audit separates consistent, inverted and non-physical labels; no null replicate was flagged. Coding the label the other way round, as in the first version of this study, produced the mirror-image gap and a flag, so the audit also detects a wrong label convention. Mixed tail evidence and the synthetic origin limit what the labels reveal about real engines.

Keywords: Engine fault detection; Marine diesel engine; Physical consistency; Monotonic constraints; Explainable boosting machine; SHAP; Label convention; Positive control

Citation: Abdillah, A. A. (2026). Auditing Physical Consistency in Engine Fault Detection Using Domain-Informed Monotonic XGBoost and Explainable Boosting Machine: A Case Study on a Public Marine Diesel Engine Dataset. Recent in Engineering Science and Technology, 4(4), 224–244. Retrieved from https://www.mbi-jounals.com/index.php/riestech/article/view/180.

Academic Editor: Yudhi Ariadi

Received: 1 October 2026
Accepted: 18 October 2026
Published: 31 October 2026

Publisher's Note: MBI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2026 by the authors. Licensee MBI, Jakarta, Indonesia. This article is an open access article distributed under MBI license (https://mbi-journals.com/licenses/by/4.0/).

1. Introduction

Condition-based and predictive maintenance aim to detect deteriorating equipment before it fails, reducing unplanned downtime and maintenance cost [1,2]. On ships, the main engine is the most critical item of machinery. Routinely logged variables such as engine speed, lubrication-oil pressure and temperature, fuel pressure and coolant pressure and temperature are natural inputs for data-driven fault detection [3]. Machine-learning (ML) classifiers have become the dominant approach for mapping these signals to a health state [4–6].

Most studies judge such models by predictive metrics alone, typically accuracy, F1-score or the area under the receiver operating characteristic curve (ROC-AUC). High accuracy, however, does not guarantee that a model has learned the right relationships. In a well-known medical example, a pneumonia-risk model learned that asthma lowers mortality risk [7]. This was an artefact of how asthmatic patients were treated, not a physiological effect. Models readily exploit such shortcuts [8], and image classifiers have been shown to rely on spurious "Clever Hans" cues [9]. Ross et al. formalised the aim of

models that are "right for the right reasons" [10]. In maintenance, technicians must trust and act on alerts. A detector that treats rising oil temperature as a sign of health would therefore be unacceptable, even with a competitive ROC-AUC [11].

Explainable AI (XAI) is often proposed as the remedy. SHapley Additive exPlanations (SHAP) [12,13] are widely used in predictive-maintenance studies [6]. In many applications, however, explanations are shown as importance rankings or dependence plots without a formal test of whether each learned direction agrees with engineering expectations. Theory-guided data science lists physical consistency as a requirement for scientific ML models [14]. Physics-informed ML embeds scientific knowledge into models [15], and monotonic constraints encode directional knowledge in tree ensembles [16]. Recent work has begun to use such constraints diagnostically. Koklev quantified the "price of monotonicity" as the performance change between unconstrained and constrained gradient boosting [17]. Abbas et al. refitted an XGBoost classifier with pre-specified monotonic constraints as a direction test [18], and Hsiao et al. corrected non-physical EBM shape functions with domain knowledge [19]. These studies price or repair a model. They do not provide a calibrated statistic that indicates whether the labels of a dataset agree with the assumed physics. This is the first gap addressed here.

The second gap concerns the data themselves. Datasheets for datasets were proposed because public data are often released without documentation of how they were collected and labelled [20]. Label noise and systematic labelling errors degrade classifiers and mislead their interpretation [21], and confident learning estimates label errors from model disagreement [22]. These methods do not exploit directional physical knowledge. A wrong label convention, in which the two classes are swapped, leaves accuracy and model confidence unchanged while reversing every learned direction, so it cannot be detected in this way. Data leakage can further inflate reported performance [23,24]. The public ship main-engine dataset studied here illustrates the problem [25]. Its repository record does not state which label value denotes a fault, although the creator's articles do [3,26]. Studies of this file report accuracies or areas under the curve between about 0.74 and above 0.98 [26–28], and the label convention is not always stated.

This study addresses both gaps with a physical-consistency audit built on two complementary models. The first is a three-way comparison of XGBoost [29] configurations that differ only in their monotonic constraints: unconstrained, constrained in the direction dictated by engineering knowledge, and constrained in the reversed direction. The second is the Explainable Boosting Machine (EBM) [30,31], a glass-box generalised additive model whose shape functions show the learned directions without any constraint. Relative to the work cited above, the distinctive elements are threefold. The first is the reversed-constraint control. The second is the physical-consistency gap (PCG), a statistic with a confidence interval and explicit decision rules. The third is a semi-synthetic calibration that shows which PCG values arise under consistent, inverted, partially reversed and non-physical labels. The audit is also checked on a positive-control dataset with documented failure mechanisms [32]. All experiments follow a leakage-safe protocol with repeated cross-validation, corrected statistical tests and a single evaluation on an isolated hold-out set.

2. Materials and Experiment Methods

2.1. Dataset, Label Convention and Data Origin

The study uses the public dataset "Predictive Maintenance on Ship's Main Engine using AI" [25], deposited on IEEE DataPort (DOI 10.21227/g3za-v415) and downloaded on 29 June 2026. The file engine_data.csv has 19,535 records with six continuous sensor variables (Table 1) and a binary label, Engine Condition. The repository record does not define the label values. The dataset creator defines them in two articles. Mohakul et al. state that "if engine condition = 0, the engine will fail and if engine condition = 1, no failure

will take place" [3]. Nurdin et al., in an article co-authored by the creator, define the label as healthy (1) and unhealthy (0) and state that the data describe an MTU Series 1400 marine diesel engine, with pressures in bar and temperatures in °C [26]. The 20 example records printed in that article match our file exactly, including their labels.

The models in this study therefore predict the fault class, defined as Fault = 1 − Engine Condition. Faults make up 36.9% of the cleaned records. The first version of this study used Engine Condition = 1 as the fault class. That coding is the reverse of the documented one, and Section 3.9 shows that the audit detects it.

The records are synthetic. Mohakul et al. describe how the data were produced [3]. Readings of the six sensors were transcribed from the logs of a ship's main engine, giving 174 records. Only three of them had Engine Condition = 0. These three were randomly oversampled to 51, and a Gaussian copula fitted with the Synthetic Data Vault [33] then generated 200,000 records. A companion conference paper also refers to "simulated data" [34]. The public file is smaller (19,535 records, 36.9% with Engine Condition = 0, against 34.3% in the 200,000 records), and how it was derived from the generated data is not documented. Its inter-sensor correlations are close to zero (Table 2). All pairwise Pearson correlations are below 0.074 in magnitude, and the correlation between oil pressure and engine speed is +0.025. In a running engine, oil pressure depends on speed and oil temperature, so the sensors were probably generated almost independently of one another. Two consequences follow. First, the audit tests the data-generating process and its labelling rather than the physics of a real engine. Second, the fault class rests on three genuine fault observations.

2.2. Preprocessing and Cleaning

The data contained no missing values and no duplicate records. One fixed domain rule was applied before data partitioning: records with a coolant temperature above 110 °C were removed. Two records were affected, with 118.4 °C (labelled fault) and 195.5 °C (labelled no fault). The next-highest value is 95.9 °C. A reading of 195.5 °C is far above the boiling range of a pressurised cooling circuit and was treated as a sensor error. A reading of 118.4 °C could reflect genuine overheating; it was removed only so that one consistent rule is applied, and Appendix Table A2 shows the audit with either record kept. Because the rule is not estimated from data, applying it before the split cannot leak information.

Minimum pressures are close to zero (Table 1), and 118 records report an oil pressure below 0.5 bar, which is incompatible with a running engine. Their other sensor readings are typical of the whole file, and their fault rate (0.356) is close to the base rate. This pattern suggests generated values outside the physical range rather than lubrication failures. Excluding these records did not change the audit result (Table A2). No feature scaling was applied to tree-based models; standardisation for logistic regression was embedded in a pipeline and fitted on training folds only. No resampling was used, because the class ratio (63:37) is only moderately imbalanced and resampling would alter the conditional distributions that the audit examines.

Table 1. Sensor variables after cleaning (n = 19,533), their association with the fault class and the expected direction of fault risk used for the monotonic constraints.

Variable	Mean	SD	Min	Max	Mean (healthy)	Mean (fault)	r_{pb} (t)	Expected fault risk
Engine speed (rpm)	791.26	267.61	61.000	2239.00	736.32	885.01	+0.268 (+38.9)	unconstrained
Lub. oil pressure (bar)	3.30	1.02	0.003	7.27	3.35	3.22	−0.061 (−8.6)	↑ as pressure ↓ (−1)
Fuel pressure (bar)	6.66	2.76	0.003	21.14	6.90	6.24	−0.116 (−16.3)	unconstrained
Coolant pressure (bar)	2.34	1.04	0.003	7.48	2.32	2.37	+0.024 (+3.4)	unconstrained
Lub. oil temperature (°C)	77.64	3.11	71.322	89.58	77.42	78.02	+0.094 (+13.1)	↑ as temp. ↑ (+1)
Coolant temperature (°C)	78.42	6.14	61.673	95.86	78.20	78.80	+0.047 (+6.6)	↑ as temp. ↑ (+1)

Units follow Nurdin et al. [26]. Healthy: Engine Condition = 1; fault: Engine Condition = 0. r_{pb}: point-biserial correlation with the fault class; t: corresponding t statistic (df = 19,531); $|t| > 3.3$ corresponds to $p < 0.001$.

Table 2. Pearson correlation matrix of the six sensors and the fault class (cleaned data).

	Speed	Oil p.	Fuel p.	Cool. p.	Oil T	Cool. T	Fault
Speed	1						
Oil p.	0.025	1					
Fuel p.	−0.002	0.044	1				
Cool. p.	−0.025	−0.009	0.033	1			
Oil T	0.052	−0.008	−0.025	−0.021	1		
Cool. T	0.031	−0.061	−0.043	0.035	0.074	1	
Fault	0.268	−0.061	−0.116	0.024	0.094	0.047	1

Largest absolute inter-sensor correlation: 0.074. Spearman correlations (supplementary results file) give the same picture.

2.3. Leakage-Safe Experimental Protocol

Figure 1 summarises the protocol. After cleaning, the data were split once into a stratified training set (80%, 15,626 records) and a hold-out test set (20%, 3,907 records). Both retain the 36.9% fault prevalence. The test set was isolated immediately and scored once, at the end. All operations that learn from data were carried out on the training set only [23,35]: scaling, hyper-parameter tuning, model comparison and threshold selection. Hyper-parameters were tuned by randomised search with three-fold stratified cross-validation (CV) on the training set, maximising ROC-AUC (Table A1). Models were then compared with 5 × 3 repeated stratified CV on the same training set.

This benchmark design is not nested: the tuning data are reused for the comparison, which can bias CV estimates optimistically [35]. The hold-out results are therefore treated as the primary performance evidence. For the three XGBoost configurations that form the audit, a fully nested CV was also run, in which each configuration was re-tuned inside every outer training fold. For threshold-dependent metrics, the threshold of each model was chosen by maximising Youden's J [36] on five-fold out-of-fold training predictions and then frozen before the test set was scored.

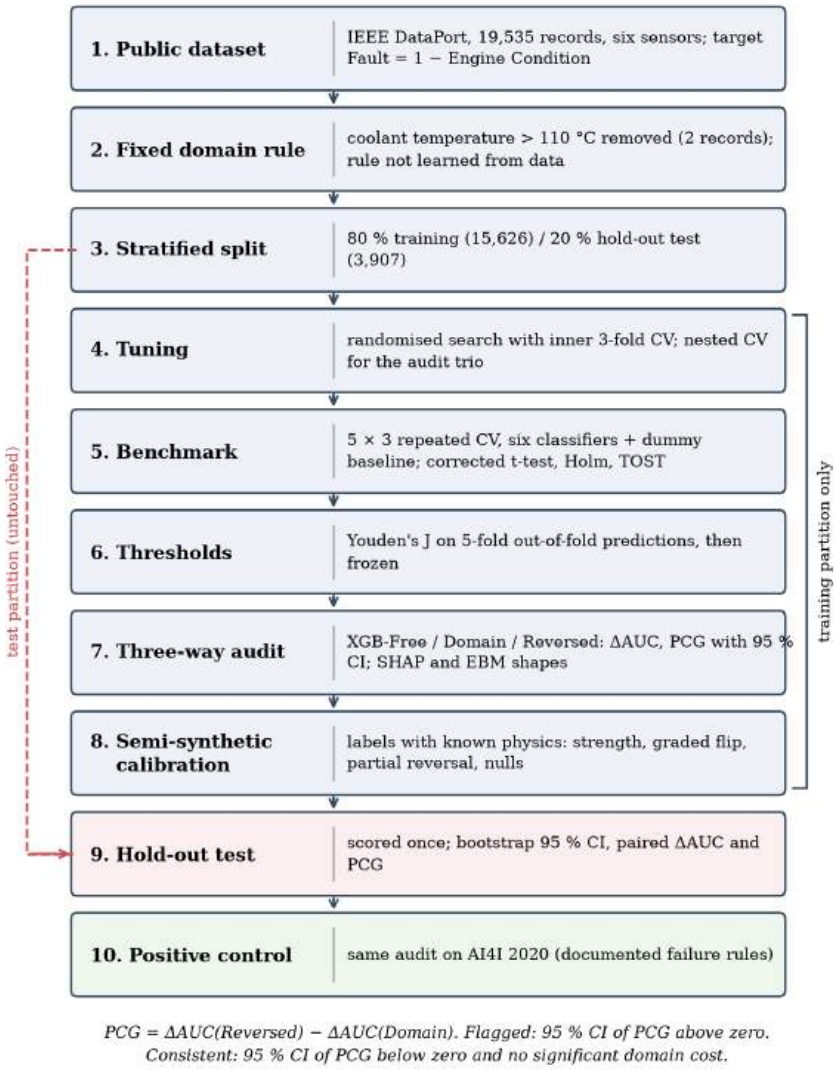


Figure 1. Leakage-safe experimental protocol. The hold-out test set is separated at step 3 and scored once (step 9). Tuning, benchmarking, threshold selection, the audit and the semi-synthetic calibration use the training partition only; the positive control repeats the audit on the AI4I 2020 dataset.

2.4. Benchmark Classifiers

Six classifiers were benchmarked with scikit-learn [37] and the respective libraries: logistic regression (LR) with L2 regularisation; random forest (RF) [38]; LightGBM [39]; CatBoost [40]; unconstrained XGBoost (XGB-Free) [29]; and EBM [30,31]. A majority-class dummy classifier, which ignores the sensors and always predicts the most frequent class (healthy), served as the minimum reference that any useful model must exceed; it is not part of any other model. The two constrained XGBoost configurations of the audit (Section 2.5) were evaluated alongside the benchmark. The tuned hyper-parameters are listed in Table A1. The two models central to the audit are described below.

2.5. Domain-Informed Monotonic XGBoost

XGBoost builds an additive ensemble of regression trees by gradient boosting. A monotonic constraint $d_f \in \{-1, 0, +1\}$ on feature f restricts every split on that feature. The ensemble output $F(\mathbf{x})$, the log-odds of a fault, is then non-decreasing ($d_f = +1$) or non-increasing ($d_f = -1$) in x_f when all other features $\mathbf{x}_{-f}$ are held fixed:

$$d_f = +1: \quad x_f \le x'_f \;\Rightarrow\; F(x_f, \mathbf{x}_{-f}) \le F(x'_f, \mathbf{x}_{-f}) \tag{1a}$$

$$d_f = -1: \quad x_f \le x'_f \;\Rightarrow\; F(x_f, \mathbf{x}_{-f}) \ge F(x'_f, \mathbf{x}_{-f}) \tag{1b}$$

The domain directions encode the dominant failure mechanisms of the lubrication and cooling systems of a diesel engine [41]. Low lubrication-oil pressure indicates insufficient oil flow to the bearings ($d = -1$), and high oil and coolant temperatures indicate overheating ($d = +1$). These directions are not universal. A low coolant temperature can indicate a thermostat stuck open, and a high oil pressure can indicate a stuck relief valve or cold, viscous oil. Engine speed, fuel pressure and coolant pressure were left unconstrained because both abnormally low and abnormally high values can indicate faults. The same argument applies, more weakly, to the three constrained sensors. In addition, the observed oil-temperature range (71–90 °C) and most of the coolant-temperature range (62–96 °C) lie inside normal operating windows. Manufacturer alarm set points for this engine are not publicly available and were therefore not used.

Monotonicity is therefore treated as a hypothesis about the dominant mechanism, not as a law. Three analyses test it: constraints applied to one sensor at a time, a tail analysis restricted to physically unambiguous extremes (Section 2.7), and EBM shape functions that are free to be non-monotone. Three configurations were defined: XGB-Free (no constraints), XGB-Domain (the directions above) and XGB-Reversed (all directions negated). In the main audit all three reuse the hyper-parameters tuned for XGB-Free, so that differences are attributable to the constraints alone. Because shared hyper-parameters may favour XGB-Free, the constrained configurations were also tuned separately with the same search, and the audit was repeated under nested CV.

2.6. Explainable Boosting Machine

EBM is a generalised additive model with pairwise interactions (GA²M) fitted by cyclic gradient boosting [30,31]. Each boosting round fits a shallow tree to a single feature in round-robin fashion with a small learning rate. The effects of different features are therefore learned separately and summed:

$$g(\mathrm{E}[y]) = \beta_0 + \Sigma_f\, s_f(x_f) + \Sigma_{(i,j)}\, s_{ij}(x_i, x_j) \tag{2}$$

Here g is the logit link, s_f is the shape function of feature f, and s_{ij} are pairwise interaction terms. Because each s_f can be plotted directly, the direction of every learned main effect can be read without post-hoc explanation. EBM serves as a benchmark model and as an unconstrained view of the learned directions. The tuned EBM included 10 interaction terms. With six features there are only 15 possible pairs, so interactions could absorb part of the main effects; an EBM without interaction terms was therefore also fitted.

