Doctoral theses at NTNU, 2023:419

Muhammad Gibran Alfarizi

Data-driven design for fault prognosis

Application to industrial components, subsystems, and systems

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of Science and Technology

Muhammad Gibran Alfarizi

Data-driven design for fault prognosis

Application to industrial components, subsystems, and systems

Thesis for the Degree of Philosophiae Doctor

Trondheim, December 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering
Department of Mechanical and Industrial Engineering

ISBN 978-82-326-7538-8 (printed ver.)
ISBN 978-82-326-7537-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:419

Printed by NTNU Grafisk senter

"Essentially, all models are wrong, but some are useful." - George Box.

Preface

This thesis is submitted in partial fulfillment of the requirements for the degree of Philosophiae Doctor at the Department of Mechanical and Industrial Engineering - Norwegian University of Science and Technology (NTNU).

One of my childhood dreams was to become a Professor. I do not really remember it, as it was told by my mother and grandmother. As time went by, I never envisioned myself becoming a Professor. My goal was only to get a well-paid job and climb the career ladder. Unfortunately, when I got my bachelor's degree in 2017, I did not get a job, even though I was applying everywhere. At that time the oil price was still quite low, so not many oil and gas companies recruited graduates. I was then thinking of getting a master's degree, and my parents are supportive of it. It was their intention for me to get the highest education degree, though the decisions are left to me. I was then accepted to study master's degree in petroleum engineering at NTNU.

Again, I finished my master's degree at an unfortunate time, as we are in the middle of a pandemic. But this time I am actually considering taking a doctorate degree. Fortunately, I was offered a Ph.D. position in RAMS before the pandemic. Now here I am, writing my Ph.D. thesis, many years later after I said I wanted to become a Professor. I might not become a Professor, but getting a Ph.D. degree is pretty close to that. I'm always interested in science and math, and hopefully, I will continue to do so.

Professor Shen Yin and Professor Jørn Vatn have been the main supervisor and co-supervisor, respectively. This Ph.D. position was funded by NTNU and carried out from August 2020 until August 2023. The target audience of this thesis includes researchers and practitioners interested in the areas of fault prognosis, accident prevention, and machine learning applications in industry.

Trondheim, September 2023
Muhammad Gibran Alfarizi

Acknowledgement

This section is dedicated to everyone who has supported my Ph.D. study over the last three years.

First, I want to express my gratitude to my supervisor, Professor Shen Yin, who has been very supportive and helpful during my Ph.D. journey. I was very anxious in the first year of my Ph.D. because of the unclear path forward of my study. However, with his extensive experience in academia and his guidance, the journey thereafter was smooth sailing. He helped me to find peers to collaborate with and clear guidance towards the research and publication. I am happy that we published three journals together even before the end of my study. I could not ask for a better supervisor!

I would also like to thank my co-supervisor, Professor Jørn Vatn, who has also been acting as my main supervisor while waiting for Professor Shen Yin to come to Norway. Jørn has been very critical of my work and provides insightful advice about the relationship between academic work and implementation in industry. His perspective and comments are always stimulating and awaited whenever we present our work in the RAMS seminar.

Thanks to my colleagues whom I collaborated with (Bahar, Federico, and Nicola) for your help and input towards our publication together. The work feels easier when I have somebody to ask and provide an outside perspective. Additionally, thanks to my office colleagues in RAMS (Tianqi, Xingheng, Jie, Yixin, Emefon, Dimi, Ewa, Tom Ivar, Nanda, Renny, and all the dear others) who made my time in NTNU enjoyable and not lonely. Special thanks to my friend in Discord "Paguyuban Indo Norwegia" (Mikael, Fadhil, Bonti, Ikhsan, Salma, WW, Handita, and all the dear others) for playing with me even though I have to carry them so hard in-game.

Finally, my greatest gratitude to my family who has been nothing but supportive in every step that I take. Especially to my wife, Dea Lana Asri, who patiently waited for me to finish this Ph.D. before marrying her :) This one is for you!

Abstract

Technical processes in diverse industries, like manufacturing, chemicals, and power generation, involve intricate operations that aim to achieve specific outcomes. These operations often entail complex interactions among components and systems. However, they can also carry substantial risks to health, the environment, and industry sustainability. Hence, the implementation of a robust fault prognosis system is paramount for safeguarding the safety and dependability of these intricate technical processes.

Conversely, these technical processes frequently accumulate vast quantities of historical data through routine sensor measurements, event logs, and records. This observation fuels a compelling interest in crafting fault prognosis methods solely reliant on this abundant process data. Consequently, the main objective of this thesis was to design effective data-driven fault prognosis strategies tailored to diverse operational scenarios.

This thesis explores fault prognosis across various technical process levels, including key components, subsystems, and systems. It begins with basic component-level issues and progresses to intricate system-level challenges, aiming to gain a holistic grasp of data-driven fault prognosis complexities in industrial contexts.

The first objective aimed to create a reliable fault prognosis system for critical technical process components, specifically roller bearings. A novel data-driven prediction framework was proposed, involving two phases: feature extraction via Empirical Mode Decomposition and Remaining Useful Life (RUL) prediction using an RFs-based model with hyperparameters fine-tuned through Bayesian optimization. Notably, this approach demonstrated substantial enhancements in RUL prediction accuracy compared to conventional data-driven and stochastic methods during an actual run-to-failure experiment involving roller bearings.

The second objective aimed to create an effective fault prognosis system for technical process subsystems, with a focus on preventing operational failures. The research selected an automated fuse test bench, a manufacturing line subsystem, as the subject of study. Initially, an integrated fault diagnosis system based on extreme gradient boosting was introduced, showcasing superior performance in detection and classification accuracy while achieving quicker diagnosis times than standard approaches. Subsequently, the scope of extreme gradient boosting was broadened to encompass fault prognosis through methodological enhancements and the incorporation of supplementary data streams like images.

The third objective focused on designing an accurate predictive model for anticipating future operating conditions in industrial systems to avert catastrophic ac-

cidents. The study honed in on a liquid hydrogen storage system as its research subject. A novel application of the random forests algorithm was introduced to enable early detection of hazardous incidents like liquid hydrogen spills, thereby averting catastrophic outcomes like detonation. The model demonstrated remarkable accuracy, surpassing other machine learning methods previously employed for similar experiments. This model, forged through the study, offers valuable insights for comprehensive risk analysis and the identification of prevention and mitigation measures, especially in the context of emerging liquid hydrogen technology applications.

This Ph.D. study offers the potential for enhanced industrial fault prognosis methods, fostering improved safety and sustainability within the industry. Furthermore, it may serve as a valuable reference and launching point for future academic research, offering insights into the merits and complexities of employing data-driven techniques in real-world industrial applications.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This opening chapter provides the background for this Ph.D. thesis and its objectives, research methodology, scope and limitations, and structure of the thesis.

1.1 Background

Technical processes refer to a sequence of operations or actions taken in various industries, such as manufacturing, chemical, and power generation, to achieve desired outcomes or products. These processes often involve complex interactions among different components, subsystems, and systems to accomplish specific goals efficiently and effectively. However, these processes can also pose significant risks to human health, the environment, and the sustainability of the industries themselves [5]. Consequently, there has been growing recognition of the need to improve the safety of technical processes to minimize these risks and ensure that industries are able to operate sustainably over the long term [6].

Figure 1.1 illustrates the three crucial aspects of the sustainability of the technical process: safety, security, and energy efficiency [3]. Paramount to the success of these processes is the safety and security of their operations, given that any failures in these areas can result in significant harm to people, the environment, and the economy [7]. In addition to safety and security, ensuring the use of efficient and low-carbon energy sources is vital to establish sustainable technical processes for the long term while also minimizing environmental impact [8].

Technical processes often involve hazardous chemicals, high temperatures, and complex machinery, which pose considerable risks to workers and the environment if not adequately managed [9]. Moreover, as industries become increasingly complex and interconnected, the risks associated with technical processes become

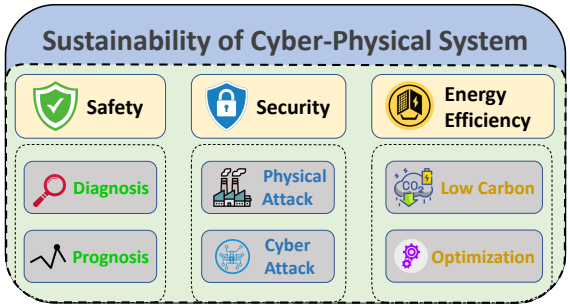

Figure 1.1: The three pillars of sustainability in technical processes [3].

more challenging to predict and mitigate [10].

Ensuring the safety of technical processes is a critical aspect of modern industrial operations. It is vital not only to protect workers and the environment but also to ensure the long-term sustainability of the industries themselves. Industrial accidents, pollution, and reputational damage can have serious economic and social consequences, impacting the health and safety of individuals, businesses, and communities [11]. In recent years, there has been an increased focus on the development of effective safety measures and the implementation of fault prognosis methods [1, 12, 13].

Fault prognosis plays a central role in ensuring the safety of technical processes. It involves predicting or forecasting a system's, machine's, or equipment's future performance based on current and past operating conditions and data. The objective of fault prognosis is to identify potential faults, defects, or failures before they occur, enabling preventive measures to minimize downtime, reduce maintenance costs, and enhance safety [14]. Additionally, the prognosis can predict a product's remaining useful life (RUL) within appropriate confidence intervals, informing decision-makers about potential cost avoidance activities and ensuring safe operation [15].

The field of fault prognosis has undergone rapid advancements in the past decade, with the development of sophisticated technologies that allow for more effective and efficient monitoring of technical processes. These technologies can be broadly classified into two categories: model-based and data-driven methods.

Model-based methods rely on mathematical models based on physical laws or a priori knowledge of the underlying system to predict faults [16]. These methods typically use sensor data to validate the model and refine the prognosis. In contrast, data-driven methods extract useful information for prognosis directly from sensor

data without making any assumptions about the system [1]. Both model-based and data-driven methods have their own strengths and weaknesses. Model-based methods can provide a more accurate prognosis when the underlying physical model is well-understood but can be limited by the complexity of the system and the accuracy of the model [17]. For example, engineers use math models to predict aircraft behavior, but when it comes to forecasting engine issues, simplifications in these models can limit accuracy due to the complexity of aircraft engines [18]. The intricate nature of these engines, with unpredictable factors, makes modeling their degradation challenging [19]. Data-driven methods, on the other hand, can handle complex systems with multiple interacting components but may struggle with identifying faults that do not have a clear signature in the sensor data [20].

Advancements in computing power have made it possible to process and analyze large volumes of sensor data generated by technical processes, leading to the prevalence of data-driven methods in academia. Integrating both model-based and data-driven methods can improve the accuracy and efficiency of fault prognosis, as demonstrated by a variety of examples in the literature [21–25]. As the volume of data continues to grow, more efficient and accurate fault prognosis methods are urgently needed to ensure the safety and sustainability of technical processes.

Although data-driven methods have the potential to significantly improve industrial operations, their adoption in industry is still limited due to their emerging nature and the need for further research and validation. This is where the present Ph.D. project contributes: it aims to implement current data-driven methods for prognosis in real-world industrial cases, validate their effectiveness, and improve them to make them more practical for industrial demands and address their limitations.

The research presented in this thesis examines fault prognosis at different levels of technical processes, including key components, subsystems, and systems. By starting with simple component-level problems and gradually progressing toward more complex system-level scenarios, this work seeks to develop a comprehensive understanding of the challenges and opportunities associated with data-driven fault prognosis in industrial settings.

This Ph.D. study holds promise for developing more effective and practical methods for industrial fault prognosis, which can ultimately lead to better safety outcomes and a more sustainable future for the industry. Moreover, this study has the potential to serve as a comprehensive reference or starting point for future research in academia, providing insights into the benefits and challenges of implementing data-driven methods in real-world industrial settings.

1.2 Objectives

The primary objective of this thesis was to design efficient fault prognosis techniques for technical processes using data-driven methods that consider the operating conditions of the process. Specifically, the goals of this thesis are stated as follows:

1. Develop a reliable fault prognosis system for crucial technical components, specifically roller bearings. The proposed data-driven framework involved two phases: feature extraction using Empirical Mode Decomposition and RUL prediction with an RFs-based model fine-tuned by Bayesian optimization.

2. Develop an efficient fault prognosis system for technical process subsystems that prevents operational failures. The research centered on an automated fuse test bench, a subsystem in manufacturing lines. It initially used extreme gradient boosting for better fault detection and classification with faster diagnosis. Later, it expanded to include fault prognosis, enhancing methods and incorporating extra data like images.

3. Develop an accurate predictive model for assessing future operating conditions within industrial systems to prevent catastrophic accidents. The study focused on a liquid hydrogen storage system. A novel application of the random forests algorithm enabled early detection of hazardous incidents like liquid hydrogen spills, preventing catastrophic outcomes like detonation.

In this thesis, one industrial benchmark process and two experiments are utilized to evaluate the effectiveness of the proposed approaches for fault prognosis purposes. The research questions and objectives are discussed in detail in chapter 3.

1.3 List of publications

This thesis is comprised of a series of papers produced during the doctoral studies, including three papers that have undergone rigorous peer-review in international journals, as well as two conference papers. Of these publications, three journal papers and one conference paper have been successfully published. Furthermore, one conference paper is planned for submission. The following section provides details regarding each of these publications.

1. Journal paper 1:
   [26] M. G. Alfarizi, B. Tajiani, J. Vatn and S. Yin, "Optimized Random

Forest Model for Remaining Useful Life Prediction of Experimental Bearings," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 6, pp. 7771-7779, 2023.

2. Journal paper 2:
   [27] M. G. Alfarizi, J. Vatn and S. Yin, "An Extreme Gradient Boosting Aided Fault Diagnosis Approach: A Case Study of Fuse Test Bench," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 661-668, 2023.

3. Journal paper 3:
   [28] M. G. Alfarizi, F. Ustolin, J. Vatn, S. Yin, and N. Paltrinieri, "Towards accident prevention on liquid hydrogen: A data-driven approach for releases prediction," *Reliability Engineering & System Safety*, vol. 236, p. 109276, 2023.

4. Conference paper 1:
   [29] M. G. Alfarizi, J. Liu, J. Vatn and S. Yin, "Sustainability of ICPS from a Safety Perspective: Challenges and Opportunities," *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*, Helsinki, Finland, pp. 1-8, 2023.

5. Conference paper 2 (to be submitted):
   [30] M. G. Alfarizi, J. Vatn and S. Yin, "Advancements in extreme gradient boosting for enhanced fault prognosis: A continuation study from fuse test bench analysis," *2024 IEEE 33rd International Symposium on Industrial Electronics (ISIE)*, Ulsan, South Korea, pp. 1-8, 2024.

1.4 Scope and limitations

This Ph.D. research focuses on the fault prognosis of technical processes, specifically utilizing data-driven approaches. The research scope does not extend to decision support functions such as maintenance optimization. Furthermore, model-based approaches will not be studied.

The primary objective of this research is to develop data-driven methods for fault prognosis that are applicable in industries with specific safety requirements. Such industries may include the oil and gas industry, manufacturing industry, and maritime industry, among others. These industries are often subject to degradation and require high reliability, limited maintenance activities, and strict safety protocols. By targeting these specific industries, this research aims to provide practical solutions to real-world problems and enhance the safety and efficiency of industrial processes.

However, a significant challenge encountered during the research is the limited availability of relevant data from industry sources. Due to the sensitivity of industrial data and privacy concerns, obtaining comprehensive and firsthand data has proven to be difficult. Consequently, the research relies on publicly available data, which may introduce the following potential limitations:

1. Data Relevance and Specificity. Publicly available data might not fully capture the intricacies and specificities of the industrial processes under investigation, potentially limiting the applicability and accuracy of the developed fault prognosis methods.
2. Limited Granularity. The granularity of publicly available data may not be sufficient to address the fine details of fault detection and prognosis in complex industrial systems, potentially impacting the precision of the proposed methodologies.
3. Lack of Domain Expertise. Publicly available data may lack the context and domain-specific insights that could be provided by industry experts. The absence of such expertise may influence the depth and accuracy of the research findings.

Despite these challenges, the research aims to navigate and mitigate these limitations effectively, striving to contribute valuable insights to the field of fault prognosis in safety-critical industrial settings, even in the face of data availability constraints.

1.5 Structure of the thesis

The thesis is organized into several chapters. Chapter 2 provides a comprehensive overview of the theoretical background that serves as the foundation for the thesis. This chapter will introduce relevant terms, concepts, and basic theory necessary to understand the formulations developed throughout the thesis. Chapter 3 outlines the research objectives and presents the main research questions that guide this study. Chapter 4 describes the research methodology and work process. Chapter 5 summarizes the main findings of the thesis, highlighting the contributions made by the submitted papers. This chapter will discuss the results in the context of the research questions and objectives presented in chapter 3. Finally, chapter 6 presents concluding remarks and proposed future works. This chapter will summarize the key findings of the thesis and provide recommendations for future research. The proposed future works will build upon the research presented in this thesis and further advance our understanding of data-driven fault prognosis in industrial settings.

Chapter 2

Data-driven fault prognosis in technical processes

Fault prognosis is a critical component in the domain of industrial technical processes. Its primary objective is to predict the time to failure of machinery, enabling proactive maintenance and preventing unexpected equipment breakdowns. This predictive maintenance technique is vital in industrial environments, where unanticipated equipment failures can lead to substantial production losses, safety risks, and elevated repair costs.

The complexity and interconnected nature of industrial processes underscore the necessity for effective techniques in fault prognosis. Accurate fault prognosis can not only enhance the safety and reliability of industrial systems but also optimize maintenance scheduling and reduce operational costs.

In the face of these challenges, data-driven techniques have emerged as a promising solution. Leveraging the power of modern computational methods and the wealth of data generated by industrial processes, these techniques offer improved accuracy in fault prognosis. They are particularly adept at handling the complexity and non-linearity of industrial systems, with methods such as machine learning and deep learning leading the way.

This chapter will delve into the current state of the art in data-driven techniques for fault prognosis in industrial technical processes. It will explore various data-driven methods, discuss their applications and limitations, and highlight recent advancements in the field.

The primary objective of this Ph.D. thesis is to enhance the current state of fault

prognosis by devising innovative methodologies and techniques that enhance the dependability and effectiveness of industrial processes. This endeavor aims to contribute to the state of the art in fault prognosis, paving the way for more reliable and efficient industrial processes.

2.1 Advancement in fault prognosis

Fault prognosis, a critical aspect of predictive maintenance in industrial technical processes, has seen significant evolution over the years. The primary objective of fault prognosis is to predict the time to failure of a system or component, enabling active maintenance and minimizing downtime.

According to the literature, the current methods for fault prognosis can be broadly categorized into three types: knowledge-based methods, model-based methods, and data-driven methods. Knowledge-based methods rely on engineering experience and past events, resulting in prognosis outcomes that are more intuitive. These methods are particularly suitable when process models are readily available or when process knowledge has been accumulated. However, developing process knowledge is a time-consuming task, and knowledge-based models heavily depend on the expertise of individuals. In some cases, it may be impossible to construct such models due to the issue of combinatorial explosion. Consequently, the prediction accuracy is significantly compromised, and the applicability of these methods is considerably constrained.

On the other hand, model-based methods utilized mathematical models, often derived from physical laws or expert knowledge, to represent the behavior of industrial systems. The models were then used to predict the system's future behavior and identify potential faults. However, these model-based techniques had their limitations. They required accurate models, which were often difficult to develop due to the complexity and non-linearity of industrial systems. Furthermore, they were not well-suited to handle the variability and uncertainty inherent in industrial processes.

With the advent of powerful computational methods and the increasing availability of data from industrial processes, a paradigm shift occurred toward data-driven techniques. Unlike model-based techniques, data-driven techniques do not require explicit mathematical models of the system. Instead, they leverage the data generated by the system to learn its behavior and predict faults. These techniques, which include machine learning and deep learning methods, have shown promise in handling the complexity of industrial systems and providing accurate fault prognoses.

The transition from model-based to data-driven techniques in fault prognosis rep-

resents a significant advancement in the field. However, it's important to note that both types of techniques have their strengths and limitations, and the choice between them often depends on the specific requirements of the industrial process in question.

2.2 Data-driven fault prognosis process

A typical data-driven fault prognosis process consists of three main steps: data acquisition, feature extraction and selection, and model selection, training, and testing. The illustration of such a process is shown in Fig. 2.1.

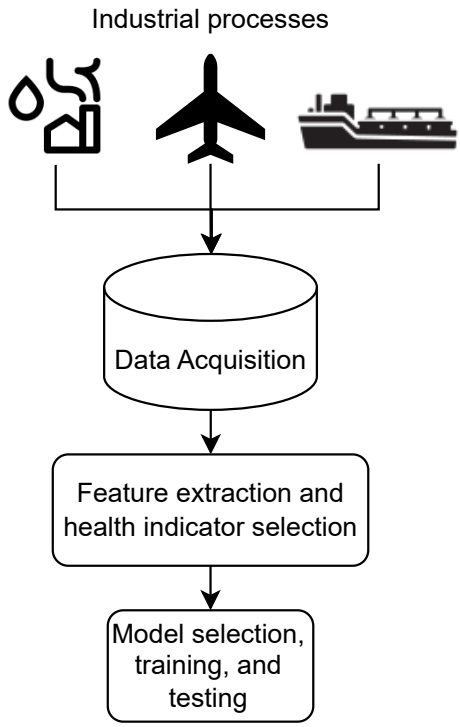

Figure 2.1: A typical data-driven fault prognosis process.

2.2.1 Data acquisition

Data acquisition refers to the process of gathering relevant data from various sources that are necessary for analyzing and predicting the occurrence of faults or failures in a system. This step involves capturing and recording data related to the system's operation, performance, and condition.

Data acquisition methods can vary depending on the nature of the system being monitored. It may involve the use of sensors, monitoring devices, or data logging systems to collect real-time measurements or historical data. The data can include various types of information such as sensor readings, operational parameters, environmental conditions, maintenance records, and other relevant variables.

The data acquisition process requires careful consideration of the sensors or data sources to be used, their placement or installation, and the sampling rate or frequency at which data is collected. It is important to ensure that the acquired data is accurate, reliable, and representative of the system's behavior and operating conditions.

Additionally, data preprocessing techniques may be applied during the data acquisition phase. This involves tasks such as data cleaning, filtering, normalization, and outlier detection to ensure the quality and integrity of the acquired data. Preprocessing steps are essential for enhancing the effectiveness of subsequent data analysis and modeling stages.

2.2.2 Feature extraction and health indicator selection

Feature extraction aims to identify and extract key information from the raw data that can be used to describe and represent the system's behavior. This process involves applying mathematical or statistical techniques to transform the data into a more concise and informative representation. The commonly used methods for feature extraction in regard to fault prognosis have been summarized by Zhong *et al.* in [1]. The techniques can be broadly categorized into two groups: statistical methods, such as Principal Component Analysis (PCA) [31], Independent Component Analysis (ICA) [32], Partial Least Squares (PLS) [33], Fisher Discriminant Analysis (FDA) [34], and subspace-aided monitoring [35]; and engineering knowledge-based methods, including Fourier transform [36] and wavelet analysis [37]. Table 2.1 provides a comparison of the common feature extraction methods for fault prognosis [1].

Health indicator selection, on the other hand, involves identifying the subset of features that are most relevant for fault prognosis. It aims to eliminate redundant or irrelevant features that may introduce noise or unnecessary complexity to the modeling process. Health indicator selection techniques can be applied to rank or score the importance of each feature based on criteria such as their predictive power, correlation with the target variable (e.g., fault occurrence), or contribution to reducing the model's complexity.

Effective feature extraction and health indicator selection can enhance the accuracy and efficiency of subsequent modeling and prediction stages in data-driven fault

Table 2.1: Comparison of the common feature extraction methods for fault prognosis [1].

Method	Advantage	Disadvantage
PCA [31]	Simplify data dimension and reduce the computational complexity	The frequency characteristics of the process data are disregarded
ICA [32]	Statistically independent and appropriate for non-Gaussian process	The training duration for multidimensional data is excessively lengthy
PLS [33]	Mitigate the effects of multicollinearity among variables	Unable to identify variables with weak correlations
FDA [34]	Sample data demonstrates optimal separability in projection subspaces	Highly susceptible to noise and reliant on labeled datasets
Subspace-aided [35]	Extracted fault subspace variables can be processed directly	Conditions for variable selection are subjective
Fourier transform [36]	To enhance display signal in spectral resolution	Not applicable to dynamic signals
Wavelet analysis [37]	Well-suited for dynamic signal processing	Challenging to select the optimal model parameters

prognosis. By focusing on the most relevant and informative features, the resulting models can better capture the patterns and relationships necessary for accurate fault detection and prediction.

2.2.3 Model selection, training, and testing

Model selection, training, and testing are critical stages in the data-driven fault prognosis process. These steps involve choosing an appropriate model, training it using the available data, and evaluating its performance to ensure accurate and reliable fault predictions.

Model selection involves choosing the most suitable algorithm or model structure that can effectively capture the relationships and patterns in the data. Various machine learning and statistical techniques can be utilized, such as decision trees, support vector machines, neural networks, or regression models. The selection process considers factors such as the complexity of the model, its interpretability, computational requirements, and the specific characteristics of the fault prognosis problem at hand.

Once a model is selected, the next step is training. Training involves feeding the model with a labeled dataset, where the input data is paired with corresponding fault labels or outcomes. The model learns from these examples and adjusts its internal parameters or weights to minimize the difference between the predicted outputs and the actual labels. This learning process allows the model to capture the underlying patterns and relationships within the data.

After training, the model performance needs to be assessed through testing. This involves evaluating the model's ability to make accurate predictions on new, unseen data. A separate dataset, often referred to as a testing or validation dataset, is used for this purpose. The model's predictions are compared against the true fault

outcomes or labels in the testing dataset, and performance metrics such as accuracy, precision, recall, or F1 score are calculated. This evaluation helps to assess the model's generalization ability and provides an indication of how well it would perform on unseen data in real-world scenarios.

Model selection, training, and testing are iterative processes. Different models can be compared based on their performance metrics, and further adjustments or refinements can be made to improve the model's predictive capabilities. This iterative approach helps select the most accurate and robust model for fault prognosis applications, ensuring reliable and effective predictions in real-time scenarios.

2.3 Data-driven fault prognosis methods

2.3.1 Statistical methods

Fault prognosis, as mentioned earlier, predicts a product's future reliability based on past and present health data. These health data, known as condition monitoring (CM) data, are obtained through ongoing inspections and serve as indicators of the system's health. Examples of CM data include tire wear, chemical concentration, fatigue fracture size, and light intensity from LEDs. As the system degrades due to usage, this degradation is reflected in the observable CM data (e.g., a decrease in the light intensity of the LED). CM data is therefore considered the signal of system deterioration, and failures are determined when a predefined threshold set by experts is reached. By modeling the degradation history and identifying when it exceeds the failure threshold, we can predict the remaining useful life (RUL) of the system.

The data used for RUL estimation can generally be categorized into event data and CM data. Event data refers to historical failure data, which may be limited to critical assets that are not allowed to operate until failure. On the other hand, CM data is a valuable source of information for RUL estimation.

Statistical data-driven methodologies are based on the availability and quality of data. According to literature [38], the observed CM data can be classified into two categories: direct CM and indirect CM. The RUL estimation essentially involves forecasting the CM data to reach a predefined threshold level, since direct CM data directly reflect the underlying condition of the system. Common examples of direct CM data include wear and crack sizes, which can be easily observed and measured. Indirect CM data refers to data that can only provide indirect or partial indications of the underlying system condition. For an accurate estimation of the RUL, it is often necessary to supplement the CM data with failure event data. Examples of this type of data include vibration and oil-based monitoring data. These

sources of information offer insights into the system's condition, albeit indirectly. By incorporating both CM data and failure event data, a more comprehensive understanding of the system's health can be obtained, enhancing the accuracy of RUL estimation.

Major categories of the statistical data-driven prognosis [39] are discussed in the following section.

Regression Based Model

Techniques within this group predominantly revolve around constructing a parametric trajectory of CM data, with random effects, that can be linear or nonlinear. The majority of prevalent methods for RUL prediction operate on the assumption that products sharing a type or batch have identical probability of failure traits. However, while overall population behavior can act as a guideline, it fails to precisely track the health progression of every individual item. This happens because each product often encounters unique usage scenarios, disparate environments, or even varying quality due to process fluctuations. Thus, it is vital to align with the health trajectories of each singular product instead of relying on interpreted group averages for more accurate reliability forecasting. Several techniques have been introduced that combine population data with insights from individual items for improved RUL prediction.

For example, a linear degradation that employs a Log-normal rate [40], where the distribution of failure time F is formulated as:

$$F(t, \beta_0, \mu, \sigma) = P(\beta_0 + \beta_1 t > L) = P\left(\beta_1 > \frac{L - \beta_0}{t}\right) \quad \text{(Eq. 2.3.1)}$$

$$= \Phi_{\text{nor}}\left(\frac{\log(t) - [\log(L - \beta_0) - \mu]}{\sigma}\right), t > 0 \quad \text{(Eq. 2.3.2)}$$

where μ is mean, σ is standard deviation, β_0 is constant, $\beta_1 \sim \text{LOGNOR}(\mu, \sigma)$, and L is the predetermined threshold.

Independent Increment Process Based Model

This model is also known as the stochastic process model, comprising two fundamental elements:

1. A stochastic process,
$$X(t), t \in \mathcal{T}, X \in \chi, \quad \text{(Eq. 2.3.3)}$$
with $X(0) = x_0$, where χ is the state space of the process and $\mathcal{T}$ is the time space.

2. A boundary set B, where $B \subset \chi$. If $X(0) = x_0$ is outside B, the first hitting time (FHT) is defined as random variable T, where $T = \inf\{t : X(t) \in B\}$. B is often reduced to threshold x^* or L and the FHT is defined as when $X(t)$ hits x^*.

There are several stochastic processes, depending on which distribution the degradation signal $X(t)$ will follow. These processes include Gamma process, Wiener process, and Inverse Gaussian process.

Gamma Process. Gamma process is used to model monotonically increasing degradation signals. In this process, the increment $\Delta X(t) = X(t + \Delta t) - X(t)$ in a specified time interval Δt has a Gamma distribution $\text{Ga}(\alpha\Delta t, \sigma)$ with scale parameter $\sigma > 0$ and shape parameter $\alpha\Delta t > 0$.

The Probability Density Function (PDF) of the first hitting time T is inverse Gaussian distributed, and formulated as $P_r(T > t) = P_r(X(t) < L)$ due to monotony. The modeling degradation steps with Gamma process are discussed in [41–43].

Wiener Process. A Wiener process $\{W(t), t \geq 0\}$ is formulated as,

$$W(t) = \lambda t + \sigma B(t) \qquad \text{(Eq. 2.3.4)}$$

where $\lambda, \sigma > 0,$ and $B(t)$ are the drift parameter, diffusion coefficient, and Brownian motion respectively. The Wiener process is characterized by $W(0) = 0$, $W(t)$ is almost sorely continuous and has independent increments, and $W(t) - W(s) \sim N(\mu(t - s), \sigma^2(t - s))$ for $0 \leq s \leq t$.

The PDF of the first hitting time T is inverse Gaussian distributed $\text{IG}((L-x_0)/\lambda, (L-x_0)^2/\sigma^2)$, which can be expanded to

$$P_r(T \leq t) = \int_0^t \frac{L}{\sqrt{2\pi\sigma^2\mu^3}} \exp\left\{-\frac{(L - \lambda\mu)^2}{2\sigma^2\mu}\right\} du \quad . \qquad \text{(Eq. 2.3.5)}$$

There are several studies that applied this model for prognostics purposes [44–48]. These models recursively update the parameters, making the prognostics history-dependent [49, 50].

Inverse Gaussian Process. The inverse Gaussian process $\{X(t), t \geq 0\}$ has specific characteristics, such as its mean function $v(t)$ and scale parameter η. In this process, the increment $\Delta X(t)$ follows an inverse Gaussian distribution $\text{IG}(\Delta t, \eta(\Delta t)^2)$. It shares similarities with the Gamma process, displaying a monotone path and approximating the failure time distribution to a Birnbaum-Saunders type distribution, which offers advantages for future computations.

Despite its potential benefits in incorporating random effects and covariates, the use of the IG process in degradation modeling is still relatively new and not widely adopted. Wang and Xu [51] introduced it to handle random effects, and Ye and Chen [52] extensively explored various aspects, including the random drift model, random volatility model, random drift-volatility model, and integration of covariates.

Markovian Process-Based Models

Another set of methodologies utilizes memoryless Markov processes. While these processes are categorized as stochastic processes, they distinguish themselves from the models previously discussed by assuming a finite degradation state and focusing on transition probabilities between these states. This group of methods includes the following major variations:

Markov Chain Model. This model assumes that the degradation process $\{X_n, n \geq 0\}$ occurs within a finite state space $\Phi = \{0, 1, ..., N\}$, where 0 denotes the perfect healthy state and N represents the failed state of the observed system. The RUL at time instant T is defined as $T = \inf\{t : X_{n+t} = N | X_n \neq N\}$. Historical data are used to estimate the probability transition matrix and the number of states. This approach categorizes the health status into distinct states such as "Perfect," "Fault detected," "Maintenance needed," and "Break down," thereby offering practical outcomes that can be readily comprehended by field engineers.

Semi-Markov Processes. The semi-Markov process $\{X(t), t \geq 0\}$ extends the Markov chain model by incorporating the random time the process spends in each state. Although this extension generally results in the loss of the Markov property, the model remains highly practical. In this model, the first hitting time indicates the duration during which the process remains in the initial and subsequent states before it enters one of the states that define set B for the first time.

Hidden Markov Model (HMM). This model comprises of two stochastic processes – a hidden Markov chain $\{Z_n, n \geq 0\}$, representing the unobservable true degradation state, and an observable process $\{Y_n, n \geq 0\}$, which is the monitored signal. This approach shares a similarity with Markovian-based models, as it assumes that the degradation process advances through a finite state space using a Markov chain. $P(Y_n | Z_n = i), i \in \Phi$, is the conditional probability measure that connects $\{Y_n, n \geq 0\}$ and $\{Z_n, n \geq 0\}$. Thus, the RUL at time n is $T = \inf\{t : Z_{n+t} = N | Z_n \neq N, Y_j, 0 \leq j \leq n\}$. This model is particularly favored when only indirect observations are available [53].

Filtering-Based Models

The Kalman filtering model doesn't directly utilize the CM as the actual degradation signal, much like the HMM. Instead, it assumes that the true degradation state is not directly observable but is somehow linked to the CM data. This model considers the CM data y_t and the unobserved condition x_t are linked by $x_t = \alpha x_{t-1} + \epsilon_t$ and $y_t = \beta_{x_t} + \eta_t$, where α, β are the parameters of the state space model and ϵ_t, η_t represent Gaussian noises. Unlike some methods that rely solely on the most recent CM status, the Kalman filtering model leverages all historical data. However, its applicability is constrained by assumptions of linearity and Gaussian noise, and researchers have made efforts to address these issues [38,54].

Proportional Hazard Model

The Proportional Hazard Model [55] has been extensively researched across various fields. When incorporating time-dependent variable(s), this model can effectively combine event data and CM data, which is advantageous when dealing with hard failures or uncertain failure thresholds [56–58]. The model's hazard rate is formulated as follows:

$$h(t) = h_0(t) \exp\left(\boldsymbol{\gamma}\mathbf{z}(t)\right), \qquad \text{(Eq. 2.3.6)}$$

where $h_0(t)$ represents the baseline hazard rate, $\mathbf{z}(t)$ is a vector of time-dependent variables, and $\boldsymbol{\gamma}$ is a vector of coefficients. The $h_0(t)$ can be either nonparametric or parametric, and the model parameters can be determined using the maximum likelihood method. The condition monitoring data are regarded as time-dependent covariates within $\mathbf{z}(t)$. By employing Eq. (Eq. 2.3.6), the system failure distribution can be calculated.