2.7. Physical-Consistency Audit

The audit has three quantitative components. The first is the cost of each constraint configuration c relative to the unconstrained model, and the resulting physical-consistency gap (PCG):

$$\Delta\mathrm{AUC}_c = \mathrm{AUC}_c - \mathrm{AUC}_{\mathrm{Free}}, \quad c \in \{\mathrm{Domain}, \mathrm{Reversed}\} \tag{3}$$

$$\mathrm{PCG} = \Delta\mathrm{AUC}_{\mathrm{Reversed}} - \Delta\mathrm{AUC}_{\mathrm{Domain}} = \mathrm{AUC}_{\mathrm{Reversed}} - \mathrm{AUC}_{\mathrm{Domain}} \tag{4}$$

If the labels follow the assumed physics, domain constraints should cost little ($\Delta\mathrm{AUC}_{\mathrm{Domain}} \approx 0$) and reversed constraints should be costly, giving $\mathrm{PCG} < 0$. A positive PCG indicates that the labels are better described by the reversed physics. In CV, ΔAUC and PCG were computed in each of the $k = 15$ folds. Their 95% confidence intervals use the corrected resampled variance of Nadeau and Bengio [42]:

$$\overline{\mathrm{PCG}} \pm t_{0.975,\,k-1} \cdot [(1/k + n_2/n_1)\, s^2]^{1/2} \tag{5}$$

where n_1 and n_2 are the training and validation sizes of a fold and s^2 is the sample variance of the fold-wise PCG values. On the test set, 2,000 bootstrap resamples of the test records were drawn once and shared by all models (paired bootstrap). ΔAUC and PCG

were recomputed on every resample, and the 2.5th and 97.5th percentiles give the 95% CI [43]. **Decision rules** use the CV estimates. A dataset is *flagged* as inconsistent with the assumed physics when the lower bound of the 95% CI of PCG lies above zero. It is read as *consistent* when the upper bound lies below zero and, in addition, the domain constraints cost no significant discrimination (upper bound of the 95% CI of $\Delta AUC_{Domain} \geq 0$). The second condition is needed because a non-physical rule can make both constraint sets costly and so also produce a negative PCG. All other results are *unresolved*. The error rates of both rules were estimated from null scenarios (Section 2.8).

The second component is directional agreement. For each feature, the learned direction was measured by the Spearman correlation ρ_f between the feature value and its contribution: the SHAP value of XGB-Free computed with TreeSHAP [13], or the EBM term contribution. Because ρ captures only the sign of a monotone trend, an effect size was added:

$$\Delta_f = \text{mean}\{\varphi_f(\mathbf{x}) : x_f \geq P_{95}\} - \text{mean}\{\varphi_f(\mathbf{x}) : x_f \leq P_5\} \quad (6)$$

where φ_f is the contribution in log-odds of a fault and P_5 and P_{95} are the 5th and 95th percentiles of the feature on the test set. Confidence intervals for ρ and Δ come from 1,000 bootstrap resamples of the test records. A feature is consistent when $\text{sign}(\rho_f) = d_f$ for SHAP, EBM and EBM without interactions, and inconsistent when all three signs oppose d_f. Changes in each feature's share of mean absolute SHAP value were also recorded.

The third component is model-free. A standardised logistic regression with Wald 95% CIs was fitted to all cleaned records as a simple baseline. In addition, the fault rate was computed in physically unambiguous extremes: oil pressure below 1.0 and 0.5 bar or above 6.0 bar, oil temperature above 85 °C, coolant temperature above 93 °C or below 65 °C, fuel pressure below 2.0 bar and engine speed below 300 rpm. These limits are conservative extremes of the observed ranges, not manufacturer alarm set points. If the labels are physical, each of these high-risk regions should show a fault rate above the base rate. Wilson 95% intervals [44] and exact binomial tests against the base rate were used.

2.8. Semi-Synthetic Calibration of the Audit

To establish which audit results to expect under known conditions, semi-synthetic labels were generated on the training partition only, keeping the real sensor values. With features standardised by the training mean and standard deviation ($\mathbf{z}$), fault labels were drawn as Bernoulli variables with probability

$$P(y = 1 \mid \mathbf{z}) = \sigma(\beta_0 + \boldsymbol{\beta}^T\mathbf{z}), \quad \sigma(u) = 1 / (1 + e^{-u}) \quad (7)$$

The coefficient vector was derived from the data rather than chosen freely. For the three unconstrained sensors, $\boldsymbol{\beta}$ equals the standardised logistic-regression coefficients of the real training labels: +0.583 (engine speed), −0.261 (fuel pressure) and +0.066 (coolant pressure). For the three constrained sensors, $\beta_f = k \cdot d_f \cdot |\hat{\beta}_f|$ keeps the magnitude of the real coefficient ($|\hat{\beta}|$ = 0.145, 0.170 and 0.046 for oil pressure, oil temperature and coolant temperature) but takes the sign of the domain direction, scaled by a strength factor k. The intercept β_0 was solved numerically so that the expected fault rate equals the real rate (36.9%); for $k = 1$, $\beta_0 = -0.581$.

Five groups of scenarios were examined: (a) consistent labels (k = 1) and their inversion ($1 - y$); (b) signal strengths k = 0.5, 2 and 4, consistent and inverted, which bracket the real effect size; (c) a graded label flip, in which a random fraction q = 0.1, …, 0.9 of consistent labels was inverted, for k = 1 and k = 2; (d) partial reversal, in which the physical sign was reversed for one or two of the three constrained sensors; and (e) three null scenarios with ten replicates each. In null N1, labels depend only on the unconstrained sensors. In N2, a non-physical U-shaped rule makes risk rise at both ends of every constrained sensor. In N3, the real labels are randomly permuted. Every scenario used the same 5 × 3 repeated CV and the same XGBoost hyper-parameters as the real-data audit.

2.9. Positive Control on a Dataset with Documented Failure Mechanisms

The AI4I 2020 predictive-maintenance dataset [32] contains 10,000 records and documents the rules that generate its failures (machine-failure rate 3.39%). Heat-dissipation failure occurs when the difference between process and air temperature is below 8.6 K and the rotational speed is below 1380 rpm; tool-wear and overstrain failures increase with tool wear. Fault risk is therefore non-increasing in the temperature difference ($d = -1$) and non-decreasing in tool wear ($d = +1$). Product type, air temperature, speed and torque were left unconstrained, because power failure occurs at both low and high power. The same protocol was applied: an 80/20 stratified split, XGBoost tuning on the training part, the three-way audit with 5 × 3 CV, and a paired bootstrap on the test part. The audit was run with the documented labels and with inverted labels. Unlike the ship dataset, where the truth is unknown, this dataset shows whether the audit returns the correct answer when the truth is known.

2.10. Evaluation Metrics and Statistical Testing

Threshold-independent metrics were ROC-AUC, the area under the precision-recall curve (PR-AUC) [45] computed with the fault class and, separately, the healthy class as the positive class, and the Brier score. The Brier score is compared with a prior-probability baseline that predicts the training prevalence for every record, and the Brier skill score is $BSS = 1 - \text{Brier}/\text{Brier}_{\text{prior}}$. Calibration slope and calibration-in-the-large intercept were obtained by logistic recalibration of the predicted log-odds [46]. At the frozen Youden threshold, balanced accuracy, the Matthews correlation coefficient (MCC) [47], and precision, recall and F1-score for both classes were reported.

Test-set ROC-AUC confidence intervals are percentile intervals from 2,000 bootstrap resamples [43]. Differences in CV ROC-AUC were tested with the Nadeau–Bengio corrected resampled *t*-test [42], and comparisons against the best model were adjusted with the Holm procedure [48]. Because a non-significant difference does not show equivalence, equivalence was tested with two one-sided tests (TOST) [49] using the same corrected variance, and with the Bayesian correlated *t*-test of Benavoli et al. [50]. Both use a region of practical equivalence (ROPE) of ±0.01 ROC-AUC, fixed before the analysis as a difference too small to matter for maintenance decisions at the present level of discrimination. All analyses used Python 3.11 with fixed random seed 42; package versions are listed in the supplementary environment file. The code and results are archived on Zenodo (https://doi.org/10.5281/zenodo.23728914).

3. Results and Discussion

3.1. Data Characteristics and Simple Baselines

Table 1 and Figure 2 show that the two classes overlap heavily for every sensor. Engine speed has the strongest association with the fault class ($r_{pb} = +0.268$), followed by fuel pressure (−0.116) and lubrication-oil temperature (+0.094). The marginal associations of the three constrained sensors follow the expected fault physics: fault records have slightly lower oil pressure and slightly higher oil and coolant temperatures. With 19,533 records these differences are highly significant ($|t| = 8.6$, 13.1 and 6.6). Fault records also run at higher engine speed; speed was left unconstrained.

A standardised logistic regression gives the same picture (Table 3). The coefficients of oil pressure (−0.138), oil temperature (+0.167) and coolant temperature (+0.055) all have the expected sign, and their 95% CIs exclude zero. Under the reversed coding all these signs would be reversed, so a simple baseline is also sensitive to the label convention. The audit adds three things. First, it measures consistency in predictive terms, as the discrimination lost when the physics is enforced or reversed, and it does so with a flexible model that allows non-linear and conditional effects. Second, the semi-synthetic

calibration and the null scenarios show what size of PCG to expect under known label mechanisms, which a coefficient sign cannot. Third, the reversed-constraint control separates a consistent direction from a mere lack of signal: sensors without signal give PCG ≈ 0, not PCG < 0.

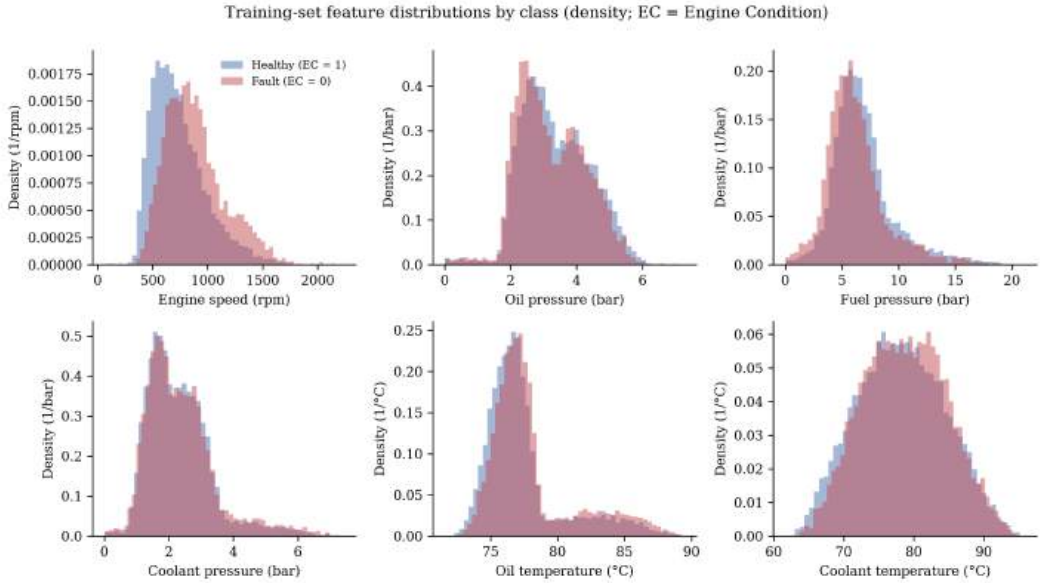


Figure 2. Training-set distributions (densities) of the six sensor variables by class. All variables overlap heavily between the two classes.

Table 3. Standardised logistic regression of the fault class on the six sensors (all cleaned records).

Variable	Coefficient per SD	95% CI	*p*	Expected sign	Verdict
Engine speed (rpm)	+0.584	[+0.552, +0.616]	<0.001	free	n/a
Lub. oil pressure (bar)	−0.138	[−0.169, −0.107]	<0.001	−	consistent
Fuel pressure (bar)	−0.268	[−0.301, −0.235]	<0.001	free	n/a
Coolant pressure (bar)	+0.078	[+0.047, +0.108]	<0.001	free	n/a
Lub. oil temperature (°C)	+0.167	[+0.137, +0.198]	<0.001	+	consistent
Coolant temperature (°C)	+0.055	[+0.024, +0.085]	<0.001	+	consistent

Wald confidence intervals; coefficients in log-odds of a fault per standard deviation of the sensor.

3.2. Tail Analysis in Physically Unambiguous Extremes

Table 4 and Figure 3 give the fault rate in extreme regions where engine physics is unambiguous (base rate 0.369). The evidence is mixed. The oil temperature > 85 °C and fuel pressure < 2.0 bar regions show a fault rate significantly above the base rate, as physics predicts. The coolant temperature < 65 °C and engine speed < 300 rpm regions show a fault rate significantly below it. The remaining regions (oil pressure < 1.0 bar, oil pressure < 0.5 bar, oil pressure > 6.0 bar and coolant temperature > 93 °C) are inconclusive. In particular, the low-oil-pressure tail, the most direct indicator of lubrication failure, is close to the base rate. These extremes contain few records and, given the synthetic origin of the data (Section 2.1), lie where the generator had the least information from the 174 source records. The tails therefore do not show that the labels are physically valid throughout the sensor range, only that the bulk evidence examined by the audit is not contradicted by them.

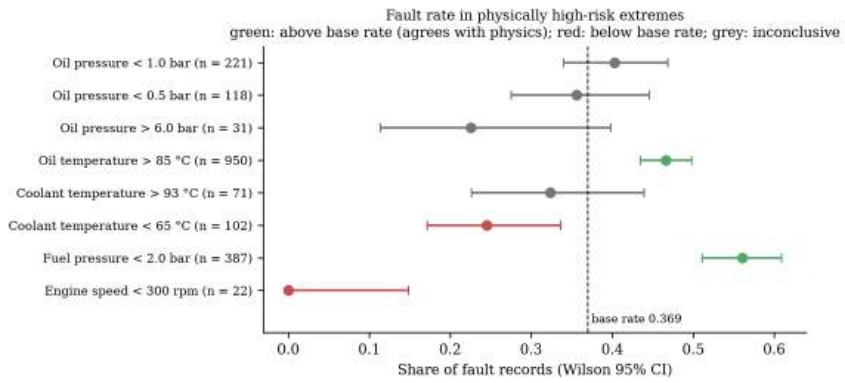


Figure 3. Fault rate in physically unambiguous extremes with Wilson 95% confidence intervals. The dashed line marks the base rate.

Table 4. Fault rate in physically unambiguous extremes (all cleaned records).

Region	*n*	Fault rate [95% CI]	*p* (vs base rate)	Physics
Oil pressure < 1.0 bar	221	0.403 [0.340, 0.469]	0.329	inconclusive
Oil pressure < 0.5 bar	118	0.356 [0.275, 0.446]	0.849	inconclusive
Oil pressure > 6.0 bar	31	0.226 [0.114, 0.398]	0.135	inconclusive
Oil temperature > 85 °C	950	0.466 [0.435, 0.498]	<0.001	agrees
Coolant temperature > 93 °C	71	0.324 [0.227, 0.439]	0.463	inconclusive
Coolant temperature < 65 °C	102	0.245 [0.172, 0.337]	0.010	contradicts
Fuel pressure < 2.0 bar	387	0.561 [0.511, 0.609]	<0.001	agrees
Engine speed < 300 rpm	22	0.000 [0.000, 0.149]	<0.001	contradicts

Base rate 0.369. Physics: Wilson interval entirely above (agrees) or below (contradicts) the base rate; otherwise inconclusive. Exact binomial test against the base rate.

3.3. Predictive Performance of the Benchmark

Table 5 and Figure 4a report the repeated-CV results on the training set. All learning models clearly outperformed the majority-class baseline (all $p < 0.001$), but the achievable discrimination is modest, with ROC-AUC around 0.70. EBM ranked first (0.7048 ± 0.0056). Non-significant differences do not establish equivalence, so the equivalence tests are decisive here. Within a margin of ±0.01, CatBoost, XGB-Free, LightGBM, XGB-Domain and Random Forest were statistically equivalent to EBM (Holm-adjusted TOST $p < 0.05$). Equivalence could not be shown for Logistic Regression and XGB-Reversed.

The hold-out results (Table 6, Figure 4b,c) agree with the CV ranking. CatBoost achieved the highest test ROC-AUC (0.708; 95% CI 0.691–0.725), and XGB-Reversed the lowest among the learning models (0.690). Test values are within about 0.01 of their CV counterparts, so there is no sign of overfitting or leakage. The glass-box EBM matches the black-box ensembles, in line with the argument that interpretable models need not sacrifice accuracy on tabular data [11].

A ROC-AUC near 0.70 limits practical use. At the chosen thresholds roughly a third of the records are misclassified (Table 7), which would be too many for alarm generation on a real engine. Published studies of the same data report much higher figures, for example 97% accuracy for a neural network trained on the 174 source records [26] and AUCs above 0.98 after oversampling [28]. Those figures are not directly comparable with ours, which use all records, the original labels and a single evaluation on an untouched hold-out set. The present study does not propose a deployable detector. Its aim is to test whether the relationships that any detector would learn from these labels are physically plausible, and this question can be answered at modest discrimination.

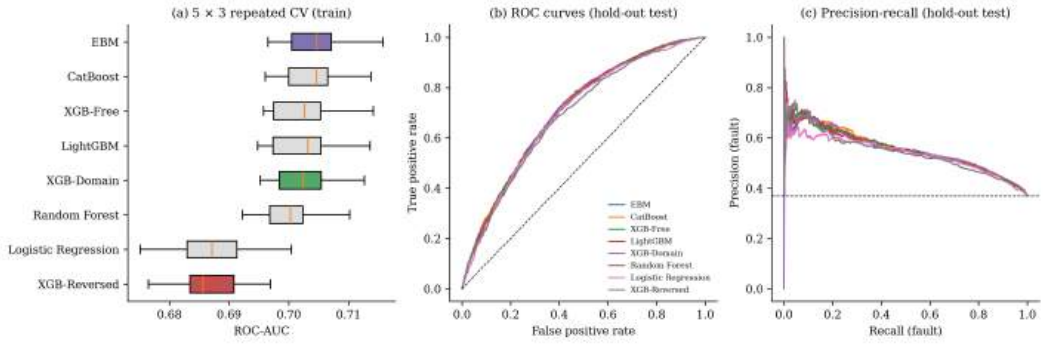


Figure 4. Predictive performance: (a) ROC-AUC over 15 repeated CV folds on the training set; (b) ROC curves and (c) precision-recall curves for the fault class on the hold-out test set. The dashed line in (c) marks the fault prevalence.

Table 5. Five-fold, three-times repeated cross-validation on the training set, ordered by mean ROC-AUC, with difference, equivalence and Bayesian tests against the best model.

Model	ROC-AUC (mean ± SD)	PR-AUC (fault)	PR-AUC (healthy)	Brier	ΔAUC vs. EBM [95% CI]	*p* (Holm)	*p* TOST (Holm)	P(ROPE)
EBM	0.7048 ± 0.0056	0.561	0.802	0.205	—	—	—	—
CatBoost	0.7040 ± 0.0054	0.561	0.800	0.206	−0.0008 [−0.0021, +0.0006]	0.241	<0.001	1.00
XGB-Free	0.7027 ± 0.0052	0.558	0.799	0.206	−0.0021 [−0.0040, −0.0002]	0.071	<0.001	1.00
LightGBM	0.7023 ± 0.0054	0.557	0.799	0.206	−0.0025 [−0.0044, −0.0007]	0.032	<0.001	1.00
XGB-Domain	0.7022 ± 0.0052	0.557	0.799	0.206	−0.0026 [−0.0042, −0.0010]	0.015	<0.001	1.00
Random Forest	0.7000 ± 0.0051	0.555	0.797	0.207	−0.0048 [−0.0073, −0.0023]	0.005	<0.001	1.00
Logistic Regression	0.6873 ± 0.0069	0.541	0.783	0.211	−0.0175 [−0.0208, −0.0143]	<0.001	1.000	0.00
XGB-Reversed	0.6868 ± 0.0062	0.537	0.787	0.210	−0.0180 [−0.0234, −0.0127]	<0.001	1.000	0.00
Dummy	0.5000 ± 0.0000	0.369	0.631	0.369	—	—	—	—

Corrected resampled *t*-test [42] with Holm adjustment; TOST and Bayesian correlated *t*-test with a ROPE of ±0.01 ROC-AUC; P(ROPE): posterior probability of practical equivalence.

Table 6. Hold-out test performance (*n* = 3,907): discrimination and calibration.

Model	ROC-AUC [95% CI]	PR-AUC (fault)	PR-AUC (healthy)	Brier	BSS	Cal. slope	Cal. intercept
CatBoost	0.708 [0.691–0.725]	0.557	0.802	0.205	+0.121	1.04	+0.012
EBM (no interactions)	0.707 [0.689–0.723]	0.553	0.801	0.205	+0.119	1.00	+0.011
LightGBM	0.706 [0.689–0.723]	0.555	0.803	0.205	+0.119	0.99	+0.011
EBM	0.706 [0.689–0.723]	0.552	0.801	0.205	+0.118	0.97	+0.011
XGB-Domain	0.706 [0.689–0.722]	0.554	0.801	0.205	+0.119	1.00	+0.011
XGB-Free	0.706 [0.689–0.722]	0.555	0.801	0.205	+0.119	1.02	+0.014
Random Forest	0.704 [0.686–0.721]	0.556	0.799	0.207	+0.113	1.25	+0.013
Logistic Regression	0.696 [0.679–0.713]	0.539	0.790	0.210	+0.098	1.08	+0.003
XGB-Reversed	0.690 [0.673–0.708]	0.542	0.790	0.209	+0.102	0.98	+0.018
Dummy	0.500	0.370	0.630	0.370	−0.586	—	—

BSS: Brier skill score relative to the prior-probability baseline (Brier = 0.233). Calibration slope and intercept from logistic recalibration (ideal values 1 and 0).

3.4. Threshold-Dependent Performance and Calibration

The dummy classifier predicts the majority class, healthy, for every record. It therefore never detects a fault (recall 0.000) yet reaches an accuracy equal to the healthy share, which is why accuracy is not reported as a headline metric. For the learning models, Table 7 reports precision, recall and F1-score for both classes. For EBM, the fault class has precision 0.509 and recall 0.713, and the healthy class precision 0.780 and recall 0.598. The Youden criterion, MCC and balanced accuracy do not depend on which class is called positive, which is why they were used for threshold selection. Criteria such as the recall-

weighted F2-score do depend on it: in the first version of this study, where the majority class was coded as the fault class, F2-optimal thresholds collapsed every model to the trivial all-majority prediction.

The dummy Brier score (0.370) reflects hard majority-class predictions. Against the fairer prior-probability baseline (Brier = 0.233), the Brier skill of the learning models is only 0.10–0.12. Calibration slopes range from 0.97 to 1.25 and intercepts are close to zero (Table 6), so the models are reasonably calibrated, but their probabilistic skill is limited (Figure 5).

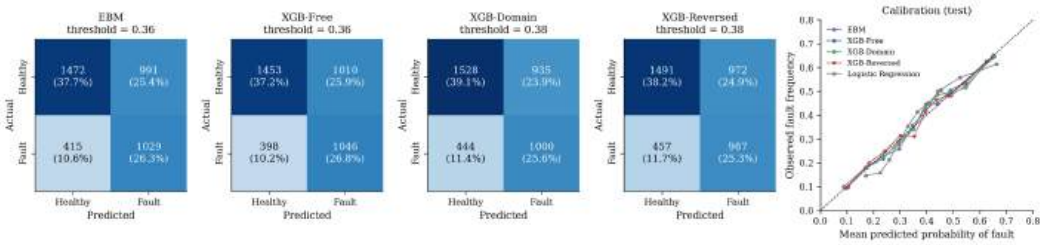


Figure 5. Confusion matrices on the hold-out test set at the frozen Youden thresholds for EBM and the three XGBoost configurations (common colour scale; fault = positive class), and calibration curves (quantile bins) for selected models.

Table 7. Hold-out test performance at the frozen Youden thresholds, for both classes.

					Fault class			Healthy class	
Model	**Thr.**	**Bal. acc.**	**MCC**	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**
CatBoost	0.37	0.659	0.307	0.514	0.712	0.597	0.782	0.606	0.683
EBM (no interactions)	0.38	0.658	0.304	0.520	0.688	0.592	0.774	0.627	0.693
LightGBM	0.36	0.659	0.308	0.509	0.732	0.601	0.789	0.586	0.673
EBM	0.36	0.655	0.300	0.509	0.713	0.594	0.780	0.598	0.677
XGB-Domain	0.38	0.656	0.302	0.517	0.693	0.592	0.775	0.620	0.689
XGB-Free	0.36	0.657	0.304	0.509	0.724	0.598	0.785	0.590	0.674
Random Forest	0.40	0.649	0.288	0.519	0.652	0.578	0.760	0.645	0.698
Logistic Regression	0.33	0.648	0.288	0.495	0.740	0.593	0.785	0.557	0.652
XGB-Reversed	0.38	0.644	0.279	0.504	0.683	0.580	0.765	0.605	0.676
Dummy	0.50	0.500	0.000	0.000	0.000	0.000	0.630	1.000	0.773

Thr.: threshold on the predicted fault probability, selected by Youden's J on out-of-fold training predictions.

3.5. Physical-Consistency Audit

The audit results are given in Table 8 and Figure 6. Imposing the engineering-consistent constraints changed ROC-AUC by only −0.0005 in CV (95% CI [−0.0023, +0.0013], p = 0.567) and by +0.0001 on the test set ([−0.0021, +0.0022]). Imposing the reversed constraints reduced it by 0.0159 in CV ([−0.0210, −0.0109], $p < 0.001$) and by 0.0153 on the test set ([−0.0233, −0.0073]). The physical-consistency gap is therefore negative: PCG = −0.0154 in CV ([−0.0208, −0.0101]) and −0.0154 on the test set ([−0.0232, −0.0073]). Under the decision rules of Section 2.7 the dataset is read as consistent with the assumed physics.

The result does not depend on shared hyper-parameters. With separately tuned constrained models, PCG = −0.0139 ([−0.0187, −0.0090]); under fully nested CV, PCG = −0.0133 ([−0.0175, −0.0091]). Constraining one sensor at a time (Table 9) gives single-sensor gaps of −0.0059 for oil pressure, −0.0077 for oil temperature, −0.0009 for coolant temperature. Their sum (−0.0145) is close to the joint PCG.

The consistency is carried by oil pressure and oil temperature, whereas coolant temperature alone gives an unresolved result; its learned direction is examined in Section 3.6.

The attribution analysis explains the mechanism (Figure 6b). Under reversed constraints, the shares of mean absolute SHAP value for oil pressure and oil temperature fell from 12.0% and 14.3% to 0.4% and 0.1%, and coolant temperature changed from 4.8% to 1.9%. The model compensated by relying more on engine speed and fuel pressure.

When forced to contradict the assumed physics, the model found no usable signal in the constrained sensors and discarded them. Under domain constraints the attribution profile was virtually identical to that of the unconstrained model, indicating that XGB-Free had already learned the physically expected directions.

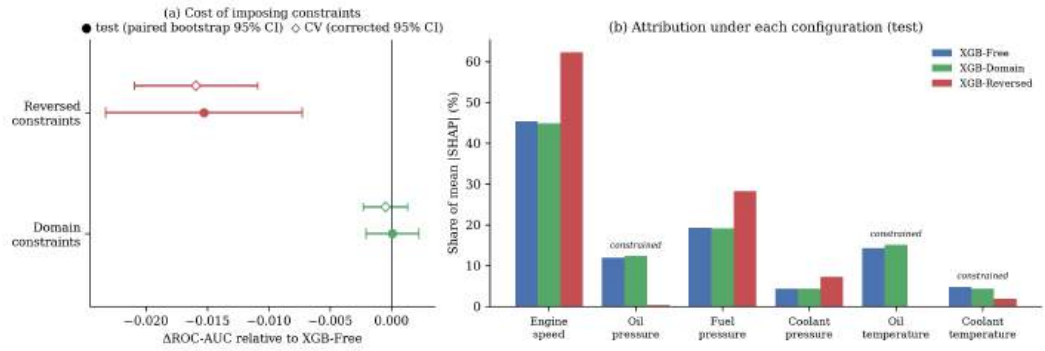


Figure 6. Physical-consistency audit. (a) Change in ROC-AUC caused by domain and reversed constraints relative to XGB-Free (filled markers: test-set paired bootstrap mean with 95% CI; open markers: CV mean with corrected 95% CI). (b) Share of mean absolute SHAP value per feature on the test set under each configuration.

Table 8. Physical-consistency audit: cost of constraints relative to XGB-Free and the physical-consistency gap (PCG, Equation (4)), under shared, separate and nested tuning.

Tuning	Quantity	CV estimate [95% CI]	CV p	Test estimate [95% CI]
Shared hyper-parameters	ΔAUC Domain	−0.0005 [−0.0023, +0.0013]	0.567	+0.0001 [−0.0021, +0.0022]
	ΔAUC Reversed	−0.0159 [−0.0210, −0.0109]	<0.001	−0.0153 [−0.0233, −0.0073]
	PCG	−0.0154 [−0.0208, −0.0101]	<0.001	−0.0154 [−0.0232, −0.0073]
Separately tuned	ΔAUC Domain	−0.0002 [−0.0024, +0.0019]	0.810	—
	ΔAUC Reversed	−0.0141 [−0.0190, −0.0093]	<0.001	—
	PCG	−0.0139 [−0.0187, −0.0090]	<0.001	−0.0125 [−0.0202, −0.0047]
Nested CV	ΔAUC Domain	−0.0004 [−0.0017, +0.0010]	0.562	—
	ΔAUC Reversed	−0.0137 [−0.0181, −0.0093]	<0.001	—
	PCG	−0.0133 [−0.0175, −0.0091]	<0.001	—

CV: 5 × 3 repeated CV on the training set with corrected 95% CI and corrected resampled t-test; test: paired bootstrap (2,000 resamples) on the hold-out set. Nested CV: each configuration re-tuned inside every outer fold.

Table 9. Audit with one sensor constrained at a time (5 × 3 CV, shared hyper-parameters).

Constrained sensor	ΔAUC Domain [95% CI]	ΔAUC Reversed [95% CI]	PCG [95% CI]	Reading
Oil pressure	−0.0001 [−0.0010, +0.0008]	−0.0061 [−0.0092, −0.0030]	−0.0059 [−0.0089, −0.0030]	consistent
Oil temperature	−0.0002 [−0.0010, +0.0006]	−0.0079 [−0.0113, −0.0044]	−0.0077 [−0.0109, −0.0045]	consistent
Coolant temperature	+0.0000 [−0.0011, +0.0011]	−0.0009 [−0.0030, +0.0013]	−0.0009 [−0.0035, +0.0016]	unresolved

Only the named sensor is constrained; all other sensors are unconstrained. Reading: decision rules of Section 2.7.

3.6. Learned Directions from SHAP and EBM

SHAP and EBM agree with the audit (Table 10, Figure 7). For all three constrained sensors, both methods learned the physically expected directions. The contribution to the log-odds of a fault decreases with oil pressure (SHAP $\rho = -0.76$, $\Delta = -0.73$) and increases with oil temperature ($\rho = +0.96$, $\Delta = +0.79$) and coolant temperature ($\rho = +0.30$, $\Delta = +0.22$). The EBM without interaction terms gives the same main-effect shape functions (Figure A1): InterpretML estimates the main effects before the pairwise terms, so the interactions do not absorb them. Because SHAP and EBM are fitted to the same labels, their agreement

shows that the learned directions are robust to the model class; it is not independent evidence about the labels.

Coolant temperature is the weakest of the three. Its SHAP and EBM contributions rise steeply up to about 72 °C, are nearly flat between 72 and 86 °C, and fall again above about 87 °C (Figure 7). The decrease at the hottest end runs against overheating physics, but this region contains few records and the tail analysis there is inconclusive (Table 4). This weak and non-monotone effect explains why constraining coolant temperature alone gives an unresolved audit (Table 9).

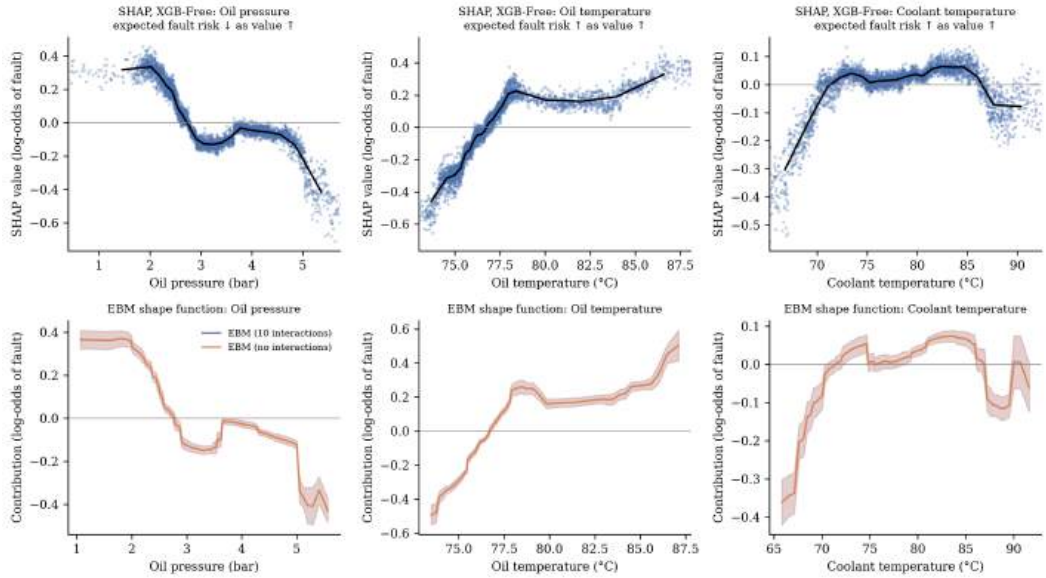


Figure 7. Learned effects of the three constrained sensors. Top: SHAP dependence of XGB-Free (points: test records; line: binned mean). Bottom: EBM shape functions with uncertainty bands, with and without interaction terms (the two curves coincide). The expected direction of fault risk is stated in each top panel title.

Table 10. Direction and size of learned effects on the test set, with bootstrap 95% confidence intervals.

Feature	Exp.	ρ SHAP [95% CI]	Δ SHAP [95% CI]	ρ EBM [95% CI]	Δ EBM [95% CI]	ρ EBM-0	Verdict
Engine speed	free	+0.99 [+0.99, +0.99]	+2.73 [+2.68, +2.78]	+1.00 [+1.00, +1.00]	+2.82 [+2.75, +2.87]	+1.00	n/a
Oil pressure	−	−0.76 [−0.78, −0.73]	−0.73 [−0.76, −0.70]	−0.76 [−0.78, −0.73]	−0.77 [−0.80, −0.73]	−0.76	Consistent
Fuel pressure	free	−0.95 [−0.96, −0.95]	−0.86 [−0.89, −0.83]	−0.98 [−0.99, −0.98]	−0.90 [−0.92, −0.88]	−0.98	n/a
Coolant pressure	free	+0.79 [+0.77, +0.81]	+0.21 [+0.18, +0.24]	+0.86 [+0.83, +0.88]	+0.20 [+0.17, +0.24]	+0.86	n/a
Oil temperature	+	+0.96 [+0.96, +0.97]	+0.79 [+0.76, +0.81]	+0.98 [+0.98, +0.98]	+0.82 [+0.80, +0.85]	+0.98	Consistent
Coolant temperature	+	+0.30 [+0.26, +0.34]	+0.22 [+0.20, +0.25]	+0.38 [+0.34, +0.42]	+0.22 [+0.19, +0.24]	+0.38	Consistent

Exp.: sign of the domain constraint; free: no directional expectation. Δ: Equation (6), change in log-odds of a fault from the lowest to the highest 5% of the feature. EBM-0: EBM without interaction terms. Consistent/Inconsistent require agreement of SHAP, EBM and EBM-0.