2.3.2 Machine learning methods

In the field of fault prognosis in technical processes, there are numerous machine learning techniques available. These methods are primarily used to forecast how much longer a system or part can be effectively used, a concept known as Remaining Useful Life (RUL). Achieving this prediction involves various approaches. One approach involves using historical data to spot patterns and estimate how much life is left. Another method starts by assessing the current damage and then projecting how it will evolve over time until it reaches a point where failure is likely.

Predicting the future state of degradation in a system relies on creating models and recognizing significant factors. However, it's important to acknowledge that there will always be uncertainties when making predictions about future conditions in monitoring systems. As time passes and we move further away from the current state, these uncertainties grow, leading to less precise predictions. In many cases,

RUL predictions are based on health indicators that provide the best insight into the condition of the asset. These indicators are typically derived from an in-depth analysis of various aspects of the system's operation.

This section will explore regression models, a fundamental type of machine learning tool used to predict RUL for systems or components. What sets regression models apart is their ability to provide a continuous prediction for RUL, offering a detailed view of potential issues. In this context, it is noteworthy that linear regression and Support Vector Regression are not discussed as extensively as Random Forest and XGBoost. This selectivity stems from the deliberate choice of Random Forest and XGBoost as the preferred algorithms for implementation in this thesis.

Linear regression

In the context of linear regression [59], a prediction denoted as $\hat{y}$ is generated through a straightforward process involving the calculation of a weighted sum of input predictors $\{x_1, x_2, \ldots, x_n\}$, along with the addition of a constant referred to as the bias term (b). This mathematical representation of the linear regression model can be articulated as follows:

$$\hat{y} = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b = \sum_{i=1}^{n} w_i x_i + b \qquad \text{(Eq. 2.3.7)}$$

where $w = \{w_1, w_2, \ldots, w_n\}$ is a weight vector.

Given that the bias term (b) may be considered negligible in certain cases, its estimation will not be described at this point. Instead, our primary objective is to estimate the weight vector (w) using a training dataset consisting of m pairs, each represented as (x_j, y_j), where x_j represents an n-dimensional training instance, and j ranges from 1 to m. To accomplish this objective, we can establish a cost function, defined as follows:

$$J(w) = \arg\min_{w} \sum_{j=1}^{m} (y_j - \hat{y}_j)^2 = \arg\min_{w} \sum_{j=1}^{m} (y_j - w x_j)^2 \qquad \text{(Eq. 2.3.8)}$$

In the pursuit of determining an optimized weight vector (w) that minimizes the squared error as expressed in Eq. (Eq. 2.3.8), a common approach involves computing the derivative of the error with respect to w and subsequently setting it equal

to zero:

$$\frac{\partial}{\partial w}\sum_{j=1}^{m}(\hat{y}_j - y_j)^2 = 2\sum_{j=1}^{m} -x_j(y_j - wx_j) \quad \text{(Eq. 2.3.9)}$$

$$2\sum_{j=1}^{m} -x_j(y_j - wx_j) = 0 \Rightarrow 2\sum_{j=1}^{m} x_j y_j - 2\sum_{j=1}^{m} wx_j x_j = 0 \quad \text{(Eq. 2.3.10)}$$

$$\Rightarrow 2 = 2\sum_{j=1}^{m} wx_j^2 \Rightarrow w = \frac{\sum_{j=1}^{m} x_j y_j}{\sum_{j=1}^{m} x_j^2} \quad \text{(Eq. 2.3.11)}$$

Support Vector Regression

In the scenario of having a training dataset characterized by multivariate sets of m instances represented as x_n, along with corresponding observed response values y_n, the objective is to determine a linear function

$$f(x) = x^T w + b \quad \text{(Eq. 2.3.12)}$$

that exhibits the flattest possible behavior. To achieve this goal, it becomes imperative to seek a function $f(x)$ with the minimal norm value $w^T w$. This particular problem can be effectively formulated as a convex optimization problem, aiming to minimize the function

$$J(w) = \frac{1}{2} w^T w \quad \text{(Eq. 2.3.13)}$$

while ensuring that all residuals possess values lower than ϵ, which is

$$|y_n - (x_n^T w + b)| \leq \epsilon, \forall n \quad \text{(Eq. 2.3.14)}$$

In situations where the existence of a function $f(x)$ to satisfy these constraints for all instances is not guaranteed, it becomes necessary to introduce slack variables ζ_n and ζ_n^* for each instance. This concept closely resembles the notion of soft margins in Support Vector Machine (SVM) [60] classification, as the introduction of slack variables allows for the accommodation of regression errors up to the values of ζ_n and ζ_n^* while still meeting the stipulated conditions. The inclusion of these slack variables results in the objective function taking the following form:

$$J(w) = \frac{1}{2} w^T w + C\sum_{n=1}^{m}(\zeta_n + \zeta_n^*) \quad \text{(Eq. 2.3.15)}$$

subject to $y_n - (x_n^T w + b) \leq \epsilon + \zeta_n, (x_n^T w + b) - y_n \leq \epsilon + \zeta_n^*, \zeta_n \geq 0$, and $\zeta_n^* \geq 0, \forall n$ where constant C is the penalty constraint and is assigned a positive

numeric value. It plays a pivotal role in regulating the penalties imposed on observations that deviate beyond the epsilon margin (ϵ), thereby aiding in the prevention of overfitting, a technique known as regularization. This value essentially dictates the balance between the flatness of the function $f(x)$ and the extent to which deviations larger than ϵ are allowed. In other words, it controls the trade-off between model complexity and the tolerance for deviations from the desired conditions.

The parameter w can be fully characterized as a linear combination of the training observations, and this relationship is succinctly expressed through the following equation:

$$w = \sum_{n=1}^{m} (\alpha_n - \alpha_n^*) x_n \qquad \text{(Eq. 2.3.16)}$$

The function employed for predicting new values then relies exclusively on the support vectors:

$$f(x) = \sum_{n=1}^{m} (\alpha_n - \alpha_n^*) x_n^T x + b \qquad \text{(Eq. 2.3.17)}$$

Random Forest

The Random Forests (RFs) algorithm relies on a collection of predictors influenced by the stochastic values associated with each tree within the forest [61]. The selection of input datasets for modeling RFs is done in a random manner from a given group. The algorithm's notable success can be attributed to its rapid computational performance, efficient handling of large datasets, and minimal issues related to overfitting of predictors [26]. For a visual representation of the general structure of the RFs model, please refer to Figure 2.2.

In the RFs algorithm, each forest takes as input a q-dimensional vector X, where $X = x_1, x_2, \ldots, x_q$. Inside the forest, a collection of L trees, represented as $\{T_1(x), T_2(x), \ldots, T_L(x)\}$, is constructed. Each tree calculates an output value, denoted as $\hat{Y}_1 = T_1(X), \ldots, \hat{Y}_m = T_m(X)$, where m ranges from 1 to L. Here, L represents the total number of trees in the forest.

In the context of a classification task, the RFs' output is determined by

$$\text{Predict}_{\text{RFs}}(X) = \text{majority vote}\{\hat{Y}_m(X)\}_{m=1}^{L}. \qquad \text{(Eq. 2.3.18)}$$

For a regression task, the prediction is computed by estimating the average of all tree predictors, expressed as

$$\text{Predict}_{\text{RFs}}(X) = \frac{1}{L} \sum_{m=1}^{L} \hat{Y}_m(X). \qquad \text{(Eq. 2.3.19)}$$

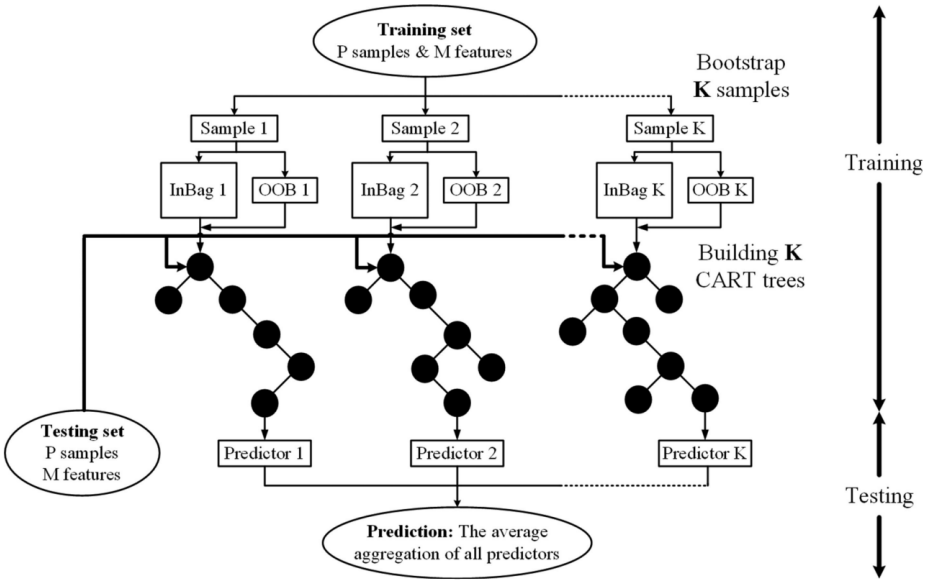

Figure 2.2: The general structure of the RFs model [4].

The training dataset is formed by inputs x_i for $i = 1, \ldots, n$ and corresponding outputs y_i for $i = 1, \ldots, n$. This training dataset is denoted as $T = T_1, T_2, \ldots, T_n = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$.

The RF algorithm employs the bootstrap resampling method to randomly create L tree sample sets from the initial training set T. Approximately two-thirds of the original training data in T are used in the bootstrap samples, which are referred to as in-bag data. The remaining data constitute the out-of-bag data (OOB).

During the construction of each tree in the forest, the algorithm needs to choose an appropriate attribute for splitting at each node. This selection is based on a measure that aims to maximize dissimilarity between classes. The Gini Index is a commonly used criterion in RFs for selecting the best split. Given a training dataset T, the mathematical formula for the Gini Index is expressed as

$$GI = \sum \sum_{j \neq i} (f(C_i, T)/|T|)(f(C_j, T)/|T|) \quad , \qquad \text{(Eq. 2.3.20)}$$

where C_i represents the class to which a randomly selected sample belongs, and $f(C_i, T)/|T|$ represents the probability that a selected case belongs to class C_i. The Gini Index is used to quantify the impurity or heterogeneity in classes within a node. Higher values of the Gini Index indicate greater class heterogeneity, while lower values indicate higher class homogeneity. A successful split is indicated

when the Gini Index of the child node is lower than that of the parent node, signifying that class homogeneity increases as the tree deepens. When the Gini Index reaches 0, it means that each terminal node contains only one class, and the tree-splitting process is complete. At this point, the decision tree has grown to its maximum depth without any pruning.

Once all the L trees in the Random Forests (RFs) are constructed, new data can be predicted by aggregating the outcomes of the predictions from these trees.

In summary, the provided pseudocode in Algorithm 1 outlines the RFs algorithm as described by Breiman in his work [62].

Algorithm 1 Random Forest

Require: Training samples $T = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, testing samples x_t
  for $m = 1$ **to** L **do**
    Using the training set T, create the bootstrap sample T_m at random with replacement
    From T_m, build a non-pruning decision tree $\hat{Y}_m$
    Randomly choose n_try features from N features
    Select the best feature to split from each node's n_try features
    Split the tree until it achieves its largest possible size
  end for
Ensure: A set of decision trees $\{\hat{Y}_m,\ m = 1, 2, \ldots, L\}$. For the testing samples, the predictor $\hat{Y}_m(x_t)$ is produced by the decision tree $\hat{Y}_m$. The RFs' output is determined using the formula in (Eq. 2.3.18) or (Eq. 2.3.19)

Due to the utilization of multiple predictors within their framework, RFs offer enhanced predictive accuracy compared to a single decision tree. They achieve this while maintaining efficient computational performance through parallel ensemble construction, creating submodels for each sample. Furthermore, RFs exhibit strong performance, particularly when applied to tabular data.

Gradient Boosting (XGBoost)

XGBoost stands out as an effective ensemble learning model. It employs a boosting strategy to iteratively construct decision trees, as described in Chen's work [63]. Each newly generated decision tree focuses on correcting the prediction residuals of the previous one. These individual decision trees are then combined to make the final prediction, resulting in a substantial enhancement in accuracy compared to the predictive capability of a single decision tree.

Let's consider a dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}$, where i ranges from 1 to n, and each $\mathbf{x}_i$

belongs to $\mathbb{R}^m$ representing the features of n observational examples, each corresponding to the target variable y_i in $\mathbb{R}$. For a given observation indexed by i, a tree ensemble model is calculated as the sum of predictions from K additive functions

$$\hat{y_i} = \phi(\mathbf{x}_i) = \sum_{k=1}^{K} f_k(\mathbf{x}_i), \quad f_k \in \mathcal{F}, \tag{Eq. 2.3.21}$$

In the context of XGBoost, the space of regression trees is denoted as $\mathcal{F} = \{f(\mathbf{x}) = w_{q(x)}\}(q : \mathbb{R}^m \to T, w \in \mathbb{R}^T)$, where:

- q represents the structure of each tree, which maps an example to the corresponding leaf index
- T stands for the number of leaves in the tree
- w signifies the weights associated with each leaf
- f_k represents an individual regression tree responsible for predicting the value of $f_k(\mathbf{x}_i)$ for the i-th example

During the training process, the objective function is minimized. This objective function incorporates both loss terms, denoted as (l), which quantify the difference between predicted and actual values, as well as regularization terms, represented as Ω, to prevent overfitting. The minimization of this objective function is crucial in the training of XGBoost models, as described in equation (Eq. 2.3.22).

$$\begin{aligned} \mathcal{L}(\phi) &= \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \\ \text{where } \Omega(f) &= \gamma T + \frac{1}{2}\lambda ||w||^2 \end{aligned} \tag{Eq. 2.3.22}$$

The hyperparameters γ and λ serve as means to penalize the model's complexity. The loss term l represents a function that quantifies the dissimilarity between the prediction $\hat{y}_i$ and the target y_i, such as the cross-entropy loss for classification problems.

The iterative minimization of the objective function (Eq. 2.3.22) involves adding a regression tree at each iteration. Specifically, the objective function for the i-th instance at the t-th iteration can be expressed as follows

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y_i}^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t) \quad . \tag{Eq. 2.3.23}$$

By eliminating terms that are independent of f_t and employing a second-order Taylor expansion, the objective function in (Eq. 2.3.23) transforms into

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^{n} [g_i f_t(\mathbf{x_i}) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i)] + \Omega(f_t) \quad , \qquad \text{(Eq. 2.3.24)}$$

where g_i and h_i are the first and second order derivatives of $l(y_i, \hat{y_i}^{(t-1)})$ with respect to $\hat{y_i}^{(t-1)}$. Additionally, I_j is defined as the instance set corresponding to leaf j.

The optimal leaf weights w_j^* and the corresponding optimal value of $\tilde{\mathcal{L}}^{(t)}$ are then determined by the following equations

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \qquad \text{(Eq. 2.3.25)}$$

$$\tilde{\mathcal{L}}^{(t)}(q) = -\frac{1}{2} \sum_{j=1}^{T} \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \qquad \text{(Eq. 2.3.26)}$$

The evaluation criteria for finding the optimal tree split is determined by Eq. (Eq. 2.3.26). A greedy algorithm is employed, initiating the split from a single leaf and gradually adding branches in accordance with (Eq. 2.3.26), thereby avoiding the exhaustive enumeration of all possible tree structures q. To gain a deeper understanding of the splitting process, I_L and I_R are introduced as the instance sets for the left and right nodes after the split. Let $I = I_L \cup I_R$, the loss reduction resulting from the split is computed as follows

$$\mathcal{L}_{split} = \frac{1}{2} \left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \qquad \text{(Eq. 2.3.27)}$$

The best split is determined by identifying the configuration that yields the maximum value for the loss reduction in (Eq. 2.3.27). Following the selection of the best split, leaf values are assigned using (Eq. 2.3.25). To illustrate the XGBoost algorithm, please refer to the pseudocode provided in Algorithm 2.

2.4 Comparative analysis

In the realm of fault prognosis, the choice of methodology can significantly impact the effectiveness of predictive maintenance strategies in various industrial sectors. This section embarks on a comparative analysis between two prominent categories of data-driven fault prognosis methods: statistical and machine learning techniques. The analysis delves into the nuances of these approaches, exploring their

Algorithm 2 XGBoost

Require:
$\mathbf{x}$: the training set
y: the label
l: the loss function
f: the base model
Steps:
Initialize: $F_0(\mathbf{x}) = 0$
for $k = 1$ **to** K **do**
for $i = 1$ **to** N **do**
$g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y_i}^{(t-1)})$
$h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y_i}^{(t-1)})$
end for
use g_i, h_i to compute objective function $\tilde{\mathcal{L}}^{(t)}$ in (Eq. 2.3.24)
greedily grow a tree $f_k(\mathbf{x})$ in (Eq. 2.3.21)
$F_k(\mathbf{x}) = F_{k-1}(\mathbf{x}) + \epsilon f_k(\mathbf{x})$
end for
Ensure: $F(\mathbf{x}) = F_k(\mathbf{x})$

respective strengths and weaknesses across critical dimensions, including accuracy, computational complexity, ease of implementation, and suitability for different fault types and industrial processes. By shedding light on the distinct advantages and trade-offs of statistical and machine learning methodologies, this section aims to provide valuable insights to guide researchers and practitioners in selecting the most appropriate approach for their specific fault prognosis needs.

Accuracy. Statistical methods often require strong assumptions about data distribution and relationships. Their accuracy heavily depends on how well these assumptions hold. If the underlying assumptions are met, statistical methods can provide accurate results. On the other hand, machine learning methods are more flexible and can handle complex relationships in data without strict assumptions. They often yield competitive or superior accuracy when dealing with non-linear or high-dimensional data.

Computational Complexity. Many statistical methods have relatively low computational complexity and can be implemented with minimal computational resources. They are suitable for real-time applications with limited computational capacity. Conversely, machine learning methods, especially complex algorithms like Gradient Boosting, can be computationally expensive, requiring significant processing power and memory. They may not be ideal for resource-constrained environments.

Ease of Implementation. Statistical methods often have well-established mathematical formulas and techniques, making them easier to implement, especially when analytical solutions exist. However, customizing them for specific applications may require advanced statistical expertise. Machine learning methods are generally easier to implement using libraries and frameworks like scikit-learn or TensorFlow. They are accessible to a broader audience and do not always require deep statistical knowledge. However, hyperparameter tuning and feature engineering may demand expertise.

Suitability for Fault Types. Statistical methods are well-suited for modeling faults that adhere to probabilistic or distributional assumptions. For instance, they can work effectively for predicting faults with well-defined statistical patterns, such as sensor drift. Machine learning methods excel in capturing complex patterns and anomalies in data. They are more adaptable to various fault types, including non-linear, time-varying, or complex interactions among variables. They are often preferred for fault prognosis in diverse industrial processes.

Suitability for Industrial Processes. Statistical methods can be suitable for processes with known statistical properties and where underlying assumptions are met. They may be preferred in situations where process dynamics are relatively stable. On the contrary, machine learning methods are versatile and can be applied to a wide range of industrial processes, including those with complex, non-linear dynamics and unknown fault patterns. They are particularly valuable when dealing with processes that evolve over time or where data is high-dimensional.

The choice between statistical and machine learning methods for data-driven fault prognosis should be based on the specific requirements of the application. Statistical methods are suitable when assumptions hold, computational resources are limited, and the fault patterns are well-understood. On the other hand, machine learning methods offer greater flexibility and accuracy, making them suitable for a broader range of applications, especially those with complex fault patterns and diverse industrial processes.

2.5 Concluding remarks

This chapter offers a brief introduction to the major technology and basic concepts of fault prognosis techniques, which include statistical and machine learning approaches. The analysis revealed a fundamental distinction between statistical and machine learning approaches, each bearing unique advantages and trade-offs. While statistical methods are grounded in well-defined assumptions and are suitable for cases where these assumptions hold true, machine learning methods offer unparalleled flexibility in handling complex, non-linear relationships and adapting

to diverse industrial processes. The choice between these approaches hinges upon the specific demands and constraints of the application at hand.

This investigation underscores the crucial role of data-driven methodologies in the progression of fault prognosis within industrial technical processes. In an era marked by data abundance and technological innovation, harnessing the power of data-driven models offers the promise of not only enhancing system reliability and minimizing downtime but also ushering in a new era of efficiency and sustainability in industrial operations.

Chapter 3

Research questions and objectives

In this chapter, the researcher outlines the key research questions and the specific objectives that will guide the study. The research question and objectives are carefully crafted to ensure that they are relevant, feasible, and aligned with the existing literature in the field. This chapter serves as a roadmap for the research project, providing a clear direction for the study and helping to ensure that the research remains focused and on track.

3.1 Research questions

The main objective of this Ph.D. thesis is to design efficient fault prognosis techniques for technical processes using data-driven methods that consider the operating conditions of the process. The research focuses on fault prognosis across various levels of technical processes, encompassing critical components, subsystems, and systems. The specific challenges of these levels are identified and research questions are raised.

3.1.1 Fault prognosis of critical components

In accordance with ISO 13381-1 [64], predicting future fault progressions requires a deep understanding of the physical underlying failure modes, as well as the relationships with future operating conditions. The fundamental goal of data acquisition and preprocessing is to create a model based on the extracted health indicators that accurately describes the degree of degradation in the monitored component. This model utilizes data from various condition monitoring sources to analyze the degradation mechanisms that will affect the system's performance while taking

into account material properties, boundary conditions, and operating and environmental factors [65]. Consequently, possessing good knowledge of degrading performance and mechanisms is crucial for establishing a reliable model. Given the fundamental goal of data acquisition and preprocessing to create an accurate model of the monitored component's degradation mechanisms, it is crucial to overcome the challenge of data interpretation into useful information through feature extraction.

Feature extraction involves identifying the most relevant information from the sensor data and transforming it into meaningful features that can be used to accurately predict the RUL of the monitored component. The process of feature extraction can be complex and time-consuming, as it requires expertise in data analysis and a deep understanding of the underlying physical processes. Extracting relevant features involves a careful selection of statistical, spectral, or other signal processing methods to obtain the most informative features that can capture the degradation mechanisms of the component. A feature set that is not well-tailored to the monitored component may result in a model that does not accurately predict the RUL. Furthermore, the selection of the appropriate feature extraction method may vary depending on the monitored component, making it essential to possess a broad range of expertise in data analysis.

Identifying relevant health indicators is another significant challenge. Determining which indicators are most relevant for predicting RUL can be difficult due to the complexity of industrial processes and the multitude of potential factors influencing component degradation. Selecting the wrong indicators can lead to inaccurate RUL predictions and ineffective maintenance planning.

Quantifying the relationship between health indicators and RUL is also a challenging task. Establishing a clear, quantifiable relationship between health indicators and RUL can be difficult, particularly when dealing with noisy or incomplete data or non-linear relationships between indicators and RUL. Failure to establish this relationship accurately can result in incorrect RUL predictions and ineffective maintenance planning.

The selection of an appropriate failure threshold is another crucial challenge. A failure threshold is the point at which a component is considered to have failed, and it can have a significant impact on RUL prediction accuracy. Selecting an incorrect failure threshold can result in incorrect RUL predictions and ineffective maintenance planning.

Finally, model selection and comparison present significant challenges in this research. Evaluating and comparing the accuracy of various data-driven methods for

RUL prediction requires a comprehensive assessment using standardized metrics, benchmarks, and validation datasets, which can be challenging to establish in a consistent manner. The selected model should accurately predict RUL to support decision-makers in maintenance planning and prevent faults and operation failure. Overcoming these challenges will be critical in achieving accurate and reliable RUL predictions for effective maintenance planning and reduced downtime.

The relevant research questions related to fault prognosis for critical components can be concluded as:

1. How to extract features from the raw data to create more interpretable data?
2. What are the most effective methods for selecting appropriate health indicators for predicting RUL of critical components?
3. How to accurately quantify the relationship between health indicators and RUL to enhance the precision of RUL predictions and improve maintenance planning?
4. How to select the appropriate failure threshold for predicting RUL?
5. Which data-driven methods are the most accurate for predicting RUL of critical components?
6. How can these methods be incorporated into a framework that is applicable in industry and overcomes the limitations of standard techniques?

3.1.2 Fault prognosis of subsystems level

The inclusion of subsystems level fault prognosis in the thesis is essential as it provides an intermediate level of complexity, capturing interactions among interconnected components and enabling more targeted interventions. This approach allows for a nuanced analysis, striking a balance between the detailed examination of individual components and the high-level overview of the entire system. By focusing on subsystems, the research gains a better understanding of how faults within specific functional areas may influence overall system performance, leading to more accurate predictions and tailored maintenance strategies for enhanced system reliability.

The successful development and implementation of prognostic models in technical processes' subsystems are a significant research challenge. A key challenge is understanding the complex interdependencies among subsystems that are inherent in many technical processes. These relationships and dependencies need to be

accurately modeled to support effective prognostic modeling. The challenge lies in the development of models that can capture these complex interdependencies while remaining computationally efficient.

Computational efficiency is a crucial challenge in developing prognostic models. While complex models can capture the intricacies of technical processes, they may require significant computational resources. This can pose a challenge in striking a balance between model accuracy and computational efficiency. The development of efficient models that can provide timely and actionable insights is a critical research challenge.

Another challenge in developing prognostic models is the integration of these models with maintenance and decision-making systems. To be useful, the prognostic results need to be seamlessly integrated with maintenance planning and decision-making systems. This requires addressing challenges in data compatibility, communication protocols, and decision-making algorithms. Successful integration of prognostic models with these systems can support effective decision-making, improve maintenance planning, and reduce downtime. However, achieving this integration is an important research challenge that requires a multidisciplinary approach.

Addressing the complex interdependencies, computational efficiency, and integration with maintenance and decision-making systems requires a multidisciplinary approach that draws upon expertise in data science, engineering, and decision-making. Overcoming these challenges will be critical in realizing the potential of prognostic modeling to improve maintenance planning and enhance safety.

The research questions pertaining to fault prognosis of subsystem level may be summarized as:

1. How can the complex interdependencies among subsystems be accurately modeled and incorporated into prognostic models to ensure robust and reliable predictions?

2. What approaches can be used to optimize the balance between computational efficiency and prediction accuracy in prognostic models for subsystems of technical processes?

3. How can prognostic models be effectively integrated with maintenance and decision-making systems to enhance overall operational efficiency and system reliability?

3.1.3 Fault prognosis of system level

The fault prognosis of system level in industrial operations presents several challenges that need to be addressed. Two significant challenges include selecting suitable prediction techniques and early detection of failure precursors.

The vast array of available techniques for predicting faults at system level presents a challenge in selecting the most appropriate one for a particular industrial operation. Given the complexity of industrial systems, it is difficult to determine which technique will be most effective. Moreover, techniques that work well for one system may not be suitable for another. This requires an evaluation of the available techniques to identify the most appropriate one.

Another significant challenge in fault prognosis is the early detection of failure precursors. The early identification of potential faults is crucial for effective maintenance planning and avoiding unexpected downtime. However, identifying these precursors can be challenging, as they may be subtle or masked by normal system variations. This requires advanced monitoring techniques that can distinguish between normal system variations and early signs of potential failure. Moreover, expertise in data analysis can be beneficial to identify the relevant signals and interpret them accurately.

Addressing these challenges is critical for the effective prognosis of system level in industrial operations. Failure to do so can result in ineffective maintenance planning, unexpected downtime, and potential safety hazards.

The research questions related to fault prognosis at system level can be summarized as:

1. What are the most effective data-driven methods for predicting future operating conditions in industrial settings?
2. What are the most effective techniques and methods for identifying and interpreting subtle precursors of failure in industrial systems?

3.2 Research objectives

The primary objective of this Ph.D. thesis is to develop effective fault prognosis approaches for technical processes utilizing data-driven methodologies that take into account the process's operating conditions. To accomplish this goal, the research questions described earlier have been translated into several sub-objectives, including:

1. Design a reliable fault prognosis system for key components of the technical

process to accurately predict the remaining useful life (RUL) using data-driven methods. The proposed method should be applicable in industry and overcome the limitations of standard techniques.

2. Develop an efficient fault prognosis system for technical process subsystems that prevents operational failures. It is also desirable to investigate the potential faults before prognosis to ensure appropriate fault prediction.

3. Develop an accurate predictive model for assessing future operating conditions within industrial systems to prevent catastrophic accidents. The proposed model should enable the selection of appropriate mitigation and risk reduction measures to enhance safety.

In this thesis, one industrial benchmark process and two experiments are utilized to demonstrate the effectiveness of the proposed approaches for fault prognosis at the component, subsystem, and system levels.

Chapter 4

Research methodology and approach

The purpose behind pursuing a Ph.D. degree has been to gain fresh insights and cultivate expertise in research by constructing research questions, devising research strategies, composing scientific papers, and delivering research outcomes. This subchapter delineates the general research principles and research methodologies.

4.1 Classification of research

It is crucial to begin by understanding the concept of research. Creswell [66] defines research as a process of steps used to collect and analyze information to increase our understanding of a topic or issue. In other words, research is a means of discovering the unknown and expanding our understanding of a particular subject. Types of research can be classified into three, namely basic research, applied research, and experimental development [67].

Basic research aims to expand our knowledge and understanding of fundamental principles and concepts in science and technology. Basic research is often curiosity-driven and seeks to answer questions about the nature of the universe, the laws of physics, and the behavior of living organisms. It can lead to breakthrough discoveries and innovations, but its practical applications are often unclear. Applied research focuses on solving specific problems and developing practical solutions to real-world issues. The goal of applied research is to develop new products, processes, or services that can be used to improve people's lives and enhance economic growth. Applied research is often interdisciplinary, involving collaboration

between researchers in different fields. Experimental development involves the creation and testing of prototypes, models, and prototypes of new products, processes, or services. Experimental development aims to turn scientific discoveries and technological innovations into practical applications that can be commercialized and brought to market. This type of research often involves significant investments of time, money, and resources, and requires close collaboration between researchers, engineers, and business professionals.

This Ph.D. thesis is explicitly classified as **applied research** based on the classification described in the Frascati Manual. The primary objective of this study is to develop effective data-driven models for fault prognosis of technical processes. The research is intentionally designed to provide practical solutions to real-world problems and is explicitly intended to be applied in industrial settings to demonstrate its relevance and applicability. By emphasizing the development and testing of these models, this research aims to contribute substantially to the improvement of industrial processes, enhancing their efficiency and reliability in a way that aligns with the fundamental characteristics of applied research.

4.2 Model evaluation and verification

In scientific research, the assessment and verification of models cannot rely solely on empirical or experimental methods. The Model Evaluation Group, established by the EU in 1992, proposed a model evaluation process for cases where other approaches are necessary [68]. The primary goal of the group was to promote a culture of model development that included voluntary model evaluation procedures based on a formalized consensus protocol. To achieve this objective, the group recommended a model evaluation process that includes:

- scientific assessment,
- verification, and
- validation

By following this model evaluation process, researchers can ensure that their models are reliable and can be used with confidence in various practical applications.

The scientific assessment stage involves a critical examination of the theoretical basis of the model and its underlying assumptions. It aims to evaluate the scientific validity of the model and its suitability for the intended application. Verification is the process of assessing whether the model is correctly implemented and performs as expected. This stage involves comparing the results generated by the model

against independent sources of information, such as analytical solutions, benchmark data, or other models. Validation is the process of assessing the model's predictive performance and its ability to represent the behavior of the system under study. This stage involves comparing the model's predictions against actual observations or experimental data obtained from the real system.

This thesis adopts a **scientific assessment** approach to scrutinize the underlying assumptions of the developed models, informed by an extensive literature review of cutting-edge data-driven approaches, specifically focusing on random forest and XGBoost, for fault prognosis in technical processes. Rigorous evaluations by experts from academic and industrial backgrounds contribute to ensuring the validity and reliability of the developed models. Notably, the methods introduced in this study represent a notable improvement over existing approaches recommended by industry standards in the field.

To **verify** and **validate** the methods proposed in this thesis, expert judgments from industry partners have been leveraged. The thesis presents relevant case studies of technical processes, namely the roller bearings component, automated manufacturing line subsystem, and liquid hydrogen storage system, to demonstrate how these methods can be deployed for decision-making support. The numerical results obtained from these methods accurately reflect the performance of the developed methods for fault prognosis of technical processes. Through this process, the accuracy and reliability of these methods have been demonstrated, providing a sound basis for practical applications in real-world industrial settings.

4.3 Scientific quality

As per the guidelines set forth by the Research Council of Norway [69], quality research in science should embody three key attributes: Originality, Solidness, and Relevance. Originality is a measure of the novelty and innovation of the research. Solidness refers to the degree to which the research's statements and conclusions are supported by strong evidence and rigorous analysis. Relevance is judged based on the research's potential for contributing to professional and societal development by providing practical and useful insights.

This thesis endeavors to strike a balance between the three key attributes of quality research - **Originality**, **Solidness**, and **Relevance**. The research introduces novel frameworks, such as the amalgamation of EMD and RF, and enhances existing methods, exemplified by the hyperparameter optimization of random forest using Bayesian optimization, thereby pioneering innovative approaches in the realm of fault prognosis in industrial settings. To underscore the practicality and relevance of these methods, the thesis presents case studies focused on diverse technical

processes (bearings, automated fuse test bench, liquid hydrogen storage system), showcasing their potential for real-world applications. Additionally, the scientific quality of the work is ensured by disseminating it in scientific peer-reviewed journals and international conferences, and the feedback received from expert reviewers has been instrumental in enhancing the quality and rigor of the research presented in this thesis. By focusing on these three attributes, this thesis seeks to make a valuable contribution to the field of data-driven fault prognosis and advance our understanding of these critical processes in industrial settings.

Chapter 5

Main results

This chapter presents and describes the main findings of the Ph.D. project, which are documented in the form of five articles. These findings are evaluated to determine the extent to which the research objectives have been achieved.

5.1 Overview

The research articles aim to address the research questions and to achieve the research objectives that have been identified in Chapter 3. Three articles were published in relevant international journals. In addition, one article has been presented in a peer-reviewed international conference and one article is currently planned for submission.

This chapter provides a summary of the key findings and contributions of the Ph.D. thesis, specifically in relation to the proposed objectives. The objective of each article is outlined in Table 5.1.

Table 5.1: Summary of contributions, objectives, and articles of this PhD thesis

Research objective	Main topic	Article
Objective 1	RUL prediction of critical components	Article I [26] Article IV [29]
Objective 2	Fault prognosis of a subsystem of manufacturing line	Article II [27] Article IV [29] Article V
Objective 3	Future operating condition prediction of a maritime operation	Article III [28] Article IV [29]

Table 5.1 also highlights each article's correlation to the overall research topics of the thesis. The subsequent sections provide detailed insights into each objective's contributions. The complete versions of the articles are included in the Articles part.

5.2 Contributions

5.2.1 Contributions to fault prognosis of critical components

The first objective of this thesis is to design a reliable fault prognosis system for key components of the technical process to accurately predict the remaining useful life (RUL) using data-driven methods. The two articles are related to this objective:

Article I: Optimized random forest model for remaining useful life prediction of experimental bearings.
Article IV: Sustainability of ICPS from a safety perspective: Challenges and opportunities.

1. Article I proposed a novel data-driven prediction framework for bearing RUL, which is a key component in many industrial technical processes, such as aero engines, high-speed trains, and wind turbines. The framework, illustrated in Figure 5.1, involves two phases: feature extraction using Empirical Mode Decomposition and RUL prediction using an RFs-based model with hyperparameters tuned by Bayesian optimization. The approach showed significant improvement compared to standard data-driven and stochastic approaches in an actual run-to-failure experiment of roller bearings.