3.7. Semi-Synthetic Calibration

Table 11 and Figure 8 compare the real-data audit with the semi-synthetic scenarios. With consistent labels at the real effect size ($k = 1$), domain constraints were free (ΔAUC = +0.0000) and reversed constraints were costly (ΔAUC = −0.0138), giving PCG = −0.0139; the result was read as consistent. Inverting these labels produced the mirror image (PCG

= +0.0139, flagged). The symmetry follows by construction, because inverting all labels negates every learned effect.

The size of the signature grows with the strength of the constrained signal. For consistent labels, $\Delta AUC_{Reversed}$ was −0.0055, −0.0138, −0.0429 and −0.1203 for k = 0.5, 1, 2 and 4. The real reversed cost (−0.0159) lies between the values for k = 1 and k = 2. The real labels therefore behave like consistent labels of realistic strength. They are also compatible with a stronger consistent signal of which a minority of labels is randomly flipped: for k = 2, a reversed cost within 0.003 of the real value arises at q = 0.4. The audit cannot separate these two generating processes. What it can say is that most of the constrained signal follows the assumed physics.

Partial reversal gave intermediate values (PCG between −0.0122 and +0.0104). One of the 6 partial scenarios was flagged (reversed: oil pressure, oil temperature). One was nevertheless read as consistent (reversed: coolant temperature). A reversal confined to a sensor with a small coefficient can therefore go undetected, so a consistent reading means that most, not necessarily all, of the constrained signal follows the physics. Across the 30 null replicates, PCG ranged from −0.0352 to +0.0084. None was flagged, and none was read as consistent (Table 12, Figure 8c). The U-shaped rule N2 produced clearly negative PCG values, but its domain cost was significant, which is why the consistent reading requires a negligible domain cost. In the graded-flip series, however, the scenario with q = 0.5 and k = 1, whose labels carry no information (Free AUC = 0.486), was read as consistent (Table A3). The readings therefore presuppose an unconstrained model that discriminates better than chance, which should be checked first; the real labels (Free AUC = 0.703) meet this condition. The semi-synthetic experiment is deliberately simple and does not model the true data-generating process. Its purpose is to calibrate how the audit should be read.

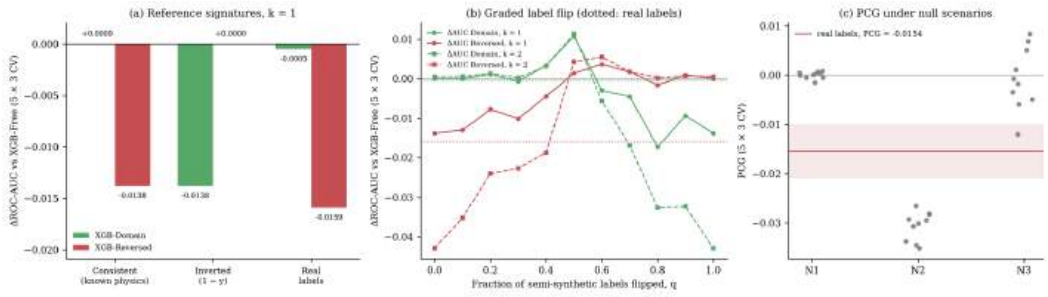


Figure 8. Semi-synthetic calibration (5 × 3 CV). (a) ΔROC-AUC relative to XGB-Free for consistent and inverted labels at the real effect size and for the real labels. (b) Graded label flip for two signal strengths; dotted lines mark the real-data values. (c) PCG in 10 replicates of each null scenario; the line and band show the real PCG and its 95% CI.

Table 11. Audit signatures under semi-synthetic labels with known physics and under the real labels (5 × 3 CV, ΔAUC relative to XGB-Free).

Scenario	Fault rate	Free AUC	ΔAUC Domain	ΔAUC Reversed	PCG [95% CI]	Reading
Consistent labels, k = 1	0.371	0.672	+0.0000	−0.0138	−0.0139 [−0.0182, −0.0095]	consistent
Inverted labels, k = 1	0.629	0.672	−0.0138	+0.0000	+0.0139 [+0.0095, +0.0182]	flagged
Consistent, k = 0.5	0.371	0.663	+0.0000	−0.0055	−0.0055 [−0.0083, −0.0026]	consistent
Inverted, k = 0.5	0.629	0.663	−0.0055	+0.0000	+0.0055 [+0.0026, +0.0083]	flagged
Consistent, k = 2	0.373	0.699	+0.0004	−0.0429	−0.0433 [−0.0506, −0.0361]	consistent
Inverted, k = 2	0.627	0.699	−0.0429	+0.0004	+0.0433 [+0.0360, +0.0506]	flagged
Consistent, k = 4	0.371	0.765	+0.0001	−0.1203	−0.1204 [−0.1303, −0.1104]	consistent
Inverted, k = 4	0.629	0.765	−0.1203	+0.0001	+0.1204 [+0.1104, +0.1303]	flagged
Reversed: Oil pressure	0.370	0.671	−0.0042	−0.0065	−0.0023 [−0.0087, +0.0041]	unresolved
Reversed: Oil temperature	0.371	0.668	−0.0054	−0.0075	−0.0021 [−0.0084, +0.0041]	unresolved
Reversed: Coolant temperature	0.370	0.671	−0.0008	−0.0130	−0.0122 [−0.0168, −0.0076]	consistent
Reversed: Oil temperature, Coolant temperature	0.372	0.667	−0.0062	−0.0058	+0.0004 [−0.0079, +0.0087]	unresolved

Scenario	Fault rate	Free AUC	ΔAUC Domain	ΔAUC Reversed	PCG [95% CI]	Reading
Reversed: Oil pressure, Coolant temperature	0.372	0.669	−0.0046	−0.0042	+0.0004 [−0.0062, +0.0069]	unresolved
Reversed: Oil pressure, Oil temperature	0.375	0.668	−0.0104	+0.0000	+0.0104 [+0.0044, +0.0165]	flagged
Real dataset labels	0.369	0.703	−0.0005	−0.0159	−0.0154 [−0.0208, −0.0101]	consistent

Partial reversal: the physical sign is reversed only for the named sensors. Reading: decision rules of Section 2.7. Graded-flip scenarios are listed in Table A3.

Table 12. PCG under null scenarios (10 replicates each, 5 × 3 CV).

Scenario	Replicates	Mean PCG	Range	Share flagged	Share consistent
N1 unconstrained sensors only	10	+0.0000	[−0.0015, +0.0009]	0.0	0.0
N2 non-monotone (U-shaped) rule	10	−0.0306	[−0.0352, −0.0266]	0.0	0.0
N3 permuted real labels	10	−0.0007	[−0.0120, +0.0084]	0.0	0.0
Real dataset labels	1	−0.0154	—	0.0	1.0

N1: labels depend only on the unconstrained sensors; N2: non-physical U-shaped rule on the constrained sensors; N3: randomly permuted real labels.

3.8. Positive Control on the AI4I 2020 Dataset

On the AI4I 2020 dataset, whose failure mechanisms are documented, the audit behaved as expected (Table 13). With the documented labels, domain constraints were almost free (ΔAUC = +0.0001) and reversed constraints were costly (ΔAUC = −0.0909), giving PCG = −0.0910 ([−0.1254, −0.0566]); the result was read as consistent. With inverted labels, the gap changed sign (PCG = +0.0910, [+0.0566, +0.1254]) and the dataset was flagged. The hold-out test gave the same picture. The audit thus recognises correct physics as correct and inverted labels as inverted on a dataset where the truth is known.

Table 13. Positive control: audit on the AI4I 2020 dataset with documented and inverted labels.

Labels	Free AUC	ΔAUC Domain [95% CI]	ΔAUC Reversed [95% CI]	CV PCG [95% CI]	Test PCG [95% CI]	Reading
Documented labels	0.978	+0.0001 [−0.0014, +0.0017]	−0.0909 [−0.1258, −0.0560]	−0.0910 [−0.1254, −0.0566]	−0.0711 [−0.1092, −0.0395]	consistent
Inverted labels (1 − y)	0.978	−0.0909 [−0.1258, −0.0560]	+0.0001 [−0.0014, +0.0017]	+0.0910 [+0.0566, +0.1254]	+0.0711 [+0.0395, +0.1092]	flagged

Constraints: temperature difference −1, tool wear +1; product type, air temperature, speed and torque unconstrained. 5 × 3 CV on the training part; test: paired bootstrap on the 20% hold-out part.

3.9. Implications for the Label Convention and the Data

Under the documented coding, the audit reads the labels as consistent with the assumed physics, and the semi-synthetic calibration, the null scenarios and the positive control show how this reading should be interpreted. Inverting all labels negates every learned effect and leaves ROC-AUC unchanged, so the opposite coding exchanges the roles of the two constraint sets. Running the complete protocol with Engine Condition = 1 as the fault class confirmed this: ΔAUC_{Domain} = −0.0142, $\Delta AUC_{Reversed}$ = −0.0007 and PCG = +0.0136 ([+0.0090, +0.0181]), which is flagged; the small differences from the mirror image of the present results come from re-tuning. This is the coding of the first version of this study. The repository record does not define the coding, and a user question on the record asking which value means good has no answer there [25]. A wrong convention is therefore an easy mistake, and it is invisible to accuracy, calibration and label-noise detectors [22], because swapping the classes leaves model confidence unchanged. The audit detected it from directional knowledge alone. This is the most practical use of the method shown here: a quick check of the label convention of a poorly documented dataset.

Three observations limit what the consistent reading means. First, the tail analysis is mixed: some extremes agree with physics, others contradict it, and the low-oil-pressure tail is uninformative. Second, the data are synthetic (Section 2.1). The fault class derives from three genuine fault observations, oversampled and expanded by a generator, and

the near-zero inter-sensor correlations show that the sensors were not generated as a coupled engine system. The consistent reading therefore shows that the generated labels preserve the expected dominant directions of oil pressure and oil temperature; it does not show that they describe real engine failures. Third, directional agreement can also arise from confounding, from sensors that are missing from the file, or from monotone assumptions that are too simple. The audit tests agreement between labels and assumed physics; it cannot by itself establish that the labels are correct.

3.10. Practical Implications and Limitations

For practitioners, the audit is inexpensive. It requires two additional model fits and a short list of directional expectations that domain experts can readily supply, and it should be run before a detector is trusted, in particular when the label definition is undocumented or taken from a secondary source. A flag should trigger a review of the label coding and of the data-collection or simulation process. Models trained on the present synthetic data, including EBM, should not be deployed on real engines; EBM is recommended here as an audit instrument, because it exposes its learned effects directly to engineers.

Several limitations apply. First, the analysis uses one synthetic maritime dataset and one positive control, and the achievable discrimination is modest (ROC-AUC ≈ 0.70). Second, the benchmark tuning is not nested; the hold-out results and the nested audit mitigate but do not remove this issue. Third, manufacturer alarm set points were not available, so the tail regions are conservative extremes. Fourth, monotonic constraints cannot express non-monotone risk such as that of engine speed, fuel pressure or very low coolant temperature; shape constraints beyond monotonicity would be needed. Fifth, the file contains independent snapshots without machine identifiers or timestamps, which restricts the task to fault detection rather than prognosis. Future work should apply the audit to measured datasets with documented labels and run-to-failure histories and extend it to non-monotone domain knowledge.

4. Conclusions

This study proposed and evaluated a physical-consistency audit for engine fault detection that combines domain-informed monotonic XGBoost with the glass-box Explainable Boosting Machine. On a public marine diesel engine dataset, the best classifiers reached a modest test ROC-AUC of about 0.71. Addressing the first research gap, the three-way audit showed that, under the label coding documented by the dataset creator, engineering-consistent constraints had no measurable cost, whereas reversed constraints reduced the test ROC-AUC by 0.015 and removed almost all reliance on oil pressure and oil temperature. The resulting physical-consistency gap (PCG = −0.0154) was robust to separate and nested tuning, and SHAP and EBM agreed on the expected directions. Addressing the second research gap, semi-synthetic calibration, null scenarios and a positive control showed how the gap should be read, and the opposite label coding, used in the first version of this study, was flagged. The audit thus detects a wrong label convention, which accuracy-based evaluation cannot. Mixed tail evidence and the synthetic origin of the data show that consistent directions do not by themselves make labels valid for real engines. Physical-consistency auditing should therefore complement accuracy-based evaluation, and label definitions should be documented and checked before fault-detection models are trusted in maintenance practice.

Supplementary Materials: The Python code of the revised experiment (audit_experiment_v3.py), the result tables (results_v3.xlsx), the summary file (summary_v3.json), the package versions (environment.txt) and all figures are provided as supplementary files and are archived on Zenodo (https://doi.org/10.5281/zenodo.23728914).

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, and visualization, A.A.A. The author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The ship main-engine dataset is publicly available on IEEE DataPort (https://doi.org/10.21227/g3za-v415) [25]; the file engine_data.csv was downloaded on 29 June 2026 (MD5 629ad357fe0e5d0e6beb7421e2ba2c32). The AI4I 2020 dataset is available from the UCI Machine Learning Repository [32]. The code needed to reproduce all results is provided in the Supplementary Materials and on Zenodo (https://doi.org/10.5281/zenodo.23728914).

Acknowledgments: The author thanks the providers of the public engine-condition dataset.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Table A1. Hyper-parameters selected by randomised search with three-fold CV on the training set.

Model	Tuned hyper-parameters	Inner-CV ROC-AUC
Logistic Regression	C: 0.006158482110660267	0.6876
Random Forest	min_samples_leaf: 10, max_features: sqrt, max_depth: 6	0.7000
LightGBM	subsample: 0.85, num_leaves: 7, n_estimators: 200, min_child_samples: 50, learning_rate: 0.05, colsample_bytree: 1.0	0.7025
CatBoost	learning_rate: 0.05, l2_leaf_reg: 10, iterations: 200, depth: 3	0.7047
XGBoost (XGB-Free; shared by the main audit)	subsample: 0.7, n_estimators: 200, min_child_weight: 20, max_depth: 4, learning_rate: 0.03, colsample_bytree: 0.85	0.7031
EBM	max_bins: 128, learning_rate: 0.03, interactions: 10	0.7057
XGB-Domain (own tuning)	subsample: 0.85, n_estimators: 500, min_child_weight: 10, max_depth: 4, learning_rate: 0.01, colsample_bytree: 0.7	0.7029
XGB-Reversed (own tuning)	subsample: 1.0, n_estimators: 100, min_child_weight: 20, max_depth: 3, learning_rate: 0.05, colsample_bytree: 0.85	0.6888
XGBoost (AI4I 2020 positive control)	subsample: 0.7, n_estimators: 300, min_child_weight: 1, max_depth: 4, learning_rate: 0.01, colsample_bytree: 1.0	—

EBM without interactions uses the EBM hyper-parameters with interactions = 0.

Table A2. Sensitivity of the audit to the cleaning rule (5 × 3 CV on the respective training partition).

Variant	Records	ΔAUC Domain	ΔAUC Reversed	PCG [95% CI]	Reading
Main analysis (coolant temperature > 110 °C removed)	19,533	−0.0005	−0.0159	−0.0154 [−0.0208, −0.0101]	consistent
No record removed	19,535	−0.0005	−0.0155	−0.0151 [−0.0209, −0.0093]	consistent
Only the 195.5 °C record removed	19,534	−0.0004	−0.0161	−0.0156 [−0.0201, −0.0112]	consistent
Also oil pressure < 0.5 bar removed	19,415	−0.0007	−0.0164	−0.0156 [−0.0202, −0.0110]	consistent

Table A3. Graded label flip: audit signatures when a fraction q of consistent semi-synthetic labels is inverted.

k	q	ΔAUC Domain	ΔAUC Reversed	PCG [95% CI]	Reading
1	0.0	+0.0000	−0.0138	−0.0139 [−0.0182, −0.0095]	consistent
1	0.1	+0.0000	−0.0130	−0.0130 [−0.0171, −0.0090]	consistent
1	0.2	+0.0011	−0.0077	−0.0088 [−0.0151, −0.0025]	consistent
1	0.3	−0.0007	−0.0101	−0.0094 [−0.0186, −0.0002]	consistent
1	0.4	+0.0034	−0.0045	−0.0078 [−0.0212, +0.0055]	unresolved
1	0.5	+0.0108	+0.0014	−0.0094 [−0.0162, −0.0025]	consistent
1	0.6	−0.0030	+0.0037	+0.0067 [−0.0047, +0.0180]	unresolved
1	0.7	−0.0045	+0.0017	+0.0062 [−0.0011, +0.0136]	unresolved
1	0.8	−0.0172	−0.0016	+0.0157 [+0.0057, +0.0257]	flagged
1	0.9	−0.0093	+0.0009	+0.0102 [+0.0023, +0.0181]	flagged
1	1.0	−0.0138	+0.0000	+0.0139 [+0.0095, +0.0182]	flagged
2	0.0	+0.0004	−0.0429	−0.0433 [−0.0506, −0.0361]	consistent
2	0.1	+0.0005	−0.0352	−0.0357 [−0.0425, −0.0289]	consistent
2	0.2	+0.0013	−0.0240	−0.0253 [−0.0341, −0.0164]	consistent
2	0.3	+0.0002	−0.0226	−0.0228 [−0.0316, −0.0140]	consistent

k	*q*	ΔAUC Domain	ΔAUC Reversed	PCG [95% CI]	Reading
2	0.4	+0.0032	−0.0187	−0.0220 [−0.0301, −0.0139]	consistent
2	0.5	+0.0114	+0.0043	−0.0070 [−0.0214, +0.0073]	unresolved
2	0.6	−0.0056	+0.0055	+0.0111 [+0.0022, +0.0200]	flagged
2	0.7	−0.0168	+0.0017	+0.0186 [+0.0093, +0.0279]	flagged
2	0.8	−0.0326	+0.0002	+0.0327 [+0.0246, +0.0409]	flagged
2	0.9	−0.0323	+0.0007	+0.0330 [+0.0247, +0.0413]	flagged
2	1.0	−0.0429	+0.0004	+0.0433 [+0.0360, +0.0506]	flagged

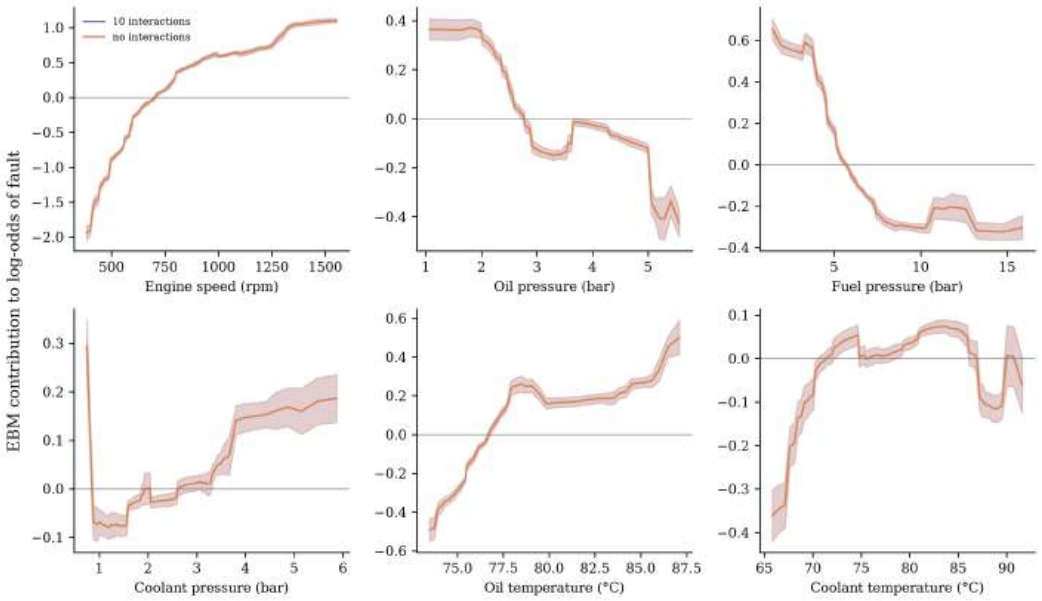


Figure A1. EBM shape functions for all six sensor variables (contribution to the log-odds of a fault), with and without interaction terms, trimmed to the 0.5–99.5th percentile range.