2. Article IV provides an overview of the present state-of-the-art in the field of remaining useful life prediction in industrial processes, thereby establishing the literature foundation for the objective of this thesis. This article also identifies and delineates the research challenges that currently prevail in the field, and thereby serves as a key source of inspiration and motivation for the development of the approach outlined in Article I of this thesis.

The main findings of these articles are summarized as follows:

1. The proposed method in Article I outperforms other data-driven and model-based approaches in terms of accuracy and error in predicting the bearing RUL, as outlined in Table 5.2 and Table 5.3 respectively. Fig. 5.2 shows the RUL prediction for every bearing with the best score.

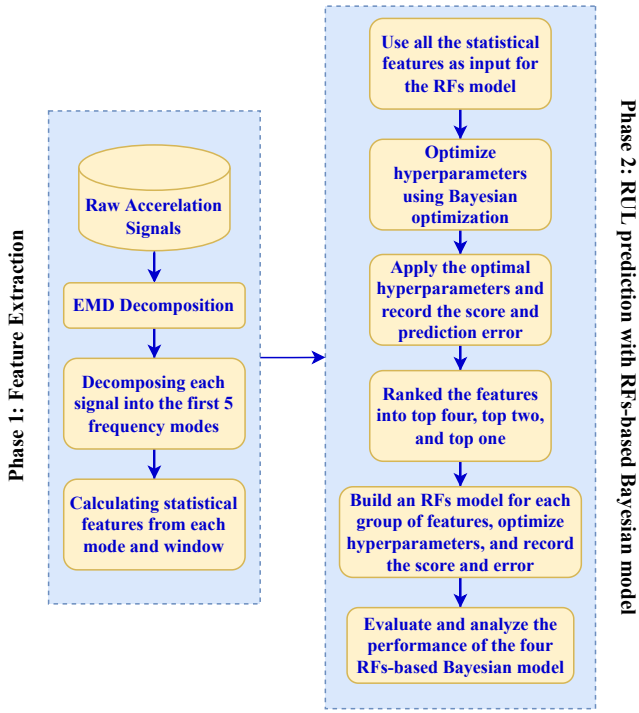

Figure 5.1: The proposed framework for bearing RUL prediction.

Table 5.2: The score of RUL predictions for every bearing.

Score	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	Average
All features	0.7621	0.7950	0.9277	0.9533	0.8760	0.9648	0.8809	0.9516	0.9094	0.9461	0.8967
4 features	0.7636	0.7853	0.9307	0.9580	0.8643	0.9634	0.8972	0.9537	0.9164	0.9487	0.8981
2 features	0.7761	0.8597	0.9362	0.9612	0.9329	0.9641	0.8447	0.9603	0.9135	0.9429	0.9092
1 feature	0.9536	0.6414	0.8151	0.8413	0.8096	0.9645	0.8532	0.8266	0.7917	0.8306	0.8328

Table 5.3: The RMSRE of RUL predictions for every bearing.

RMSRE	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	Average
All features	2.6739	2.2487	1.3654	0.7220	1.1861	0.8493	1.4983	0.8499	1.4479	0.9599	1.3801
4 features	2.6187	2.2637	1.0971	0.6671	1.4759	0.9307	1.3785	0.8547	1.4315	0.8149	1.3533
2 features	2.3878	1.3728	1.2691	0.6417	0.7147	0.8891	1.8001	0.7249	1.2908	1.1332	1.2224
1 feature	0.5278	4.2387	4.3198	3.4316	1.7146	0.8916	1.8849	5.5522	7.3177	2.9226	3.2802

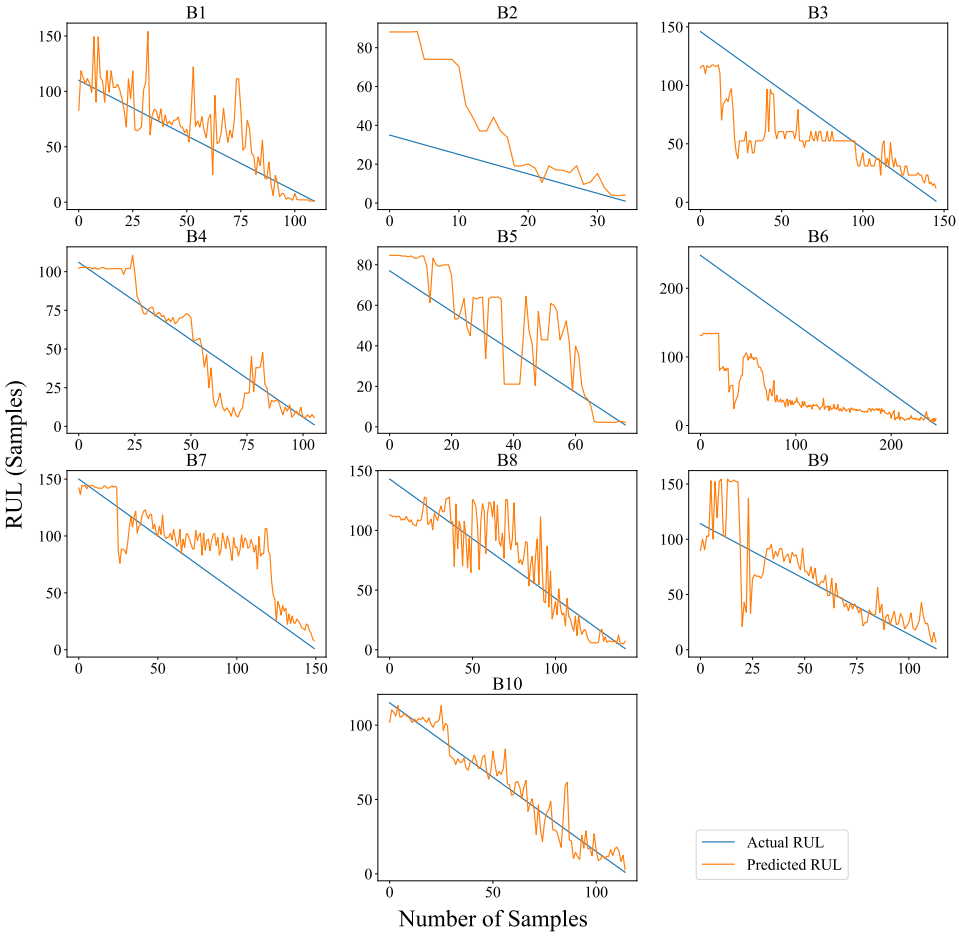

Figure 5.2: The best RUL prediction of every bearing.

The proposed model demonstrates close alignment between predicted and actual RUL values, particularly towards the end of a bearing's lifespan, where accurate predictions are crucial. The model excels with bearings having an average number of samples (around 100-110), such as those in instances like bearing 1, 4, 9, and 10. However, accuracy diminishes for bearings with a limited number of samples (short lifespan), notably bearing 2 and 5, which, according to the manufacturer, have an average lifespan of eight to twelve hours. This discrepancy suggests potential technical issues, possibly stemming from incorrectly mounted accelerometers on these bearings, leading to erroneous signals. Bearing 6, with an unusually high number of samples (248), exhibits subpar predictions, likely due to the model's

lack of exposure to such high-sample instances during training. Overall, the model's prediction error is more pronounced at the beginning of a bearing's lifespan, reflecting inherent uncertainty in the input data. As the bearing progresses towards failure, the model's predictions become more accurate, benefiting from increased exposure to degradation data, reducing uncertainty in capturing the degradation trend.

2. The proposed methodology exhibits improved RUL prediction performance in contrast to the Wiener process. Moreover, the proposed framework mitigates the need for failure threshold determination, which is a source of uncertainty in the model, thus addressing the challenge of managing uncertainty highlighted in Article IV.

5.2.2 Contributions to fault prognosis of subsystems level

The second objective of this thesis is to develop an efficient fault prognosis system for technical process subsystems that prevents operational failures. The three articles are related to this objective:

Article II: An extreme gradient boosting aided fault diagnosis approach: A case study of fuse test bench.
Article IV: Sustainability of ICPS from a safety perspective: Challenges and opportunities.
Article V: Advancements in extreme gradient boosting for enhanced fault prognosis: A continuation study from fuse test bench analysis.

1. Article II proposed an integrated fault diagnosis system based on extreme gradient boosting to detect, classify, and identify the root causes of faults in an automated fuse test bench, which is a subsystem of a manufacturing line. The flowchart of the proposed method is shown in Figure 5.3. The proposed diagnosis system is validated using the data set from PHM 2021 Data Challenge. By diagnosing faults correctly, this article serves as a foundation for fault prognosis of subsystem level for Article V.

2. In Article IV, an overview of the current state-of-the-art in the area of industrial process fault diagnosis and prognosis is presented, laying the groundwork for the objective of this thesis. Moreover, the article identifies and outlines the existing research challenges in this domain, which serves as a significant source of inspiration and impetus for the development of the approach presented in Article II.

3. Article V expanded the capability of the extreme gradient boosting into fault

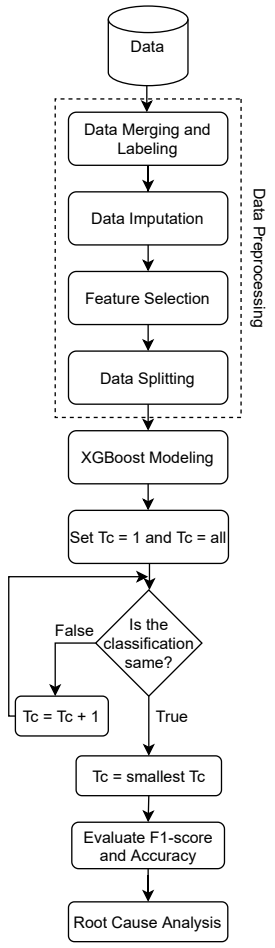

Figure 5.3: The flowchart of the proposed method in Article II.

prognosis of an automated fuse test bench. The paper introduces methodological refinements to the previously established XGBoost model. These refinements encompass enhancements to the model architecture, feature engineering techniques, and meticulous hyperparameter tuning, all aimed at bolstering the model's accuracy and efficacy in predicting faults and failures.

Building upon the foundation laid in the earlier paper, this paper discusses how the model has evolved to incorporate additional data streams, such as image data. It details the seamless integration of these new data sources and illuminates the profound impact of this integration on the model's perform-

ance, ultimately improving the accuracy and timeliness of fault prognosis.

This practical application showcases the approach's efficacy in a dynamic and operational environment, providing insights into data collection, system integration, and the tangible outcomes achieved, including streamlined maintenance processes, reduced downtime, and heightened overall system reliability. The paper underscores the ongoing evolution of data-driven fault prognosis, showcasing the continual advancements in the use of XGBoost as a potent tool in predictive maintenance and fault prognosis for industrial processes.

The main findings of these articles are summarized as follows:

1. The proposed fault diagnosis system in Article II has high classification accuracy, fast diagnosis time, and interpretable root cause analysis when verified by the real industrial data from PHM Challenge 2021. The proposed system also outperforms several standard fault diagnostic approaches in detection and classification accuracy and requires a shorter diagnosis time as outlined in Table 5.4 and Table 5.5.

Table 5.4: Comparison with standard approaches for detection and classification tasks using the fuse test bench data set.

Method	**Detection**		**Classification**	
	F1-score	**Accuracy**	**F1-score**	**Accuracy**
XGBoost	**0.9745**	**0.9764**	**0.9693**	**0.9792**
RF	0.9234	0.9302	0.7955	0.8958
MLP	0.7118	0.7274	0.7717	0.8684
Logistic Regression	0.6192	0.6942	0.6360	0.8196
LDA	0.6912	0.7396	0.6081	0.7911

Table 5.5: Time to classification for all methods.

Method	**Average T_c in Class**									**Average T_c**
	0	**2**	**3**	**4**	**5**	**7**	**9**	**11**	**12**	**Overall**
XGBoost	**1.17**	**1**	82.5	**1**	**1**	**1**	4	5	**3**	**4.77**
RF	1.17	96	57	1	1	1	**3**	**3**	3	7.58
MLP	20.76	360	360	360	6.5	11.5	31.5	360	360	76.77
Logistic Regression	9.53	1	**3.5**	7.5	7.5	5.5	5.5	8	4	8.27
LDA	5	62	360	3.5	6	7	5	360	13	29.79

2. Article II of this thesis proposes a fault diagnosis system that adeptly addresses the challenges of integrating multiple sources of data and techniques,

as well as the need for real-time implementation emphasized in Article IV, through the utilization of multiple data sources and computationally efficient algorithms that enable real-time implementation.

3. Article V highlights the continued evolution and efficacy of the XGBoost-based fault prognosis approach in industrial technical processes. Methodological refinements, including model architecture enhancements, feature engineering, and hyperparameter tuning, resulted in improved fault prediction accuracy. The integration of new data sources further bolstered the model's performance, enhancing its ability to adapt to changing operational conditions. The real-world application of the refined approach demonstrated its practical utility, showcasing streamlined maintenance processes, reduced downtime, and heightened overall system reliability. These findings underscore the ongoing advancement of data-driven fault prognosis, reaffirming the value of XGBoost as a potent tool for predictive maintenance and reliability enhancement in diverse industrial contexts.

5.2.3 Contributions to fault prognosis of system level

The third objective of this thesis is to develop an accurate predictive model for assessing future operating conditions within industrial systems to prevent catastrophic accidents. The two articles are related to this objective:

Article III: Towards accident prevention on liquid hydrogen: A data-driven approach for releases prediction.
Article IV: Sustainability of ICPS from a safety perspective: Challenges and opportunities.

1. Article III proposed a novel application of the random forests algorithm for early detection of hazardous events in liquid hydrogen (LH_2) spills to prevent catastrophic accidents, such as detonation. In this case, early detection was carried out by predicting the oxygen phase change and estimating whether the hydrogen concentration was above the lower flammability limit (LFL). The framework of the proposed application is illustrated in Figure 5.4. The proposed application is verified by real experiments of LH_2 release tests to replicate spills of LH_2 inside the ship's tank connection space and during bunkering operations.

2. Article IV provides an overview of the current state-of-the-art in the area of industrial process fault prognosis, thereby setting the foundation for the goal of this thesis. Additionally, the article appraises and elucidates the current

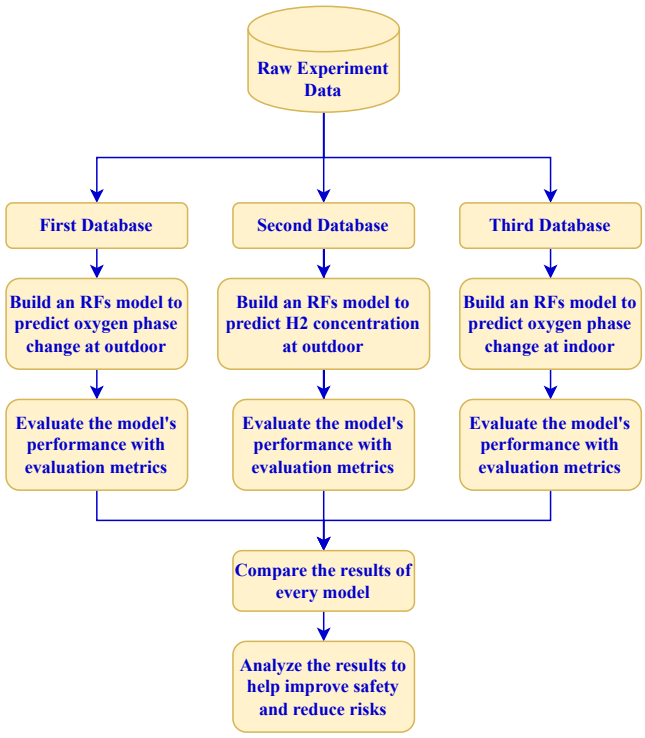

Figure 5.4: The framework of the proposed application in Article III.

research challenges within this realm, which serve as a vital source of inspiration and momentum for the development of the application proposed in Article III.

The main findings of these articles are summarized as follows:

1. The model was trained with experimental data and specific databases were created to analyze the liquefaction and solidification of oxygen during the LH_2 release and the hydrogen concentration in the air. The model accurately predicted the above-mentioned parameters as summarized in Table 5.6.

2. The random forest approach was more robust and precise than other machine learning techniques previously used to simulate similar experiments as shown in Table 5.7.

3. The model developed in this study offers valuable insights that can be utilized to conduct comprehensive risk analysis and identify suitable prevention

and mitigation measures, with the aim of mitigating the overall risk of LH_2 technologies, particularly in emerging applications.

Table 5.6: Performance metrics of RFs model for all databases.

Label	Accuracy	Precision	Recall	F1	AUC-PR
First database (liquid oxygen)	0.9984	0.9981	0.9983	0.9984	0.9985
First database (solid oxygen)	0.9995	0.9985	0.9955	0.9984	0.9972
Second database (H_2 concentration >LFL)	0.9993	0.9873	0.9583	0.9861	0.9731
Third database (liquid oxygen)	0.9993	0.9992	0.9994	0.9993	0.9994
Third database (solid oxygen)	0.9997	0.9993	0.9994	0.9996	0.9994

Table 5.7: Performance comparison of RFs model and linear model (LM) in [2] for all labels.

Label	Accuracy		Precision		Recall		AUC-PR	
	LM	RFs	LM	RFs	LM	RFs	LM	RFs
Liquid oxygen	0.9020	**0.9984**	0.8480	**0.9981**	0.9360	**0.9983**	0.9490	**0.9985**
Solid oxygen	0.9570	**0.9995**	0.830	**0.9985**	0.6130	**0.9955**	0.8070	**0.9972**
H_2 concentration >LFL	0.9880	**0.9993**	0.6490	**0.9873**	0.1840	**0.9583**	0.3660	**0.9731**

Chapter 6

Concluding remarks and outlook

The primary objective of this thesis was to design efficient fault prognosis techniques for technical processes using data-driven methods that consider the operating conditions of the process. This objective is decomposed into three sub-objectives that are addressed through five articles.

The first sub-objective was to design a reliable fault prognosis system for key components of the technical process to accurately predict the RUL using data-driven methods. Roller bearings were chosen as the key components. A novel data-driven prediction framework was proposed to predict the RUL of bearings. The framework involves two phases: feature extraction using Empirical Mode Decomposition and RUL prediction using an RFs-based model with hyperparameters tuned by Bayesian optimization. The approach showed significant improvement compared to standard data-driven and stochastic approaches in an actual run-to-failure experiment of roller bearings.

The second sub-objective was to develop an efficient fault prognosis system for technical process subsystems that prevents operational failures. An automated fuse test bench, which is a subsystem of a manufacturing line, was selected as the subsystem of interest in this research. First, an integrated fault diagnosis system based on extreme gradient boosting was proposed. The proposed fault diagnosis system outperforms several standard fault diagnostic approaches in detection and classification accuracy and requires a shorter diagnosis time. Then, the capability of the extreme gradient boosting was expanded into fault prognosis. Methodological refinements and the capability to incorporate additional data streams, such as image data, were introduced.

The third sub-objective was to develop an accurate predictive model for assessing

future operating conditions within industrial systems to prevent catastrophic accidents. A liquid hydrogen storage system was selected as the system of interest in this research. A novel application of the random forests algorithm was proposed for the early detection of hazardous events in liquid hydrogen spills to prevent catastrophic accidents, such as detonation. The model can accurately predict the events and is better than other machine learning techniques previously used to simulate similar experiments. The model developed in this study provides insights for conducting thorough risk analysis and identifying prevention and mitigation measures, particularly in emerging liquid hydrogen technology applications.

The future work inspired by this thesis should aim to advance the field of data-driven fault prognosis by expanding its scope, improving its accuracy, and facilitating its practical implementation across diverse industrial domains. Another topic that requires worthy attention is the design of real-time fault prognosis systems that can continuously monitor and assess the health of technical processes. Such efforts hold the potential to enhance system reliability, reduce downtime, and ultimately contribute to safer and more efficient industrial processes.

Bibliography

[1] K. Zhong, M. Han, and B. Han, "Data-driven based fault prognosis for industrial systems: A concise overview," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 2, pp. 330–345, 2019.

[2] F. Ustolin, F. Ferrari, and N. Paltrinieri, "Prediction of condensed phase formation during an accidental release of liquid hydrogen," *Chemical Engineering Transactions*, vol. 91, pp. 439–444, 2022.

[3] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee, and S. K. S. Gupta, "Ensuring safety, security, and sustainability of mission-critical cyber–physical systems," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 283–299, 2011.

[4] I. A. Ibrahim, M. Hossain, and B. C. Duck, "An optimized offline random forests-based model for ultra-short-term prediction of pv characteristics," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 202–214, 2019.

[5] M. Rausand, *Risk assessment: theory, methods, and applications*, vol. 115. John Wiley & Sons, 2013.

[6] J. Lee, I. Cameron, and M. Hassall, "Improving process safety: What roles for digitalization and industry 4.0?," *Process safety and environmental protection*, vol. 132, pp. 325–339, 2019.

[7] X. Lyu, Y. Ding, and S.-H. Yang, "Safety and security risk assessment in cyber-physical systems," *IET Cyber-Physical Systems: Theory & Applications*, vol. 4, no. 3, pp. 221–232, 2019.

[8] A. M. Omer, "Focus on low carbon technologies: The positive solution," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 9, pp. 2331–2357, 2008.

[9] R. Isermann and P. Balle, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control engineering practice*, vol. 5, no. 5, pp. 709–719, 1997.

[10] G. Sonnemann, M. Tsang, F. Castells, and M. Schuhmacher, *Integrated life-cycle and risk assessment for industrial processes*. CRC press, 2003.

[11] R. Tong, X. Yang, T. Parker, B. Zhang, and Q. Wang, "Exploration of relationships between safety performance and unsafe behavior in the chinese oil industry," *Journal of loss prevention in the process industries*, vol. 66, p. 104167, 2020.

[12] R. Arunthavanathan, F. Khan, S. Ahmed, and S. Imtiaz, "A deep learning model for process fault prognosis," *Process Safety and Environmental Protection*, vol. 154, pp. 467–479, 2021.

[13] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 9, pp. 2046–2054, 2016.

[14] C. Ly, K. Tom, C. S. Byington, R. Patrick, and G. J. Vachtsevanos, "Fault diagnosis and failure prognosis for engineering systems: A global perspective," in *2009 ieee international conference on automation science and engineering*, pp. 108–115, IEEE, 2009.

[15] J. Z. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical systems and signal processing*, vol. 25, no. 5, pp. 1803–1836, 2011.

[16] T. Ekanayake, D. Dewasurendra, S. Abeyratne, L. Ma, and P. Yarlagadda, "Model-based fault diagnosis and prognosis of dynamic systems: A review," *Procedia Manufacturing*, vol. 30, pp. 435–442, 2019.

[17] J. Luo, M. Namburu, K. Pattipati, L. Qiao, M. Kawamoto, and S. Chigusa, "Model-based prognostic techniques [maintenance applications]," in *Proceedings AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference.*, pp. 330–340, Ieee, 2003.

[18] B. Zhang, K. Zheng, Q. Huang, S. Feng, S. Zhou, and Y. Zhang, "Aircraft engine prognostics based on informative sensor selection and adaptive degradation modeling with functional principal component analysis," *Sensors*, vol. 20, no. 3, p. 920, 2020.

[19] J. Xu, Y. Wang, and L. Xu, "Phm-oriented integrated fusion prognostics for aircraft engines based on sensor data," *IEEE Sensors Journal*, vol. 14, no. 4, pp. 1124–1132, 2013.

[20] X. Jin, Z. Que, Y. Sun, Y. Guo, and W. Qiao, "A data-driven approach for bearing fault prognostics," *IEEE Transactions on Industry Applications*, vol. 55, no. 4, pp. 3394–3401, 2019.

[21] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2226–2238, 2013.

[22] L. Zhang, J. Lin, and R. Karim, "An angle-based subspace anomaly detection approach to high-dimensional data: With an application to industrial fault detection," *Reliability Engineering & System Safety*, vol. 142, pp. 482–497, 2015.

[23] J. Yang, Y. Guo, and W. Zhao, "Long short-term memory neural network based fault detection and isolation for electro-mechanical actuators," *Neurocomputing*, vol. 360, pp. 85–96, 2019.

[24] H. Chen, Z. Chai, O. Dogru, B. Jiang, and B. Huang, "Data-driven designs of fault detection systems via neural network-aided learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[25] X. Li, Y. Jiang, C. Liu, S. Liu, H. Luo, and S. Yin, "Playing against deep neural network-based object detectors: A novel bidirectional adversarial attack approach," *IEEE Transactions on Artificial Intelligence*, 2021.

[26] M. G. Alfarizi, B. Tajiani, J. Vatn, and S. Yin, "Optimized random forest model for remaining useful life prediction of experimental bearings," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 6, pp. 7771–7779, 2023.

[27] M. G. Alfarizi, J. Vatn, and S. Yin, "An extreme gradient boosting aided fault diagnosis approach: A case study of fuse test bench," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 661–668, 2023.

[28] M. G. Alfarizi, F. Ustolin, J. Vatn, S. Yin, and N. Paltrinieri, "Towards accident prevention on liquid hydrogen: A data-driven approach for releases

prediction," *Reliability Engineering & System Safety*, vol. 236, p. 109276, 2023.

[29] M. G. Alfarizi, J. Liu, J. Vatn, and S. Yin, "Sustainability of icps from a safety perspective: Challenges and opportunities," in *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*, pp. 1–8, 2023.

[30] M. G. Alfarizi, J. Vatn, and S. Yin, "Advancements in extreme gradient boosting for enhanced fault prognosis: A continuation study from fuse test bench analysis," in *2024 IEEE 33rd International Symposium on Industrial Electronics (ISIE)*, pp. 1–8, 2024.

[31] Q. Jiang, X. Yan, and B. Huang, "Performance-driven distributed pca process monitoring based on fault-relevant variable selection and bayesian inference," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 1, pp. 377–386, 2015.

[32] Y. Zhang, N. Yang, and S. Li, "Fault isolation of nonlinear processes based on fault directions and features," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1567–1572, 2013.

[33] S. Yin, X. Zhu, and O. Kaynak, "Improved pls focused on key-performance-indicator-related fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1651–1658, 2014.

[34] X. Jin, M. Zhao, T. W. Chow, and M. Pecht, "Motor bearing fault diagnosis using trace ratio linear discriminant analysis," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2441–2451, 2013.

[35] C. Zhao and F. Gao, "Fault subspace selection approach combined with analysis of relative changes for reconstruction modeling and multifault diagnosis," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 3, pp. 928–939, 2015.

[36] M. Riera-Guasp, M. Pineda-Sánchez, J. Pérez-Cruz, R. Puche-Panadero, J. Roger-Folch, and J. A. Antonino-Daviu, "Diagnosis of induction motor faults via gabor analysis of the current in transient regime," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 6, pp. 1583–1596, 2012.

[37] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal processing*, vol. 96, pp. 1–15, 2014.

[38] W. Wang and A. H. Christer, "Towards a general condition based maintenance model for a stochastic dynamic system," *Journal of the operational research society*, vol. 51, no. 2, pp. 145–155, 2000.

[39] K. L. Tsui, N. Chen, Q. Zhou, Y. Hai, W. Wang, *et al.*, "Prognostics and health management: A review on data driven approaches," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[40] W. Q. Meeker, L. A. Escobar, and F. G. Pascual, *Statistical methods for reliability data*. John Wiley & Sons, 2022.

[41] N. D. Singpurwalla, "Survival in dynamic environments," *Statistical science*, pp. 86–103, 1995.

[42] J. Lawless and M. Crowder, "Covariates and random effects in a gamma process model with application to degradation and failure," *Lifetime data analysis*, vol. 10, pp. 213–227, 2004.

[43] Z.-S. Ye, M. Xie, L.-C. Tang, and N. Chen, "Semiparametric estimation of gamma processes for deteriorating products," *Technometrics*, vol. 56, no. 4, pp. 504–513, 2014.

[44] G. Whitmore, "Estimating degradation by a wiener diffusion process subject to measurement error," *Lifetime data analysis*, vol. 1, pp. 307–319, 1995.

[45] X. Wang, "Wiener processes with random effects for degradation data," *Journal of Multivariate Analysis*, vol. 101, no. 2, pp. 340–351, 2010.

[46] W. Wang, M. Carr, W. Xu, and K. Kobbacy, "A model for residual life prediction based on brownian motion with an adaptive drift," *Microelectronics Reliability*, vol. 51, no. 2, pp. 285–293, 2011.

[47] Z. Wang, C. Hu, W. Wang, Z. Zhou, and X. Si, "A case study of remaining storage life prediction using stochastic filtering with the influence of condition monitoring," *Reliability Engineering & System Safety*, vol. 132, pp. 186–195, 2014.

[48] Z.-S. Ye, Y. Wang, K.-L. Tsui, and M. Pecht, "Degradation data analysis using wiener processes with measurement errors," *IEEE Transactions on Reliability*, vol. 62, no. 4, pp. 772–780, 2013.

[49] X.-S. Si, W. Wang, M.-Y. Chen, C.-H. Hu, and D.-H. Zhou, "A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution," *European Journal of Operational Research*, vol. 226, no. 1, pp. 53–66, 2013.

[50] X.-S. Si, W. Wang, C.-H. Hu, M.-Y. Chen, and D.-H. Zhou, "A wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 219–237, 2013.

[51] X. Wang and D. Xu, "An inverse gaussian process model for degradation data," *Technometrics*, vol. 52, no. 2, pp. 188–197, 2010.

[52] Z.-S. Ye and N. Chen, "The inverse gaussian process as a degradation model," *Technometrics*, vol. 56, no. 3, pp. 302–311, 2014.

[53] A. Ghasemi, S. Yacout, and M.-S. Ouali, "Parameter estimation methods for condition-based maintenance with indirect observations," *IEEE Transactions on reliability*, vol. 59, no. 2, pp. 426–439, 2010.

[54] Z.-Q. Wang, C.-H. Hu, W. Wang, and X.-S. Si, "An additive wiener process-based prognostic model for hybrid deteriorating systems," *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 208–222, 2014.

[55] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–202, 1972.

[56] H. Liao, W. Zhao, and H. Guo, "Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model," in *RAMS'06. Annual Reliability and Maintainability Symposium, 2006.*, pp. 127–132, IEEE, 2006.

[57] Q. Zhou, J. Son, S. Zhou, X. Mao, and M. Salman, "Remaining useful life prediction of individual units subject to hard failure," *IIE Transactions*, vol. 46, no. 10, pp. 1017–1030, 2014.

[58] J. Bogdanoff, F. Kozin, and H. Saunders, "Probabilistic models of cumulative damage," 1988.

[59] J. Buckley and I. James, "Linear regression with censored data," *Biometrika*, vol. 66, no. 3, pp. 429–436, 1979.

[60] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273–297, 1995.

[61] X. Shu, J. Shen, Z. Chen, Y. Zhang, Y. Liu, and Y. Lin, "Remaining capacity estimation for lithium-ion batteries via co-operation of multi-machine learning algorithms," *Reliability Engineering & System Safety*, vol. 228, p. 108821, 2022.

[62] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[63] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.

[64] ISO 13381-1, "Condition monitoring and diagnostics of machines – prognostics – part 1: General guidelines," standard, International Standard Organization, 2015.

[65] P. Vaidya and M. Rausand, "Remaining useful life, technical health, and life extension," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 225, no. 2, pp. 219–231, 2011.

[66] J. W. Creswell, *Educational research: Planning, conducting, and evaluating quantitative*, vol. 7. Prentice Hall Upper Saddle River, NJ, 2002.

[67] OECD, *Frascati manual 2015: Guidelines for collecting and reporting data on research and experimental development*. OECD Publishing, 2015.

[68] K. E. Petersen, "The eu model evaluation group," *Journal of hazardous materials*, vol. 65, no. 1-2, pp. 37–41, 1999.

[69] The Research Council of Norway, *Kvalitet i norsk forskning (quality in norwegian research)*. The Research Council of Norway, 2000.

Articles

Article 1

Optimized Random Forest Model for Remaining Useful Life Prediction of Experimental Bearings

M. G. Alfarizi, B. Tajiani, J. Vatn and S. Yin, "Optimized Random Forest Model for Remaining Useful Life Prediction of Experimental Bearings," in *IEEE Transactions on Industrial Informatics*, vol. 19, no. 6, pp. 7771-7779, 2023.

Optimized Random Forest Model for Remaining Useful Life Prediction of Experimental Bearings

Muhammad Gibran Alfarizi, Bahareh Tajiani, Jørn Vatn, and Shen Yin, *Senior Member, IEEE*

***Abstract*—Bearings are essential to the reliable operation of rotating machinery in manufacturing processes. There is a rising demand for accurate bearing remaining useful life (RUL) predictions. The data-driven technique for predicting RUL of bearing has demonstrated promising prospects to facilitate intelligent prognostics. This article proposes a new data-driven prediction framework for bearing RUL utilizing an integration of empirical mode decomposition, random forest (RF), and Bayesian optimization. The proposed framework consists of two main phases: 1) feature extraction and 2) RUL prediction. The first phase of this framework focused on decomposing the empirical input signals using empirical mode decomposition into distinct frequency bands to filter out irrelevant frequencies and determine the fault characteristics of the signals. In the second phase, the RUL prediction is then carried out by an RFs-based model with its hyperparameters tuned by Bayesian optimization. The proposed approach is validated using datasets obtained from an actual run-to-failure experiment of roller bearings. The experiment results significantly improved compared to the standard data-driven and stochastic approaches.**

***Index Terms*—Bayesian optimization, empirical mode decomposition (EMD), random forest (RF), remaining useful life (RUL) prediction, roller bearings.**

I. Introduction

ROLLING bearings are vital in ensuring a stable operation in many industrial mechanical drive systems, such as aero engines, high-speed trains, and wind turbines. However, the operation of bearings may encounter many unexpected failures due to lack of lubrication, overload, and inappropriate installation. To avoid undesired consequences due to bearings failure, researchers have developed several remaining useful life (RUL) prediction methods to estimate the failure of the rolling bearings in advance and to schedule proper maintenance actions [1], [2]. RUL prediction is an efficient tool to improve the overall system's reliability and reduce maintenance expenses. However, the challenge remains in describing the degradation trend of rolling bearings due to unknown external circumstances, such as ambient noise and rotational speed, which can produce a significant discrepancy between actual rolling bearing life and theoretical calculating value.

In recent years, many approaches have been proposed to predict the RUL of bearings. These approaches can be mainly classified into model-based, data-driven, or a combination of both. Model-based approaches use mathematical functions or mappings to represent the physical nature of the failure [3]. Most model-based approaches use the crack growth curves to describe the degradation of rolling bearing, such as the Marco–Starkey theory [4], where statistical estimation methods are applied, e.g., extended Kalman filtering (EKF) [5], to update the prediction over time. Some researchers have favored stochastic processes within model-based approaches because random errors in measurements, working environment uncertainties, and stochastic dynamics in the degradation process could be considered [6]. However, several other variables will influence the final prediction findings, such as fatigue theory, computation method, materials, fracture, and vibration, resulting in a lack of extensibility. Therefore, an increasing number of researchers pay more attention to the data-driven approaches, which use historical failure data to build the degradation model without knowing the physical failure mechanism of the system. A recent review of data-driven methods for RUL prediction can be found in [7].

Data-driven approaches for bearing RUL prediction can be broadly divided into health indicator (HI)-based and direct prediction methods. Using similarity-based interpolation techniques, the first category generates a synthetic HI that depicts the level of system degradation from sensor data. For instance, Huang et al. [8] employed a similarity weighting technique based on Euclidean distance for RUL estimation. They also produced confidence intervals for the predicted RULs using an adaptive kernel density estimation method for managing uncertainty. Likewise, Wang et al. [9] suggested HI curve modeling for online RUL prediction based on sparse Bayesian learning. Yu et al. [10] used a similarity-based HI curve matching method for RUL prediction and a bidirectional recurrent-neural-network-based autoencoder to construct HIs. Although the capability of this group of ways to incorporate new instances is beneficial, the system degradation patterns could be distorted by the data-fusion techniques used to simulate HIs [11].

Machine learning models are explicitly utilized for RUL prediction in direct prediction methods. According to recent studies, HIs are just rescaled RULs and are directly predicted rather than RULs [12], [13]. Recent advances in deep learning have led to the development of numerous direct prediction algorithms.

Manuscript received 22 June 2022; revised 9 August 2022 and 22 August 2022; accepted 4 September 2022. Date of publication 13 September 2022; date of current version 24 May 2023. Paper no. TII-22-2667. *(Corresponding author: Muhammad Gibran Alfarizi.)*

The authors are with the Department of Mechanical and Industrial Engineering, Faculty of Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: muhammad.g.alfarizi@ntnu.no; bahareh.tajiani@ntnu.no; jorn.vatn@ntnu.no; shen.yin@ntnu.no).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2022.3206339.

Digital Object Identifier 10.1109/TII.2022.3206339

Using a deep belief network, Deutsch and He [14] have created RUL prediction models for gears. Li et al. [15] showed that convolutional neural networks perform better than conventional shallow machine learning models when used to predict the RULs of aircraft engines. Wu et al. [16] show that long short-term memory can predict the RULs of aviation engines.

Despite the effectiveness of deep learning models in various prognostics and health management tasks, the lack of datasets makes it difficult to increase the RUL prediction performance [17]. A significant number of parameters must be modified during the training phases of deep learning models. Deep learning models can offer excellent prediction performance on test samples with good generalization abilities when sufficient data are available. In the case of inadequately labeled historical data, training may result in unsatisfactory model parameters that only apply to the information upon which models were trained [18]. In addition, deep learning requires relatively high computational cost and lengthy training time [19]. Other algorithms, such as the random forest (RF) method, can be utilized to overcome the drawbacks of deep learning.

Developed by Breiman (in 2001) [20], RF is an ensemble machine-learning technique. It is a bagging-based approach that integrates many predictors (regression trees) generated randomly using the bootstrapping method. Multiple predictors can minimize complexity and yield superior performance compared to a single predictor. RFs can handle discrete and continuous variables without making assumptions about their distribution and are not susceptible to overfitting [21]. Moreover, the RFs approach minimizes complexity in the following three different ways.

1) It keeps complexity low by constructing submodels from sample subsets using the bootstrap.
2) It keeps computation speedy by using a parallel ensemble to create submodels on each sample.
3) It has fewer hyperparameters to tune in comparison to deep learning [22].

The hyperparameters must be determined optimally to predict the RUL of bearings using RFs. Hyperparameter optimization identifies a set of hyperparameters that minimizes a predetermined loss function when applied to a group of independent data. One approach to optimizing hyperparameters is the grid search approach, which is an exhaustive search over a manually chosen subset of a learning algorithm's hyperparameters space. Another approach, random search, is an improvement relative to grid search because it replaces the exhaustive enumeration of all combinations by randomly selecting them. When only a small number of hyperparameters influences the machine learning algorithm's performance, it can outperform grid search [23]. However, these two approaches are relatively inefficient since they do not select the next hyperparameters to assess depending on the outcomes of the prior evaluations. Previous evaluations do not influence the grid and random searches. They frequently spend a lot of time analyzing "bad" hyperparameters.

For noisy black-box functions, there is a global optimization technique called Bayesian optimization [24]. The Bayesian approach remembers the outcomes of earlier analyses, unlike random or grid search. Using these data, they create a probabilistic model that connects hyperparameters to the probability of a score on the objective function $P(\text{score} \mid \text{hyperparameters})$. Bayesian hyperparameters optimization is used because it can identify better model configurations in fewer iterations than random search by assessing hyperparameters that appear more promising based on the previous iteration.

The main objective of this article is to assess how accurately the proposed approach can predict the RUL of rolling bearings. The Bayesian optimization algorithm is used to optimize the hyperparameters of the RFs model. To validate the performance of the proposed method, the authors conducted a run-to-failure experiment of rolling bearings and collected the raw acceleration signals. The natural signals are then decomposed into several HIs using the empirical mode decomposition (EMD) [25]. The most significant advantage of the EMD method is that it is adaptive and data-driven, without the need for *a priori* basis function selection for signal decomposition. In addition, EMD, as a widely used signal processing approach, can also analyze nonlinear and nonstationary vibration signals. These HIs are then selected using feature importance from the RFs algorithm and inputted to the RFs-based Bayesian model. Its performance is compared with support vector regression (SVR), LASSO regression, artificial neural networks (ANNs), and conventional RFs. Additionally, the performance of the proposed model is compared with a model-based approach using the Wiener process.

The main contributions of this article can be summarized as follows.

1) A framework to construct HIs is proposed by applying EMD and selecting them using feature importance to determine the best feature for RUL prediction.
2) A novel data-driven approach is proposed by utilizing the proposed HI framework, RF, and Bayesian optimization to offer precise RUL prediction to overcome the drawbacks of deep learning (e.g., high computational cost and lengthy training time) and improve prediction accuracy.
3) The proposed RUL prediction approach is verified by real-world datasets obtained from a run-to-failure experiment of rolling bearings. Furthermore, the performance of the proposed model has been investigated and compared with other data-driven and model-based approaches for RUL prediction of bearings.

The remainder of this article is organized as follows. Section II presents the proposed framework and the preliminaries of our approach. Section III describes the data used in the case study. Section IV discusses the performance of the proposed model in real-world datasets. Section V concludes this article.

II. Proposed Method

A. Framework of the Proposed Method

Fig. 1 depicts the framework of the proposed method. The initial phase of this framework is focused on feature extraction, in which each of the gathered raw acceleration signals is decomposed into distinct frequency bands to filter out irrelevant frequencies and determine the fault characteristics of the signals. The signal decomposition is performed with EMD. The EMD decomposition is based on the variation in frequency over the bearings' lifespan. The statistical features as indicators (i.e.,

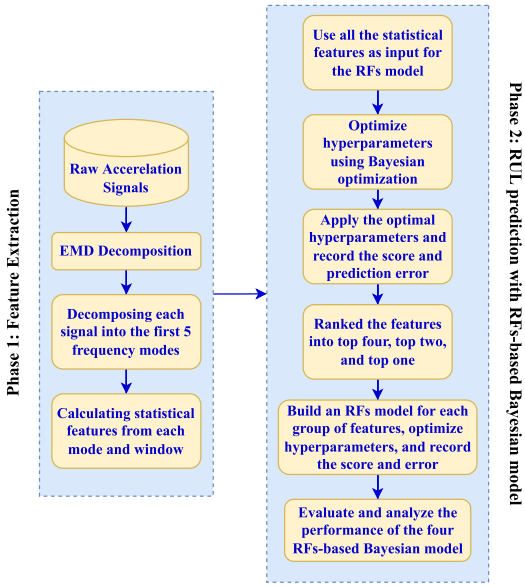

Fig. 1. Framework of the proposed method.

Algorithm 1: Empirical Mode Decomposition.

Input: Acceleration signal $y(t)$
Identify all local maxima and minima of $y(t)$
for $i = 1$ **to** n **do**
while $j \leq k$ **do**
The upper and lower envelopes, y_u^t and y_l^t, are formed by connecting all local maxima and minima with a cubic spline line, respectively
Calculate the mean of y_u^t, y_l^t, as $m_{i1}^t = (y_u^t + y_l^t)/2$, and $h_{i1}^t = y(t) - m_{i1}^t$
The above sifting process is repeated several times as $h_{ij}^t = h_{i(j-1)}^t - m_{ij}^t$
end while
h_{1k}^t satisfies the definition of IMF and the stopping condition, SD, where $SD = \sum_{t=0}^{T} \left(\frac{|h_{1(k-1)}^t - h_{1k}^t|}{h_{1(k-1)}^t} \right)^2$, it becomes IMF, that is, $c_1 = h_{jk}^t$
A new signal, x_1, is generated by $x_1 = y(t) - c_1$
Repeat above steps as $x_i = x_{i-1} - c_i$
end for
Output: The original signal is decomposed into n IMFs, and one residual x_n

kurtosis, RMS, crest factor, shape factor, impulse factor, skewness, and mean value) are calculated for each intrinsic mode function (IMF) to compare the frequency bands and determine which feature from which band is suitable for predicting RUL of bearings. The EMD decomposition is further elaborated in Section II-B.

The second phase is related to the RUL prediction. The bearing datasets have been obtained by an experimental laboratory with the same operating condition. The statistical features from other bearings' datasets are used as input basis for the RFs-based model to predict the RUL of the current bearing of interest. The hyperparameters of concern (the number of trees in the forest, the maximum depth of the tree, and the minimum number of samples required to be at a leaf node) of every RFs-based model are optimized using Bayesian optimization. The features are then ranked and divided into four groups based on their importance, which results in groups containing all, top four, top two, and top one features. The RFs-based model considers every feature group as the input to predict the RUL of the current bearing of interest. The RUL prediction of every bearing is evaluated based on its score and prediction error. The details of each step will be further elaborated in Section II-F.

B. Empirical Mode Decomposition

Traditional signal-processing techniques assume that the signals are linear or stationary. Thus, they may result in false information and erroneous results. However, among various time-frequency methods, the Hilbert–Huang transform (HHT) is an empirical approach that can analyze stationary and nonstationary signals for fault detection, condition monitoring, and failure prognosis of bearings [25]. HHT is an adaptive data-driven approach that can also handle the nonlinearity of the signals [26]. The first step in HHT is EMD, which decomposes the signals into different segments called IMFs based on their time-domain characteristics. The process of carrying out an EMD approach on an acceleration signal is summarized in Algorithm 1.

C. Random Forests

The RFs algorithm is based on a collection of predictors influenced by each forest tree's random value. To observe the modeling of RFs, input datasets are extracted from a group and then randomized. The higher success of the RFs algorithm is attributable to its efficient handling of big datasets, rapid operational speed, and lack of predictor overfitting [27].

The forest in RFs is built using the p dimension input vector of $X = x_1, x_2, \ldots, x_p$. Then, a set of K trees $\{T_1(x), T_2(x), \ldots, T_k(x)\}$ is developed inside the forest. Each tree determines its output value, represented as $\hat{Y}_1 = T_1(X), \ldots, \hat{Y}_m = T_m(X)$, where $m = 1, \ldots, K$. The overall output of the RFs is calculated by the estimation of an average of all tree predictors, mathematically expressed as

$$\text{Predict}_{\text{RFs}}(X) = \frac{1}{K} \sum_{m=1}^{K} \hat{Y}_m(X). \tag{1}$$

The training dataset $D = D_1, D_2, \ldots, D_n = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ is taken independently from input and output, where $x_i, i = 1, \ldots, n$ denotes the input vector training dataset and $y_i, i = 1, \ldots, n$ denotes the output vector training dataset. From the original sample set D, the RF randomly creates K tree sample sets using the bootstrap resampling technique. Approximately one-third of the data in the original sample set D is not drawn in bootstrap samples;

Algorithm 2: Random Forest.

Input: Training samples $D = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, testing samples x_t
for $m = 1$ **to** K **do**
Construct the bootstrap sample D_m randomly from the original training set D with replacement
Create a nonpruning decision tree $\hat{Y}_m$ through D_m
Pick n_try features at random from N features
Choose the best feature from each node's n_try features to split
Split the tree until it reaches its maximum size
end for
Output: A set of decision trees $\{\hat{Y}_m, m = 1, 2, \ldots, K\}$. The decision tree $\hat{Y}_m$ generates the predictor $\hat{Y}_m(x_t)$ for the testing samples. The overall output of the RFs is calculated as in (1)

these data are known as out-of-bag (OOB) data, whereas the remaining data are known as in-bag data. Pseudocode of the RF algorithm [20] is described in Algorithm 2.

As stated previously, the growth of each tree depends on a bootstrap sample containing almost two-thirds of the training data. Testing is done with the remaining training data (OOB). As shown in the following equation, the MSE of the OOB error is derived by deducting the predicted values from the reference values

$$\text{MSE} \approx \text{MSE}^{\text{OOB}} = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}(X_i) - Y_i)^2 \tag{2}$$

where $\hat{Y}(X_i)$ denotes the predicted output, Y_i denotes the observed output, and N represents the total number of samples.

D. Bayesian Optimization

Bayesian optimization produces a probabilistic model of the function mapping from hyperparameter values to the objective as evaluated by a validation set when used for hyperparameter optimization. Bayesian optimization is an iterative procedure that involves two fundamental components: 1) a probabilistic surrogate model and 2) an acquisition function to determine the next point to examine. In each iteration, the surrogate model is fitted to all previous observations of the target function. After that, the acquisition function, which utilizes the probabilistic model's predictive distribution, decides the utility of various prospective spots while balancing exploitation and exploration. The acquisition function can be thoroughly tuned because it is inexpensive to compute compared to assessing the costly black-box process.

A visual representation of Bayesian optimization is shown in Fig. 2. After two evaluations, the black line represents the surrogate model's initial estimate, with the related uncertainty shown in gray. The surrogate model falls short of the actual objective function shown in red.

The surrogate function after eight evaluations is shown in Fig. 3. The real objection function is now matched with the surrogate function. In light of this, the hyperparameters that

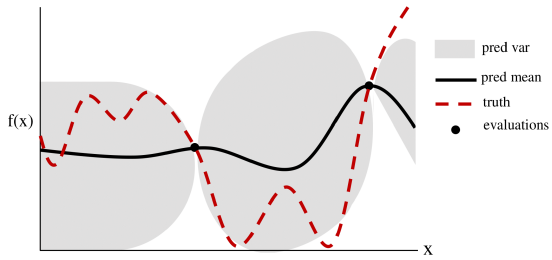

Fig. 2. Bayesian optimization run after two evaluations. Modified from Snoek et al. [28].

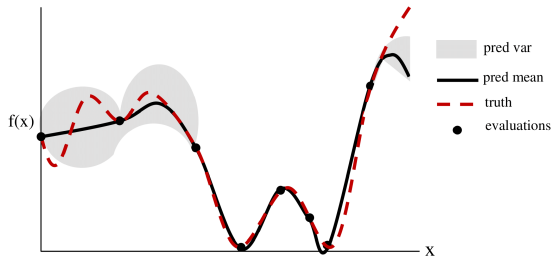

Fig. 3. Bayesian optimization run after eight evaluations. Modified from Snoek et al. [28].

maximize the surrogate will also optimize the actual objective function.

One of the acquisition functions is expected improvement (EI)

$$\text{EI}_{y^*}(x) = \int_{-\infty}^{y^*}(y^* - y)p(y|x)dy \tag{3}$$

where y^* is the objective function's threshold value, the objective function value y is calculated using the hyperparameters x and the surrogate probability model $p(y|x)$ expresses the probability that y will occur given x. The goal is to maximize the EI concerning x.

The tree-structured Parzen estimator applies the Bayes rule, which is expressed as

$$p(y|x) = \frac{p(x|y) * p(y)}{p(x)} \tag{4}$$

to construct a model where the probability of the hyperparameters given the score on the objective function is denoted by $p(x|y)$, which may be written as

$$p(x|y) = \begin{cases} l(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases} \tag{5}$$

where $y < y^*$ denotes an objective function value below the threshold. According to this equation, there are two possible distributions for the hyperparameters: One in which the objective function value is lower than the threshold ($l(x)$) and another in which it is greater ($g(x)$).

Since the $l(x)$ distribution only includes x that resulted in scores below the threshold, draw values of x from this distribution are superior. Bayes' rule also supports it since the EI

Algorithm 3: Bayesian Hyperparameter Optimization.

Input: (hyperparameter space Θ, Target score function $H(\theta)$, max number of evaluation $n_{\max}$)
Select an initial configuration $\theta_0 \in \Theta$
Evaluate the initial score $y_0 = H(\theta_0)$
Set $\theta^* = \theta_0, y^* = H(\theta_0)$, and $S_0 = \{\theta_0, y_0\}$
for $n = 1$ **to** $n_{\max}$ **do**
  Select a new hyperparameter configuration $\theta_n \in \Theta$ by maximizing EI in (6)
  Evaluate H in θ_n to obtain a new numeric score $y_n = H(\theta_n)$
  Augment the data $S_n = S_{n-1} \cup \{\theta_n, y_n\}$
  Update the surrogate model
  if $y_n < F^*$ **then**
    $\theta^* = \theta_n$ and $y^* = y_n$
  end if
end for
Output: θ^* and y^*

equation becomes

$$\mathrm{EI}_{y^*}(x) = \frac{\gamma y^* l(x) - l(x)\int_{-\infty}^{y^*} p(y)dy}{\gamma l(x) + (1-\gamma)g(x)} \propto \left(\gamma + \frac{g(x)}{l(x)}(1-\gamma)\right)^{-1} \tag{6}$$

which states that the EI is inversely proportional to $\frac{l(x)}{g(x)}$. Therefore, this ratio will be maximized when values for x are taken from $l(x)$, increasing the EI. The variable γ represents the quantile of the negative accuracy scores (observed thus far), used as a cutoff point, and defines the threshold y^*. This user-defined cutoff point is often set at 15% by default.

Before saving the results and scores, the algorithm compares each new set of hyperparameters it proposes with the real objective function. The algorithm builds a probabilistic model that gets better with each iteration by computing $l(x)$ and $g(x)$ using the account of the objective function.

In conclusion, a preliminary estimate of the surrogate function is developed and modified as new information becomes available. When there are enough evaluations of surrogate functions, the objective function will eventually be adequately reflected. The objective function would also be maximized by the same hyperparameters that maximize the EI. The pseudocode for Bayesian hyperparameter optimization is presented in Algorithm 3.

E. Model Evaluation

A score function is defined to evaluate the model's performance by comparing the actual and predicted RUL of bearings. The score function is based on the percent errors of predictions, which is defined as

$$\%\mathrm{Error} = 100 \times \frac{\text{Act RUL} - \widehat{\text{RUL}}}{\text{Act RUL}}. \tag{7}$$

The early prediction of RUL (i.e., cases where $\%\mathrm{Error} > 0$) and late prediction of RUL (i.e., cases where $\%\mathrm{Error} \leq 0$) will be treated differently. Thus, a penalty function, expressed in (8), is introduced, which penalizes late RUL predictions more than early ones. Late RUL predictions might result in more severe consequences and are more crucial than early ones [29]

$$A_i = \begin{cases} \exp\left(-\ln(0.5)\cdot(\%\mathrm{Error}/5)\right) & \text{if } \%\mathrm{Error} \leq 0 \\ \exp\left(+\ln(0.5)\cdot(\%\mathrm{Error}/20)\right) & \text{if } \%\mathrm{Error} > 0. \end{cases} \tag{8}$$

Moreover, accurate RUL predictions at the later stage of a bearing's lifetime are more vital than predictions at the early stage. So, higher weights are given to the predictions near the bearing's time to failure. The weight is assigned linearly as the first prediction data point weights 1, and the nth prediction weights n. The score function is then defined as

$$\mathrm{Score} = \frac{\sum_{i=1}^{n} w_i \times A_i}{\sum_{i=1}^{n} w_i} \tag{9}$$

where n is the number of samples, and w_i is the weight of ith prediction.

In addition to the score function, the root-mean-square relative error (RMSRE) is also calculated. RMSRE is expressed as

$$\mathrm{RMSRE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\text{Act RUL} - \widehat{\text{RUL}}}{\text{Act RUL}}\right)^2}. \tag{10}$$

A higher score and lower RMSRE values indicate that the model's predictive abilities for RUL prediction are superior.

TABLE I
INITIAL HYPERPARAMETERS AND THEIR SEARCH RANGE

Hyperparameter	Initial Value	Search Space
The number of trees	100	[10:500]
The maximum depth of tree	50	[2:100]
The minimum leaf per tree	1	[1:50]

F. Framework of the Proposed RFs-Based Bayesian Model

The proposed model's objective is to accurately predict the RUL of rolling bearings. The RFs regression model is responsible for predicting the RUL while Bayesian optimization will optimize the RFs model's hyperparameters to increase its prediction accuracy. The RUL prediction of bearings begins with specifying the input sample, the features, and the hyperparameters' initial values. The initial hyperparameters and their search range are summarized in Table I.

The optimal hyperparameter settings will decrease prediction error [30]. According to the literature, no analytical formula exists to find the optimum set of hyperparameters [31]. Thus, Bayesian optimization is combined with the RFs model to determine the optimal hyperparameter settings, resulting in lower prediction error. The objective of the optimization is to minimize the MSE of the RFs model. The flowchart of the proposed RFs-based Bayesian model is shown in phase 2 of Fig. 1.

The steps involved in phase 2 of the flowchart are explained as follows.

TABLE II
BEARINGS AND THEIR NUMBER OF SAMPLES

Bearing	B1	B2	B3	B4	B5
Number of samples	110	35	146	106	77
Bearing	B6	B7	B8	B9	B10
Number of samples	248	150	143	114	115

1) The samples and features are obtained from the feature extraction phase with EMD decomposition using other bearings' data for each bearing. The features used for input are kurtosis, RMS, crest factor, shape factor, impulse factor, skewness, and mean value. The output is the RUL of bearings for every timestep.
2) The hyperparameters (see Table I) are set to the initial values and then optimized using Bayesian optimization. It was observed that most of the time, the value of the surrogate function does not improve after 40 evaluations; the maximum number of evaluations is set to 50.
3) The optimal set of hyperparameters is applied to the RFs model, and the score and the RMSRE are recorded for model evaluation.
4) The features are ranked based on their importance in predicting the RUL of bearings using RFs feature importance and SHAP importance analysis. The best 50% of the features are selected and then ranked again. The process is repeated, and as a result, the features are divided into the top four, the top two, and the top one.
5) Build an RFs model for each group of features and repeat steps 2 and 3. In total, there are four different RFs models for each bearing.
6) Evaluate and analyze the performance of the four RFs-based Bayesian models.

III. CASE STUDY

A. Data Description

A set of run-to-failure experimental tests on roller bearings is designed to illustrate the effectiveness of the proposed approach. The bearings are operated under similar laboratory conditions with the same motor speed of 2975 r/min. The original data are a batch of acceleration signals in the horizontal direction collected from the bearings' healthy state until the acceleration amplitude crosses the level of $10g$ where the experiments ceased. The bearings are degraded by contamination, i.e., pouring a mixture of Silicium carbide solid particles and lubricant onto the bearing at regular intervals until the stopping threshold (i.e., $10g$) is reached. Table II lists the number of samples of different bearings from a healthy state to a failed state. The details of the data and the vibration setup used to conduct the experiments can be found in [32].

B. Health Indicator

In this section, EMD is implemented to decompose each horizontal acceleration signal over different frequency ranges according to the signal's time-scale characteristics resulting in a series of IMFs. Comparing different IMFs with their feature scales for each signal gives insights into the relevant or informative frequencies and features concerning the bearings' faults. The statistical features will then be calculated from each of the first 5 IMFs of the acceleration signals. The first IMF in EMD is the dominant feature among all with the highest value of fitness function (i.e., metric that combines monotonicity, trendability, and prognosability). It is thus chosen for the HIs. Table III shows how to calculate the statistical features where M is the number of measurements in one sample, c_i is the processed signal measurements obtained by the EMD algorithm, and the mean and standard deviation are m and σ, respectively.

In every bearing, the RMS feature has the highest value in monotonicity, prognosability, and trendability, which means it is the essential feature for RUL prediction. The RFs feature importance and SHAP importance analysis also corroborate this finding, as RMS is the critical feature for every bearing. RMS is also the only feature that visually trends upward as time passes, indicating higher degradation over time until bearing failure. However, the value of RMS in one bearing cannot be compared to other bearings, as the RMS value at failure time is different in every bearing. Thus, the RMS feature must be considered individually.

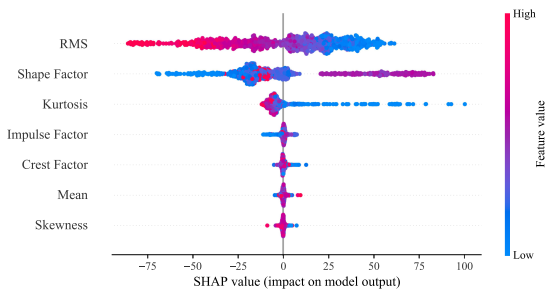

Fig. 4. SHAP variable importance plot for B10.

IV. EXPERIMENT RESULTS AND ANALYSIS

A. Model Results

Fig. 4 shows the SHAP variable importance plot for B10. The plot shows that the RUL prediction depends on the RMS feature. Moreover, the RMS is negatively correlated with the RUL value, which means that the higher the RMS value, the lower the RUL value. This correlation makes sense since higher RMS indicates higher degradation, which means lower the RUL of the bearing. The SHAP variable importance plot shows similar results for other bearings.

Tables IV and V show the score and RMSRE of the RFs-based Bayesian model for every bearing after hyperparameters optimization is applied to every group of features. Fig. 5 shows the RUL prediction for every bearing with the best score.

Tables IV and V show that the proposed model, on average, produces the highest score and lowest RMSRE when using only the top two features. However, the differences in score and error are only significant when using only one feature and insignificant

TABLE III
FORMULA FOR STATISTICAL FEATURES

Feature	Kurtosis	RMS	Crest factor	Skewness
Formula	$K = \frac{\sum_{i=1}^{M}(c_i - m)^4}{(M-1)\sigma^4}$	$RMS = \sqrt{\frac{1}{M}\sum_{i=1}^{M} c_i^2}$	$CF = \frac{\max(\lvert c_i \rvert)}{\sqrt{\frac{1}{M}\sum_{i=1}^{M} c_i^2}}$	$S = \frac{\sum_{i=1}^{M}(c_i - m)^3}{(M-1)\sigma^3}$
Feature	Mean	Shape factor	Impulse factor	
Formula	$Mean = \frac{1}{M}\sum_{i=1}^{M} c_i$	$SF = \frac{\sqrt{\frac{1}{M}\sum_{i=1}^{M} c_i^2}}{\frac{1}{M}\sum_{i=1}^{M} \lvert c_i \rvert}$	$IF = \frac{\max \lvert c_i \rvert}{\frac{1}{M}\sum_{i=1}^{M} \lvert c_i \rvert}$	

TABLE IV
SCORE OF RUL PREDICTIONS FOR EVERY BEARING

Score	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	Average
All features	0.7621	0.7950	0.9277	0.9533	0.8760	0.9648	0.8809	0.9516	0.9094	0.9461	0.8967
4 features	0.7636	0.7853	0.9307	0.9580	0.8643	0.9634	0.8972	0.9537	0.9164	0.9487	0.8981
2 features	0.7761	0.8597	0.9362	0.9612	0.9329	0.9641	0.8447	0.9603	0.9135	0.9429	0.9092
1 feature	0.9536	0.6414	0.8151	0.8413	0.8096	0.9645	0.8532	0.8266	0.7917	0.8306	0.8328

TABLE V
RMSRE OF RUL PREDICTIONS FOR EVERY BEARING

RMSRE	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	Average
All features	2.6739	2.2487	1.3654	0.7220	1.1861	0.8493	1.4983	0.8499	1.4479	0.9599	1.3801
4 features	2.6187	2.2637	1.0971	0.6671	1.4759	0.9307	1.3785	0.8547	1.4315	0.8149	1.3533
2 features	2.3878	1.3728	1.2691	0.6417	0.7147	0.8891	1.8001	0.7249	1.2908	1.1332	1.2224
1 feature	0.5278	4.2387	4.3198	3.4316	1.7146	0.8916	1.8849	5.5522	7.3177	2.9226	3.2802

otherwise. The top feature here is the RMS feature. This difference suggests that using multiple HIs is beneficial for predicting RUL values. Many HIs can provide better information about the bearing's health condition than only one. The proposed model also runs efficiently due to the use of Bayesian optimization to update parameters, which avoids the high computational cost and lengthy training time.

There are differences between the proposed method as a data-driven approach and the Wiener process as a stochastic approach. In the Wiener process, the underlying randomness or the probabilistic behavior in the bearings' degradation is considered while the uncertainty of the failure threshold is also considered at each timestep. On the other hand, the proposed method can provide better RUL prediction results regarding score and error without determining a failure threshold.

Fig. 5 shows that the prediction results of the proposed model are close to the actual RUL values in most of the bearings, especially near the end of the bearings' lifetime, where the RUL prediction matters most. Generally, the proposed model performs better when the bearing has an average number of samples (i.e., 100–110 samples) like bearing 1, 4, 9, and 10. However, the model does not give an accurate RUL prediction when the bearing only has a small number of samples (i.e., short lifetime), such as bearing 2 and 5. According to the manufacturer, the average lifespan of such bearings is between 8 and 12 hours. Consequently, bearing 2 and 5 are atypical, and there may be technical problems in conducting experiments. The technical issue may be caused by the accelerometers not being correctly mounted to the bearings, resulting in erroneous signals. In addition, the predictions of bearing 6 perform poorly compared to the other bearings. Its inferior prediction could be due to its higher number of samples (248) than others. Because the RFs model learned from other bearings datasets to make predictions, it has not seen samples higher than 200 and, therefore, makes outrageous predictions.

Generally, at the beginning of a bearing's lifespan, the model has a more considerable prediction error, and this is due to some uncertainty in the input data. At the middle and deterioration stages, when the bearings are approaching failure, the uncertainty has decreased because more degradation data have been seen and utilized to capture the degradation trend, resulting in a more accurate RUL prediction.

B. Results Validation

A comparison with other data-driven algorithms is made to corroborate the proposed model's performance, including conventional RFs, ANNs [33], SVR [34], and LASSO regression [35]. The hyperparameters of ANN (number of layers, hidden unit, and learning rate), SVR (gamma and C), and LASSO (alpha) are optimized with Bayesian hyperparameters optimization. Moreover, the performance of the proposed model is also compared with a model-based approach using the Wiener process. The average score and RMSRE are recorded for every method and summarized in Table VI. The proposed method results are bolded in Table VI.

Based on the score and RMSRE in Table VI, the proposed method outperforms all other ways to predict the RUL of bearings. Compared to conventional RFs, the proposed model has improved the average score by 3% and reduced the average error by 9%. Furthermore, compared to the Wiener process, the proposed model has enhanced the average score by 6% and reduced the average error by 54%. A higher score and lower

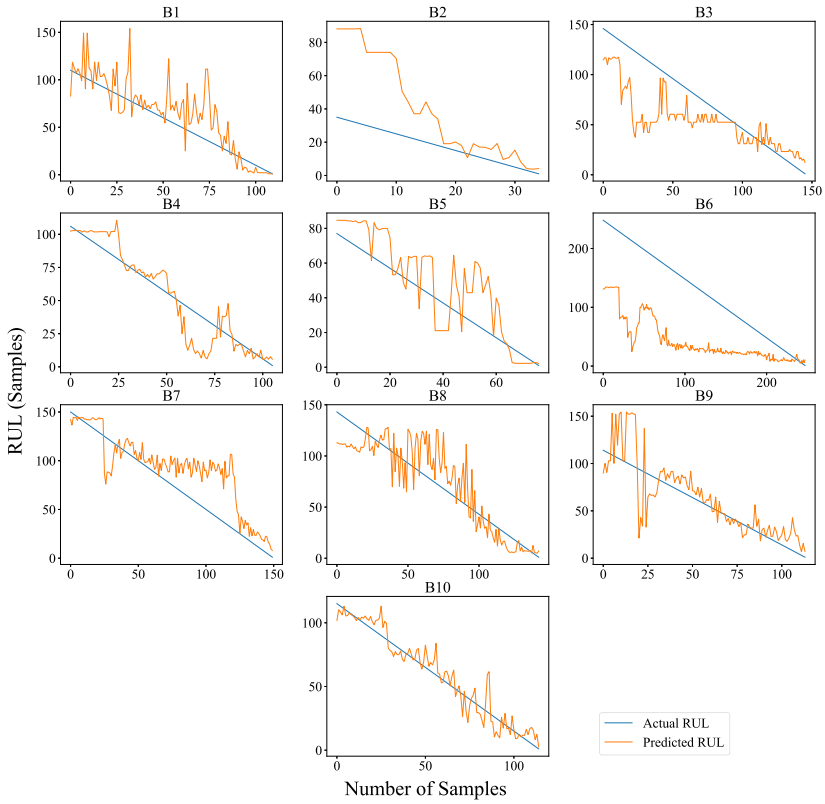

Fig. 5. Best RUL prediction of every bearing.

TABLE VI
SCORE AND RMSRE OF ALL METHODS

Method	Avg. Score	Avg. RMSRE
Proposed method	**0.9092**	**1.2224**
Conventional RFs	0.8856	1.3375
ANN	0.7735	6.5496
SVR	0.7581	7.5216
LASSO	0.8196	3.6325
Wiener process	0.8606	2.6301

RMSRE values imply that the model's predictive capabilities for RUL prediction are superior.

V. CONCLUSION

This article proposes a data-driven framework for predicting the RUL of bearings. The framework consists of the feature extraction phase and RUL prediction phase. EMD performed the feature extraction to decompose input signals into different frequency bands to identify signal fault characteristics. Then, an RFs-based model combined with Bayesian hyperparameters optimization was developed to predict the RUL of bearings. A run-to-failure experiment of roller bearings was performed to validate the proposed method. Score function and RMSRE were used to evaluate the method's performance. The proposed method outperforms other data-driven and model-based approaches in terms of score and error. Additionally, the proposed approach requires relatively low computational cost and fast training time compared to the deep learning approach. However, like any RFs-based method, one has to find a tradeoff between the training time and prediction accuracy. RF is also a black box algorithm, so the results are not easily interpretable, i.e., it does not provide complete visibility into the coefficients like linear regression.

For further research, a more detailed investigation is required to compare the proposed method and the stochastic Wiener process depending on the decision context, the length of the dataset or the bearing's lifetime, accuracy of the prediction results on the two distinct healthy stages, and deterioration stage of bearings, as well as their mathematical complexity and computational time.

REFERENCES

[1] W. Mao, J. Chen, J. Liu, and X. Liang, "Self-supervised deep domain-adversarial regression adaptation for online remaining useful life prediction of rolling bearing under unknown working condition," *IEEE Trans. Ind. Informat.*, 2022, doi: 10.1109/TII.2022.3172704.

[2] Y. Qin, D. Chen, S. Xiang, and C. Zhu, "Gated dual attention unit neural networks for remaining useful life prediction of rolling bearings," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 6438–6447, Sep. 2021.