References

1. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510.
2. 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.
3. Mohakul, D.; Kumar, C.R.S.; Singh, S.; Katti, S. Condition based predictive maintenance on ship's major equipment using AI. *Int. J. Sci. Res. Eng. Manag.* **2023**, *7*(2), 1–8. https://doi.org/10.55041/IJSREM17699.
4. Carvalho, T.P.; Soares, F.A.A.M.N.; Vita, R.; Francisco, R.P.; Basto, J.P.; Alcalá, S.G.S. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024.
5. Zhang, W.; Yang, D.; Wang, H. Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Syst. J.* **2019**, *13*, 2213–2227.
6. Cummins, L.; Sommers, A.; Ramezani, S.B.; Mittal, S.; Jabour, J.; Seale, M.; Rahimi, S. Explainable predictive maintenance: A survey of current methods, challenges and opportunities. *IEEE Access* **2024**, *12*, 57574–57602. https://doi.org/10.1109/ACCESS.2024.3391130.
7. Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; Elhadad, N. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 1721–1730.
8. Geirhos, R.; Jacobsen, J.-H.; Michaelis, C.; Zemel, R.; Brendel, W.; Bethge, M.; Wichmann, F.A. Shortcut learning in deep neural networks. *Nat. Mach. Intell.* **2020**, *2*, 665–673.
9. Lapuschkin, S.; Wäldchen, S.; Binder, A.; Montavon, G.; Samek, W.; Müller, K.-R. Unmasking Clever Hans predictors and assessing what machines really learn. *Nat. Commun.* **2019**, *10*, 1096.
10. Ross, A.S.; Hughes, M.C.; Doshi-Velez, F. Right for the right reasons: Training differentiable models by constraining their explanations. In Proceedings of the 26th International Joint

Conference on Artificial Intelligence (IJCAI-17), Melbourne, Australia, 19–25 August 2017; pp. 2662–2670. https://doi.org/10.24963/ijcai.2017/371.
11. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215.
12. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 4765–4774.
13. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67.
14. Karpatne, A.; Atluri, G.; Faghmous, J.H.; Steinbach, M.; Banerjee, A.; Ganguly, A.; Shekhar, S.; Samatova, N.; Kumar, V. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Trans. Knowl. Data Eng.* **2017**, *29*, 2318–2331.
15. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440.
16. Gupta, M.; Cotter, A.; Pfeifer, J.; Voevodski, K.; Canini, K.; Mangylov, A.; Moczydlowski, W.; van Esbroeck, A. Monotonic calibrated interpolated look-up tables. *J. Mach. Learn. Res.* **2016**, *17*, 1–47.
17. Koklev, P. What's the price of monotonicity? A multi-dataset benchmark of monotone-constrained gradient boosting for credit PD. *arXiv* **2025**, arXiv:2512.17945.
18. Abbas, S.; Thomas, D.; Soomro, N.; Shafik, R.; Heer, R.; Adhikari, K. Fixing a model that learned worse cancer means lower risk: Monotonic constraints in bladder cancer recurrence prediction. *arXiv* **2026**, arXiv:2610.00858.
19. Hsiao, C.-H.; Kumar, K.; Rathje, E.M. Domain-informed explainable boosting machines for trustworthy lateral spread predictions. *arXiv* **2026**, arXiv:2603.17175.
20. Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J.W.; Wallach, H.; Daumé III, H.; Crawford, K. Datasheets for datasets. *Commun. ACM* **2021**, *64*, 86–92.
21. Frénay, B.; Verleysen, M. Classification in the presence of label noise: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 845–869.
22. Northcutt, C.; Jiang, L.; Chuang, I. Confident learning: Estimating uncertainty in dataset labels. *J. Artif. Intell. Res.* **2021**, *70*, 1373–1411.
23. Kaufman, S.; Rosset, S.; Perlich, C.; Stitelman, O. Leakage in data mining: Formulation, detection, and avoidance. *ACM Trans. Knowl. Discov. Data* **2012**, *6*, 15.
24. Kapoor, S.; Narayanan, A. Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* **2023**, *4*, 100804.
25. Mohakul, D. Predictive Maintenance on Ship's Main Engine using AI [Dataset]. IEEE DataPort, 16 November 2022. https://doi.org/10.21227/g3za-v415 (file engine_data.csv, accessed on 29 June 2026).
26. Nurdin, F.; Effendi, M.K.; Mohakul, D. Machine learning-based prediction of diesel engine health using operational parameters: Comparison of SVM and BPNN models. *Int. J. Mech. Eng. Sci.* **2025**, *9*, 69–80. https://doi.org/10.12962/j25807471.v9i2.10511.
27. Chukwudi, I.J.; Zaman, N.; Rahim, M.A.; Rahman, M.A.; Alenazi, M.J.F.; Pillai, P. An ensemble deep learning model for vehicular engine health prediction. *IEEE Access* **2024**, *12*, 63433–63451. https://doi.org/10.1109/ACCESS.2024.3395927.
28. Harrou, F.; Dorbane, A.; Sun, Y. Explainable machine learning with SMOTE for predictive maintenance of ship engines. *Proc. Inst. Mech. Eng. Part M J. Eng. Marit. Environ.* **2025**, online ahead of print. https://doi.org/10.1177/14750902251400373.
29. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
30. Lou, Y.; Caruana, R.; Gehrke, J.; Hooker, G. Accurate intelligible models with pairwise interactions. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 623–631.
31. Nori, H.; Jenkins, S.; Koch, P.; Caruana, R. InterpretML: A unified framework for machine learning interpretability. *arXiv* **2019**, arXiv:1909.09223.
32. Matzka, S. Explainable artificial intelligence for predictive maintenance applications. In Proceedings of the 2020 Third International Conference on Artificial Intelligence for Industries (AI4I), Irvine, CA, USA, 21–23 September 2020; pp. 69–74. https://doi.org/10.1109/AI4I49448.2020.00023.

33. Patki, N.; Wedge, R.; Veeramachaneni, K. The Synthetic Data Vault. In Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, 17–19 October 2016; pp. 399–410. https://doi.org/10.1109/DSAA.2016.49.
34. Mohakul, D.; Kumar, C.; Singh, S.; Katti, S.; Chougule, S. Health monitoring of ship's engine with simulated data by using classifiers—preliminary result. In Proceedings of the 2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 5–7 April 2023; pp. 1–5.
35. Varma, S.; Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* **2006**, *7*, 91.
36. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35.
37. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
38. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
39. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A highly efficient gradient boosting decision tree. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 3146–3154.
40. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 3–8 December 2018; pp. 6638–6648.
41. Woodyard, D. *Pounder's Marine Diesel Engines and Gas Turbines*, 9th ed.; Butterworth-Heinemann: Oxford, UK, 2009.
42. Nadeau, C.; Bengio, Y. Inference for the generalization error. *Mach. Learn.* **2003**, *52*, 239–281.
43. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall: New York, NY, USA, 1993.
44. Wilson, E.B. Probable inference, the law of succession, and statistical inference. *J. Am. Stat. Assoc.* **1927**, *22*, 209–212.
45. Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE* **2015**, *10*, e0118432.
46. Van Calster, B.; McLernon, D.J.; van Smeden, M.; Wynants, L.; Steyerberg, E.W. Calibration: The Achilles heel of predictive analytics. *BMC Med.* **2019**, *17*, 230.
47. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6.
48. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
49. Schuirmann, D.J. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J. Pharmacokinet. Biopharm.* **1987**, *15*, 657–680.
50. Benavoli, A.; Corani, G.; Demšar, J.; Zaffalon, M. Time for a change: A tutorial for comparing multiple classifiers through Bayesian analysis. *J. Mach. Learn. Res.* **2017**, *18*, 1–36.

RiESTech
Recent in Engineering Science and Technology

MBI

Article

Analysis of the Effect of Air Inlet Filter Replacement on Gas Turbine Performance at Unit 1A of Bekasi Power Combined Cycle Power Plant

Gilang Fitra Ardiansyah [1], Arifia Ekayuliana [2,*]

[1] Bachelor of Applied Energy Generation Engineering Technology Program, Department of Mechanical Engineering, Politeknik Negeri Jakarta, Depok 16424, Indonesia,
[2] Department of Mechanical Engineering, Politeknik Negeri Jakarta, Depok 16424, Indonesia
* Correspondence: arifia.ekayuliana@mesin.pnj.ac.id

Abstract: Air inlet filters protect gas turbine compressors from fouling, yet quantitative evidence of how filter replacement changes the performance of an operating unit remains limited. This study evaluates the effect of air inlet filter replacement on the GT1A gas turbine at Unit 1A of Bekasi Power Combined Cycle Power Plant. Secondary data from the control system and operator check sheets were compared before (11 November 2025) and after (7 January 2026) replacement, using snapshots at 00:00, 12:00 and 16:00. Compressor, combustion-chamber, turbine and thermal efficiencies were calculated from Brayton-cycle relations, and the sensitivity to hot-gas properties was tested. Filter differential pressure fell from 76–82 to 45–46 mmH_2O. Compressor efficiency rose from 78.85 – 81.76% to 88.71–90.88%, and thermal efficiency from 18.23–21.58% to 27.92–30.82%, while generator load increased from 37.4 to 38.5–38.9 MW. A fishbone (6M) analysis supported by failure mode and effect analysis identified environmental factors, mainly airborne dust, as the dominant cause of filter replacement. The improvement coincides with filter replacement and a compressor wash performed in the same period, so the effect of the filter alone cannot be separated.

Keywords: Engine fault detection; Marine diesel engine; Physical consistency; Monotonic constraints; Explainable boosting machine; SHAP; Label convention; Positive control

Citation: Ardiansyah, G. F., Ekayulianan, A. (2026). Analysis of the Effect of Air Inlet Filter Replacement on Gas Turbine Performance at Unit 1A of Bekasi Power Combined Cycle Power Plant. Recent in Engineering Science and Technology, 4(4), 245–254. Retrieved from https://www.mbi-jounals.com/index.php/riestech/article/view/175.

Academic Editor: Yudhi Ariadi

Received: 22 July 2026
Accepted: 19 October 2026
Published: 31 October 2026

Publisher's Note: MBI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2026 by the authors. Licensee MBI, Jakarta, Indonesia. This article is an open access article distributed under MBI license (https://mbi-journals.com/licenses/by/4.0/).

1. Introduction

Gas turbines are widely used in simple-cycle and combined-cycle plants because of their fast start-up and operational flexibility. The air drawn from the atmosphere passes through an inlet filter house before it enters the compressor, so the condition of the filter governs the pressure, cleanliness and mass flow of the air supplied to the whole cycle [1,2]. Particles that penetrate the filter cause fouling, erosion and corrosion of the compressor blades, which lowers airflow, pressure ratio and compressor efficiency and, in turn, reduces power output and thermal efficiency [3–5]. At the same time, dust loading raises the pressure drop across the filter house, which acts as an inlet pressure loss [2,6].

Earlier studies have addressed parts of this problem. Simanjuntak et al. [7] compared the compressor of a gas turbine before and after air inlet filter replacement at a power plant in Batam, Indonesia, and reported compressor efficiency gains of about 2 and 3 percentage points at 38 and 39 MW. Purwanto [8] analysed how the compressor inlet temperature affects compressor, combustion-chamber and thermal efficiency of a gas turbine at the Grati power plant. Effiom et al. [9] investigated inlet air filtration for offshore gas turbines experimentally, and Jin et al. [6] proposed a model-based approach to diagnose filter condition in gas turbine air intake systems. Köse and Yılmazoğlu [10] examined the effect of inlet-filter blockage within a study of periodic gas turbine maintenance and determined a cost-optimal filter replacement period. Kurz and Brun [3] and Meher-Homji and Bromley [4] reviewed degradation and fouling mechanisms in

general. These works concentrate on a single performance indicator, on ambient-temperature effects, or on laboratory and model-based diagnosis.

To the authors' knowledge, field evidence that links the measured filter differential pressure with the compressor, combustion-chamber, turbine and thermal efficiencies of one operating industrial unit before and after a filter replacement, together with a structured analysis of why the replacement was needed, is still limited. The present study addresses this gap for the GT1A gas turbine at Unit 1A of Bekasi Power Combined Cycle Power Plant. Control-system and check-sheet data from before and after replacement are used to recompute the four efficiencies from Brayton-cycle relations, including a sensitivity check on hot-gas properties, and fishbone (6M) and failure mode and effect analysis (FMEA) are applied to identify the causes of filter deterioration [11–13]. The results show a clear improvement of all efficiencies after replacement, but also that the filter pressure loss alone cannot explain its full size, which is discussed in Section 3.

2. Materials and Experiment Methods

2.1. Research Object

The object is the GT1A gas turbine and its air inlet filter system at Unit 1A of Bekasi Power Combined Cycle Power Plant (Cikarang, West Java, Indonesia). The unit is a General Electric PG6581B gas turbine generator with a nominal output of 37.97 MW, fired with natural gas and monitored by a Mark VI control system (DCS). The filter elements were installed on 1 November 2022 and replaced on 19 November 2025.

2.2. Data and Sampling

The study is quantitative and comparative and uses secondary data from the DCS and the operators' two-hourly check sheets. Purposive sampling was applied to select days with relatively stable operation: 11 November 2025 (before replacement, eight days before the physical replacement) and 7 January 2026 (after replacement). On each day three snapshots were taken: night (00:00), afternoon (12:00) and evening (16:00). Each value is a single observation logged at the stated hour (n = 1 per snapshot), not an average. The inputs are compressor inlet temperature (T_1) and pressure (P_1), compressor discharge temperature (T_2) and pressure (P_2), turbine inlet temperature (T_3), exhaust temperature (T_4), fuel gas flow ($\dot{m}_f$), generator load, and the differential pressure of the filter house (ΔP_f, "DP Inlet Filter" on the check sheet, in mmH_2O; 1 mmH_2O = 9.80665 Pa). Pressures are absolute (kPa A) as logged by the control system. All inputs are listed in Table 1.

2.3. Assumptions

The analysis assumes steady operation at each snapshot, ideal-gas behaviour, no pressure loss in the combustor, and expansion of the hot gas from P_2 to P_1 (exhaust at compressor inlet pressure). Because the real expansion ratio is lower (combustor and exhaust-duct losses), the ideal expansion work is overestimated and the calculated turbine efficiency is a conservative, lower estimate. T_3 is the combustion-chamber temperature measured and logged by the DCS. In the base case, air properties are used throughout (γ_a = 1.4), following the original calculation. Because this overestimates γ on the hot side, a sensitivity case uses γ_g = 4/3, $c_{p,g}$ = 1.148 kJ kg^{-1} K^{-1} for combustion products and $c_{p,a}$ = 1.005 kJ kg^{-1} K^{-1} for air [1].

2.4. Performance Calculation

2.4.1. Compressor and Combustion Chamber

The isentropic compressor outlet temperature and the compressor efficiency are [14,15]:

$$T_{2s} = T_1 \left(\frac{P_2}{P_1}\right)^{\frac{\gamma_a - 1}{\gamma_a}} \quad (1)$$

$$\eta_c = \frac{T_{2s} - T_1}{T_2 - T_1} \quad (2)$$

The combustion-chamber efficiency is defined, as in the original calculation, as a temperature-based index comparing the actual temperature rise with the rise that would follow from isentropic compression:

$$\eta_{cc} = \frac{T_3 - T_2}{T_3 - T_{2s}} \quad (3)$$

2.4.2. Turbine

$$T_{4s} = T_3 \left(\frac{P_1}{P_2}\right)^{\frac{\gamma_g - 1}{\gamma_g}} \quad (4)$$

$$\eta_t = \frac{T_3 - T_4}{T_3 - T_{4s}} \quad (5)$$

In the base case $\gamma_g = \gamma_a = 1.4$.