[3] M. G. Alfarizi, J. Vatn, and S. Yin, "An extreme gradient boosting aided fault diagnosis approach: A case study of fuse test bench," *IEEE Trans. Artif. Intell.*, 2022, doi: 10.1109/TAI.2022.3165137.
[4] L. Cortese, F. Nalli, and M. Rossi, "A nonlinear model for ductile damage accumulation under multiaxial non-proportional loading conditions," *Int. J. Plast.*, vol. 85, pp. 77–92, 2016.
[5] L. Cui, X. Wang, Y. Xu, H. Jiang, and J. Zhou, "A novel switching unscented Kalman filter method for remaining useful life prediction of rolling bearing," *Measurement*, vol. 135, pp. 678–684, 2019.
[6] Z. Zhang, X. Si, C. Hu, and Y. Lei, "Degradation data analysis and remaining useful life estimation: A review on wiener-process-based methods," *Eur. J. Oper. Res.*, vol. 271, no. 3, pp. 775–796, 2018.
[7] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, 2018.
[8] C.-G. Huang, H.-Z. Huang, W. Peng, and T. Huang, "Improved trajectory similarity-based approach for turbofan engine prognostics," *J. Mech. Sci. Technol.*, vol. 33, no. 10, pp. 4877–4890, 2019.
[9] P. Wang, B. D. Youn, and C. Hu, "A generic probabilistic framework for structural health prognostics and uncertainty management," *Mech. Syst. Signal Process.*, vol. 28, pp. 622–637, 2012.
[10] W. Yu, I. Y. Kim, and C. Mechefske, "An improved similarity-based prognostic algorithm for RUL estimation using an RNN autoencoder scheme," *Rel. Eng. Syst. Saf.*, vol. 199, 2020, Art. no. 106926.
[11] R. Khelif, S. Malinowski, B. Chebel-Morello, and N. Zerhouni, "RUL prediction based on a new similarity-instance based approach," in *Proc. IEEE 23rd Int. Symp. Ind. Electron.*, 2014, pp. 2463–2468.
[12] J. Zhang, Y. Jiang, S. Wu, X. Li, H. Luo, and S. Yin, "Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism," *Rel. Eng. Syst. Saf.*, vol. 221, 2022, Art. no. 108297.
[13] J. Zhang, Y. Jiang, X. Li, M. Huo, H. Luo, and S. Yin, "An adaptive remaining useful life prediction approach for single battery with unlabeled small sample data and parameter uncertainty," *Rel. Eng. Syst. Saf.*, vol. 222, 2022, Art. no. 108357.
[14] J. Deutsch and D. He, "Using deep learning-based approach to predict remaining useful life of rotating components," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 1, pp. 11–20, Jan. 2018.
[15] X. Li, Q. Ding, and J.-Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Rel. Eng. Syst. Saf.*, vol. 172, pp. 1–11, 2018.
[16] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla LSTM neural networks," *Neurocomputing*, vol. 275, pp. 167–179, 2018.
[17] C. Sun, M. Ma, Z. Zhao, S. Tian, R. Yan, and X. Chen, "Deep transfer learning based on sparse autoencoder for remaining useful life prediction of tool in manufacturing," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2416–2425, Apr. 2019.
[18] D. Yoo, H. Fan, V. Boddeti, and K. Kitani, "Efficient k-shot learning with regularized deep networks," in *Proc. AAAI Conf. Artif. Intell.*, vol. 32, no. 1, 2018, pp. 4382–4389.
[19] C. Ferreira and G. Gonçalves, "Remaining useful life prediction and challenges: A literature review on the use of machine learning methods," *J. Manuf. Syst.*, vol. 63, pp. 550–562, 2022.
[20] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[21] C. Lei et al., "A comparison of random forest and support vector machine approaches to predict coal spontaneous combustion in gob," *Fuel*, vol. 239, pp. 297–311, 2019.
[22] I. A. Ibrahim and T. Khatib, "A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm," *Energy Convers. Manage.*, vol. 138, pp. 413–425, 2017.
[23] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 2, pp. 281–305, Feb. 2012.
[24] J. Mockus, *Bayesian Approach to Global Optimization: Theory and Applications*. Berlin, Germany: Springer, 2012, vol. 37.
[25] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process.*, vol. 35, no. 1/2, pp. 108–126, 2013.
[26] Q. Miao, D. Wang, and M. Pecht, "Rolling element bearing fault feature extraction using EMD-based independent component analysis," in *Proc. IEEE Conf. Prognostics Health Manage.*, 2011, pp. 1–6.
[27] M. H. Lipu et al., "Real-time state of charge estimation of lithium-ion batteries using optimized random forest regression algorithm," *IEEE Trans. Intell. Veh.*, 2022, doi: 10.1109/TIV.2022.3161301.
[28] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, Red Hook, NY, USA, 2012, vol. 2, pp. 2951–2959.
[29] P. Nectoux et al., "Pronostia: An experimental platform for bearings accelerated degradation tests," in *Proc. IEEE Int. Conf. Prognostics Health Manage.*, 2012, pp. 1–8.
[30] I. A. Ibrahim, M. Hossain, and B. C. Duck, "An optimized offline random forests-based model for ultra-short-term prediction of PV characteristics," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 202–214, Jan. 2020.
[31] A. Liaw et al., "Classification and regression by RandomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
[32] B. Tajiani and J. Vatn, "RUL prediction of bearings using empirical wavelet transform and Bayesian approach," in *Proc. 31st Eur. Saf. Rel. Conf.*, 2021, pp. 2006–2013.
[33] A. Krogh, "What are artificial neural networks?," *Nature Biotechnol.*, vol. 26, no. 2, pp. 195–197, 2008.
[34] M. Awad and R. Khanna, "Support vector regression," in *Efficient Learning Machines*. Berlin, Germany: Springer, 2015, pp. 67–80.
[35] L. Meier, S. Van De Geer, and P. Bühlmann, "The group LASSO for logistic regression," *J. Roy. Stat. Soc.: Ser. B (Stat. Methodol.)*, vol. 70, no. 1, pp. 53–71, 2008.

Muhammad Gibran Alfarizi received the B.Sc. degree from the Bandung Institute of Technology, Bandung, Indonesia, in 2017, and the M.Sc. degree from the Norwegian University of Science and Technology, Trondheim, Norway, in 2020, both in petroleum engineering. He is currently working toward the Ph.D. degree in reliability, availability, maintainability, and safety with the Norwegian University of Science and Technology.

His research interests include safety and reliability, data-driven monitoring and optimization, and predictive maintenance.

Bahareh Tajiani received the bachelor's degree in industrial engineering from the University of Tehran, Tehran, Iran, in 2016, and the master's degree in reliability, availability, maintenance, and safety (RAMS) from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2019. She is currently working toward the Ph.D. degree with RAMS group, NTNU.

Her main research interests include stochastic modeling, accelerated life tests, and signal processing.

Jørn Vatn received the M.Sc. degree in mathematical statistics and the Ph.D. degree with a dissertation on maintenance optimization from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 1986 and 1996, respectively.

He has more than 30 years of experience as a Researcher with SINTEF. He is currently a Professor with Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway.

Shen Yin (Senior Member, IEEE) received the M.Sc. degree in control and information system and Ph.D. (Dr.-Ing.) degree in electrical engineering and information technology from the University of Duisburg-Essen, Duisburg, Germany, in 2007 and 2012, respectively.

He is currently a Full Professor (DNV-GL Professor) with the Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway. His research interests include fault diagnosis/prognosis and fault-tolerance strategy, safety, reliability, and security of complicated systems, system and control theory, data-driven and machine learning approaches, and applications in health diagnosis and cyber-physical systems.

Article 2

An Extreme Gradient Boosting Aided Fault Diagnosis Approach: A Case Study of Fuse Test Bench

M. G. Alfarizi, J. Vatn and S. Yin, "An Extreme Gradient Boosting Aided Fault Diagnosis Approach: A Case Study of Fuse Test Bench," in *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 661-668, 2023.

An Extreme Gradient Boosting Aided Fault Diagnosis Approach: A Case Study of Fuse Test Bench

Muhammad Gibran Alfarizi, *Member, IEEE*, Jørn Vatn, and Shen Yin, *Senior Member, IEEE*

***Abstract*—The health status of a fuse test bench is essential to monitor to ensure quality control of the fuse. A system failure during operation will lead to significant impacts on the final quality of fuses. Thus, it is important to have a fault diagnosis system to detect, classify, and identify the root causes of faults to prevent operation failure. An effective fault diagnosis system should have high accuracy, fast diagnosis time, and interpretable root cause analysis. This article proposes an integrated fault diagnosis system based on extreme gradient boosting for an automated fuse test bench to solve those challenges. The proposed diagnosis system is then validated using the dataset from PHM 2021 Data Challenge. Performance comparison of the fault diagnosis system with other standard approaches in practice is also carried out. Experimental results show that the diagnostic accuracy of the proposed system outperforms several standard fault diagnostic approaches.**

***Impact Statement*—An appropriate fault diagnosis system is needed to ensure a safe operation of the fuse test bench. This paper proposes a fault diagnosis system based on extreme gradient boosting for a fuse test bench. The proposed system in this paper reaches 97% accuracy in both fault detection and classification and gives interpretable root cause analysis. A diagnosis system that can accurately detect when there is a fault and identify which type of fault can lead to swift and proper maintenance action. Early maintenance, which prevents a fault turns into a system failure, will ensure the quality control of fuses continues and reduce the probability of mislabelling the quality of the fuses.**

***Index Terms*—Extreme gradient boosting, fault classification, fault detection, fault diagnosis, fuses, root cause analysis.**

I. Introduction and Case Study

FUSES have been used as essential safety devices from the early days of electrical engineering. A fuse operates to provide overcurrent protection of an electrical circuit. It is crucial to let them pass the quality control pipeline to ensure the quality of the fuses. With the rapid digital transformation of the manufacturing/production process in the era of Industry 4.0, a fuse test bench serves as an automated quality control pipeline for electrical fuses.

The fuse test bench generally consists of a four-axis SCARA-robot to pick up electrical fuses with a vacuum gripper, a thermal camera to monitor the temperature, and a camera to detect fuses on a feeder. The indicated system first assesses the fuse conductivity. If it is conductive, the fuse is then heated up by applying 200-mA current in a 1.5-s time interval, in which a thermal camera measures the heating process. After the test completion, the fuse is moved back into the feeder with two conveyor belts.

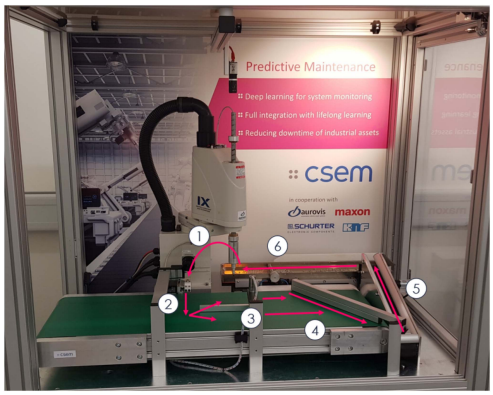

Fig. 1. Quality control pipeline for electrical fuses system [27].

The automated quality control process is depicted in Fig. 1. The fuses are first picked up by a robotic arm. They are then carried to the visual field of a thermal camera responsible for finding signs of overheating or degradation. After the analysis is terminated, the fuses are placed on a conveyor belt and sorted by a robotic bar. Finally, the fuses are moved by the large conveyor belt to a small conveyor belt that transports them back to the feeder, where the fuses are stored before restarting the cycle.

A fault diagnosis system is necessary to ensure a safe operation of the fuse test bench. Faults in the test bench happen when there is an alteration of the behavior of one or more components. Typical abnormalities include changes in the speed of the conveyor belts and pressure leakage on the pneumatic system. The fault categories in the test bench are generally known in practice. Therefore, it is essential to have a system to detect and classify faults accordingly. Subsequently, to help the operator understand the operation status, the system should be able to identify the abnormal behavior, i.e., the root cause identification. Furthermore, the system should be able to achieve diagnosis tasks as soon as possible to avoid further problems arising due to the abnormalities.

Manuscript received 26 November 2021; revised 3 February 2022; accepted 2 April 2022. Date of publication 6 April 2022; date of current version 21 July 2023. This article was recommended for publication by Associate Editor H. Luo upon evaluation of the reviewers' comments. *(Corresponding author: Muhammad Gibran Alfarizi.)*

The authors are with the Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: muhammad.g.alfarizi@ntnu.no; jorn.vatn@ntnu.no; shen.yin@ntnu.no).

Digital Object Identifier 10.1109/TAI.2022.3165137

The fault diagnosis methods can be roughly classified as either physics model-based or data-driven ones [1]–[3]. The model-based approach requires the knowledge of the first principles of the item of interest, e.g., the system mechanisms, structural integrity, and material properties [4], [5]. This method is highly accurate on the component level but may not be as accurate on the system level due to complex interactions intrasystem that are not easy to model with first principles [6]. The data-driven methods are trying to find hidden patterns and knowledge from empirical data without knowing the physics beforehand [7]–[10]. With the ever-increasing data and computing power in the industry, more examples of data-driven methods applications are found in the big data era [11]–[15].

Many sensors equipped with the fuse test bench provide a huge amount of data. When dealing with such an amount of data, engineers usually resort to data-driven analysis techniques. An example of such popular techniques is deep learning approaches. In the era of big data, deep learning algorithms have attracted intense interest in fault detection and diagnosis because of their enormous representing power and performance in solving complex problems [16]. The recurrent neural network (RNN) and autoencoder (AE) are examples of the popular deep learning algorithms used for fault detection. Dinaki et al. [17] used an LSTM network (a variant of RNN) to localize the network faults using only the video quality of experience metrics from the video service providers. Dewangan and Maurya [18] employed the AE to detect anomalies in rotating machinery. The deep learning approaches are also powerful tools for fault diagnosis. CNN can be implemented for fault diagnosis in a semiconductor manufacturing [19], unmanned aerial vehicle [20], and cyber-physical systems [21]. Restricted Boltzmann machine (RBM) and its variants are especially useful for fault diagnosis in structured data, some of which are shown in [22]–[25], and [26].

However, the aforementioned deep learning approaches are typically separately applied to fault detection, fault classification, or root cause analysis. A combination of several methods is commonly used in order to achieve detection, classification, and root cause analysis simultaneously. For example, Chen et al. [28] applied single shot multibox detector [29] and a fast single-shot detection method [30] to localize defects and then applied deep convolutional neural network to classify defects in fasteners. This is typically referred to as separate learning, in which independent methods are used to feature learning and pattern recognition. In industrial practice, a preferred diagnosis system would be end-to-end learning, which can take raw signals and directly generate the desired outcome, all in one integrated system.

There are many benefits of using end-to-end learning. First, instead of being constrained to human prejudices, it can directly capture anything implied in the data. Moreover, accuracy can be improved by optimizing the parameters of the whole system. Compared to separate learning, the global optimum might not be reached because optimization happens individually at each stage. Last but not least, the system is generic, which means it is adaptable or transferable to different but similar problems.

Some of the deep learning algorithms can be implemented as end-to-end learning, e.g., CNN. CNN can automatically extract features, learn, and classify them accurately, which are helpful for fault diagnosis systems in the industry. These learned features, however, are part of a larger "black box." This will worsen the interpretability of the features. Therefore, it does not help find the root cause of the faults, which is one of the objectives in this case study. The conflict between interpretability and accuracy is an important issue in machine learning. High-accuracy results are typically achieved by deep learning, but their decision-making processes are not interpretable, as the algebraic complexity of the functions tends to lose meaning with respect to the original set of feature variables. On the other hand, a highly interpretable algorithm, such as linear regression, performs unsatisfactorily when faced with real-world data.

Random forest is an algorithm that can defy the interpretability-accuracy tradeoff, or at least push it to its limit. Random forest can also be implemented as end-to-end learning because of its feature importance analysis. The feature importance analysis is performed directly in the algorithm training, which is essential to identify which signals represent anomalous behavior. A similar technique to random forest but with improved accuracy is gradient tree boosting [31]. Gradient tree boosting combines weak learners sequentially so that each new tree corrects the errors of the previous one. This is especially helpful in this study due to the small amount of faulty data, making the fault classifier the weak learner. Tree boosting has been shown to produce state-of-the-art results on a variety of standard classification benchmarks [32]. Extreme gradient boosting [33] (XGBoost) is one of the tree boosting techniques that has been frequently used in many machine learning and data mining challenges. The success of XGBoost is due to its scalability in various scenarios and faster running times [33]. XGBoost is highly accurate and can be directly implemented to detect and classify anomalies, while also identifying signals representing an abnormal operation, all in one integrated system and fast running time.

To this end, the practical demands of fuse test bench can be summarized into three main points as follows.

1) An integrated fault diagnosis system to detect, classify, and identify the root causes of faults in a fuse test bench.
2) The fault diagnosis system is expected to have high classification accuracy, fast diagnosis time, and interpretable root cause analysis.
3) Performance comparison of the designed fault diagnosis system with other standard approaches in practice.

An extreme gradient boosting aided fault diagnosis approach is then proposed to solve these practical demands. PHM Challenge 2021 dataset [27] is used for comparison analysis.

The rest of this article is organized as follows. Section II presents the preliminaries of our approach. Section III describes the data and the proposed fault diagnosis system approaches. Section IV discusses the performance of the proposed system. Finally, Section V concludes this article.

II. Preliminaries

A. XGBoost Algorithm

XGBoost is an effective ensemble learning model. It employs a boosting strategy to generate decision trees repeatedly [33]. The newly generated decision tree can correct the previous decision tree's prediction residuals. Finally, multiple decision

trees are combined to make the final prediction, which greatly improves accuracy when compared to a single decision tree.

Consider a dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}(i = 1...n, \mathbf{x}_i \in \mathbb{R}^m, y_i \in \mathbb{R})$, where m represent features for each of n observation examples which correspond to the target variable y. A tree ensemble model for a given observation i is calculated by sum of predictions from K additive functions

$$\hat{y}_i = \phi(\mathbf{x}_i) = \sum_{k=1}^{K} f_k(\mathbf{x}_i), \quad f_k \in \mathcal{F} \tag{1}$$

where $\mathcal{F} = \{f(\mathbf{x}) = w_{q(x)}\}(q : \mathbb{R}^m \to T, w \in \mathbb{R}^T)$ is the space of regression trees; q denotes the structure of each tree that maps an example to the corresponding leaf index; T represents the number of leaves in the tree; w is the leaf weights; and f_k is a regression tree predicting the value of $f_k(\mathbf{x}_i)$ for the ith example. The objective function with loss terms (l) and regularization terms Ω in (2) is minimized in the training process

$$\mathcal{L}(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k)$$

$$\text{where } \Omega(f) = \gamma T + \frac{1}{2}\lambda ||w||^2. \tag{2}$$

The γ and λ are the hyperparameters to penalize the model complexity. The loss term l is a function that measures the difference between the prediction $\hat{y}_i$ and the target y_i, e.g., the cross-entropy loss for classification problems.

The minimization of objective function (2) is done iteratively by adding a regression tree at each iteration. The objective function at tth iteration for ith instance becomes

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y_i}^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t). \tag{3}$$

By removing the terms independent of f_t and applying a second-order Taylor expansion, the objective function in (3) becomes

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^{n} \left[g_i f_t(\mathbf{x_i}) + \mathbf{\frac{1}{2} h_i f_t^2(x_i)} \right] + \Omega(f_t) \tag{4}$$

where g_i and h_i are the first and second-order derivatives of $l(y_i, \hat{y_i}^{(t-1)})$ with respect to $\hat{y_i}^{(t-1)}$. Define I_j as the instance set of leaf j. The optimal leaf weights w_j^* and the corresponding optimal value of $\tilde{\mathcal{L}}^{(t)}$ are then computed by the following equations:

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{5}$$

$$\tilde{\mathcal{L}}^{(t)}(q) = -\frac{1}{2} \sum_{j=1}^{T} \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \tag{6}$$

Equation (6) is used as the evaluation criteria to find an optimal tree split. A greedy algorithm starts splitting from a single leaf and adds branches according to (6) to avoid enumerating all possible tree structure q. To better understand the splitting procedure, consider I_L and I_R are the instance sets of left and right nodes after the split. Let $I = I_L \cup I_R$, the loss reduction caused by the split is then calculated by

$$\mathcal{L}_{split} = \frac{1}{2} \left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma. \tag{7}$$

The configuration which gives the maximum value of the loss reduction in (7) is considered the best split. The leaf values are then assigned by (5). The pseudocode of XGBoost is shown in Algorithm 1.

Algorithm 1: XGBoost.

Input:
- $\mathbf{x}$: The training set
- y: The label
- l: The loss function
- f: The base model

Steps:
- Initialize: $F_0(\mathbf{x}) = 0$
- **for** $k = 1$ **to** K **do**
  - **for** $i = 1$ **to** N **do**
    - $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y_i}^{(t-1)})$
    - $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y_i}^{(t-1)})$
  - **end for**
  - use g_i, h_i to compute objective function $\tilde{\mathcal{L}}^{(t)}$ in (4)
  - greedily grow a tree $f_k(\mathbf{x})$ in (1)
  - $F_k(\mathbf{x}) = F_{k-1}(\mathbf{x}) + \epsilon f_k(\mathbf{x})$
- **end for**

Output: $F(\mathbf{x}) = F_k(\mathbf{x})$

B. Health Indicator

XGBoost can be applied to classification problems. In case of a fault detection task, which is a binary classification problem, the loss terms (l) in (2) becomes

$$l(y_i, \hat{y}_i) = -y_i \log(\hat{y}_i) - (1 - y_i) \log(\hat{y}_i). \tag{8}$$

For a fault classification task, i.e., multiclass classification problem, the loss terms (l) in (2) becomes

$$l(y_i, \hat{y}_i) = -\sum_{c=1}^{M} y_{i,c} \log(\hat{y}_{i,c}) \tag{9}$$

where M is the number of classes. The best split at the end, which gives minimum error $\mathcal{L}^{(t)}$, will determine the classification for a given observation i.

III. Experiment Setup and Methodology

A. Data Description

The fuse test bench dataset is provided by the Swiss Centre for Electronics and Microtechnology (CSEM) [27]. Data have been acquired with different fault modes during the operation by an automated data acquisition system. The fault modes consist of fault-free instances and a variety of seeded faults under

TABLE I
DATASETS DESCRIPTION

Label	0	2	3	4	5	7	9	11	12
Quantity	70	4	4	3	4	4	4	3	3

TABLE II
STATISTICAL FEATURES DESCRIPTION

Name	**Description**
vCnt	Number of samples recorded in the time window
vFreq	Sampling frequency
vMax	Maximum value of the samples
vMin	Minimum value of the samples
vStd	Standard deviation of the samples
vTrend	Derivative-based trend of the samples
value	Mean value of the samples

TABLE III
HYPERPARAMETERS CONFIGURATION FOR XGBOOST MODEL

Hyperparameter	**Value**
max_depth	3
learning_rate	0.1
n_estimators	100
objective	multi:softmax or binary:logistic
booster	gbtree
n_jobs	1
gamma	0
min_child_weight	1
max_delta_step	0
subsample	1
colsample_bytree	1
colsample_bylevel	1
reg_alpha	1
reg_lambda	1
scale_pos_weight	2.3
base_score	0.5
random_state	18

controlled conditions. Each mode influences one or more sensor signals.

There are 99 datasets (experiments) provided for model training and validation, consisting fault-free instances with class label 0 and faulty instances with eight different class labels. Table I provides an overview of the datasets.

Every dataset is composed of a set of 50 signals, describing the evolution over time of a quantity of interest, and runs between one hour and three hours. These signals can be mainly categorized into health monitoring signals, such as *pressure* and *vacuum*, environmental monitoring signals, such as *humidity* and *temperature*, and general process monitoring signals, such as *CPUTemperature* and *ProcessMemoryConsumption*.

For every signal, up to seven statistical features are calculated over a time window of 10 s. Only these statistical features are considered for further analysis, instead of the raw signals. Table II shows the description of every statistical features derived from the sensor signals.

B. Data Preprocessing

1) Data Merging and Labeling: The data collected in every experiment are summarized in a single dataset. Thus, to analyze the whole experiment, the individual datasets for training and validation are merged (99 experiments). Additionally, every data point in the datasets is labeled to their respective classes. This allows for fault classification using only one time window. However, the strategy of choosing the optimal time window will be explained in a later subsection.

The merged dataset contains reading from sensors. The particular set of fields of every sensor are considered as features.

2) Missing Value Imputation: There are two linearly dependent statistical features in the dataset, which are *vCnt* and *vFreq*, as described in Table II. Thus, the feature *vCnt* is chosen, and *vFreq* is excluded from the analysis. In addition, when *vCnt* equals 0, all other features are missing, which makes sense since there is no recorded data in that time window. Therefore, all missing value is replaced by 0 before inputted into training.

3) Feature Selection: The first step in the feature selection is to remove all features whose value is only 0. Many features are removed with this rule, such as *ErrorFrame* and *FeederAction*. Furthermore, features with zero-variance are removed, i.e., features that have the same value in all samples. These steps are carried out to remove features that have minimum information and simplify model interpretation.

4) Data Splitting: The merged datasets are divided into training sets and test sets. Training sets consist of 80% of the datasets, and the remaining 20% is allocated into test sets. The evaluation is done in the test datasets.

C. Hyperparameters Configuration

The hyperparameters configuration used to build the XGBoost model are summarized in Table III.

D. Time to Classification

The optimal time to classification for each experiment should be determined to do the classification in the shortest time possible. First, classification is done using all the data points in that experiment, and the class is determined by choosing the most predicted class. Then, initialize the time to classification (T_c) to 1. If by using $T_c = 1$ yield the same classification as the one that uses all the data point, then the T_c for that experiment is 1. Otherwise, add the T_c incrementally by one until the classification is the same as the one that uses all the data points. The classification for every T_c is also done by using the most predicted class in that time window.

E. Evaluation Metrics

F1-score is used to evaluate the system performance because the dataset is imbalanced. The F1-score is suitable for imbalanced dataset because it takes into account the precision and recall of the system, and it is defined as the harmonic mean of the system's precision and recall. The calculation of the F1-score is given in as

$$F1 = 2 \times \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{10}$$

TABLE IV
COMPARISON WITH STANDARD APPROACHES FOR DETECTION AND CLASSIFICATION TASKS USING THE FUSE TEST BENCH DATASET

Method	Detection		Classification	
	F1-score	Accuracy	F1-score	Accuracy
XGBoost	**0.9745**	**0.9764**	**0.9693**	**0.9792**
RF	0.9234	0.9302	0.7955	0.8958
MLP	0.7118	0.7274	0.7717	0.8684
Logistic Regression	0.6192	0.6942	0.6360	0.8196
LDA	0.6912	0.7396	0.6081	0.7911

where TP, FP, and FN are the number of true positives, false positives, and false negative, respectively.

To give additional insight, accuracy of the system is also calculated by using the formula stated as

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}. \tag{11}$$

F. Root Cause Analysis

The root cause analysis for each class is done by using the feature importance analysis. For a single decision tree, importance is calculated by the amount that each feature split point improves the Gini index, weighted by the number of observations for which the node is responsible. The feature importance is then averaged across all of the system's decision trees. In general, importance provides a score indicating how useful or important each feature was in the creation of the system's boosted decision trees. The more a feature is used to make significant decisions with decision trees, the more important it is. This importance is explicitly calculated for each feature in the dataset, allowing them to be ranked and compared to one another.

A specific model for each class is built to determine the most important signals in every class. For instance, to determine the most important signals for class 2, only the data in faulty experiment class 2 and healthy data in class 0 were used. As a result, the model is now doing binary classification, and the feature importance from this model is used to identify the root cause for class 2. A similar procedure is also done for other classes.

To visualize the steps involved in our approach, Fig. 2 shows the flowchart of the proposed method.

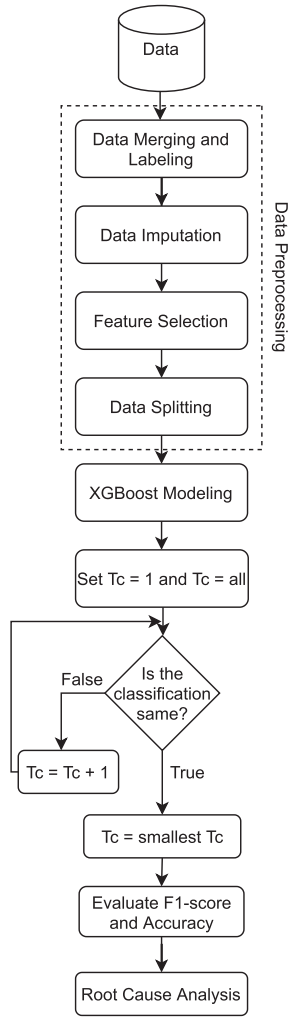

Fig. 2. Flowchart of the proposed method.

IV. EXPERIMENT RESULTS AND ANALYSIS

A. Fault Detection and Classification

In the fault detection problem, all of the faulty classes are labeled as 1. This becomes a binary classification problem. Fig. 3 shows the confusion matrix for fault detection system in test dataset. The values in every row are normalized by the total number of experiments per row (true label).

The detection system yields high accuracy in class 0 (0.99) but performed slightly worse for class 1 (0.95). This might happen due to the higher amount of data available for class 0 than class 1. F1-score and accuracy of the detection system can be seen from Table IV, which is 0.9745 for F1-score and 0.9764 for accuracy. These results prove that the system can detect fault accurately. In the next step, the system will classify the faults according to their classes. Fig. 4 shows the confusion matrix for the fault classification system in the test dataset.

The classification system yields high accuracy in classes 0, 9, 11, and 12, but the accuracy decreased slightly in classes 2, 3, 4, 5, and 7. When the system misclassifies classes 2, 3, 4, 5, and 7, the majority of them are classified into class 0. Again, this might happened because of the class imbalance between class 0 and other classes. However, this happens rarely, considering the accuracy in those classes is still high. The F1-score and accuracy of the classification system are 0.9792 and 0.9693, respectively. A high F1 score indicates high precision and recall, meaning that the system has a low false-positive and false-negative rate.

The proposed system's capability to detect and classify faults is proven satisfactory when validated in the fuse test bench datasets. However, to corroborate the system's performance, a comparison with other algorithms is made, which include

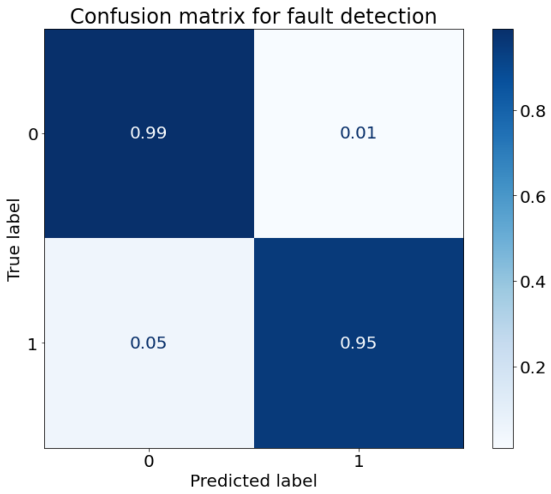

Fig. 3. Confusion matrix for fault detection system in test dataset with normalized value.

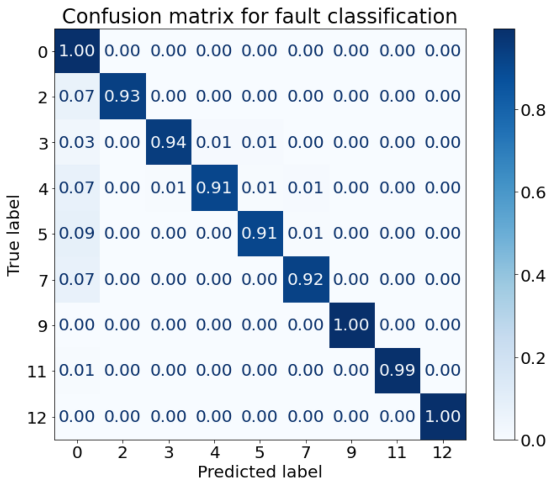

Fig. 4. Confusion matrix for fault classification system in test dataset with normalized value.

linear discriminant analysis (LDA) [34], logistic regression, multilayer perceptron (MLP), and random forest (RF) [35]. Table IV presents the F1-score and accuracy of each system for detecting and classifying faults in the fuse test bench datasets.

The proposed system outperforms all other methods to detect and classify faults in terms of F1-score and accuracy on the test dataset. The RF method has the closest performance with the proposed system, due to their similarity in the algorithm. However, RF suffers in classification F1-score despite having high accuracy. Because the data consisted of mostly the healthy data, RF got high accuracy by classifying most of the data into healthy data and got a lower F1-score by misclassifying the faulty data into the healthy data. Nevertheless, the RF method still performs better than MLP, logistic regression, and LDA. MLP may fall off in F1-score and accuracy due to overfitting the training data (F1-score of 0.9055 and accuracy of 0.9429 on the training data, compared to F1-score of 0.7717, and accuracy of 0.8684 on the test data). Logistic regression achieved unsatisfactory results due to its linear decision boundaries assumption in this not-linearly separable data. LDA performs worst due to its normal distribution assumption on features, which usually is not the case for many of the real-world data.

The proposed system performs superiorly in these imbalanced datasets due to its capability in tuning each class' weight. A higher weight can be assigned to the less represented class, i.e., the faulty classes, and a lower weight to the more represented class, i.e., the healthy class. This will penalize misclassification made in the minority class more than in the majority class. Additionally, the proposed approach has in-built L1 (Lasso regression) and L2 (Ridge regression) regularization, which prevents the system from overfitting. Moreover, it utilizes the power of parallel and distributed computing which enables quicker execution.

B. Time to Classification

As explained in Section III-D, the optimal time to classification for each experiment was determined. Table V summarizes the average time to classification of all classes using the proposed system and several standard fault diagnostic approaches in test dataset.

In classes 2, 4, 5, and 7, the proposed system is able to classify the faults by only using one time window, with a length of 10 s. These results are desired for a fault diagnosis system because it can diagnose faults as soon as it happens. Similar results can be seen in class 0, but one experiment needs seven time windows for classification, bringing down the average time to classification to 1.17. In classes 9, 11, and 12, a sufficient amount of time is needed for the fault classification. However, in class 3, a considerably long time is required to make a judgment. The system requires 131 time windows in one of the experiments in class 3. On the one hand, this is an undesired result for a fault diagnosis system, as the fault remains undetected for a long time. On the other hand, the system's accuracy is very high in classifying faults. In class 3, one might find the tradeoff between the system's accuracy and time to classification. Another solution is to gather more training data in class 3, so the system can classify faster without sacrificing accuracy.

Table V also provides the average time to classification for any class for every method. The average time is computed by weighted average with class 0 having the most weight because it has the most amount of experiments. The proposed system has the shortest average time to classification compared to other methods. Again, the RF method performs closest to the proposed system due to their similarity in the algorithm. Logistic regression also able to classify the faults relatively quickly. However, given the low accuracy results from logistic regression, this method may be undesired for fault diagnosis system. MLP and LDA need a noticeably long time to classify compared to others. This is because the whole duration of the experiment is required

TABLE V
TIME TO CLASSIFICATION FOR ALL METHODS

Method	Average T_c in Class									Average T_c
	0	**2**	**3**	**4**	**5**	**7**	**9**	**11**	**12**	**Overall**
XGBoost	**1.17**	**1**	82.5	**1**	**1**	**1**	4	5	**3**	**4.77**
RF	1.17	96	57	1	1	1	**3**	**3**	3	7.58
MLP	20.76	360	360	360	6.5	11.5	31.5	360	360	76.77
Logistic Regression	9.53	1	**3.5**	7.5	7.5	5.5	5.5	8	4	8.27
LDA	5	62	360	3.5	6	7	5	360	13	29.79

TABLE VI
MOST IMPORTANT SIGNALS FOR EVERY CLASS

Class	Most Important Signals
2	NumberFuseEstimated NumberFuseDetected FuseCycleDuration
3	SharpnessImage FeederBackgroundIlluminationIntensity
4	Temperature Pressure
5	VacuumValveClosed Vacuum
7	FusePicked Vacuum
9	SmartMotorSpeed SmartMotorPositionError
11	SmartMotorSpeed SmartMotorPositionError
12	DurationRobotFromFeederToTestBench DurationRobotFromTestBenchToFeeder

TABLE VII
FAULT INTERPRETATION OF EVERY CLASS

Class	Fault Interpretation
2	Fewer fuses are estimated and detected compared to normal operation.
3	Less image sharpness and background illumination LED may suggest the feeder is dirty.
4	Change in environmental conditions and pressure used for the feeder barrier.
5	Defect of the vacuum gripper.
7	Fewer fuses are picked may suggest the vacuum gripper or the robotic arm is defective.
9	Fluctuation in speed of the big conveyor belt is observed.
11	Slowing down in speed of the big conveyor belt is observed.
12	The time needed by the robot to move from the test bench to the feeder and vice versa is increasing, indicating a deceleration in robot speed.

for some of the evaluations. This implies that those two methods cannot find the unique characteristics in some of the classes from small sets of data. Regardless, the accuracy of MLP and LDA is still inadequate despite of the long evaluation time. The proposed system is proven to offer the best of both worlds, a high accuracy and short evaluation time.

C. Root Cause Analysis

The root cause of a fault is determined by the most important signals in the feature importance analysis. The most important signals for each class are reviewed in Table VI.

Deeper investigations into the data, specifically in the most critical signals, are carried out to understand better the causes of the faults and their physical interpretation. For example, a constant value in *SmartMotorSpeed.value* is observed in healthy data, implying a constant speed in the big conveyor belt. However, fluctuations in *SmartMotorSpeed.value* are observed in class 9 data, indicating an abnormal behavior compared to healthy data. This suggests that the fault happened in the big conveyor belt. Therefore, the proper maintenance action would fix the conveyor belt. A similar problem is also observed in class 11. In classes 2, 5, 7, and 12, fewer fuses are estimated and detected compared to normal operation. This may suggest defects in the robotic arm or the vacuum gripper. The right maintenance actions should fix the robotic arm and the vacuum gripper. In class 3, less image sharpness and background illumination LED are observed. Thus, this may suggest the feeder is dirty, and the appropriate maintenance response is to clean the feeder. In class 4, there are changes in environmental conditions and pressure used for the feeder barrier. Therefore, a suitable maintenance plan would change the environmental condition and pressure back to normal. The faults interpretation of all classes is summarized in Table VII.

V. CONCLUSION

This article proposes an integrated fault diagnosis system based on extreme gradient boosting for a fuse test bench line to achieve superior performance in fault detection, fault classification, and root cause analysis. The proposed fault diagnosis system has high classification accuracy, fast diagnosis time, and interpretable root cause analysis when verified by the real industrial data from PHM Challenge 2021. The proposed system also outperforms several standard fault diagnostic approaches in detection and classification accuracy and requires a shorter diagnosis time.

REFERENCES

[1] S. X. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. London, U.K.: Springer, 2008.
[2] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6418–6428, Nov. 2014.

[3] S. Yin, X. Li, H. Gao, and O. Kaynak, "Data-based techniques focused on modern industry: An overview," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 657–667, Jan. 2015.
[4] S.-P. Zhu, H.-Z. Huang, W. Peng, H.-K. Wang, and S. Mahadevan, "Probabilistic physics of failure-based framework for fatigue life prediction of aircraft gas turbine discs under uncertainty," *Rel. Eng. Syst. Saf.*, vol. 146, pp. 1–12, 2016.
[5] M. Pecht and J. Gu, "Physics-of-failure-based prognostics for electronic products," *Trans. Inst. Meas. Control*, vol. 31, no. 3/4, pp. 309–322, 2009.
[6] H. Li, H.-Z. Huang, Y.-F. Li, J. Zhou, and J. Mi, "Physics of failure-based reliability prediction of turbine blades using multi-source information fusion," *Appl. Soft Comput.*, vol. 72, pp. 624–635, 2018.
[7] S. X. Ding, *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*. Berlin, Germany: Springer, 2014.
[8] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques-part II: Fault diagnosis with knowledge-based and hybrid/active approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3768–3774, Jun. 2015.
[9] Z. Hosseinpoor, M. M. Arefi, R. Razavi-Far, N. Mozafari, and S. Hazbavi, "Virtual sensors for fault diagnosis: A case of induction motor broken rotor bar," *IEEE Sensors J.*, vol. 21, no. 4, pp. 5044–5051, Feb. 2021.
[10] N. Gharesi, M. M. Arefi, R. Razavi-Far, J. Zarei, and S. Yin, "A neuro-wavelet based approach for diagnosing bearing defects," *Adv. Eng. Inform.*, vol. 46, 2020, Art. no. 101172.
[11] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2226–2238, Nov. 2013.
[12] L. Zhang, J. Lin, and R. Karim, "An angle-based subspace anomaly detection approach to high-dimensional data: With an application to industrial fault detection," *Rel. Eng. System Saf.*, vol. 142, pp. 482–497, 2015.
[13] J. Yang, Y. Guo, and W. Zhao, "Long short-term memory neural network based fault detection and isolation for electro-mechanical actuators," *Neurocomputing*, vol. 360, pp. 85–96, 2019.
[14] H. Chen, Z. Chai, O. Dogru, B. Jiang, and B. Huang, "Data-driven designs of fault detection systems via neural network-aided learning," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2021.3071292.
[15] X. Li, Y. Jiang, C. Liu, S. Liu, H. Luo, and S. Yin, "Playing against deep neural network-based object detectors: A novel bidirectional adversarial attack approach," *IEEE Trans. Artif. Intell.*, vol. 3, no. 1, pp. 20–28, Feb. 2022.
[16] L. Zhang, J. Lin, B. Liu, Z. Zhang, X. Yan, and M. Wei, "A review on deep learning applications in prognostics and health management," *IEEE Access*, vol. 7, pp. 162415–162438, 2019.
[17] H. E. Dinaki, S. Shirmohammadi, E. Janulewicz, and D. Côté, "Deep learning-based fault localization in video networks using only client-side QoE," *IEEE Trans. Artif. Intell.*, vol. 1, no. 2, pp. 130–138, Oct. 2020.
[18] G. Dewangan and S. Maurya, "Fault diagnosis of machines using deep convolutional beta-variational autoencoder," *IEEE Trans. Artif. Intell.*, vol. 3, no. 2, pp. 287–296, Apr. 2021.
[19] K. B. Lee, S. Cheon, and C. O. Kim, "A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes," *IEEE Trans. Semicond. Manuf.*, vol. 30, no. 2, pp. 135–142, May 2017.
[20] D. Guo, M. Zhong, H. Ji, Y. Liu, and R. Yang, "A hybrid feature model and deep learning based fault diagnosis for unmanned aerial vehicle sensors," *Neurocomputing*, vol. 319, pp. 155–163, 2018.
[21] Z. Wu, Y. Guo, W. Lin, S. Yu, and Y. Ji, "A weighted deep representation learning model for imbalanced fault diagnosis in cyber-physical systems," *Sensors*, vol. 18, no. 4, 2018, Art. no. 1096.
[22] Y. Guo *et al.*, "Deep learning-based fault diagnosis of variable refrigerant flow air-conditioning system for building energy saving," *Appl. Energy*, vol. 225, pp. 732–745, 2018.
[23] Y. Zhang *et al.*, "A cable fault recognition method based on a deep belief network," *Comput. Elect. Eng.*, vol. 71, pp. 452–464, 2018.
[24] D. Yu, Z. Chen, K. Xiahou, M. Li, T. Ji, and Q. Wu, "A radically data-driven method for fault detection and diagnosis in wind turbines," *Int. J. Elect. Power Energy Syst.*, vol. 99, pp. 577–584, 2018.
[25] J. Dai, H. Song, G. Sheng, and X. Jiang, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Trans. Dielectrics Elect. Insul.*, vol. 24, no. 5, pp. 2828–2835, Oct. 2017.
[26] J. Yin and W. Zhao, "Fault diagnosis network design for vehicle on-board equipments of high-speed railway: A deep learning approach," *Eng. Appl. Artif. Intell.*, vol. 56, pp. 250–259, 2016.
[27] L. Biggio, M. Russi, S. Bigdeli, I. Kastanis, D. Giordano, and D. Gagar, "PHME data challenge," in *Proc. Eur. Conf. Prognostics Health Manage. Soc.*, 2021. [Online]. Available: https://github.com/PHME-Datachallenge/Data-Challenge-2021#citation
[28] J. Chen, Z. Liu, H. Wang, A. Núñez, and Z. Han, "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 257–269, Feb. 2018.
[29] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.
[30] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, 2016, pp. 779–788.
[31] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.
[32] G. Ke *et al.*, "Lightgbm: A highly efficient gradient boosting decision tree," *Adv. Neural Inf. Process. Syst.*, vol. 30, pp. 3146–3154, 2017.
[33] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 785–794.
[34] A. J. Izenman, "Linear discriminant analysis," in *Modern Multivariate Statistical Techniques*, New York, NY, USA: Springer, 2013, pp. 237–280.
[35] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

Muhammad Gibran Alfarizi (Member, IEEE) received the B.Sc. degree in petroleum engineering from the Bandung Institute of Technology, Bandung, Indonesia, in 2017, and the M.Sc. degree in petroleum engineering in 2020 from Norwegian University of Science and Technology, Trondheim, where he is currently working toward the Ph.D. degree in reliability, availability, maintainability, and safety.

Jørn Vatn received the M.Sc. degree in mathematical statistics and the Ph.D. degree in industrial mathematics with a thesis on maintenance optimization, from the Norwegian University of Science and Technology, Trondheim, Norway, in 1986 and 1996, respectively.

He has more than 35 years of experience as a Researcher with SINTEF Safety and Reliability. He is currently a Professor with Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway.

Shen Yin (Senior Member, IEEE) received the M.Sc. degree in control and information system and the Ph.D. (Dr.-Ing.) degree in electrical engineering and information technology from the University of Duisburg-Essen, Duisburg, Germany, in 2007 and 2012, respectively.

He is currently a Full Professor (DNV-GL Professor) with the Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway. His research interests include fault diagnosis/prognosis and fault-tolerance strategy, safety, reliability, and security of complicated systems, system and control theory, data-driven and machine learning approaches, and applications in health diagnosis and cyber-physical systems.

Article 3

Towards Accident Prevention on Liquid Hydrogen: A Data-driven Approach for Releases Prediction

M. G. Alfarizi, F. Ustolin, J. Vatn, S. Yin, and N. Paltrinieri, "Towards accident prevention on liquid hydrogen: A data-driven approach for releases prediction," *Reliability Engineering & System Safety*, vol. 236, p. 109276, 2023.

Reliability Engineering and System Safety 236 (2023) 109276

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Towards accident prevention on liquid hydrogen: A data-driven approach for releases prediction

Muhammad Gibran Alfarizi *, Federico Ustolin, Jørn Vatn, Shen Yin, Nicola Paltrinieri

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Keywords:
Accident prevention
Hydrogen safety
Liquid hydrogen
Machine learning
Random forests
Risk analysis

ABSTRACT

Hydrogen is a clean substitute for hydrocarbon fuels in the marine sector. Liquid hydrogen (LH_2) can be used to move and store large amounts of hydrogen. This novel application needs further study to assess the potential risk and safety operation. A recent study of LH_2 large-scale release tests was conducted to replicate spills of LH_2 inside the ship's tank connection space and during bunkering operations. The tests were performed in a closed and outdoor facility. The LH_2 spills can lead to detonation, representing a safety concern. This study analyzed the aforementioned LH_2 experiments and proposed a novel application of the random forests algorithm to predict the oxygen phase change and to estimate whether the hydrogen concentration is above the lower flammability limit (LFL). The models show accurate predictions in different experimental conditions. The findings can be used to select reliable safety barriers and effective risk reduction measures in LH_2 spills.

1. Introduction

Liquid hydrogen (LH_2) is one of the best solutions for storing and transporting large amounts of hydrogen. LH_2 can be used in the maritime sector to store large amounts of hydrogen due to the limited space onboard ships [1]. However, LH_2 is a cryogenic gas that is more difficult to handle than conventional fuels (e.g., gasoline). Hydrogen is liquefied by decreasing its temperature below −253 °C, which corresponds to its normal boiling point [2], and is stored at atmospheric pressure in very well thermally insulated tanks. Different insulation types (e.g., perlite, multi-layer insulation (MLI) [3]) can be installed in the tank vacuum jacket. Additional details on the LH_2 tank's design for automotive applications can be found in [4–7].

Although hydrogen is potentially clean and renewable and has a high energy content, it is a hazardous substance since it is highly flammable and has low minimum ignition energy [8]. Additional hazards are added when it is liquefied since it becomes a cryogenic substance. Therefore, it must be handled appropriately, and loss of containment of LH_2 storage components must be avoided. Many consequences of an LH_2 release must still be fully understood. For instance, air components (nitrogen and oxygen) can condensate or solidify when in contact with LH_2 due to its extremely low-temperature [9,10]. Condensed phase explosions are "detonations" that can be provoked by the ignition of the mixture composed of LH_2 and liquid (LOX) or solid oxygen (SOX). It is critical to investigate and comprehend under which circumstances this phenomenon might manifest to prevent or mitigate it. Only in this fashion, the overall risk of LH_2 technologies and applications can be reduced.

Prediction of oxygen phase change and hydrogen concentration during LH_2 spills is vital to prevent unwanted consequences. Data-driven approaches are popular techniques for system prognosis, with some researchers employing deep learning methods for remaining useful life prediction [11–13]. Other researchers use the random forest algorithm as a key component in their prediction models, enhancing the accuracy and performance of ship collision severity prediction and lithium-ion battery capacity estimation [14,15]. The random forests (RFs) algorithm has demonstrated notable results in prediction and classification tasks [16]. RF is a type of machine learning that integrates numerous predictors (decision trees) into an ensemble [17]. These multiple predictors offer higher accuracy in predicting outcomes compared to a single decision tree and maintain low computational time by using a parallel ensemble to create submodels on each sample. Consequently, the RFs algorithm is well-suited for predicting potential changes in

Abbreviations: AUC-PR, Area Under the Precision–Recall Curve; CFD, Computational Fluid Dynamics; DDT, Deflagration to Detonation Transition; FN, False Negative; FP, False Positive; LFL, Lower Flammability Limit; LH_2, Liquid Hydrogen; LNG, Liquefied Natural Gas; LOX, Liquid Oxygen; MLI, Multi-Layer Insulation; OOB, Out-Of-Bag; PFPs, Passive Fire Protections; RFs, Random Forests; RPT, Rapid Phase Transition; SOX, Solid Oxygen; TCS, Tank Connection Space; TN, True Negative; TP, True Positive; VCE, Vapor Cloud Explosion.

* Corresponding author.
E-mail addresses: muhammad.g.alfarizi@ntnu.no (M.G. Alfarizi), federico.ustolin@ntnu.no (F. Ustolin), jorn.vatn@ntnu.no (J. Vatn), shen.yin@ntnu.no (S. Yin), nicola.paltrinieri@ntnu.no (N. Paltrinieri).

https://doi.org/10.1016/j.ress.2023.109276
Received 11 November 2022; Received in revised form 24 March 2023; Accepted 27 March 2023
Available online 31 March 2023

oxygen phase and hydrogen concentration during LH_2 spills to aid in accident prevention.

This study aims to shed light on LH_2 releases and condense phase explosions by utilizing machine learning algorithms, namely the random forests, to predict their consequences. This paper proposes a novel application of the random forests algorithm to predict the occurrence of oxygen phase change and to estimate whether the hydrogen concentration is above the lower flammability limit (LFL) in the event of LH_2 leakage. The proposed application was investigated using the LH_2 release tests carried out by DNV in 2019 and 2020 as part of a project led by the Norwegian Defence Research Establishment (FFI) [18]. To the authors' knowledge, the application of machine learning is one of the pioneering works to predict LH_2 release consequences. This study is a follow-up of the investigation performed by Ustolin et al. [19] where a linear model was adopted as part of a machine learning approach, namely a supervised learning method with the classification task. This paper's novelty involves employing a more accurate and robust methodology and analyzing all sequences of the LH_2 release test experiments. The models provide accurate and reliable predictions in different experimental conditions. The outcome of the proposed application can be exploited to select effective mitigation measures as part of a risk assessment and increase the overall safety of LH_2 technologies.

The following is a summary of this paper's main contributions.

1. A novel application of machine learning, specifically the random forests algorithm, is proposed to predict the development of condensation or solidification of oxygen in the air and to estimate whether the hydrogen concentration is higher than LFL in the event of LH_2 leakage to prevent disastrous accidents.
2. The proposed application is verified by real experiments of LH_2 release tests conducted by DNV as part of a project led by the Norwegian Defence Research Establishment [18] and showed accurate and reliable predictions in different experimental settings.
3. The mitigation and risk reduction measures of accidents caused by LH_2 spills are discussed.

The remainder of this paper is organized as follows. Section 2 describes the consequences of liquid hydrogen loss of containment. Section 3 presents the measures of this study. Section 4 describes the experimental setup and methodology used in this study. Section 5 discusses the experiment results and their analysis. Section 6 concludes this paper.

2. Liquid hydrogen loss of containment

Several causes may provoke a loss of containment of LH_2 storage or transfer components. Since hydrogen is a hazardous substance, its unintended release must be prevented. On the other hand, it is critical to understand the consequences that an LH_2 loss of containment may have in order to mitigate them. The main consequences of an LH_2 loss of containment are described in Section 2.1.

2.1. Liquid hydrogen release consequences

Usually, LH_2 flashes to gas when released due to different factors such as pressure difference between the LH_2 storage and the atmosphere, heat transfer between LH_2, air, and the ground where LH_2 is spilled. The flashing fraction depends on the released hydrogen conditions (temperature and pressure) and should increase when both pressure and temperature are higher. For instance, a high fraction can be expected when LH_2 is superheated [20]. If the pressure inside the LH_2 storage equipment is high enough, a high-momentum jet can be generated. This jet may be hazardous for personnel in case of contact. Frostbites can have lethal consequences on humans as well as asphyxiation in the event the jet creates a hydrogen atmosphere with absence or limited oxygen concentration [21]. Moreover, metal structures impinged by the cryogenic jet can become brittle due to the extremely low LH_2 temperature, resulting in component failure. It is still unclear how passive fire protections (PFPs) used to protect pipes and tanks from flames behave under cryogenic conditions. Therefore, cryogenic jets may severely affect humans and structures near the release.

Due to the flashing phenomenon, the hydrogen jet during the release is usually two-phase (liquid and gaseous). Rainout may occur depending on the release flow rate, distance from the ground, and orientation of the leakage (e.g., horizontal, vertical downward). Hence, an LH_2 pool might form on the released surface (e.g., ground, water). It was concluded by the experiments carried out by Aaneby et al. [18] that there is a higher probability of pool formation in case the release is oriented vertically downward. Condensation or even solidification of air components (nitrogen and oxygen) may occur since the LH_2 temperature at atmospheric pressure is way below the boiling and melting point of both nitrogen (−195.8 °C, −210.0 °C [2]) and oxygen (−182.9 °C, −218.8 °C [2]). However, nitrogen and oxygen can lose significant amounts of heat since their latent vaporization heat is 199.6 and 213.1 kJ/kg, respectively. At the same time, hydrogen requires 448.5 kJ/kg to change phase from liquid to vapor [22]. Moreover, the heat necessary to cool nitrogen and oxygen from ambient temperature (288 K) down to their boiling point at atmospheric pressure is 221.5 and 182.0 kJ/kg, respectively [2]. In addition, the water vapor in the air and the surface, where LH_2 is spilled (e.g., ground), cede heat to LH_2. This means that relatively large amounts of LH_2 are necessary to liquefy oxygen. Royle and Willoughby [23] observed during their LH_2 release experiments that a solid deposit appeared on top and close to the LH_2 pool formed on the ground. Although it was impossible to analyze this deposit, it has been hypothesized that it is composed of solidified air components.

Before describing the condensed phase explosion phenomenon, it is important to distinguish between two types of chemical explosions, namely deflagration and detonation. A deflagration is a type of explosion that generates subsonic blast waves, i.e. the flame front has a speed lower than the speed of sound, and it may occur even for low hydrogen concentration (4.0% vol–11% vol in air). On the other hand, detonation is a rapid combustion process that is characterized by supersonic combustion waves that travel through the flammable cloud at speeds of up to 2,000 m/s [24]. Usually, a hydrogen concentration between 11 and 59% vol in air is required to achieve a detonation [25]. This explosion can generate high temperatures and overpressures that can cause structural damage and injury to people in the area.

The cryogenic pool might be formed during an LH_2 release and is thus composed of liquid hydrogen, liquid or solid nitrogen, and oxygen, and can be enriched with oxygen because it condenses at a temperature higher than nitrogen. The enrichment occurs because the LOX density is three orders of magnitude higher than the gaseous nitrogen one. Gill et al. [26] experimentally demonstrated that the mixture of LH_2 and solid air created on top of the cryogenic pool may even transition from deflagration to detonation (DDT) in the presence of an ignition source and if the SOX concentration in the solid deposit is higher than 50%. Therefore, the combustion process is initiated in a condensed or solidified phase mixture, meaning that the hydrogen is in the liquid state and the oxygen can be either liquid or solid. For this reason, this type of phenomenon is often called a condensed phase explosion and can result in severe consequences. On the other hand, the mixture barely burns without exploding for a higher concentration of nitrogen in the solid compound. Moreover, a mixture of LH_2 and LOX or SOX is also sensitive to pressure variation. This means that an increase in pressure (e.g., blast wave) might have sufficient energy to provoke a detonation [27].

Finally, LH_2 continuously evaporates, generating a dense flammable cloud that can remain close to the ground and slowly rises and disperses depending on the presence of natural or forced ventilation. Hydrogen has a very low minimum ignition energy, especially at the

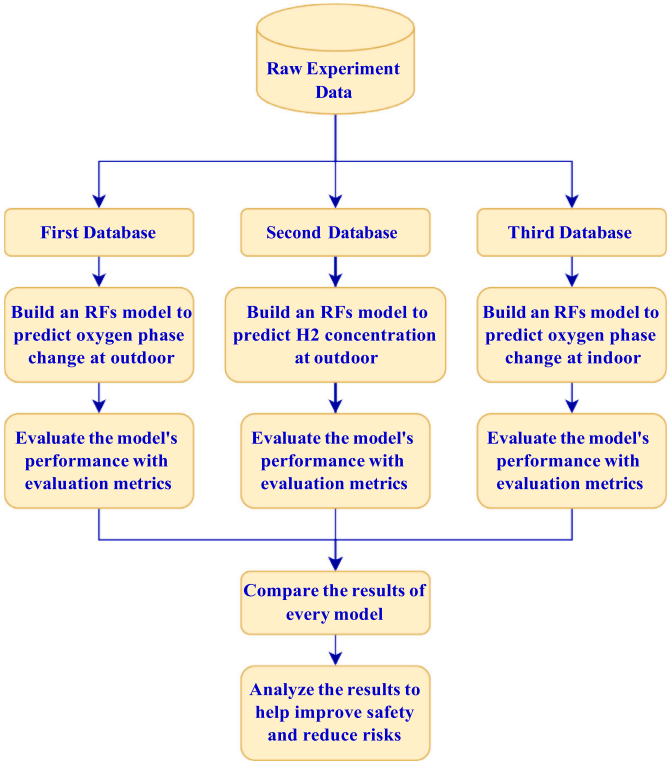

Fig. 1. The framework of the proposed application.

stoichiometric concentration (0.017 mJ [28]), and a wide flammability range (4.0%vol - 75.0%vol in air [29]). Therefore, the flammable cloud formed from the LH_2 pool or the cryogenic jet would easily be ignitable in the presence of an ignition source. If the cryogenic two-phase jet is ignited, a jet fire can develop. Similarly, a pool fire results from the ignition of the gaseous hydrogen evaporated on top of the LH_2 pool. Instead, flash fire or a vapor cloud explosion (VCE) are two consequences of hydrogen-flammable cloud ignition. The main distinction between these two phenomena is the flame front velocity which is higher than 40 m/s for VCE and is thus able to generate explosive effects [30]. In densely congested areas, the cryogenic hydrogen cloud combustion may even undergo DDT [31]. Finally, to summarize, the main consequences of an LH_2 release from a hydrogen component are as follows:

- LH_2 or cryogenic hydrogen jet;
- material embrittlement impacted by LH_2 jet;
- pool formation;
- air component condensation or solidification;
- dispersion and flammable cloud formation,

and in case of ignition:

- fire;
- jet fire;
- pool fire;
- flash fire;
- vapor cloud explosion (VCE);
- deflagration to detonation transition (DDT);
- condensed phase explosion (detonation).

It must be mentioned that other types of consequences such as rapid phase transition (RPT) were investigated for LH_2 spills onto or into water [32–34]. RPT is a physical explosion that may occur when liquefied natural gas (LNG) or other cryogenic fluids are spilled onto the water. Few studies concluded that RPT is unlikely for LH_2. Nevertheless, a high momentum jet of LH_2 direct into the water can lead to a self-ignition of the flammable cloud developed during the spillage above the water surface as demonstrated by recent experimental studies [35].

Many other experiments where an LH_2 release was simulated were carried out in the past. To the authors' knowledge, the first time that LH_2 spill tests were carried out was at the end of the 1950s, as described by Zabetakis and Burgess [36]. Afterward, a few more tests were performed in 1981 by NASA at its White Sands Test Facility in New Mexico [37], in 1994 by Federal Institute for Materials Research and Testing (BAM) in collaboration with the Research Center Juelich in Drachhausen, Germany [38,39], in 2012 by Health and Safety Laboratory (HSL) at its facility in Buxton, UK [23,40], in 2019 and 2020 by DNV at Spadeadam facility in UK [18], and in 2021 by BAM at its facility in Horstwalde, Germany [35]. A more detailed description of these LH_2 release experiments can be found in [41]. Moreover, additional experiments were recently carried out with cryogenic (gaseous) hydrogen and LH_2 during the European project PRESLHY [42]. During these experimental campaigns, different parameters such as release rate, spill duration, nozzle diameter, and surface (e.g., gravel, sand, concrete, water) were considered to comprehend and investigate the consequences of the loss of containment for LH_2 technologies.

Different authors attempted to simulate most of the abovementioned phenomena through theoretical and numerical models. These models were, in most cases, validated against the outcomes of various experiments. Most of these studies adopted a computational fluid dynamics

(CFD) approach to assess the consequences of LH_2 releases. A list of these investigations can be found in [41]. The developed models can be employed for the consequence analysis as part of a risk assessment. Despite CFD tools being widely used in industries and the computational power of commercial machines exponentially increasing in the last decades, a few drawbacks, such as the need for highly qualified personnel, remain. Furthermore, a long computational time may still be a disadvantage when investigating large geometrical domains to achieve accurate results.

2.2. Prevention and mitigation measures

As previously described, the consequences of an LH_2 release might severely affect structure and humans. Prevention of accident scenarios is critical to reduce the probability of failure and, consequently, the overall risk of the hydrogen system. Standard prevention measures such as training and safety devices may be employed to prevent accidents. On the other hand, mitigation measures are also required in the undesired event of an LH_2 release. To increase the response of the mitigation system, automatic activation of safety systems should be adopted. Sensors must be installed close to the hydrogen components to detect an undesired LH_2 leakage. The detectors may measure parameters such as hydrogen concentration or temperature and automatically activate alarms or specific safety devices. If the hydrogen concentration is measured, the threshold to trigger the mitigation measures must be set at 1%vol in air, which is 25% of the hydrogen lower flammability limit (LFL). All the sensors and devices must perform well at temperatures close to the LH_2 one. Examples of safety devices to be activated are shutdown valves to stop the LH_2 flow (e.g., from pipes), sprinklers, water curtains, inert gases (e.g., nitrogen), and ventilation.

3. Measures

3.1. The framework of the proposed application

Fig. 1 shows the flowchart of the proposed application. First, three databases are built based on the data from experiments conducted by DNV [18]. Each database is built based on different experimental settings and for different prediction purposes. After that, an RFs model is created for each database and fed with the input data, which consists of many parameters to predict oxygen phase change or the possibility of hydrogen concentration above LFL. Then, every model's performance is evaluated with several metrics to assess its accuracy. In the end, the models' results are discussed and analyzed to help select reliable safety barriers and effective risk reduction measures. Every step will be detailed in Section 4.

3.2. Random forests

The foundation of the RFs algorithm is a set of predictors affected by the random values of each forest tree [15]. Input datasets are selected randomly from a group to model the RFs. The RFs algorithm's relatively high success rate can be attributed to its quick operating speed, effective handling of large datasets, and minimal overfitting of predictors [43]. The general structure of the RFs model is illustrated in Fig. 2.

The input for forest in RFs is a q-dimensional input vector of X, in which $X = x_1, x_2, \dots, x_q$. Within the forest, a collection of L trees $\{T_1(x), T_2(x), \dots, T_L(x)\}$ are formed. Output value of each tree is calculated, denoted as $\hat{Y}_1 = T_1(X), \dots, \hat{Y}_m = T_m(X)$, where $m = 1, \dots, L$. L is the number of trees. In a classification task, the RFs' output is determined by

$$\text{Predict}_{\text{RFs}}(X) = \text{majority vote}\{\hat{Y}_m(X)\}_{i=1}^{L}. \tag{1}$$

The inputs, x_i, $i = 1, \dots, n$, and outputs, y_i, $i = 1, \dots, n$, form the training dataset $T = T_1, T_2, \dots, T_n = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. The RF then uses the bootstrap resampling method to randomly generate L tree sample sets from the initial training set T. About two-thirds of the original training set T data are taken in bootstrap samples, referred to as in-bag data. In contrast, the remaining data are out-of-bag data (OOB).

At each node, the algorithm needs to select a suitable attribute using the appropriate measure, maximizing dissimilarity between classes. The Gini Index, which assesses an element's impurity in relation to the other classes, is frequently used by RFs to select the best split. Given a training dataset T, the mathematical formula of the Gini Index is

$$GI = \sum \sum_{j \neq i} (f(C_i, T)/|T|)(f(C_j, T)/|T|) \quad , \tag{2}$$

where C_i is the class a randomly selected sample belongs to, and $f(C_i, T)/|T|$ is the probability that a selected case belongs to class C_i. The class heterogeneity increases as the Gini Index increases, and vice versa. A successful split is indicated with the Gini Index of the child node being less than a parent node, meaning the class homogeneity increases in that class as the tree grows deeper. When the Gini Index is 0, which indicates that only one class is present at each terminal node, tree splitting is completed. Consequently, a decision tree is made to grow to its maximum depth with no pruning. When every L tree in the forest is grown, the new data can be predicted using the outcome of the predictions of L trees. To summarize, Algorithm 1 outlines the pseudocode of the RFs algorithm [17].

Algorithm 1 Random Forest

Require: Training samples $T = \{(x_1, y_1), \dots, (x_n, y_n)\}$, testing samples x_t
  for $m = 1$ **to** L **do**
    Using the training set T, create the bootstrap sample T_m at random with replacement
    From T_m, build a non-pruning decision tree $\hat{Y}_m$
    Randomly choose *n_try* features from N features
    Select the best feature to split from each node's *n_try* features
    Split the tree until it achieves its largest possible size
  end for
Ensure: A set of decision trees $\{\hat{Y}_m,\ m = 1, 2, \dots, L\}$. For the testing samples, the predictor $\hat{Y}_m(x_t)$ is produced by the decision tree $\hat{Y}_m$. The RFs' output is determined using the formula in (1)

Because RFs utilized multiple predictors in its structure, it can provide higher accuracy in predicting outcomes over a single decision tree. It also keeps the computational time low by using a parallel ensemble to create submodels on each sample. Additionally, it works very well with tabular data. The experiment data investigated in this study is tabular and has many data points. Therefore, the RFs algorithm is well-suited for application in this study because it has high accuracy, low computational time for large datasets, and works well with tabular data. The RFs algorithm will be used to predict potential changes in oxygen phase and hydrogen concentration for accident prevention. To the authors' knowledge, the application of RFs to predict LH_2 release consequences is one of the pioneering works in this field of study.

4. Experiment setup and methodology

4.1. Liquid hydrogen release experiments

The LH_2 release experiments carried out by the Norwegian Defence Research Establishment (FFI) at the DNV Spadeadam facility in 2019 and 2020 [18] were considered as a case study in this investigation. The tests aimed to simulate a realistic release accident scenario during bunkering operations or onboard LH_2 ships and comprehend the LH_2 behavior during these events. A total of 15 trials were executed by varying LH_2 flow rate, release duration and orientation (horizontal and vertical downward), and location: outdoor and inside a closed room. In the following, the outdoor and indoor tests are briefly described.

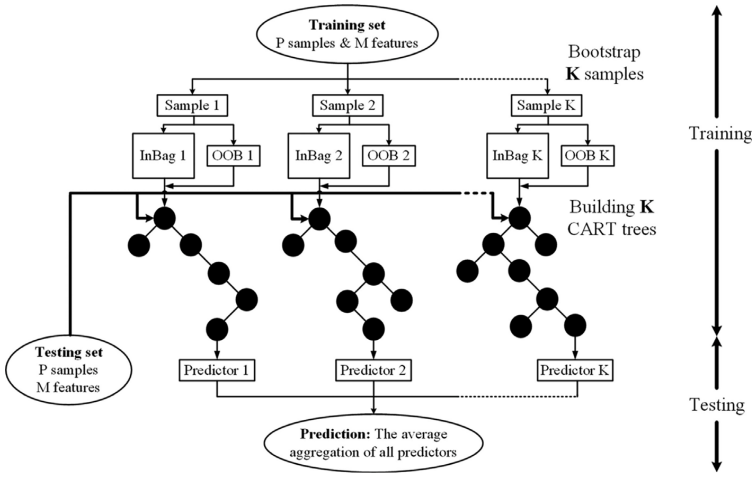

Fig. 2. The general structure of the RFs model [44].

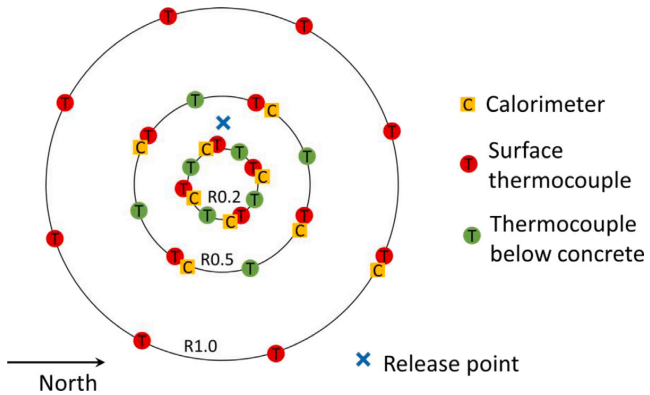

Fig. 3. Top view of the placement of thermocouples and calorimeters on the concrete pad. Dimensions are in meters. The green dots denote the positions of the measurements beneath the concrete surface, while the red dots denote the locations of the surface measurements. The release point of LH_2 is indicated by the blue cross, while the calorimeters are represented by the yellow squares.
Source: Adapted from [18].

4.1.1. Outdoor leakage studies

The experimental setup built for this experiment was positioned on a concrete pad, (see Fig. 3). It is composed of an LH_2 line ending with a nozzle with an internal diameter of 25.4 mm located in the center of the pad and surrounded by a few obstacles (two containers and metal frames to install the equipment). Moreover, several sensors (40 thermocouples, 30 oxygen sensors, 12 radiometers, and ten calorimeter blocks) are installed, as shown in Fig. 3, to measure temperature, hydrogen concentration, pressure, ambient conditions, radiation, and heat flux from the ignited release. Cameras were used to record the experiments as well. Additional information on the setup can be found in [45]. A total of seven trials were carried out as part of these studies. The objectives of this type of experiment were [18]:

- investigate the LH_2 pool behavior (formation, propagation, and duration);
- estimate of the flammable cloud developed from the release;
- analyze the cloud's combustion type (fire, deflagration, detonation).

As previously mentioned, it was observed that the LH_2 pool formation depends on the release orientation. With the set release rate and vertical downward direction, the maximum LH_2 pool diameter was 1.0 m. Furthermore, frozen air components were noted around the release point. As pointed out in the previous section, a mixture of LH_2 and LOX or SOX can lead to severe consequences, and its formation must be prevented or mitigated. The outdoor releases studies were already analyzed by Ustolin et al. [19] using a machine learning approach, namely a supervised learning method with the classification task and a linear model.