2.4.3. Thermal Efficiency and Mass Flows

Three mass flows are distinguished: the air mass flow $\dot{m}_a$ through the compressor, the fuel mass flow $\dot{m}_f$ (natural gas, from the DCS), and the gas mass flow through the turbine $\dot{m}_g$, with $\dot{m}_g = \dot{m}_a + \dot{m}_f$ (Equation (6)). The air mass flow is not recorded at this unit; the fuel flow is about 2% of the exhaust flow in the manufacturer's data, so $\dot{m}_g \approx \dot{m}_a$. Only efficiencies, which are ratios in which the mass flow cancels, are reported; absolute powers are therefore not computed from $\dot{m}_f$.

$$\dot{m}_g = \dot{m}_a + \dot{m}_f \quad (6)$$

$$W_c = \dot{m}_a c_{p,a}(T_2 - T_1), \quad W_t = \dot{m}_g c_{p,g}(T_3 - T_4) \quad (7)$$

$$Q_{in} = \dot{m}_g c_{p,g} T_3 - \dot{m}_a c_{p,a} T_2 \quad (8)$$

$$\eta_{th} = \frac{W_t - W_c}{Q_{in}} \quad (9)$$

In the base case (constant properties, $\dot{m}_g = \dot{m}_a$) Equation (9) reduces to

$$\eta_{th} = \frac{(T_3 - T_4) - (T_2 - T_1)}{T_3 - T_2} \quad (10)$$

This cycle efficiency is a temperature-based indicator for comparing operating conditions; it is not the fuel-based electrical efficiency of the plant.

2.5. Filter Service Life

The service life is $T = t_2 - t_1$ (Equation (11)), with t_1 the installation date and t_2 the replacement date, and the total operating hours are $H = T \times H_o$ (Equation (12)), with H_o = 4,032 h per year the annual operating hours from plant records.

$$T = t_2 - t_1, \quad H = T \times H_o \quad (11, 12)$$

2.6. Fishbone and FMEA Analysis

Causes of filter deterioration were collected by field observation and interviews with five operation and maintenance respondents and grouped into the six categories Man,

Method, Machine, Material, Measurement and Environment [11,12]. For FMEA, severity (S), occurrence (O) and detection (D) were scored from 1 to 10 for the dominant failure mode of each category, and the risk priority number was calculated as [13]:

$$RPN = S \times O \times D \tag{13}$$

3. Results and Discussion

3.1. Operating Data and Filter Pressure Drop

Table 1 lists the inputs of all six snapshots. The differential pressure across the filter house was 76–82 mmH_2O (745–804 Pa) before and 45–46 mmH_2O (441–451 Pa) after replacement, a reduction of 294–363 Pa, or about 0.3 kPa. Relative to an ambient pressure of about 101 kPa, the inlet pressure loss therefore decreased by only about 0.3%. The compressor inlet pressure logged by the control system (103.5–110.0 kPa) does not resolve a change of this size: the January values are lower, not higher, than the November values, and all values exceed the standard atmospheric pressure of 101.3 kPa, which suggests a transmitter offset that could not be verified. It was used only as the recorded input of the pressure ratio.

Table 1. Measured inputs of GT1A before (11 November 2025) and after (7 January 2026) filter replacement.

Parameter	Unit	11 November 2025 (before)			7 January 2026 (after)		
		00:00	12:00	16:00	00:00	12:00	16:00
Generator load	MW	37.42	37.40	37.40	38.85	38.50	38.65
T_1	K	301.05	304.35	303.65	299.95	305.25	305.05
P_1 (abs.)	kPa	108.5	110.0	109.7	103.5	109.2	108.0
T_2	K	633.65	646.05	645.05	625.95	636.75	635.25
P_2 (abs.)	kPa	1032	1012	1015	1146	1158	1153
T_3	K	1338.15	1323.15	1326.15	1378.15	1355.15	1358.15
T_4	K	853.55	858.00	857.15	820.34	823.05	822.15
$\dot{m}_f$ (fuel gas)	kg/s	2.74	2.80	2.78	2.36	2.49	2.45
Filter ΔP_f [1]	mmH_2O	82	76	76	45	46	46
P_2/P_1	–	9.51	9.20	9.25	11.07	10.60	10.68

[1] Read from the operators' check sheet ("DP Inlet Filter"); values transcribed from the original handwritten log.

Ambient conditions were comparable: T_1 differed by −1.1 to +1.4 K between the two days at the same hour. Generator load was 1.1–1.4 MW (3–4%) higher after replacement. The compressor discharge pressure rose from 1012–1032 to 1146–1158 kPa, so that the pressure ratio increased from 9.20–9.51 to 10.60–11.07, while the compressor discharge temperature was 7.7–9.8 K lower, the exhaust temperature 33–35 K lower and T_3 32–40 K higher. Fuel gas flow fell from 2.74–2.80 to 2.36–2.49 kg/s.

Physically, the filter house is a pressure loss in series with the compressor. For a single-shaft machine at constant speed the compressor swallows an almost fixed volume flow, so the air mass flow scales with the inlet density, i.e. with P_1/T_1, and the discharge pressure scales with P_1 for a given pressure ratio [1]. A 0.3 kPa smaller inlet loss therefore raises air mass flow and P_2 by roughly 0.3% only. The measured rises in discharge pressure (11–14%) and pressure ratio (15–16%) are far larger. A heavily loaded filter also lets more fine particles reach the compressor, where deposits reduce flow capacity and efficiency [3,4]; such deposits are not removed by the filter change itself, although a new filter limits further fouling. Consequently, the changes in Table 1 cannot be attributed to the lower filter pressure loss alone. A compressor wash was carried out in the period between the two datasets during the maintenance window around 19 November 2025. Because fouling deposits are removed by washing and not by changing the filter [4], the

wash most likely accounts for a large part of the recovery of compressor efficiency, and the effect of the filter replacement cannot be separated from it. Differences in inlet guide vane position or firing-temperature control and instrument calibration may also have contributed. Filter condition itself is characterised by the relation between flow resistance and dust loading described in ISO 29461-1 [16].

3.2. Compressor Efficiency

Table 2 summarises the four calculated efficiencies and Figure 1 compares them for matching times of day.

Table 2. Calculated efficiencies of GT1A (base case, $\gamma = 1.4$).

Snapshot	Load (MW)	η_c (%)	η_{cc} (%)	η_t (%)	η_{th} (%)
11 Nov 2025, 00:00	37.42	81.76	92.07	76.31	21.58
11 Nov 2025, 12:00	37.40	78.85	90.36	74.87	18.23
11 Nov 2025, 16:00	37.40	79.01	90.48	75.18	18.73
7 Jan 2026, 00:00	38.85	90.88	96.20	81.45	30.82
7 Jan 2026, 12:00	38.50	88.71	95.05	80.02	27.92
7 Jan 2026, 16:00	38.65	89.35	95.36	80.27	28.47

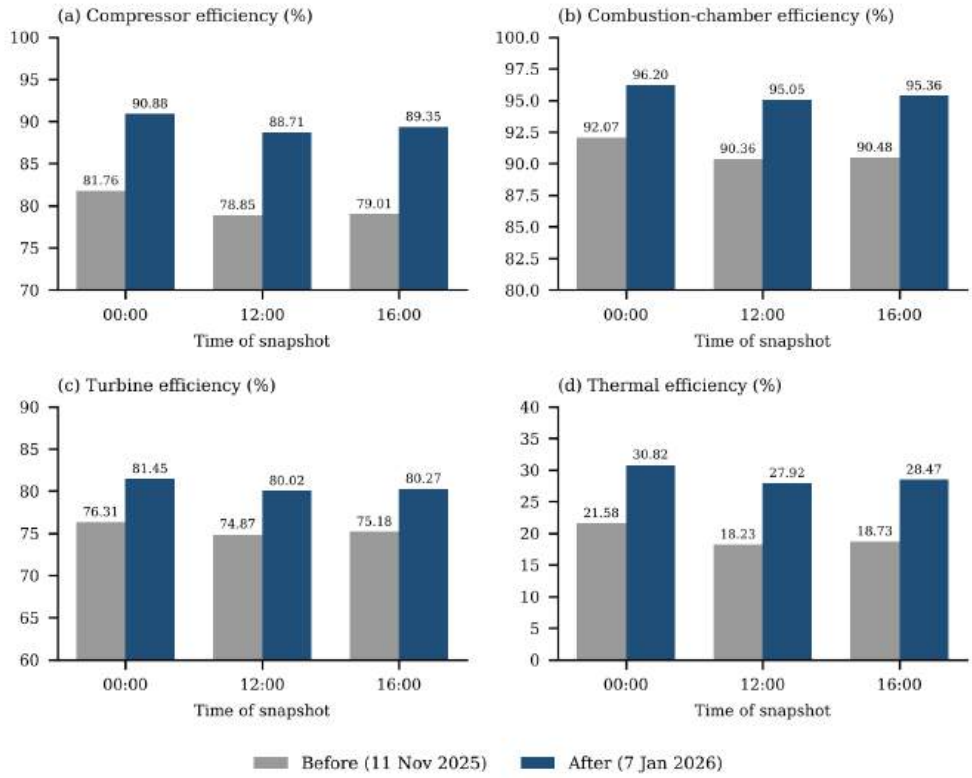


Figure 1. Efficiencies of GT1A before and after air inlet filter replacement: (a) compressor, (b) combustion chamber, (c) turbine and (d) thermal efficiency.

Compressor efficiency rose from 78.85–81.76% to 88.71–90.88%, an increase of 9.1–10.3 percentage points at matching times (Table 2, Figure 1a). Equation (2) shows the origin: after replacement the pressure ratio is 15–16% higher, so the ideal temperature rise $T_{2s} - T_1$ grows by 24–25 K, whereas the actual rise $T_2 - T_1$ shrinks by 7–11 K. A higher pressure ratio reached with a smaller temperature rise is the signature of a less degraded compressor, since fouling lowers airflow, pressure ratio and efficiency [4]. Simanjuntak et

al. [7] found gains of about 2 and 3 percentage points at 38 and 39 MW after filter replacement; the gain here is three to five times larger, which is consistent with the additional effect of the compressor wash. Isentropic compressor efficiencies of the order of 0.85–0.90 are usual for axial machines [1], so the post-replacement values are plausible and the pre-replacement values indicate degraded compression. Ambient temperature cannot explain the change: Purwanto [8] reported a drop of only 0.09 percentage points (92.46% to 92.37%) for a 7.4 K higher inlet temperature, whereas T_1 differed here by at most 1.4 K.

3.3. Combustion-Chamber and Turbine Efficiency

Combustion-chamber efficiency increased from 90.36–92.07% to 95.05–96.20% and turbine efficiency from 74.87–76.31% to 80.02–81.45% (Table 2, Figure 1b,c). Equation (3) is a temperature-based index: it rises when compression is more efficient even if the combustor itself is unchanged, so it shows how the compressor improvement propagates through the cycle and is not a measure of combustion completeness. For the turbine, the actual temperature drop $T_3 - T_4$ increased by 14–15%, while the ideal drop increased by only 7–8%, because the higher T_3 and expansion ratio enlarge both, but the exhaust temperature fell.

Table 3 tests the sensitivity to the property assumption of Section 2.3. With hot-gas properties the isentropic exhaust temperature is higher, the ideal drop smaller, and the turbine efficiency 7.7–8.1 percentage points higher than in the base case; the thermal efficiency is 2.0–3.7 points higher. The absolute levels therefore depend on the assumption (the base-case turbine efficiency is probably underestimated), but the change between the two datasets is robust: about +5.4 points for the turbine and +8.1 to +8.5 points for the thermal efficiency, against +5.1 to +5.2 and +9.2 to +9.7 points in the base case.

Table 3. Sensitivity of turbine and thermal efficiency to gas properties.

Snapshot	η_t base (%)	η_t hot gas (%)	η_{th} base (%)	η_{th} hot gas (%)
11 Nov 2025, 00:00	76.31	84.11	21.58	24.69
11 Nov 2025, 12:00	74.87	82.56	18.23	21.91
11 Nov 2025, 16:00	75.18	82.90	18.73	22.34
7 Jan 2026, 00:00	81.45	89.59	30.82	32.81
7 Jan 2026, 12:00	80.02	88.07	27.92	30.32
7 Jan 2026, 16:00	80.27	88.33	28.47	30.79

Base: $\gamma = 1.4$ throughout. Hot gas: $\gamma_g = 4/3$, $c_{p,g} = 1.148$ and $c_{p,a} = 1.005$ kJ kg^{-1} K^{-1} (Equations (4) and (7)–(9)).

3.4. Thermal Efficiency

The thermal efficiency indicator rose from 18.23–21.58% to 27.92–30.82% (Figure 1d), an increase of 9.2–9.7 points in the base case and 8.1–8.5 points with hot-gas properties (21.91–24.69% to 30.32–32.81%). Both levels are below the value implied by the manufacturer's nominal heat rate of 11,400 kJ/kWh (3600/11,400 = 31.6%), because the indicator neglects mechanical and generator losses and depends on T_3, which is not an independent measurement; it should be used to compare conditions, not as an absolute efficiency. The direction agrees with the finding that a growing inlet pressure loss from filter degradation reduces power and thermal efficiency [6]. Load also differs between the datasets (+1.1 to +1.4 MW). A study of a combined cycle plant reported a rise of gas turbine cycle efficiency from 31.28% at 23 MW to 45.56% at 33 MW [17]; applying that average slope (about 1.4 points per MW) as an indication only, the load difference could account for roughly 1.6–2.0 of the 9–10 points. The fuel gas flow fell by 11–14% while the load rose (Table 1), consistent in direction with the higher efficiency, but larger than the inlet pressure loss alone would suggest.

3.5. Causes of Filter Replacement: Fishbone and FMEA

Figure 2 shows the fishbone diagram. In the *Man* category, filter condition was monitored mainly by visual inspection, abnormal pressure drop could be recognised late, and installation errors and an uneven understanding of the actual filter condition were mentioned. In *Method*, replacement followed a fixed time interval and was not yet fully based on differential pressure, and the inspection procedure did not prevent late replacement. In *Machine*, dust and contaminants accumulated on the filter, and housing-seal leakage, differential-pressure sensor inaccuracy and clogging of the drain were possible additional contributors. In *Material*, filter quality governs the service life, and the media degrades with long use and with humidity or contamination. In *Environment*, the air around the plant carries high dust and particle loads; a dirty working environment, weather and humidity accelerate clogging and raise the replacement frequency. In *Measurement*, pressure-drop monitoring was not always the basis for replacement, a decline in filter performance could be detected late, and condition trending was not fully exploited.

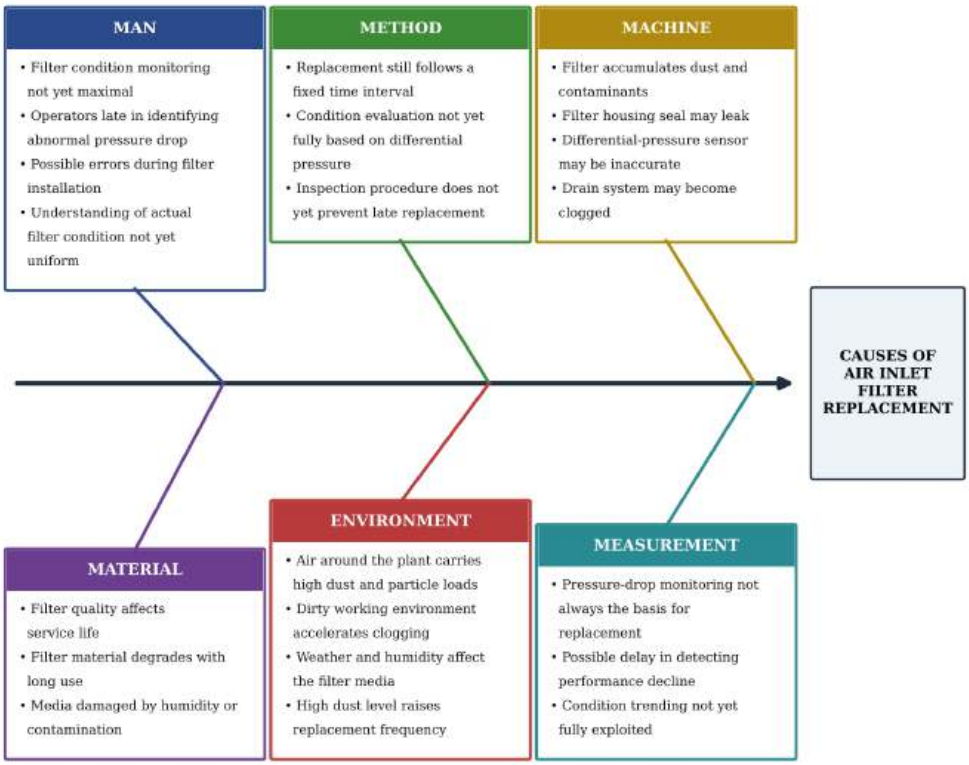


Figure 2. Fishbone (6M) diagram of the causes of air inlet filter replacement at GT1A.

Table 4 ranks the categories by RPN (Equation (13)). Environment has the highest priority (RPN = 126), followed by Material (80), Machine (64), Method (60), Measurement (45) and Man (36). The scores were assigned by the authors from interviews and observation, so they are subjective, and no dust concentration was measured at the site. The dominance of the environmental factor is therefore a qualitative finding.

Table 4. FMEA of the dominant failure mode in each fishbone category.

Rank	Category	Dominant failure mode and cause	*S*	*O*	*D*	RPN
1	Environment	Filter clogged by dust and contaminants; dusty surroundings, long cleaning interval	6	7	3	126
2	Material	Degraded filter media; service life, high humidity, material quality	5	4	4	80

Rank	Category	Dominant failure mode and cause	*S*	*O*	*D*	RPN
3	Machine	Leaking housing or seal; worn seal, imprecise installation, housing corrosion	4	4	4	64
4	Method	Replacement not matched to actual condition; interval-based, visual inspection	5	4	3	60
5	Measurement	Inaccurate DP sensor; irregular calibration, sensor contaminated by dust	5	3	3	45
6	Man	Lack of inspection thoroughness; limited experience, visual identification	4	3	3	36

3.6. Filter Service Life and Replacement Criterion

The filter was installed on 1 November 2022 and replaced on 19 November 2025, giving $T = 3.05$ years (Equation (11)) and $H = 3.05 \times 4{,}032 \approx 12{,}297$ h (Equation (12)). Over this period the filter differential pressure rose from 45 to 82 mmH_2O (441 to 804 Pa, +82%). The final value exceeds the final pressure drop of 600 Pa (≈61 mmH_2O) given in the specification of the air inlet filter. A condition-based criterion, in which replacement is triggered when ΔP_f approaches 61 mmH_2O instead of after a fixed 3 years or 12,300 h, is consistent with guidelines that base filter changeout on pressure drop and operating conditions [2]; Köse and Yılmazoğlu [10] found for a GT13E2 unit that replacement every 2,500 operating hours or at $\Delta p = 5$ mbar (≈51 mmH_2O) was the cost-optimal option, a threshold of the same order as the specified limit used here. If the pressure drop rose monotonically, such a criterion would have prompted replacement before 82 mmH_2O was reached. It is based on one replacement cycle and should be validated with continuous ΔP trending.

3.7. Limitations and Outlook

The conclusions are subject to the following limitations: (i) only two days with one reading per snapshot were compared, without uncertainty analysis, and November and January differ in season; (ii) the compressor wash carried out between the datasets confounds the effect of the filter replacement, so the two cannot be separated (Section 3.1); (iii) T_3 and the logged P_1 are DCS readings whose measurement uncertainty is not known; (iv) constant-property and no-pressure-loss assumptions affect absolute values; and (v) the efficiencies are temperature-based, because the air mass flow was not measured. Future work should include continuous ΔP and particulate measurements, compressor-wash records, a fuel-based efficiency using $\dot{m}_f$ and the lower heating value, uncertainty propagation, and several replacement cycles.

4. Conclusions

This study evaluated the effect of air inlet filter replacement on the GT1A gas turbine of Unit 1A of Bekasi Power Combined Cycle Power Plant. The main conclusions are:

1. After replacement, filter differential pressure fell from 76–82 to 45–46 mmH_2O (about 0.3 kPa), while compressor efficiency rose from 78.85–81.76% to 88.71–90.88%, combustion-chamber efficiency from 90.36–92.07% to 95.05–96.20%, turbine efficiency from 74.87–76.31% to 80.02–81.45% and thermal efficiency from 18.23–21.58% to 27.92–30.82%.
2. The direction and approximate size of the efficiency changes are robust to hot-gas properties (turbine about +5 points, thermal +8 to +10 points), but the 0.3 kPa pressure-loss reduction explains only a small part of them; a compressor wash performed in the same period is a major concurrent factor, so the effect of the filter replacement alone cannot be isolated.
3. Fishbone and FMEA analyses identified environmental factors, mainly airborne dust (RPN = 126), as the dominant cause of filter deterioration.

4. The filter served 3.05 years (about 12,300 h) and reached 82 mmH_2O, above the specified final limit of about 61 mmH_2O; a condition-based replacement criterion using ΔP_f is recommended.

Author Contributions: Conceptualization, G.F.A. and A.E.; methodology, G.F.A. and A.E.; formal analysis, G.F.A.; investigation, G.F.A.; data curation, G.F.A.; writing—original draft preparation, G.F.A.; writing—review and editing, A.E.; supervision, A.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to express their sincere gratitude to Bekasi Power Combined Cycle Power Plant for providing operational data and technical support during this research. The authors also thank the Department of Mechanical Engineering, Politeknik Negeri Jakarta, for academic guidance and support throughout the completion of this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Sample Calculation

Snapshot 11 November 2025, 12:00: $T_1 = 304.35$ K, $P_1 = 110$ kPa, $T_2 = 646.05$ K, $P_2 = 1012$ kPa, $T_3 = 1323.15$ K, $T_4 = 858.00$ K. Equation (1): $T_{2s} = 304.35\ (1012/110)^{0.2857} = 573.77$ K. Equation (2): $\eta_c = (573.77 - 304.35)/(646.05 - 304.35) = 78.85\%$. Equation (3): $\eta_{cc} = (1323.15 - 646.05)/(1323.15 - 573.77) = 90.36\%$. Equation (4): $T_{4s} = 1323.15\ (110/1012)^{0.2857} = 701.84$ K. Equation (5): $\eta_t = (1323.15 - 858.00)/(1323.15 - 701.84) = 74.87\%$. Equation (10): $\eta_{th} = (465.15 - 341.70)/677.10 = 18.23\%$. Hot-gas case: $T_{4s} = 1323.15\ (110/1012)^{0.25} = 759.7$ K, $\eta_t = 82.56\%$, $\eta_{th} = 21.91\%$.

References

1. Saravanamuttoo, H.I.H.; Rogers, G.F.C.; Cohen, H.; Straznicky, P.V.; Nix, A.C. *Gas Turbine Theory*, 7th ed.; Pearson Education: Harlow, UK, 2017.
2. Wilcox, M.; Baldwin, R.; Garcia-Hernandez, A.; Brun, K. *Guidelines for Gas Turbine Inlet Air Filtration Systems*; Gas Machinery Research Council, Southwest Research Institute: 2010.
3. Kurz, R.; Brun, K. Degradation in gas turbine systems. *J. Eng. Gas Turbines Power* **2001**, *123*, 70–77. https://doi.org/10.1115/1.1340629.
4. Meher-Homji, C.B.; Bromley, A. Gas turbine axial compressor fouling and washing. In Proceedings of the 33rd Turbomachinery Symposium, Houston, TX, USA; Texas A&M University: College Station, TX, USA, 2004; pp. 163–192.
5. Meher-Homji, C.B.; Chaker, M.; Motiwala, A.F. Gas turbine performance deterioration. In Proceedings of the 30th Turbomachinery Symposium, Houston, TX, USA; Texas A&M University: College Station, TX, USA, 2001; pp. 139–175.
6. Jin, Y.; et al. A novel integrated modeling approach for filter diagnosis in gas turbine air intake system. *Proc. Inst. Mech. Eng. Part A J. Power Energy* **2022**, *236*, 435–449. https://doi.org/10.1177/09576509211044392.
7. Simanjuntak, C.B.; Putra, I.Z.; Toar, H.; Aryeni, I. Analisis pengaruh penggantian air inlet filter terhadap kerja kompresor di PT. Energi Listrik Batam. *J. Rekayasa Energi* **2024**, *3*(1). https://doi.org/10.31884/jre.v3i1.53.
8. Purwanto, D. Analysis of the effect of inlet compressor temperature on the thermal efficiency of PLTG Unit 3.2 in Grati. *J. Mech. Des. Test.* **2024**, *6*, 101–114. https://doi.org/10.22146/jmdt.102262.
9. Effiom, S.O.; et al. Experimental study on the optimal performance of gas turbine (GT) inlet air filtration system for offshore application. *J. Eng. Appl. Sci.* **2023**, *70*. https://doi.org/10.1186/s44147-023-00303-8.
10. Köse, O.Ö.; Yılmazoğlu, M.Z. Gaz türbini periyodik bakımlarının sistem performansına etkilerinin incelenmesi [Investigation of the effects of gas turbine periodic maintenance on system performance]. *Mühendis ve Makina* **2025**, 718, 129–172. https://doi.org/10.46399/muhendismakina.1564458.

11. Liliana, L. A new model of Ishikawa diagram for quality assessment. *IOP Conf. Ser. Mater. Sci. Eng.* **2016**, *161*, 012099. https://doi.org/10.1088/1757-899X/161/1/012099.
12. Kumah, A.; et al. Cause-and-effect (fishbone) diagram: A tool for generating and organizing quality improvement ideas. *Glob. J. Qual. Saf. Healthc.* **2024**, *7*, 85–87. https://doi.org/10.36401/JQSH-23-42.
13. Stamatis, D.H. *Failure Mode and Effect Analysis: FMEA from Theory to Execution*, 2nd ed.; ASQ Quality Press: Milwaukee, WI, USA, 2003.
14. Moran, M.J.; Shapiro, H.N.; Boettner, D.D.; Bailey, M.B. *Fundamentals of Engineering Thermodynamics*; John Wiley & Sons: Hoboken, NJ, USA, 2018.
15. Çengel, Y.A.; Boles, M.A.; Kanoğlu, M. *Thermodynamics: An Engineering Approach*; McGraw-Hill Education: New York, NY, USA, 2019.
16. ISO 29461-1:2021. *Air Intake Filter Systems for Rotary Machinery—Test Methods—Part 1: Static Filter Elements*; International Organization for Standardization: Geneva, Switzerland, 2021.
17. Yohana, E.; Herriza, R.M. Analisis efisiensi siklus combine cycle power plant (CCPP) gas turbine generator terhadap beban operasi PT Krakatau Daya Listrik. *ROTASI* **2016**, *28*(1).

RiESTech
Recent in Engineering Science and Technology

MBI

Article

Analysis of the Development of an Electric Mini Excavator Based on Environmentally Friendly Technology

Fuad Zainuri [1,*], Rahmat Noval [2], M. Todaru [3]

[1] Department of Mechanical Engineering, Politeknik Negeri Jakarta, Depok 16424, Indonesia
* Correspondence: fuad.zainuri@pnj.ac.id

Abstract: This paper reports the development and preliminary experimental evaluation of a battery-electric mini-excavator prototype using an electro-hydraulic powertrain. The diesel prime mover was replaced by a 5 kW, 48 V brushless DC (BLDC) motor supplied by a 48 V, 100 Ah lithium-ion battery pack through a 48 V inverter. The motor drives a hydraulic pump rated at 16 L/min, while the original hydraulic actuators are retained for boom, arm, and bucket movement. The available experimental record includes boom motion times under no-load pressure settings of 125 and 250 psi, a 12 kg load condition, and a maximum-load test. Boom extend time decreased from 2.72 s at 125 psi to 1.97 s at 250 psi under the reported no-load condition. Under a 12 kg load, the reported boom extend time was 8.50 s at 125 psi and 3.28 s at 250 psi. At a 50 kg load, the reported extend time was 9.91 s; hydraulic leakage was observed when the applied load reached 57 kg. These results demonstrate preliminary actuation feasibility, but do not establish comparative energy efficiency, noise reduction, operating cost, or performance superiority over a diesel excavator because the corresponding measurements and diesel baseline are not available. The nominal battery energy is 4.8 kWh; actual operating duration requires measured duty-cycle power and battery-management data. The prototype is therefore evaluated as a preliminary low-voltage electric conversion demonstrator, with further instrumentation and repeatability testing required.

Keywords: Electric mini excavator; Electro-hydraulic system; BLDC drive; Hydraulic actuator

Citation: Zainuri, F., Noval, R., Todaru, M. (2026). Analysis of the Development of an Electric Mini Excavator Based on Environmentally Friendly Technology. Recent in Engineering Science and Technology, 4(4), 255–262. Retrieved from https://www. mbi-jounals.com/index.php/riestech /article/view/138.

Academic Editor: Yudhi Ariadi

Received: 9 September 2026
Accepted: 25 October 2026
Published: 31 October 2026

Publisher's Note: MBI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2026 by the authors. Licensee MBI, Jakarta, Indonesia. This article is an open access article distributed under MBI license (https://mbi-journals.com/licenses/by/4.0/).

1. Introduction

Electrification of compact construction equipment is being pursued to reduce exhaust emissions at the point of use, enable operation in enclosed or noise-sensitive environments, and improve the controllability of machine power demand. Battery-electric compact excavators are already commercially available. For example, Volvo Construction Equipment lists the ECR25 Electric with a 48 V, 20 kWh battery in its Asia product information [1] and has since upgraded the model to a 40 kWh battery [2], while JCB markets the 19C-1E as a fully electric mini excavator with a 48 V driveline [3]. These products demonstrate that compact excavator electrification is an established engineering direction rather than a new concept in itself.

Research on electric hydraulic excavators has investigated matching electric motor speed and pump displacement, independent metering, and coordinated pump control to reduce throttling and standby losses. He et al. examined the energy consumption characteristics of a pump-control system for electric excavator arms and reported that a compound speed–displacement strategy improved efficiency by approximately 18.9% relative to a conventional variable-speed pump-controlled cylinder system [4]. Ge et al. showed that a speed- and displacement-variable pump with independent metering reduced idle electric power from 2.05 kW to 0.7 kW compared with a displacement-only concept [5], and a later study of a speed-variable power source reported efficiencies of 40%–71%, which were 3%–10% higher than those of a fixed-speed source [6]. Reviews of hydraulic excavator energy-regeneration systems describe hydraulic accumulators and electro-hydraulic approaches for capturing energy that would otherwise be dissipated

through throttling and heat generation [7], [8]. Hybrid hydraulic excavators with electric–hydraulic variable powertrains and regeneration systems have likewise shown test-bench reductions in fuel consumption and emissions [9], [10].

In parallel, the heavy-equipment maintenance literature documents the failure and troubleshooting patterns of conventional diesel-driven excavators. Low-power complaints in excavator hydraulic systems have been investigated through structured troubleshooting, including pressure and cycle-time checks on a 302.5 mini excavator [11] and on other excavator units [12], [13]; cylinder leakage has been treated as a root-cause diagnosis problem [14]; and engine-side topics such as starting failure, wiring-harness maintenance, fuel-pressure faults, and crankcase blow-by recur as case subjects [15]–[19]. This body of work defines the conventional-drive baseline that electrification changes, but it does not address the battery-electric conversion of the hydraulic power unit. The present paper therefore reads the prototype's actuation data alongside these diagnostic practices.

The contribution of this work is a preliminary conversion and test study of a compact excavator prototype using a 48 V battery–inverter–BLDC motor arrangement and locally developed integration. Its scope is limited to describing the architecture and the boom actuation data currently documented. It does not claim novelty in battery-electric excavators generally, nor does it claim measured energy or noise improvements.

2. Materials and Experiment Methods

2.1. Prototype Architecture

The prototype uses a battery-electric electro-hydraulic architecture. A 48 V, 100 Ah lithium-ion battery supplies a 48 V inverter rated at 100 A. The inverter drives a 5 kW BLDC motor coupled to the main hydraulic pump. The pump is reported to deliver up to 16 L/min. The hydraulic cylinders and linkage for the boom, arm, and bucket are retained. A custom electronic control unit (ECU) is identified in the original design as the system coordination and monitoring unit.

Tabel 1. Reported specifications and functions of the main components of the electro-hydraulic prototype.

Component	Reported specification	Function
Electric motor	BLDC, 5 kW, 48 V	Drives hydraulic pump
Battery	Li-ion, 48 V, 100 Ah	Electrical energy storage
Inverter	48 V input, 100 A rated	Motor drive and protection
Hydraulic pump	Variable displacement; reported flow 16 L/min	Hydraulic power supply
Controller	Custom ECU; detailed control law not reported	System coordination and monitoring

The hydraulic power delivered to the fluid can be estimated from $P_h = pQ$, where p is pressure and Q is volumetric flow rate. For the reported maximum flow of 16 L/min (2.667×10^{-4} m³/s) and pressure of 250 psi (1.724 MPa), the ideal hydraulic power is approximately 0.460 kW. This is substantially below the 5 kW motor nameplate rating. The difference does not, by itself, establish the motor's actual operating efficiency: actual pump flow, pressure at the pump outlet, motor input power, transient duty, and losses must be measured simultaneously.

2.2. Testing Procedure

Experimental The original manuscript reports the following conditions: no-load operation at 125 psi and 250 psi; a 12 kg load condition; and a maximum-load condition of 50 kg. It reports boom extend and retract times for the no-load tests, boom extend times for the 12 kg and 50 kg tests, and leakage at 57 kg. The original document states that

electrical power and motor/inverter temperature were measured, but does not provide their values, measurement instruments, or uncertainty. Therefore, those quantities are not treated as available results in this revision.

Tabel 2. Boom extension and retraction times reported in the original manuscript for the no-load and loaded test conditions at system pressures of 125 psi (0.862 MPa) and 250 psi (1.724 MPa).

Test condition	Boom extend (s)	Boom retract (s)	Other recorded observation
No load; 125 psi	2.72	1.77	-
No load; 250 psi	1.97	1.39	-
12 kg; 125 psi	8.50	-	Extend time only
12 kg; 250 psi	3.28	-	Extend time only
50 kg	9.91	-	Hydraulic leakage reported at 57 kg

Noted : The number of repetitions (n) for each condition and the raw observations. Report mean ± standard deviation only after the raw repeated measurements are available.

2.3. Electrical and hydraulic calculations

The nominal stored battery energy is estimated as $E_{nom} = V_{nom} \times C = 48\ V \times 100\ Ah = 4{,}800\ Wh = 4.8\ kWh$. At a continuous electrical demand equal to the 5 kW motor rating, the ideal energy-to-power ratio is $E_{nom}/P = 0.96\ h$. This is an ideal upper-bound arithmetic estimate, not a measured runtime. Actual runtime will be lower because usable capacity is limited by the BMS depth-of-discharge window and because conversion, motor, pump, and hydraulic losses and duty-cycle variations are present.

The hydraulic power relation is $P_h = pQ$. At $p = 250\ psi = 1.724\ MPa$ and $Q = 16\ L/min = 2.667 \times 10^{-4}\ m^3/s$, $P_h \approx 0.460\ kW$. Overall battery-to-hydraulic efficiency can only be calculated when simultaneous electrical input and hydraulic output are measured: $\eta_{total} = (pQ)/(V_{DC} I_{DC})$, with consistent time averaging over the same operating interval. If flow varies during a cycle, the cycle energy should be calculated by integrating power over time rather than multiplying nominal maximum flow by a single pressure value.

The reported pressure settings are interpreted as pressure-limit or relief settings, not as direct flow commands. Actuator speed is primarily related to delivered flow and effective actuator area ($v = Q/A$). Pressure determines available force ($F = pA$) and influences whether the pump flow is bypassed through a relief path or delivered to the loaded actuator.

The cited IEC 62619:2022 standard specifies safety requirements and tests for secondary lithium cells and batteries used in industrial applications [20]. The original manuscript does not document evidence of design verification against individual IEC 62619 clauses. Accordingly, this revision does not claim IEC 62619 compliance.