4.1.2. Closed room and ventilation mast studies

An enclosure composed of a metal box with a volume of 24 m^3 with internal dimensions (height, width, and depth) of $2.26 \times 2.96 \times 2.69$ m, respectively, was built to simulate a tank connection space (TCS). The TCS was connected to a ventilation mast with a diameter of 0.45 m and horizontal and vertical lengths of approx. 3 and 10 m, respectively, separated by a 90° bend. A vent panel was simulated by an opening (dimensions 1.6×2.3 m) on one side of the TCS and covering it with a polyethylene sheet. Also, in this case, ambient

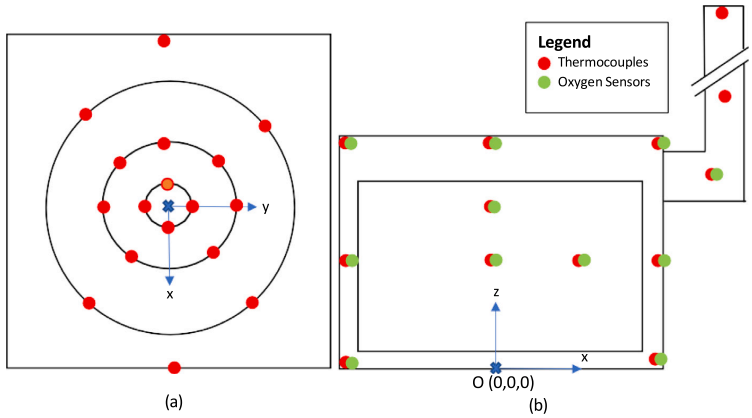

Fig. 4. Locations of thermocouples and oxygen sensors are indicated by red dots and green dots, respectively, while blue crosses indicate the LH_2 release points. The schematic displays (a) the top view of the test pad, and (b) the cross-section of the TCS connected to the ventilation mast. The test pad radius is 0.2, 0.5, and 1.0 m from the release point, respectively. The TCS has internal dimensions (height, width, and depth) of $2.26 \times 2.96 \times 2.69$ m, respectively. The x and y axes indicate the ground plane, while z is the vertical axis.
Source: Adapted from [18].

conditions, temperature, pressure, and gas concentration were recorded inside and outside the TCS ventilation mast. An illustration of the TCS is shown in Fig. 4. A total of eight release tests were carried out inside the TCS by positioning the leakage point in the middle of the room, oriented vertically downward. In this case, the diameter of the release nozzle was 25.4 mm in the first two trials and then reduced to 12.7 mm in the remaining tests. The first five tests had been carried out without ignition, while the flammable release was intentionally ignited in the last three. In this series, the maximum flow rate was 40.1 kg/min. The maximum extension of the LH_2 pool was 1.0 m from the release point (2.0 m diameter). Liquid or frozen air components were identified on the TCS floor and remained on the ground longer than LH_2, which evaporated 30–40 s after the leakage stopped [18]. When an ignition source was activated at the top of the ventilation mast, it took 30 min for an explosion to occur in case the TCS was sealed. On the other hand, when one of the two vents of the enclosure was opened, a severe explosion manifested 10–15 s after ignition on the same position. No spontaneous ignition was observed during the entire experimental campaign. Additional information on the outcomes of these experiments can be found in [18].

4.2. Database creation

The database is built based on the data collected from the case study. For the duration of the experimental test, each row represents the value that each thermocouple or sensor measures at each given moment in time. This latter corresponds to the sample rate of 0.1 s. Here, there are three different databases created:

1. The first database is built based on outdoor leakage studies to predict when air components may condense or freeze.
2. The second database is built based on outdoor leakage studies to predict the hydrogen concentration within the gas cloud.
3. The third database is built based on closed room studies, also called indoor leakage studies, to predict when air components may condense or freeze.

Hydrogen concentration was not investigated in the indoor studies since the box is quickly saturated with hydrogen after the beginning of the release. As the experiments previously described, if TCS is sealed, the enclosure is saturated with hydrogen, and it takes a long time to have flammable mixtures. Purging inert gas (e.g., nitrogen) can avoid the entrance of air into the room and the formation of a flammable atmosphere and can be adopted as a mitigation measure. The features in all of the databases are the following parameters:

Table 1
The list of features.

Feature	Feature
Time	PT_04
Ambient_Pressure	Wind_Direction_High
Release_Rate	Wind_Direction_Low
Humidity	Wind_Speed_High
Release_Orientation	Wind_Speed_Low
P01	TT
P02	x
P03	y
P04	z
PT_01	liquid_oxygen
PT_02	solid_oxygen
PT_03	H2_concentration_above_LFL

- the timestamp that indicates the sensors' sampling rate;
- atmospheric conditions: relative humidity, pressure, speed, and wind direction;
- liquid hydrogen release flowrate;
- temperature and hydrogen concentration;
- internal liquid hydrogen tank temperature and pressure;
- spatial coordinates of the thermocouples (x, y, z).

The parameters the model will predict are the labels specified in the database. For each sensor location and time in this study, the labels were defined as follows:

- liquid oxygen formation;
- solid oxygen formation;
- hydrogen concentration above the LFL.

In total, there are 24 features in these databases. Table 1 lists all features. For the first and third databases, two labels have been assigned. Label 1 is assigned if oxygen condensation or solidification occurs after 200 s; otherwise, label 0 is assigned. The oxygen boiling and melting point are considered at the pressure of 0.96 bar and correspond to $T_b = -183.5$ °C and $T_m = -218.79$ °C [22]. For the second database, two labels have been assigned. Label 1 is assigned

Table 2
The description of the first database.

Experiment	Duration (s)	No. of sensors locations	Datapoints
1	585	32	187,232
2	301.2	32	96,417
3	700.9	32	224,320
4	360.9	32	115,552
5	298.1	32	95,424
		Total datapoints	718,945

Table 3
The description of the second database.

Experiment	Duration (s)	No. of sensors locations	Datapoints
1	900.9	30	270,300
2	700.9	30	210,300
3	190.9	30	57,300
4	424	30	127,230
		Total datapoints	665,130

Table 4
The description of the third database.

Experiment	Duration (s)	No. of sensors locations	Datapoints
1	460.9	25	115,250
2	460.9	25	115,250
3	400.9	25	100,250
4	340.9	25	85,250
5	200.9	25	50,250
		Total datapoints	466,250

if the hydrogen concentration is above the LFL after 200 s; otherwise, label 0 is assigned. The LFL of hydrogen at atmospheric pressure is 4% vol. [46]. The 200 s time frame was chosen based on the water systems' response time after a hydrogen release.

The first, second, and third databases have 718,945 data points, 665,130 data points, and 466,250 data points, respectively. These databases only include the unignited release tests because the ignited one cannot provide any indication of the hydrogen dispersion and condensation of air components due to the hydrogen combustion. The first database was generated from outdoor studies and consisted of 5 experiments with varying duration, a sample rate of 0.1 s, and 32 different sensor locations. The description of the first database is tabulated in Table 2. The second database was generated from outdoor studies and consisted of 4 experiments with varying duration, a sample rate of 0.1 s, and 30 different sensor locations. The description of the second database is tabulated in Table 3. Finally, the third database was generated from indoor studies and consisted of 5 experiments with varying duration, a sample rate of 0.1 s, and 25 different sensor locations. The description of the third database is tabulated in Table 4. Before being inputted into the model, the databases are split into training and test sets. The databases are shuffled randomly before 75% of their data are assigned to the training sets and the remaining 25% to the test sets. Therefore, in the test sets, the first, second, and third databases have 179,736 data points, 166,282 data points, and 116,562 data points, respectively.

4.2.1. Impact of LFL changes

The LFL of hydrogen at atmospheric pressure and room temperature is approximately 4% by volume in air and assumed constant in this study to create a conservative model in terms of safety. This means that hydrogen must be present in a concentration of at least 4% in the air before it can ignite. The LFL of hydrogen varies depending on the pressure and temperature of the surrounding environment, the mixture composition ratio, and the orientation of the hydrogen release [47–50]. The LFL of hydrogen increases if the pressure is raised, while it decreases when the temperature is increased.

It is important to note that hydrogen has a large volume expansion ratio (1:850) [51] when comparing its density in liquid state at boiling point at atmospheric pressure with hydrogen gas at atmospheric conditions. So, a small LH_2 leak can quickly create a large volume of hydrogen gas in the atmosphere. Therefore, it is important to detect and isolate leaks as quickly as possible in order to prevent the concentration of hydrogen from reaching the flammable limit.

In summary, the lower flammability limit of hydrogen at atmospheric pressure and room temperature is around 4% by volume in air, but this can vary depending on the pressure and temperature of the surrounding environment and release orientation. It is important to note that hydrogen is highly flammable and must be handled with care and proper safety measures in place. For these reasons, a conservative model was developed in this study.

Table 5
Hyperparameters setup for RFs model.

Hyperparameter	Value
n_estimators	100
criterion	gini
max_depth	None
min_samples_split	2
min_samples_leaf	1
min_weight_fraction_leaf	0
max_features	sqrt
max_leaf_nodes	None
min_impurity_decrease	0
max_samples	None
random_state	25

4.3. Hyperparameters configuration

Table 5 summarizes the hyperparameters' value to build the RFs model.

4.4. Evaluation metrics

Performance metrics are used to evaluate the model's performance. In binary classification problems, the confusion matrix is typically used to summarize the predictions by breaking them down into four alternative outcomes: 1) true negative (TN) occurs when the accurate label and the predicted label are both negative (0); 2) true positive (TP) is when the accurate label and the predicted label are both positive (1); 3) false negative (FN) is when the accurate label is positive, but the predicted label is negative, and 4) false positive (FP) is when the accurate label is negative, but the predicted label is positive. These four measurements can derive the following performance metrics: accuracy, precision, recall, and F1-score. Accuracy describes the ratio of correct prediction and all predictions as expressed in (3). Precision represents the fraction of accurate positive predictions, as stated in (4), while recall indicates the percentage of accurate positive labels correctly predicted, as stated in (5). The F1-score, defined as the harmonic mean of the system's precision and recall, is appropriate for imbalanced data sets since it considers the precision and recall of the system. The formula of the F1-score is expressed in (6).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad , \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad , \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad , \tag{5}$$

$$F1 = 2 \times \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad . \tag{6}$$

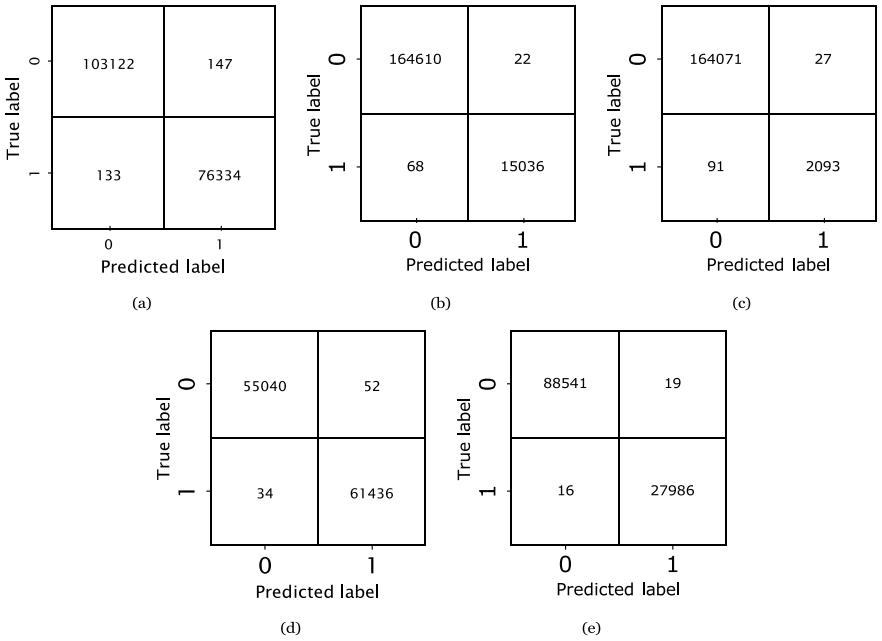

Fig. 5. Confusion matrices for the labels (a) first database - liquid oxygen, (b) first database - solid oxygen, (c) second database - H_2 concentration >LFL, (d) third database - liquid oxygen, and (e) third database - solid oxygen.

Table 6
Performance metrics of RFs model for all databases.

Label	Accuracy	Precision	Recall	F1	AUC-PR
First database (liquid oxygen)	0.9984	0.9981	0.9983	0.9984	0.9985
First database (solid oxygen)	0.9995	0.9985	0.9955	0.9984	0.9972
Second database (H_2 concentration >LFL)	0.9993	0.9873	0.9583	0.9861	0.9731
Third database (liquid oxygen)	0.9993	0.9992	0.9994	0.9993	0.9994
Third database (solid oxygen)	0.9997	0.9993	0.9994	0.9996	0.9994

Finally, the precision–recall curves can be plotted, and the area under the precision–recall curve (AUC-PR) can be calculated. High classifier performance is indicated by high accuracy, precision, recall, F1, and AUC-PR. All of the performance metrics values range from 0 to 1.

5. Prediction results and analysis

5.1. Model results

Table 6 summarizes the performance of the RFs model to predict various labels for all databases. The RFs model achieved very high scores in terms of accuracy, precision, recall, F1, and AUC-PR in every label. This indicates that the developed model can accurately predict air components condensing or solidifying on the ground in outdoor and indoor environments and whether the H_2 concentration will be above the LFL.

Fig. 5 depicts the obtained confusion matrices. In every label, the RFs model generates many correct predictions (a high number of true negatives and true positives) and minimal amounts of false predictions (a low number of false negatives and false positives). This means the probability of false alarms, i.e., predicting that there will be air components condensing or solidifying on the ground, or the H_2 concentration will be higher than the LFL while it will not, is very low, indicated by the low number of false positives. Similarly, the probability of non-detection, i.e., predicting that there will not be air components condensing or solidifying on the ground, or the H_2 concentration will not be higher than the LFL while it will, is also very low, denoted by the low number of false negatives.

The calculated AUC-PR for every database is summarized in Table 6. The scores indicate that the RFs model has high precision and recall for every label, except in the second database, where the precision and recall are relatively lower than others. This might happen because the dataset in the second database is very imbalanced, which means there are considerably much more data points for label 0 than for label 1. Thus, it can complicate the model's prediction as it has only a few data points to learn for label 1. Nevertheless, the precision and recall for the second database are still satisfactory for predicting if the hydrogen concentration within the gas cloud is above the LFL.

Furthermore, from the models' prediction, the extension of predicted LOX and SOX deposits can be plotted in spatial dimensions (x and y), as shown in Fig. 6. From Fig. 6, it can be seen that the predicted LOX and SOX deposits happen only within 1 m from the release point. These results are also observed in the actual experiments by FFI [18], thus confirming the models' prediction. The LOX deposits extend farther than the SOX deposits and follow the wind direction.

5.2. Performance comparison

Ustolin et al. [19] have conducted a similar study to predict condensation and solidification during an accidental release of LH_2. Their study investigates the release of LH_2 on the ground on a pad, comparable to this study's first and second databases. Table 7 compares the

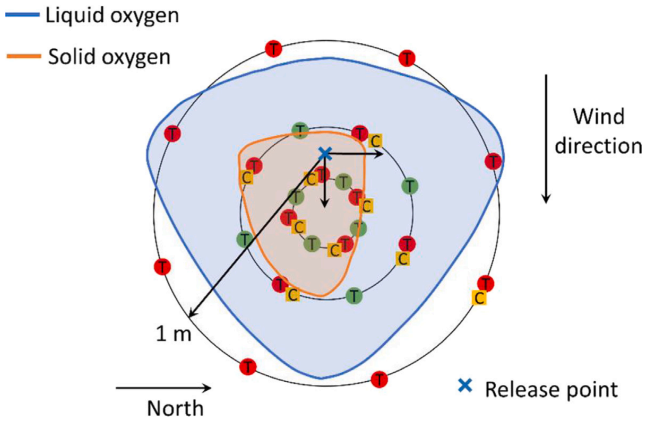

Fig. 6. Spatial extension of predicted LOX and SOX deposits on the test pad for outdoor leakage experiments.

Table 7
Performance comparison of RFs model and linear model (LM) in [19] for all labels.

Label	Accuracy		Precision		Recall		AUC-PR	
	LM	RFs	LM	RFs	LM	RFs	LM	RFs
Liquid oxygen	0.9020	**0.9984**	0.8480	**0.9981**	0.9360	**0.9983**	0.9490	**0.9985**
Solid oxygen	0.9570	**0.9995**	0.830	**0.9985**	0.6130	**0.9955**	0.8070	**0.9972**
H_2 concentration >LFL	0.9880	**0.9993**	0.6490	**0.9873**	0.1840	**0.9583**	0.3660	**0.9731**

performance of the RFs model in this study with the linear model used in [19] to predict the occurrence of LOX, SOX, and H_2 concentration above the LFL. The bold numbers denote the best results between the two methods.

As seen in Table 7, the RFs model in this study significantly improved the prediction results compared to the linear model in [19], especially in predicting the H_2 concentration above LFL. Although the hydrogen concentration, in this case, is rarely higher than the LFL, it could still have profound implications. This improvement of the model's predictive capabilities is crucial to provide accurate early warnings, which can prevent severe consequences from happening. In addition, the prediction of liquid and solid oxygen formation is also improved. Thus, one can adopt prevention and mitigation measures to reduce the overall risk of the LH_2 storage system.

5.3. Risk reduction measures

The developed model can aid the selection of measures necessary to decrease the risk of LH_2 storage and transfer components (e.g., tanks, pipes, valves). One can adopt appropriate preventive measures such as training and safety devices by knowing the extension of the LH_2 release consequences under different conditions. First, people who do not wear protective equipment must be restricted from the spill area. Then, all the ignition sources must be safely removed. Installing ATEX devices can remove the ignition sources from the most critical areas. This type of equipment is usually expensive, and its limited employment can drastically decrease the high costs of LH_2 technologies, thus facilitating their deployment.

Once the model has been trained, the engineers can change the operating conditions to predict the consequences of different release scenarios. In this fashion, the extension of the LH_2 and LOX/SOX pool area and the distance reached by the flammable cloud can be estimated. Afterward, one can determine when the mitigation measures must be activated to lessen the release consequences. On the other hand, 200 s were considered in this study based on the activation time of standard safety devices such as sprinklers [52]. Moreover, the best mitigation measure (e.g., sprinkler, water curtains, inert gases) can be selected depending on the type and yield of consequence. To avoid the formation of condensed phase explosions, releasing inert gases such as nitrogen is considered a good option since this phenomenon manifests if the oxygen concentration in the mixture is above 50% as demonstrated by Atkinson [10]. However, installing a water system can also be adopted. For instance, it was verified during a series of tests of the European project PRESLHY [42] that the contact between LH_2 and water does not lead to rapid phase transition (RPT) explosion, hence making sprinklers and water curtains effective mitigation measures to prevent ignition.

Furthermore, the estimation of the extension of the flammable hydrogen cloud could be used as input for other consequence analysis models (explosion models) and support the determination of precise and effective separation distances. The separation distances can be determined by considering the threshold values of the physical effects of fires and explosions. The distance should be far enough such that the thermal radiation or overpressure is sufficiently low to not cause any harm to the operators. However, the use of mitigation strategies may reduce the separation distances. For instance, the use of sprinklers can reduce the thermal radiation effects, and installing safety barriers such as walls can protect from explosion-related blasts and projectiles. It is crucial to ensure that the mitigation strategies will not worsen the effects of other risks. For example, barriers built to withstand blasts and missiles may prevent hydrogen from dispersing, lengthening the time flammable concentrations are present, and raising the risk of ignition. A combustible cloud explosion may become more likely as a result of the wall's partial containment [53].

5.4. Future studies

This study investigated condensed phase explosions and hydrogen dispersion (concentration during the release). However, the developed model can provide critical indications on each parameter included in the database. Therefore, additional simulations may be conducted to analyze other parameters and phenomena. For instance, the radiation from the jet fire provoked by the LH_2 release ignition can be predicted to indicate separation distances or sprinkler activation. In addition, the

model might be further developed to predict few of the consequences described in Section 2.1.

Experimental studies involving liquid hydrogen pose many economic and technological challenges mainly due to the high cost of LH_2 equipment and low LH_2 availability. Despite additional tests being necessary further to study the loss of containment of LH_2 technologies, it is suggested to extensively exploit the available experimental data to develop advanced models further as performed in this work. If there are more experimental studies about liquid hydrogen leakage in the future, it is then better to test the models on different test data so that the robustness and reliability of the models can be further assessed.

Since the risk is given by the combination of probabilities and consequences of failure, additional indications on the probabilities must be provided, especially for hydrogen technologies deployed for emerging technologies. Different studies investigate this issue and provide essential insights on risk-based inspection and maintenance methodologies for storage components where only hydrogen is contained [54].

6. Conclusion

Unique LH_2 release experiments were analyzed in this work. The focus was on investigating condensed phase explosions for LH_2 and LOX or SOX mixtures and the hydrogen dispersion in the air during the release. An advanced machine learning approach based on the random forest algorithm was employed to develop a model capable of predicting the consequences of the LH_2 loss of containment. Specific databases were created to analyze the liquefaction and solidification of oxygen during the LH_2 release and the hydrogen concentration in the air. The model was initially trained with the experimental data. Accurate results were obtained by the model predicting the above-mentioned parameters. The random forest approach was more robust and precise than other machine learning techniques previously used to simulate similar experiments. The insights provided by the model developed in this study can be further exploited to carry out detailed risk analysis and select appropriate prevention and mitigation measures to reduce the overall risk of LH_2 technologies, especially in emerging applications. Future studies, such as investigating different parameters and phenomena, were suggested.

CRediT authorship contribution statement

Muhammad Gibran Alfarizi: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Federico Ustolin:** Writing – review & editing, Writing – original draft, Validation, Investigation, Formal analysis, Data curation. **Jørn Vatn:** Writing – review & editing, Supervision, Resources. **Shen Yin:** Writing – review & editing, Validation, Supervision, Methodology. **Nicola Paltrinieri:** Writing – review & editing, Validation, Supervision, Resources, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would greatly acknowledge the Norwegian Defence Research Establishment (FFI) and DNV for data sharing of the large-scale leakage experiments of liquid hydrogen (LH_2). The authors are also grateful to Federica Ferrari for her initial research that started this study.

References

[1] Taccani R, Ustolin F, Zuliani N, Pinamonti P, Pietra A. Fuel cells and shipping emissions mitigation. In: Technology and science for the ships of the future. IOS Press; 2018, p. 885–92.
[2] Linstorm P. NIST chemistry webbook, NIST standard reference database number 69. J Phys Chem Ref Data, Monograph 1998;9:1–1951.
[3] Barthélémy H, Weber M, Barbier F. Hydrogen storage: Recent improvements and industrial perspectives. Int J Hydrogen Energy 2017;42(11):7254–62.
[4] Rüdiger H. Design characteristics and performance of a liquid hydrogen tank system for motor cars. Cryogenics 1992;32(3):327–9.
[5] Hua T, Ahluwalia R, Peng J-K, Kromer M, Lasher S, McKenney K, et al. Technical assessment of compressed hydrogen storage tank systems for automotive applications. Int J Hydrogen Energy 2011;36(4):3037–49.
[6] Noh Y, Chang K, Seo Y, Chang D. Risk-based determination of design pressure of LNG fuel storage tanks based on dynamic process simulation combined with Monte Carlo method. Reliab Eng Syst Saf 2014;129:76–82.
[7] Ustolin F, Campari A, Taccani R. An extensive review of liquid hydrogen in transportation with focus on the maritime sector. J Mar Sci Eng 2022;10(9):1222.
[8] Kurtz J, Sprik S, Peters M, Bradley TH. Retail hydrogen station reliability status and advances. Reliab Eng Syst Saf 2020;106823.
[9] Dinesh A, Benson C, Holborn P, Sampath S, Xiong Y. Performance evaluation of nitrogen for fire safety application in aircraft. Reliab Eng Syst Saf 2020;202:107044.
[10] Atkinson G. Condensed phase explosions involving liquid hydrogen. In: International conference on hydrogen safety. 2021.
[11] Wang C, Dou M, Li Z, Outbib R, Zhao D, Zuo J, et al. Data-driven prognostics based on time-frequency analysis and symbolic recurrent neural network for fuel cells under dynamic load. Reliab Eng Syst Saf 2023;233:109123.
[12] Zhang J, Li X, Tian J, Luo H, Yin S. An integrated multi-head dual sparse self-attention network for remaining useful life prediction. Reliab Eng Syst Saf 2023;109096.
[13] Zhang J, Li X, Tian J, Jiang Y, Luo H, Yin S. A variational local weighted deep sub-domain adaptation network for remaining useful life prediction facing cross-domain condition. Reliab Eng Syst Saf 2023;231:108986.
[14] Lan H, Ma X, Qiao W, Deng W. Determining the critical risk factors for predicting the severity of ship collision accidents using a data-driven approach. Reliab Eng Syst Saf 2023;230:108934.
[15] Shu X, Shen J, Chen Z, Zhang Y, Liu Y, Lin Y. Remaining capacity estimation for lithium-ion batteries via co-operation of multi-machine learning algorithms. Reliab Eng Syst Saf 2022;228:108821.
[16] Alfarizi MG, Vatn J, Yin S. An extreme gradient boosting aided fault diagnosis approach: A case study of fuse test bench. IEEE Trans Artif Intell 2022;1–8. http://dx.doi.org/10.1109/TAI.2022.3165137.
[17] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[18] Aaneby J, Gjesdal T, Voie ØA. Large scale leakage of liquid hydrogen (LH2)–tests related to bunkering and maritime use of liquid hydrogen. Tech. rep., Forsvarets forskningsinstitutt (FFI) Norwegian Defence Research Establishment; 2021.
[19] Ustolin F, Ferrari F, Paltrinieri N. Prediction of condensed phase formation during an accidental release of liquid hydrogen. Chem Eng Trans 2022;91:439–44.
[20] Kim KO, Roh T, Lee J-W, Zuo MJ. Derating design for optimizing reliability and cost with an application to liquid rocket engines. Reliab Eng Syst Saf 2016;146:13–20.
[21] Marcoulaki EC, Venetsanos A, Papazoglou IA. Quantitative safety analysis of cryogenic liquid releases in a deep underground large scale installation. Reliab Eng Syst Saf 2017;162:51–63.
[22] Green DW, Southard MZ. Perry's chemical engineers' handbook. McGraw-Hill Education; 2019.
[23] Royle M, Willoughby D. Releases of unignited liquid hydrogen-RR986-health and safety laboratory. 2014, Harpur Hill, Buxton, UK.
[24] Tretsiakova-McNally S. LECTURE. Dealing with hydrogen explosions. 2014, http://www.hyresponse.eu/files/Lectures/Dealing_with_hydrogen_explosions_notes.pdf. [Online; accessed 13-February-2023].
[25] Alcock J, Shirvill L, Cracknell R. Compilation of existing safety data on hydrogen and comparative fuels. Deliverable report, EIHP2, May, 2001.
[26] Gill J, Atkinson G, Cowpe E, Phylaktou H, Andrews G. Experimental investigation of potential confined ignition sources for vapour cloud explosions. Process Saf Environ Prot 2020;135:187–206.
[27] NASA. Safety standard for hydrogen and hydrogen systems guidelines for hydrogen system design, materials selection, operations, storage and transportation. Tech. rep., NASA; 2005.
[28] Ono R, Nifuku M, Fujiwara S, Horiguchi S, Oda T. Minimum ignition energy of hydrogen–air mixture: Effects of humidity and spark duration. J Electrost 2007;65(2):87–93.
[29] McCarty RD, Hord J, Roder HM. Selected properties of hydrogen (engineering design data), vol. 168. US Department of Commerce, National Bureau of Standards; 1981.
[30] Thomas JK, Eastwood C, Goodrich M. Are unconfined hydrogen vapor cloud explosions credible? Process Saf Progr 2015;34(1):36–43.

[31] Lyons K. Summary of experiment series E5.5 (Congestion) results - PRESLHY project. 2021.
[32] Odsæter LH, Skarsvåg HL, Aursand E, Ustolin F, Reigstad GA, Paltrinieri N. Liquid hydrogen spills on water—risk and consequences of rapid phase transition. Energies 2021;14(16):4789.
[33] Ustolin F, Odsaeter L, Reigstad G, Skarsvag H, Paltrinieri N. Theories and mechanism of rapid phase transition. Chem Eng Trans 2020;82:253–8.
[34] Aursand E, Odsæter LH, Skarsvåg H, Reigstad G, Ustolin F, Paltrinieri N. Risk and consequences of rapid phase transition for liquid hydrogen. In: Proceedings of the 30th european safety and reliability conference and the 15th probabilistic safety assessment and management conference, Venice, Italy. 2020, p. 21–6.
[35] van Wingerden K, Kluge M, Habib AK, Skarsvåg HL, Ustolin F, Paltrinieri N, et al. Experimental investigation into the consequences of release of liquified hydrogen onto and under water. Chem Eng Trans 2022;90:541–6.
[36] Zabetakis MG, Burgess DS. Research on the hazards associated with the production and handling of liquid hydrogen, vol. 5707. US Department of the Interior, Bureau of Mines; 1961.
[37] Witcofski RD, Chirivella J. Experimental and analytical analyses of the mechanisms governing the dispersion of flammable clouds formed by liquid hydrogen spills. Int J Hydrogen Energy 1984;9(5):425–35.
[38] Schmidtchen U, Marinescu-Pasoi L, Verfondern K, Nickel V, Sturm B, Dienhart B. Simulation of accidental spills of cryogenic hydrogen in a residential area. Cryogenics 1994;34:401–4.
[39] Verfondern K, Dienhart B. Experimental and theoretical investigation of liquid hydrogen pool spreading and vaporization. Int J Hydrogen Energy 1997;22(7):649–60.
[40] Hall JE, Hooker P, Willoughby D. Ignited releases of liquid hydrogen: Safety considerations of thermal and overpressure effects. Int J Hydrogen Energy 2014;39(35):20547–53.
[41] Ustolin F, Asholt HØ, Zdravistch F, Niemi R, Paltrinieri N. Computational fluid dynamics modeling of liquid hydrogen release and dispersion in gas refuelling stations. Chemical Engineering Transactions 2021;86:223–8.
[42] PRESLHY. Prenormative Research for Safe Use of Liquid Hydrogen, Research and Innovation Action Supported by the FCH JU 2.0. 2019, https://preslhy.eu/. Accessed: 2019-06-22.
[43] Alfarizi MG, Tajiani B, Vatn J, Yin S. Optimized random forest model for remaining useful life prediction of experimental bearings. IEEE Trans Ind Inf 2022;1–10. http://dx.doi.org/10.1109/TII.2022.3206339.
[44] Ibrahim IA, Hossain M, Duck BC. An optimized offline random forests-based model for ultra-short-term prediction of PV characteristics. IEEE Trans Ind Inf 2019;16(1):202–14.
[45] DNV G. Oil and Gas, Liquid hydrogen safety data report: Outdoor leakage studies. Tech. rep., Forsvarets forskningsinstitutt (FFI) Norwegian Defence Research Establishment; 2020.
[46] Hochgraf C. Applications – transportation | electric vehicles: Fuel cells. In: Garche J, editor. Encyclopedia of electrochemical power sources. Amsterdam: Elsevier; 2009, p. 236–48.
[47] Jeon J, Choi W, Kim SJ. A flammability limit model for hydrogen-air-diluent mixtures based on heat transfer characteristics in flame propagation. Nucl Eng Technol 2019;51(7):1749–57.
[48] Jeon J, Shin D, Choi W, Kim SJ. Identification of the extinction mechanism of lean limit hydrogen flames based on lewis number effect. Int J Heat Mass Transfer 2021;174:121288.
[49] Jeon J, Kim YS, Jung H, Kim SJ. A mechanistic analysis of H2O and CO2 diluent effect on hydrogen flammability limit considering flame extinction mechanism. Nucl Eng Technol 2021;53(10):3286–97.
[50] Zhang K, Luo T, Li Y, Zhang T, Li X, Zhang Z, et al. Effect of ignition, initial pressure and temperature on the lower flammability limit of hydrogen/air mixture. Int J Hydrogen Energy 2022;47(33):15107–19.
[51] Association CG, et al. Handbook of compressed gases. Springer Science & Business Media; 2012.
[52] Wade C, Spearpoint M, Bittern A, Tsai K. Assessing the sprinkler activation predictive capability of the BRANZFIRE fire model. Fire Technol 2007;43(3):175–93.
[53] Pohanish R. Sittig's handbook of toxic and hazardous chemicals and carcinogens. EngineeringPro collection, 5th ed. Elsevier Science; 2008.
[54] Campari A, Darabi MA, Ustolin F, Alvaro A, Paltrinieri N. Applicability of risk-based inspection methodology to hydrogen technologies: A preliminary review of the existing standards. In: Proceedings of the 32nd european safety and reliability conference. ESREL 2022, 2022, p. 574–81.

Article 4

Sustainability of ICPS from a Safety Perspective: Challenges and Opportunities

M. G. Alfarizi, J. Liu, J. Vatn and S. Yin, "Sustainability of ICPS from a Safety Perspective: Challenges and Opportunities," *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*, Helsinki, Finland, pp. 1-8, 2023.

Sustainability of ICPS from a Safety Perspective: Challenges and Opportunities

Muhammad Gibran Alfarizi
Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
muhammad.g.alfarizi@ntnu.no

Jie Liu
Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
jie.liu@ntnu.no

Jørn Vatn
Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
jorn.vatn@ntnu.no

Shen Yin
Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
shen.yin@ntnu.no

2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE) | 979-8-3503-9971-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIE51358.2023.10228030

***Abstract*—The growth of Industrial Cyber-Physical Systems (ICPS) has boosted industrial processes' effectiveness, security, and sustainability. Nevertheless, preserving the sustainability of ICPS is a complex challenge, especially in dynamic and unpredictable conditions. Fault diagnosis, prognosis, and remaining useful life prediction are crucial for ICPS sustainability in terms of safety, as they enable timely and effective maintenance to extend the system's lifespan. This paper presents a perspective on the current state-of-the-art in the safety of ICPS, emphasizing fault diagnosis, prognosis, and remaining useful life prediction. Key findings from existing literature are summarized, academic gaps are identified, and research challenges are highlighted. The paper offers suggestions for addressing the gaps, advancing ICPS sustainability, and ensuring secure and reliable operation.**

***Index Terms*—Fault Diagnosis, Industrial Cyber-Physical Systems (ICPS), Prognosis, Remaining Useful Life Prediction, Sustainability**

I. The State of the Art

Industrial processes play a crucial role in our daily lives, from producing goods and services to generating energy. The sustainability of Industrial Cyber-Physical Systems (ICPS) relies on three aspects: safety, security, and energy efficiency, which are illustrated in Figure 1. The safety and security of these processes are of paramount importance, as a failure in these systems can have serious consequences, including harm to people, the environment, and the economy. In addition, ensuring these processes run on efficient and low-carbon energy is also critical to protect the environment and establish a sustainable ICPS in the long run. In recent years, there have been several high-profile accidents in industrial processes, which have raised serious concerns about the safety and security of these systems. One of the examples is the attack on the Nord Stream pipeline, which caused a temporary disruption of gas supplies to Europe. The incident highlights the vulnerability of critical infrastructure to physical attacks and the importance of ensuring the security of these systems. This has led to increased attention from researchers, industry, and governments, who are working to improve the safety and security of industrial processes.

Ensuring the safety and security of industrial processes is complex and requires a multidisciplinary approach encompassing various fields, including engineering, computer science, and social sciences. The use of modern technologies and methodologies, such as machine learning [1]–[3], the Internet of Things [4], [5], and cybersecurity [6]–[8], have the potential to enhance the safety and security of these processes significantly. However, there are also challenges and limitations associated with these technologies, and it is essential to identify and address these issues to ensure industrial processes' safe and secure operation.