3. Results and Discussion

3.1. Effect of pressure setting on boom motion

Under the reported no-load condition, increasing the pressure setting from 125 psi to 250 psi was associated with a reduction in boom extend time from 2.72 s to 1.97 s (27.6%) and retract time from 1.77 s to 1.39 s (21.5%). These are descriptive differences from the reported values; no repeat count or variability estimate is available. The result should not be interpreted as evidence that higher pressure directly increases fluid velocity. Actuator velocity is governed mainly by flow rate and actuator area. A higher pressure limit may instead reduce flow diverted through a relief valve or provide sufficient force margin under the tested load, depending on the actual hydraulic circuit and pump-control strategy.

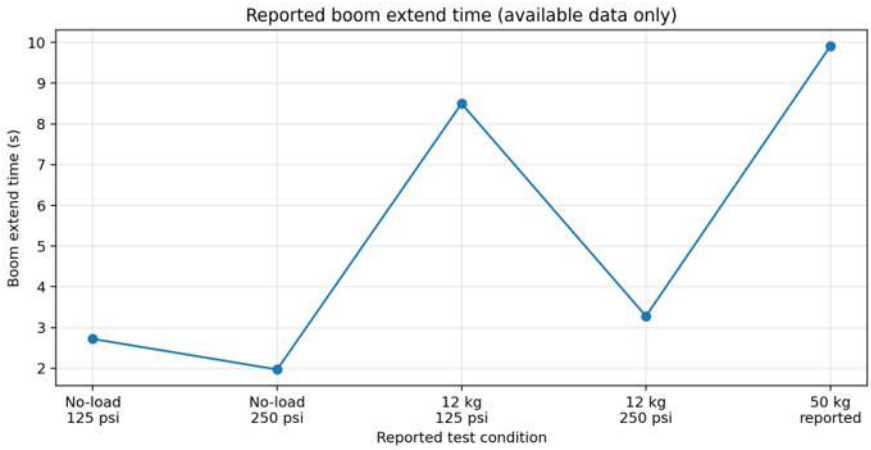


Figure 1. Boom extend times explicitly reported in the source manuscript

3.2. Effect of applied load

For the 12 kg condition, boom extend time was reported as 8.50 s at 125 psi and 3.28 s at 250 psi, a descriptive reduction of approximately 61.4%. The original manuscript attributed this change to torque control. However, the documented test variable is pressure setting, and no controller command, motor-speed trace, torque estimate, pump displacement trace, or pressure-control logic is supplied. The defensible interpretation is therefore limited to an association between the pressure setting and the measured boom extend time.

The 50 kg test produced a reported boom extend time of 9.91 s. Hydraulic leakage was observed when the applied load reached 57 kg. This observation indicates a limitation of the tested prototype configuration, but it does not establish a certified mechanical load limit or safe working load. Further tests should identify the leakage location, pressure at failure, structural stability, and repeatability.

The following table defines the dataset that should be supplied before publication.

Tabel 3. Extension and retraction times (mean ± SD) for each actuator and test condition

Actuator	Condition	Extend time (s), mean ± SD	Retract time (s), mean ± SD
Boom	No load; 125 psi	2.72; n/SD unavailable	1.77; n/SD unavailable
Boom	No load; 250 psi	1.97; n/SD unavailable	1.39; n/SD unavailable
Boom	12 kg; 125 psi	8.50; n/SD unavailable	-
Boom	12 kg; 250 psi	3.28; n/SD unavailable	-
Boom	50 kg	9.91; n/SD unavailable	-
Arm	All conditions	-	-
Bucket	All conditions	-	-

The 5 kW motor rating and the calculated ideal hydraulic power of approximately 0.46 kW at 250 psi and 16 L/min indicate a substantial nominal power mismatch at that operating point. This comparison highlights the need to characterize the actual motor operating point and part-load efficiency. It is not sufficient to infer low efficiency solely from the ratio of hydraulic power to motor nameplate rating; measured DC input power, actual flow, pressure, and motor speed are required.

3.3. Interpretation of the pressure- and load-dependent cycle times

For a cylinder of fixed stroke, mean piston speed is proportional to the flow reaching the actuator, so the ratio of cycle times is an indirect indicator of the ratio of effective flows. The no-load times therefore imply an effective boom extension flow at the 250 psi setting about 1.38 times that at 125 psi (2.72 s/1.97 s) and a retraction flow about 1.27 times higher (1.77 s/1.39 s). Under the 12 kg load the same ratio rises to about 2.59 (8.50 s/3.28 s). Comparing loaded and unloaded extend times at a fixed setting shows the same trend: the 12 kg load lengthened extension by a factor of about 3.1 at 125 psi (8.50 s/2.72 s) but only about 1.7 at 250 psi (3.28 s/1.97 s). These ratios are calculated from single reported values without uncertainty and indicate a trend, not a flow estimate.

Three mechanisms, not mutually exclusive, could produce this behavior, and each is testable. First, if the pressure needed to move the boom linkage and load approaches the 125 psi setting, part of the pump flow would pass over the relief valve and less would reach the cylinder; raising the setting restores margin. Second, if the pump is a variable-displacement unit with a pressure compensator, displacement would be reduced as system pressure approaches the setting, which also slows the actuator. Third, the motor speed may droop under load because of torque or inverter current limits (see Section 3.5), reducing pump flow. The same symptoms, a pressure deviation and a degraded cycle time, were among the hydraulic-system causes identified for low power on a 302.5 mini excavator by Kurniawan et al. [11], and low-power troubleshooting of other excavator units follows a similar approach [12], [13]. Boom cycle time under a fixed load, together with a verified relief pressure, is therefore a suitable commissioning check for the converted machine. Simultaneous logging of pump-outlet pressure, cylinder-port pressures, flow, and motor speed would allow the three mechanisms to be separated.

3.4. Benchmarking against commercial battery-electric compact excavators

The table below places the prototype beside two commercial machines using manufacturer-published nominal data. The prototype has about one quarter of the battery energy (4.8 kWh versus 20 kWh, 24%), motor power (5 kW versus 18 kW peak, 28%), and pump flow (16 L/min versus 58 L/min, 28%) of the Volvo ECR25 Electric [1], while sharing its 48 V bus voltage. This proportionality indicates that the prototype is a scaled-down demonstrator of the same low-voltage architectural class rather than a higher-voltage design.

Table 4. Nominal comparison of the prototype with two commercial battery-electric compact excavators (manufacturer data; prototype values as reported in Section 2).

Parameter	Prototype (this work)	Volvo ECR25 Electric [1], [2]	JCB 19C-1E [3]
Bus voltage	48 V	48 V	48 V driveline
Battery energy	4.8 kWh (nominal)	20 kWh; 40 kWh in updated model	Optional 20 kWh pack
Motor power	5 kW (rated)	18 kW (peak)	Up to 27 hp (≈20 kW), peak
Maximum pump flow	16 L/min	58 L/min	Not stated in cited source
Runtime	Not measured; 0.96 h ideal at 5 kW	Up to 4 h (20 kWh)	Manufacturer states a full shift

The commercial data also indicate what runtime is achievable. Volvo quotes up to 4 h from 20 kWh [1], which implies a mean draw of at least about 5 kW, or roughly 28% of the 18 kW peak. Applying the same ratio to the prototype's 5 kW would give a mean draw near 1.4 kW and at most about 3.5 h from 4.8 kWh before usable-capacity limits. This is a scaling illustration only; it is not a measurement and does not replace a measured duty

cycle. Volvo has since doubled the battery of the updated ECR25 to 40 kWh for up to eight hours of typical operation [2], and JCB offers an optional 20 kWh pack on the 19C-1E [3]. Battery energy is therefore likely to be the primary limit on the prototype's ability to cover a full working shift.

3.5. Electrical–hydraulic power matching

`The nominal electrical data allow a consistency check of the drive. At 48 V, the 5 kW rated motor power corresponds to a DC current of about 104 A (5,000 W/48 V) even at 100% conversion efficiency, which already exceeds the 100 A inverter rating; at an assumed combined motor and inverter efficiency of 90%, the DC current would be near 116 A. The inverter therefore limits sustained DC input power to about 4.8 kW at nominal voltage (48 V × 100 A), and the limit is lower in practice because battery voltage sags under load and with decreasing state of charge. The ratio of rated motor power to nominal battery energy is about 1.04 h^{-1} for the prototype, compared with about 0.9 h^{-1} for the peak motor power of the ECR25 Electric [1], so the prototype battery would be discharged at a comparable or slightly higher relative rate at full power. These limits do not affect the boom motions tested, whose ideal hydraulic power is only about 0.46 kW, but they matter for sizing if the machine is to be operated near the full motor rating.

The ideal hydraulic output of 0.46 kW is about 9% of the 5 kW motor rating, so the motor and inverter would operate at light load during the reported tests. Part-load operation is where fixed-speed drives are known to be inefficient: Ge et al. reported that a speed-variable power source reduced electric power consumption at idle and partial load [5], [6], and He et al. showed that matching motor speed to load power through pump-displacement control improved efficiency in a pump-controlled boom cylinder [4]. Because the ECU control law of the prototype is not documented, it is not known whether the motor runs at a fixed speed or adapts to demand. If it runs at a fixed speed, the idle and part-load losses described in [5], [6] would be expected to apply. This is a hypothesis to be tested by measuring DC input power at idle and during boom cycles. Boom-lowering energy, which is otherwise dissipated as heat, is a further candidate for regeneration through the electric drive [7]–[10].

3.6. Leakage, maintainability, and shifted failure modes

The leakage observed at 57 kg should be converted from a single event into a root-cause finding. Mubarok et al. treated cylinder leakage in a wheel-loader steering cylinder as a diagnosis problem solved by structured inspection [14]; the same approach applies here, recording the pressure at which leakage starts, the leaking component (cylinder seal, hose, fitting, or valve), and the oil condition. The tested pressure limit of 250 psi (1.72 MPa) is well below the working pressures typically used in excavator hydraulics, which are generally of the order of 10–30 MPa. If the retained cylinders, hoses, and seals come from the original machine, leakage at this low pressure would point first to seal or assembly condition rather than to over-pressure. This inference remains to be confirmed by inspection.

Electrification removes the fuel-system, starting, and crankcase failure categories that recur in diesel-machine case studies [15], [17]–[19], but it introduces battery, inverter, and wiring failure categories. Wiring-harness maintenance is already a documented topic on the diesel CAT 302.5 mini excavator [16]. In a 48 V system carrying currents of the order of 100 A, connector resistance and harness integrity become correspondingly more important for heating and voltage drop, so the maintenance plan of the converted machine should include connector torque and temperature checks. Battery safety verification against IEC 62619 [20] and documented BMS limits are prerequisites for any claim of compliance.

3.7. Limitations and directions for further work

Three limitations bound the conclusions. First, no repeat counts or variability are available, so differences between conditions cannot be tested statistically. Second, electrical power and temperature values are not reported, so efficiency, runtime, and thermal margin cannot be calculated. Third, there is no diesel baseline, so no comparison of performance, energy use, noise, or cost can be made. The next test phase should therefore log DC voltage and current, motor speed and torque, pump-outlet pressure, flow, and component temperatures during repeated boom, arm, and bucket cycles, with enough repetitions per condition to report mean ± standard deviation, followed by variable-speed control and boom-energy regeneration studies guided by [4]–[10].

4. Conclusions

A preliminary battery-electric conversion of a mini-excavator hydraulic system was described using a 48 V, 100 Ah lithium-ion battery, a 48 V inverter, a 5 kW BLDC motor, and a hydraulic pump with a reported flow rate of 16 L/min. The available boom measurements show shorter extend and retract times at the 250 psi setting than at 125 psi under the reported no-load condition. For the 12 kg load, the reported boom extend time was 8.50 s at 125 psi and 3.28 s at 250 psi. At 50 kg, the reported extend time was 9.91 s, while leakage was observed at 57 kg.

These findings support only preliminary actuation feasibility under the documented test conditions. They do not establish higher energy efficiency, 80% noise reduction, lower operating cost, or superiority to a diesel excavator. The nominal battery energy is 4.8 kWh, corresponding to an ideal 0.96 h at a continuous 5 kW draw before accounting for usable SOC limits and system losses.

References

[1] Volvo Construction Equipment, "ECR25 Electric: Electric compact excavator," product information (Asia). [Online]. Available: https://www.volvoce.com/asia/en-as/products/electric-machines/ecr25-electric/ (accessed Nov. 16, 2026).

[2] Volvo Construction Equipment, "Revamped Volvo compact electric machines offer improved runtime," news release, Apr. 8, 2025. [Online]. Available: https://www.volvoce.com/europe/en/about-us/news/2025/revamped-volvo-compact-electric-machines-offer-improved-runtime/ (accessed Nov. 16, 2026).

[3] JCB, "JCB brings electric mini-excavator to North American market," For Construction Pros. [Online]. Available: https://www.forconstructionpros.com/equipment/earthmoving-compact/mini-excavators/product/21087944/jcb-americas-jcb-brings-electric-miniexcavator-to-north-american-market (accessed Nov. 16, 2026).

[4] A. He, L. Wei, Q. Lu, and P. He, "An Investigation of Energy Consumption Characteristics of the Pump-Control System for Electric Excavator Arms," Applied Sciences, vol. 14, no. 23, art. 10791, 2024, doi: 10.3390/app142310791.

[5] L. Ge et al., "Efficiency improvement and evaluation of electric hydraulic excavator with speed and displacement variable pump," Energy Conversion and Management, vol. 150, pp. 62–71, 2017, doi: 10.1016/j.enconman.2017.08.010.

[6] L. Ge, L. Quan, X. Zhang, Z. Dong, et al., "Power matching and energy efficiency improvement of hydraulic excavator driven with speed and displacement variable power source," Chinese Journal of Mechanical Engineering, vol. 32, pp. 1–12, 2019.

[7] T. C. Do, T. D. Dang, T. Q. Dinh, and K. K. Ahn, "Developments in energy regeneration technologies for hydraulic excavators: A review," Renewable and Sustainable Energy Reviews, vol. 145, art. 111076, 2021, doi: 10.1016/j.rser.2021.111076.

[8] T. Lin, Q. Chen, H. Ren, W. Huang, Q. Chen, and S. Fu, "Review of boom potential energy regeneration technology for hydraulic construction machinery," Renewable and Sustainable Energy Reviews, vol. 79, pp. 358–371, 2017, doi: 10.1016/j.rser.2017.05.131.

[9] Y. Yu, T. C. Do, Y. Park, and K. K. Ahn, "Energy saving of hybrid hydraulic excavator with innovative powertrain," Energy Conversion and Management, vol. 244, art. 114447, 2021, doi:

10.1016/j.enconman.2021.114447.

[10] Y. Yu, T. C. Do, B. Yin, and K. K. Ahn, "Improvement of energy saving for hybrid hydraulic excavator with novel powertrain," Int. J. Precis. Eng. Manuf.-Green Technol., vol. 10, no. 2, pp. 521–534, 2023, doi: 10.1007/s40684-022-00437-9.

[11] D. A. Kurniawan, Azwardi, and R. Noval, "Analisa low power pada system hydraulic mini hydraulic excavator 302.5 di workshop alat berat," in Prosiding Seminar Nasional Teknik Mesin, Politeknik Negeri Jakarta, 2022, pp. 1083–1087.

[12] I. Leody, M. H. Tullah, and R. Noval, "Analisa troubleshooting cylinder bucket low power pada unit excavator Hyundai HX210S," in Prosiding Seminar Nasional Teknik Mesin, Politeknik Negeri Jakarta, 2022.

[13] F. Ramadhan, R. Noval, and F. Zainuri, "Analisis penyebab low power pada unit excavator Liugong 922E," Prosiding Seminar Nasional Teknik Mesin, vol. 14, no. 1, pp. 288–296, 2024.

[14] R. Mubarok, A. Maksum, and R. Noval, "Case study of leakage in the steering cylinder unit of wheel loader CLG855N at PT XYZ," Prosiding Seminar Nasional Teknik Mesin, Politeknik Negeri Jakarta, 2024.

[15] M. R. Ichsan, M. Hidayatullah, and R. Noval, "Studi kasus pengaruh fuel shutoff solenoid dalam permasalahan engine failure to start pada unit mini excavator 302.5," Prosiding Seminar Nasional Teknik Mesin, vol. 15, no. 2, pp. 2202–2208, 2026.

[16] F. Hibatulloh, R. Noval, and M. Ridwan, "Pemeliharaan engine wiring harness pada mini hydraulic excavator CAT 302.5 di workshop alat berat PNJ," Prosiding Seminar Nasional Teknik Mesin, vol. 15, no. 2, pp. 2504–2513, 2026.

[17] M. A. Saklil, R. Noval, and A. Y. Yusyama, "Analisis penyebab low fuel pressure pada engine C27 di unit track dozer D10T," Prosiding Seminar Nasional Teknik Mesin, Politeknik Negeri Jakarta, 2025, pp. 779–785.

[18] M. R. Ibrahim Batistuta, A. Maksum, and R. Noval, "Analisis penyebab munculnya warning low fuel pressure pada monitor unit excavator Liugong 922E," Prosiding Seminar Nasional Teknik Mesin, vol. 14, no. 2, pp. 1574–1583, 2025.

[19] R. G. Muzib, R. Noval, and A. Maksum, "Analisa penyebab keausan crankshaft yang mengakibatkan high blow-by pada engine CAT C18," Prosiding Seminar Nasional Teknik Mesin, vol. 15, no. 2, pp. 2350–2357, 2026.

[20] IEC, IEC 62619:2022, Secondary cells and batteries containing alkaline or other non-acid electrolytes—Safety requirements for secondary lithium cells and batteries, for use in industrial applications, 2nd ed., 2022.

PT. Mencerdaskan Bangsa Indonesia

PT MENCERDASKAN BANGSA INDONESIA
(MBI), 4th Floor Gedung STC Senayan Room
31-34, Jl. Asia Afrika Pintu IX, Jakarta 10270,
Indonesia.

ISSN 2985-704X
9 772985 704006