ICPS have been increasingly integrated into various industrial processes to enhance the efficiency and effectiveness of production, as well as provide greater flexibility in adapting to dynamic and uncertain conditions [9]–[13]. However, this integration has also increased security and reliability concerns, especially when dealing with sensitive industrial processes such as those in the chemical and petrochemical industries [14]–[16]. This paper aims to explore the current state of sustainability in ICPS, with a particular focus on maintaining secure and reliable operations in a dynamic and uncertain environment.

The current advancements in the safety of ICPS encompass a range of subjects, including fault diagnosis, prognosis, and remaining useful life (RUL) prediction. In terms of fault diagnosis, various techniques have been developed to monitor the performance of ICPS and identify potential faults before they result in system failure. For instance, several studies have utilized data-based methods such as Artificial Neural Networks (ANNs) [17]–[19] and Support Vector Machines (SVMs) [20]–[22] to detect faults in ICPS. Additionally, other studies have utilized model-based methods, including Model Predictive Control (MPC) [23]–[25] and Hybrid System models [26]–[28], to perform fault diagnosis in ICPS.

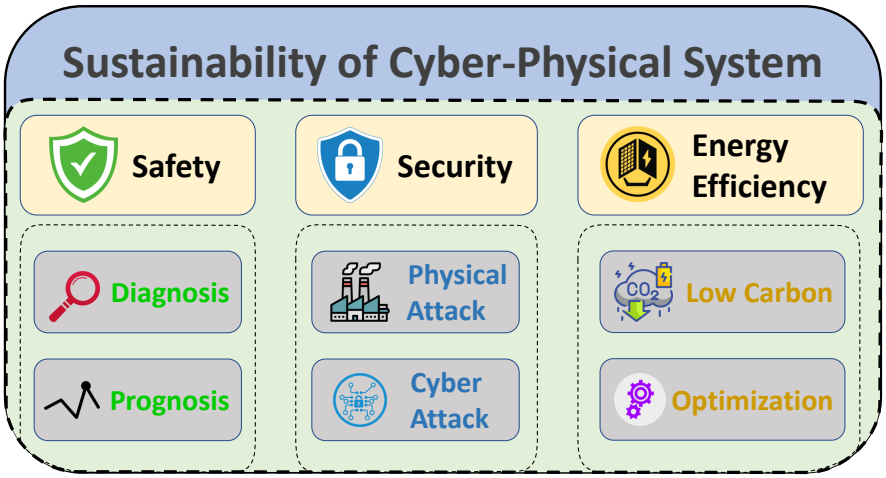

Fig. 1. The three pillars of sustainability in ICPS.

Regarding prognosis and RUL prediction, various approaches have been developed to predict the remaining lifespan of ICPS components. These methods take consider multiple factors, such as the system's historical performance, operating conditions, and the physical properties of the components. Some studies have utilized probabilistic models such as Bayesian networks [29], [30] and Markov chains [31], [32] to predict the RUL of ICPS components. Other studies have utilized data-based methods, including machine learning techniques such as Artificial Neural Networks (ANNs) [33]–[35], to make the RUL prediction of ICPS components. These advancements in the safety of ICPS are summarized in Figure 2.

Despite the advancements in the field of ICPS safety, several academic gaps still need to be addressed. One significant gap is the need for integrating multiple techniques to enhance the reliability and security of ICPS. Currently, most studies in the field focus on either fault diagnosis or RUL prediction and do not integrate these approaches to enhance the overall safety of ICPS. Additionally, the fault diagnosis, prognosis, and RUL prediction algorithms for ICPS need to be real-time implementable, but many current algorithms are too computationally expensive. There is a need for research to make the algorithms more efficient for real-time use.

To address these problems, the field of ICPS safety requires further research that integrates multiple techniques to improve reliability and security. This research should aim to holistically address fault diagnosis and RUL prediction, considering algorithms' computational efficiency for real-time implementation. Additionally, efforts should be made to improve the computational efficiency of existing algorithms so that they can be used in real-time industrial applications.

In conclusion, this paper aims to provide a comprehensive perspective of the current state of the art in ICPS safety, focusing on maintaining secure and reliable operation in dynamic and uncertain environments. The identified academic gaps and potential solutions provide valuable insights for future research and development in the field of ICPS safety.

II. Safety I: Fault Diagnosis

Fault diagnosis techniques in ICPS are methods used to identify, isolate, and diagnose faults in the system. Some literature also refers to fault diagnosis as fault detection and isolation (FDI). The significant technologies can be classified into:

1) Model-based fault diagnosis: This technique uses a mathematical model of the system to predict its behavior. Any deviation from the expected behavior is considered a fault. For example, detecting faults in a control loop is possible with a Kalman filter.
2) Data-based fault diagnosis: This technique can be divided further into stochastic-based, signal-based, and machine learning-based. Stochastic-based fault diagnosis uses statistical techniques to detect changes in the process behavior indicative of faults. For example, a control chart can see changes in a process's mean and variance. Signal-based fault diagnosis involves monitoring the signals generated by the system and using algorithms to detect anomalies in these signals. For example, wavelet analysis can detect faults in signals generated by sensors. Machine learning-based fault diagnosis uses machine learning algorithms such as neural networks, support vector machines, or decision trees to detect faults in the system.
3) Fusion: This technique uses both model-based and data-based methods to diagnose faults in the system. For example, an aircraft engine's expected performance can be represented by a model-based method, but it may not

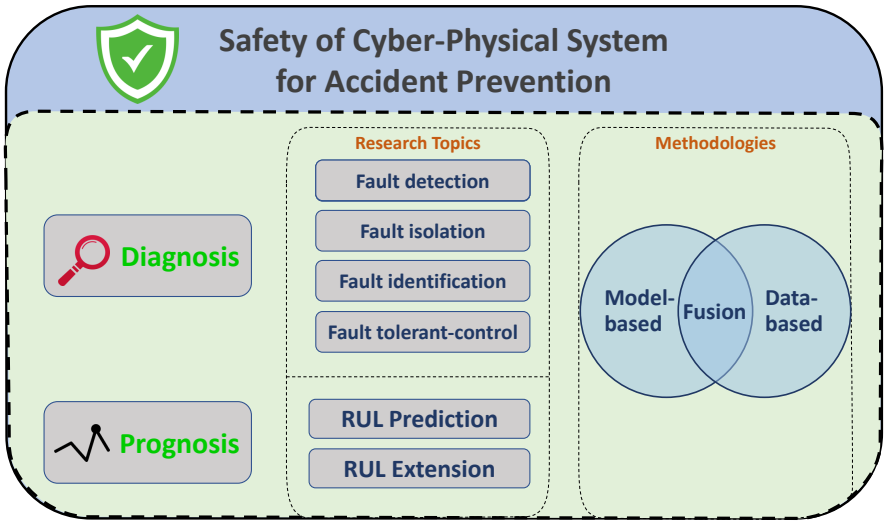

Fig. 2. Overview of sustainability in ICPS from a safety perspective.

consider all factors that can cause faults. Therefore, a data-driven method can be used to monitor the engine's actual performance during operation and detect any deviations using machine learning algorithms. This can help identify faults or failures.

A. Model-based Fault Diagnosis

Model-based fault diagnosis in ICPS is a technique that uses mathematical models of the system to detect faults. The basic idea is to use a system model to predict its behavior and compare this prediction to the system's actual behavior. Any deviation from the expected behavior is considered a fault. Several types of mathematical models can be used for model-based fault diagnosis in ICPS, including:

1) State-space models: These models describe the system's state using a set of state variables and equations. The system's state can be estimated using the measurement data and compared with the expected state to detect faults.
2) Transfer function models describe the relationship between the inputs and outputs of a system. The system's outputs can be compared with the expected outputs based on the information to detect faults.
3) Parametric models: These models describe the system's behavior in terms of parameters that can be estimated from the data. The parameters can be calculated and compared with the expected values to detect faults.

A common model-based fault diagnosis approach is a Kalman filter, a state-space model. A Kalman filter uses a mathematical model of the system and measurement data to estimate the system's state [36]. The estimated state can then be compared with the expected state to detect faults. For example, a wind turbine model was developed in [37] using a closed-loop identification technique and fault detection is done through residuals generated by dual Kalman filters. In contrast, dual sensor redundancy is used for fault isolation.

Another technique that can be classified as model-based fault diagnosis is the Hidden Markov Models (HMMs). HMMs are probabilistic graphical models that represent the underlying states of a system and can be used to detect faults. For example, a study applied HMMs to see the fault of bearings [38]. The scalar probabilities of different bearing operating conditions are used as input to the HMMs for fault diagnosis, showing promising results through experimental analysis.

B. Data-based Fault Diagnosis

1) Stochastic-based Fault Diagnosis: Stochastic-based fault diagnosis is one of a method of data-based fault diagnosis in ICPS that uses statistical techniques to monitor the behavior of a process and detect deviations from normal behavior. This approach involves collecting and analyzing data generated by a system over time, intending to see patterns or anomalies that may indicate the presence of a fault. Some standard techniques of this approach include:

- Statistical process control (SPC): A quality control method that uses statistical techniques to monitor and control a process to ensure it operates within specified limits [39].
- Multivariate Statistical Analysis (MSA): A set of techniques that use multiple variables to describe the behavior of a system and can be used to detect anomalies and faults [40].
- Principal Component Analysis (PCA): A statistical technique that uses orthogonal transformations to convert

a set of correlated variables into a set of uncorrelated variables, which can be used for residual generation and fault diagnosis [41].
- Partial Least Squares (PLS): A statistical technique that uses linear regression to model the relationship between two sets of variables and can be used for fault diagnosis [42].

Researchers widely use PCA for fault diagnosis in ICPS [43]–[47]. The advantage of using PCA for fault diagnosis is that it can reduce the dimensionality of the data, making it easier to detect faults by analyzing residuals. Additionally, PCA can handle highly correlated data, making it suitable for complex and high-dimensional data sets that are common in many industrial applications.

2) Signal-based Fault Diagnosis: Signal-based fault diagnosis is a method of detecting faults in ICPS by analyzing the signals generated by the system. The basic idea is to monitor the signals generated by the system, compare the signals to a reference model, and detect deviations from the reference model that could indicate a fault. Signal-based fault diagnosis can be performed using a variety of techniques, including:

- Frequency-domain analysis: This technique involves analyzing the frequency content of a signal. This approach transforms the signal from the time domain to the frequency domain using a method such as the Fast Fourier Transform (FFT) [48]. This allows for a more detailed analysis of the signal's frequency components, which can help identify faults such as resonances or frequency-specific disturbances.
- Time-domain analysis: This method involves examining the behavior of a system over time. In this approach, the signal is plotted against time and analyzed for any trends, changes, or anomalies that may indicate a fault. Time-domain analysis can identify faults in a system by looking for changes in the amplitude, phase, or other signal characteristics over time.
- Time-Frequency analysis: For machines experiencing changes in their operating conditions, such as varying load or unbalanced supply voltages, the signals they produce are dynamic and challenging to monitor using traditional time-domain or frequency-domain analysis methods. To effectively diagnose faults in real-time, using time-frequency decomposition tools is necessary to analyze the time-varying frequency spectrum of the transient signals [49].

For instance, a signal-based fault diagnosis method for power converters of switched reluctance motors was developed in [50] by examining the changes in the root-mean-square current characteristics between health conditions and single-/dual-transistor short circuits or open circuits situations. In [51], an improved frequency-domain blind deconvolution flow-based acoustic fault detection approach for gearboxes was proposed.

3) Machine learning-based Fault Diagnosis: Machine learning (ML) based fault diagnosis is a method of detecting faults in ICPS using ML algorithms. ML-based fault diagnosis can be performed using various techniques, including:

- Artificial Neural Networks (ANNs): ANNs are widely used in fault diagnosis due to their ability to model complex relationships between inputs and outputs. They can be trained using historical data to identify patterns in the system's behavior indicative of faults.
- Support Vector Machines (SVMs): SVMs are a type of machine learning algorithm that are particularly well-suited for solving classification problems. They can differentiate between normal and abnormal system behavior and classify faults into different categories.
- Decision Trees: Decision trees are a versatile machine learning algorithm that can be applied to both classification and regression problems. They can be used to identify the most critical factors contributing to the occurrence of faults in a system [52]–[54].
- Deep Learning: Deep learning is a type of machine learning that uses ANNs with multiple layers to learn from large amounts of data. Deep learning algorithms have been used in various industrial applications for fault diagnosis due to their ability to understand complex patterns in the data [55]–[57].

Researchers in [58] developed a fault diagnosis system to prevent operational failure in a fuse test bench. The system, based on extreme gradient boosting, which is an advancement of decision trees, was capable of detecting, classifying, and identifying the root causes of faults. The results of the experiment demonstrate that the system has a high level of accuracy, a fast diagnosis time, and provides interpretable root cause analysis. In [59], multiple autoencoders (AE), a type of deep learning algorithm, were stacked to extract features from raw bearing vibration signals. The extracted features were then processed through a softmax regression to identify the different bearing faults.

III. Safety II: Prognosis

Prognosis techniques in ICPS are methods used to predict the remaining time before a component or system fails. Prognostics approaches can be broadly divided into three categories: model-based, data-based, and fusion models.

1) Model-based prognosis: These approaches use mathematical models based on physical laws and principles to simulate the behavior of a system and predict its future behavior. Examples include physics-based models and first-principles models.
2) Data-based prognosis: These approaches rely on historical data to predict a system's behavior. Examples include statistical methods such as regression analysis and machine learning techniques such as neural networks and support vector machines (SVMs).
3) Fusion prognosis: These approaches combine the strengths of both data-based and model-based approaches by using historical data and mathematical models to make predictions.

A. Model-based Prognosis

Model-based prognosis in ICPS involves using mathematical models to predict the remaining life of components or systems. The goal is to estimate the RUL accurately and identify the degradation onset before it leads to failure. Model-based prognosis can be classified into two main approaches: physics of failure and system model approaches.

1) Physics of Failure (PoF) approach: This approach is based on the idea that the degradation of a system is driven by physical mechanisms such as wear and tear, corrosion, and fatigue. PoF models use physical laws and principles to describe the behavior of a system as it degrades over time. For example, a PoF model of a component might explain the relationship between its stress and strain and predict the point at which it will fail based on the accumulation of damage.
2) System model approach: This approach uses mathematical models to describe the behavior of a system as a whole without focusing on the individual components or degradation mechanisms. For example, a system model might explain the behavior of an engine based on its speed, torque, and fuel consumption and predict its remaining useful life based on its performance data.

PoF models can provide accurate predictions of the timing of component failure based on physical models of the degradation process. However, it can be complex and challenging to develop, particularly for systems with multiple degradation mechanisms. On the other hand, system models are often more straightforward to create than PoF models, particularly for complex systems with various components and interactions. But may not provide as accurate predictions of component failure as PoF models.

In [60], the PoF approach was used to predict the failure of the power supply. The power supply was divided into parts based on their material characteristics. By predicting the degradation of one or a combination of these components, the overall reliability of the entire power supply system could be estimated.

In [61], a system model approach was applied for gas pipeline system integrity management to prevent or reduce the likelihood of failures. The framework considers all possible failure modes of the pipeline using real-time field data and a combination of Hybrid Causal Logic and Dynamic Bayesian Networks predictive models to suggest cost-effective and optimal mitigation actions.

B. Data-based Prognosis

A data-based prognosis predicts a system's health and remaining useful life based on historical data. Data-based prognosis is based on the idea that patterns can describe the system's behavior in the data and that these patterns can be used to predict the RUL. The data used for RUL prediction can include operational parameters, performance metrics, or degradation indicators. It can be further classified into two main categories: statistical approaches and machine learning approaches.

1) Statistical approaches: These techniques use mathematical models and statistical methods to analyze large amounts of historical data. These methods are often based on traditional statistical techniques such as regression, principal component analysis (PCA), and survival analysis. These methods require high-quality data to produce accurate predictions, but they are simple to implement and interpret.
2) Machine learning approaches: These techniques use advanced algorithms such as artificial neural networks, decision trees, and support vector machines to model the relationships between input variables and system health. These approaches can learn from data without relying on explicit mathematical models, making them well-suited for complex and non-linear systems. However, they can also be more computationally intensive and require more expertise to implement and interpret.

A study in [62] developed an optimized random forest model, a machine learning approach, to predict the RUL of experimental bearings. The proposed framework integrates a signal processing technique and machine learning to improve RUL prediction accuracy. The experiment results significantly improved compared to the standard data-based and stochastic techniques.

Another study in [63] also incorporates machine learning for accident prevention by predicting condensed phase formation during an accidental release of liquid hydrogen. A model was developed to predict the likelihood and location of an oxygen phase change during refueling based on the operating conditions. The results of the model were accurate and reliable, and can be used to select appropriate safety measures, like a water drench system, to prevent an oxygen phase change.

C. Fusion Prognosis

Fusion approaches in prognostics use a combination of data-based and model-based methods to achieve more accurate and reliable predictions. These approaches leverage the strengths of both data-based and model-based methods to overcome their limitations. For example, data-based methods can provide good results with large amounts of data, while model-based methods can provide better physical insights and interpretability. However, data-based methods can be limited by the quality and quantity of data, while the complexity and accuracy of the model can determine model-based methods.

In a fusion approach, data-based models can be used to identify patterns and relationships in the data. Then these patterns can be used to inform and improve the model-based models. Similarly, model-based models can provide context and physical understanding for data-based models. This combination of methods can improve prediction accuracy and robustness compared to using either technique alone.

Some researchers in [64] propose a hybrid prediction model, which integrates random forest (RF), Artificial Bee Colony (ABC), and general regression neural network (GRNN), called RF-ABC-GRNN, to accurately predict the RUL of lithium-ion batteries in the early-cycle stage. The comparison results reveal

that the suggested model effectively identifies the key features and makes precise predictions much sooner.

Another study in [65] developed a hybrid prediction model based on a combination of a hidden semi-Mark model (HSMM) and an empirical model to predict the RUL of the Solid Oxide Fuel Cell (SOFC). The results indicate that the developed prognostic model outperforms existing approaches for SOFC prognostics with improved prediction accuracy and faster forecasting speed.

IV. Challenges from Safety Perspective: Fault Diagnosis and Prognosis

Implementing fault diagnosis and prognosis in ICPS presents various technical and operational challenges that must be addressed to ensure their reliable and efficient operation. Overcoming these challenges is critical to ensuring industrial systems' safe and reliable operation and continued growth and development. Some of the significant challenges in fault diagnosis and prognosis are:

1) Integration of multiple sources of data and techniques: In ICPS, various sources of data can be used for fault diagnosis and prognosis, such as sensor data, operational data, and maintenance data. However, there needs to be more research on effectively integrating and using these multiple data sources for improved fault diagnosis and prognosis. In addition, most studies only focus on fault diagnosis or prognosis and do not integrate these approaches to enhance the overall safety of ICPS.
2) Handling uncertainty: There is often uncertainty in the data used for fault diagnosis and prognosis, such as measurement noise and uncertainty in the model parameters. This uncertainty can lead to inaccurate results and false alarms. Research is needed to handle uncertainty in the data and models used for fault diagnosis and prognosis.
3) Real-time implementation: The fault diagnosis and prognosis algorithms used in ICPS must be implemented in real-time to be helpful in industrial applications. However, many existing algorithms are computationally expensive and run offline. There is a need for research on how to make the algorithms more computationally efficient and implement them in real-time.
4) Incorporating domain knowledge: Learning about the system and its behavior in many industrial applications can improve the accuracy of fault diagnosis and prognosis. However, more research must be conducted on effectively incorporating this domain knowledge into the algorithms.

To overcome these challenges, a variety of solutions that address each of these issues can be proposed. These solutions range from integrated approaches using multiple data sources to efficient computation techniques and approximation methods to knowledge-based methods incorporating domain knowledge. Using these solutions, researchers aim to improve the accuracy and efficiency of fault diagnosis and prognosis in ICPS and ensure these systems' safe and reliable operation.

1) Data fusion: Developing algorithms that can effectively integrate and use multiple data sources for improved fault diagnosis and prognosis.
2) Uncertainty quantification: Developing algorithms that can effectively handle uncertainty in the data and models used for fault diagnosis and prognosis. Examples of such algorithms include Bayesian filtering and Kalman filtering. These techniques can estimate the uncertainty in the data and models and provide a probabilistic assessment of the system's behavior.
3) Real-time optimization: Developing computationally efficient algorithms that can be implemented in real-time for industrial applications. Researchers can also use efficient computation techniques, such as parallel processing and hardware acceleration. Additionally, approximation techniques, such as reduced-order modeling and model pruning, can reduce the algorithms' computational complexity.
4) Incorporating domain knowledge: Researchers can use knowledge-based methods, such as rule-based and expert systems, to incorporate domain knowledge into the algorithms. These methods can use the knowledge about the system and its behavior to guide the fault diagnosis and prognosis processes and improve their accuracy. Additionally, machine learning techniques can be used to learn from the domain knowledge and improve the algorithms' performance.

In addition to the above solutions, some exciting future research topics can improve safety and reliability in ICPS:

1) Transfer learning and domain adaptation: Transfer learning and domain adaptation techniques allow the algorithms developed for one industrial system to be adapted and used in another industrial design. This approach can reduce the data and computational resources required to develop algorithms for each new industrial system. By generalizing the algorithms, transfer learning and domain adaptation techniques can improve the accuracy of the fault diagnosis and prognosis algorithms and make them more widely applicable in different industrial systems.
2) Multi-task learning: Multi-task learning algorithms can perform multiple tasks in a single framework, such as fault diagnosis and prognosis. By sharing information and resources between these tasks, multi-task learning algorithms can improve the accuracy of both tasks. For example, the fault diagnosis results can improve the prognosis and vice versa. This approach can also reduce the data and computational resources required to develop separate algorithms for each task.
3) Active learning: Active learning algorithms can actively select and label the most relevant data for improved fault diagnosis and prognosis. By focusing on the most critical data, active learning algorithms can reduce the amount of data required for training and improve the accuracy of the algorithms. This approach can also incorporate new

data into the algorithms over time, making them more robust and up-to-date.

4) Explainable AI (XAI): These algorithms provide clear and interpretable explanations of their results, making them more transparent and trustworthy in industrial applications. By giving descriptions of the algorithms' behavior, explainable AI algorithms can increase the users' confidence in the results and help them understand the underlying decision-making processes. This approach can also help identify and address potential biases in the algorithms and improve their overall accuracy.

Overall, future research should develop algorithms that can effectively use multiple data sources, handle uncertainty, apply in real-time, and incorporate domain knowledge to improve the accuracy of fault diagnosis and prognosis in ICPS.

V. Remarks

The safety and security of industrial processes are of critical importance, and there is a growing need for innovative and efficient solutions to ensure the sustainable operation of ICPS. The current advancements in ICPS safety, including fault diagnosis and prognosis, provide a promising foundation for future research and development. However, several challenges still need to be addressed, including the integration of multiple techniques and the computational efficiency of algorithms.

To address these challenges, future research in the field of ICPS safety should aim to integrate multiple techniques, handle uncertainty, and incorporate domain knowledge, to improve the reliability and security of these systems. Additionally, efforts should be made to enhance the computational efficiency of existing algorithms so that they can be used in real-time industrial applications.

This paper provides a comprehensive perspective of the current state of the art in ICPS safety, focusing on maintaining secure and reliable operation in dynamic and uncertain environments. The identified challenges and potential solutions provide valuable insights for future research and development in the field of ICPS safety. The continued efforts to improve the safety and security of ICPS will play a critical role in ensuring a sustainable and secure industrial future.

References

[1] P. H. Winston, *Artificial intelligence*. Addison-Wesley Longman Publishing Co., Inc., 1984.

[2] H. Sedjelmaci, F. Guenab, S.-M. Senouci, H. Moustafa, J. Liu, and S. Han, "Cyber security based on artificial intelligence for cyber-physical systems," *IEEE Network*, vol. 34, no. 3, pp. 6–7, 2020.

[3] Z. Lv, Y. Han, A. K. Singh, G. Manogaran, and H. Lv, "Trustworthiness in industrial iot systems based on artificial intelligence," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1496–1504, 2020.

[4] H. Ning, H. Liu, and L. T. Yang, "Cyberentity security in the internet of things," *Computer*, vol. 46, no. 4, pp. 46–53, 2013.

[5] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on industrial informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.

[6] F. Khorrami, P. Krishnamurthy, and R. Karri, "Cybersecurity for control systems: A process-aware perspective," *IEEE Design & Test*, vol. 33, no. 5, pp. 75–83, 2016.

[7] H. Abie, "Cognitive cybersecurity for cps-iot enabled healthcare ecosystems," in *2019 13th International Symposium on Medical Information and Communication Technology (ISMICT)*. IEEE, 2019, pp. 1–6.

[8] J. Zhang, L. Pan, Q.-L. Han, C. Chen, S. Wen, and Y. Xiang, "Deep learning based attack detection for cyber-physical system cybersecurity: A survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 3, pp. 377–391, 2021.

[9] A. W. Colombo, S. Karnouskos, and T. Bangemann, "Towards the next generation of industrial cyber-physical systems," *Industrial cloud-based cyber-physical systems: The IMC-AESOP Approach*, pp. 1–22, 2014.

[10] C. Lu, A. Saifullah, B. Li, M. Sha, H. Gonzalez, D. Gunatilaka, C. Wu, L. Nie, and Y. Chen, "Real-time wireless sensor-actuator networks for industrial cyber-physical systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1013–1024, 2015.

[11] P. Leitao, S. Karnouskos, L. Ribeiro, J. Lee, T. Strasser, and A. W. Colombo, "Smart agents in industrial cyber–physical systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1086–1101, 2016.

[12] Y. Jiang, S. Yin, and O. Kaynak, "Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond," *IEEE Access*, vol. 6, pp. 47 374–47 384, 2018.

[13] H. Kayan, M. Nunes, O. Rana, P. Burnap, and C. Perera, "Cybersecurity of industrial cyber-physical systems: a review," *ACM Computing Surveys (CSUR)*, vol. 54, no. 11s, pp. 1–35, 2022.

[14] L. M. Oliveira, R. Dias, C. M. Rebello, M. A. Martins, A. E. Rodrigues, A. M. Ribeiro, and I. B. Nogueira, "Artificial intelligence and cyber-physical systems: A review and perspectives for the future in the chemical industry," *AI*, vol. 2, no. 3, p. 27, 2021.

[15] D. Li, H. Suo, and W. Liu, "System structure and network computing architecture of petrochemical cyber-physical system: Overview and perspective," *The Canadian Journal of Chemical Engineering*, vol. 97, no. 8, pp. 2176–2188, 2019.

[16] A. S. Mohammed, P. Reinecke, P. Burnap, O. Rana, and E. Anthi, "Cybersecurity challenges in the offshore oil and gas industry: an industrial cyber-physical systems (icps) perspective," *ACM Transactions on Cyber-Physical Systems (TCPS)*, vol. 6, no. 3, pp. 1–27, 2022.

[17] D. Wu, H. Zhu, Y. Zhu, V. Chang, C. He, C.-H. Hsu, H. Wang, S. Feng, L. Tian, and Z. Huang, "Anomaly detection based on rbm-lstm neural network for cps in advanced driver assistance system," *ACM Transactions on Cyber-Physical Systems*, vol. 4, no. 3, pp. 1–17, 2020.

[18] X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, "Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5790–5798, 2020.

[19] J. Pang, N. Zhang, Q. Xiao, F. Qi, and X. Xue, "A new intelligent and data-driven product quality control system of industrial valve manufacturing process in cps," *Computer Communications*, vol. 175, pp. 25–34, 2021.

[20] F. Liu, S. Zhang, W. Ma, and J. Qu, "Research on attack detection of cyber physical systems based on improved support vector machine," *Mathematics*, vol. 10, no. 15, p. 2713, 2022.

[21] R. Sethuraman, S. Sellappan, J. Shunmugiah, N. Subbiah, V. Govindarajan, and S. Neelagandan, "An optimized adaboost multi-class support vector machine for driver behavior monitoring in the advanced driver assistance systems," *Expert Systems with Applications*, vol. 212, p. 118618, 2023.

[22] D. M. Sharma and S. K. Shandilya, "An efficient cyber-physical system using hybridized enhanced support-vector machine with ada-boost classification algorithm," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 21, p. e7134, 2022.

[23] Q. Sun, K. Zhang, and Y. Shi, "Resilient model predictive control of cyber–physical systems under dos attacks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4920–4927, 2019.

[24] J. Chen and Y. Shi, "Stochastic model predictive control framework for resilient cyber-physical systems: review and perspectives," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2207, p. 20200371, 2021.

[25] Q. Sun and Y. Shi, "Model predictive control as a secure service for cyber–physical systems: A cloud-edge framework," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22 194–22 203, 2021.

[26] H. Li, A. Dimitrovski, J. B. Song, Z. Han, and L. Qian, "Communication infrastructure design in cyber physical systems with applications in smart grids: A hybrid system framework," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1689–1708, 2014.

[27] C. Kwon and I. Hwang, "Cyber attack mitigation for cyber–physical systems: hybrid system approach to controller design," *IET Control Theory & Applications*, vol. 10, no. 7, pp. 731–741, 2016.

[28] R. G. Sanfelice *et al.*, "Analysis and design of cyber-physical systems. a hybrid control systems approach," *Cyber-physical systems: From theory to practice*, pp. 3–31, 2016.

[29] S. Krishnamurthy, S. Sarkar, and A. Tewari, "Scalable anomaly detection and isolation in cyber-physical systems using bayesian networks," in *Dynamic Systems and Control Conference*, vol. 46193. American Society of Mechanical Engineers, 2014, p. V002T26A006.

[30] X. Lyu, Y. Ding, and S.-H. Yang, "Bayesian network based c2p risk assessment for cyber-physical systems," *IEEE Access*, vol. 8, pp. 88 506–88 517, 2020.

[31] D. Shi, R. J. Elliott, and T. Chen, "On finite-state stochastic modeling and secure estimation of cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 65–80, 2016.

[32] Z. Qu, Q. Xie, Y. Liu, Y. Li, L. Wang, P. Xu, Y. Zhou, J. Sun, K. Xue, and M. Cui, "Power cyber-physical system risk area prediction using dependent markov chain and improved grey wolf optimization," *IEEE Access*, vol. 8, pp. 82 844–82 854, 2020.

[33] Z. Tian, "An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring," *Journal of intelligent Manufacturing*, vol. 23, pp. 227–237, 2012.

[34] Z. Tian, L. Wong, and N. Safaei, "A neural network approach for remaining useful life prediction utilizing both failure and suspension histories," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1542–1555, 2010.

[35] J. B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network," *Mechanical Systems and Signal Processing*, vol. 56, pp. 150–172, 2015.

[36] G. F. Welch, "Kalman filter," *Computer Vision: A Reference Guide*, pp. 1–3, 2020.

[37] X. Wei, M. Verhaegen, and T. van Engelen, "Sensor fault detection and isolation for wind turbines based on subspace identification and kalman filter techniques," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 8, pp. 687–707, 2010.

[38] W. Zhao, T. Shi, and L. Wang, "Fault diagnosis and prognosis of bearing based on hidden markov model with multi-features," *Applied Mathematics and Nonlinear Sciences*, vol. 5, no. 1, pp. 71–84, 2020.

[39] J. S. Oakland, *Statistical process control*. Routledge, 2007.

[40] W. Härdle, L. Simar *et al.*, *Applied multivariate statistical analysis*. Springer, 2007, vol. 22007.

[41] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[42] J. Cha, "Partial least squares," *Adv. Methods Mark. Res*, vol. 407, pp. 52–78, 1994.

[43] S. Ding, P. Zhang, E. Ding, A. Naik, P. Deng, and W. Gui, "On the application of pca technique to fault diagnosis," *Tsinghua Science and Technology*, vol. 15, no. 2, pp. 138–144, 2010.

[44] H. Zhao, J. Zheng, J. Xu, and W. Deng, "Fault diagnosis method based on principal component analysis and broad learning system," *IEEE Access*, vol. 7, pp. 99 263–99 272, 2019.

[45] J. Zhu, T. Hu, B. Jiang, and X. Yang, "Intelligent bearing fault diagnosis using pca–dbn framework," *Neural Computing and Applications*, vol. 32, pp. 10 773–10 781, 2020.

[46] X. Wen and Z. Xu, "Wind turbine fault diagnosis based on relieff-pca and dnn," *Expert Systems with Applications*, vol. 178, p. 115016, 2021.

[47] S. Cao, Z. Hu, X. Luo, and H. Wang, "Research on fault diagnosis technology of centrifugal pump blade crack based on pca and gmm," *Measurement*, vol. 173, p. 108558, 2021.

[48] E. O. Brigham, *The fast Fourier transform and its applications*. Prentice-Hall, Inc., 1988.

[49] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—part i: Fault diagnosis with model-based and signal-based approaches," *IEEE transactions on industrial electronics*, vol. 62, no. 6, pp. 3757–3767, 2015.

[50] H. Chen and S. Lu, "Fault diagnosis digital method for power transistors in power converters of switched reluctance motors," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 2, pp. 749–763, 2012.

[51] N. Pan, X. Wu, Y. Chi, X. Liu, and C. Liu, "Combined failure acoustical diagnosis based on improved frequency domain blind deconvolution," in *Journal of Physics: Conference Series*, vol. 364, no. 1. IOP Publishing, 2012, p. 012078.

[52] I. Abdallah, V. Dertimanis, H. Mylonas, K. Tatsis, E. Chatzi, N. Dervili, K. Worden, and E. Maguire, "Fault diagnosis of wind turbine structures using decision tree learning algorithms with big data," in *Safety and Reliability–Safe Societies in a Changing World*. CRC Press, 2018, pp. 3053–3061.

[53] G. Li, H. Chen, Y. Hu, J. Wang, Y. Guo, J. Liu, H. Li, R. Huang, H. Lv, and J. Li, "An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators," *Applied Thermal Engineering*, vol. 129, pp. 1292–1303, 2018.

[54] O. Kherif, Y. Benmahamed, M. Teguar, A. Boubakeur, and S. S. Ghoneim, "Accuracy improvement of power transformer faults diagnostic using knn classifier with decision tree principle," *IEEE Access*, vol. 9, pp. 81 693–81 701, 2021.

[55] D.-T. Hoang and H.-J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, 2019.

[56] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics—a comprehensive review," *IEEE Access*, vol. 8, pp. 29 857–29 881, 2020.

[57] B. A. Tama, M. Vania, S. Lee, and S. Lim, "Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals," *Artificial Intelligence Review*, pp. 1–43, 2022.

[58] M. G. Alfarizi, J. Vatn, and S. Yin, "An extreme gradient boosting aided fault diagnosis approach: A case study of fuse test bench," *IEEE Transactions on Artificial Intelligence*, pp. 1–1, 2022. doi: 10.1109/TAI.2022.3165137

[59] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical systems and signal processing*, vol. 72, pp. 303–315, 2016.

[60] L. Nasser and M. Curtin, "Electronics reliability prognosis through material modeling and simulation," in *2006 IEEE Aerospace Conference*. IEEE, 2006, pp. 7–pp.

[61] W. Chalgham, K.-Y. Wu, and A. Mosleh, "System-level prognosis and health monitoring modeling framework and software implementation for gas pipeline system integrity management," *Journal of Natural Gas Science and Engineering*, vol. 84, p. 103671, 2020.

[62] M. G. Alfarizi, B. Tajiani, J. Vatn, and S. Yin, "Optimized random forest model for remaining useful life prediction of experimental bearings," *IEEE Transactions on Industrial Informatics*, pp. 1–10, 2022. doi: 10.1109/TII.2022.3206339

[63] F. Ustolin, F. Ferrari, and N. Paltrinieri, "Prediction of condensed phase formation during an accidental release of liquid hydrogen," *Chemical Engineering Transactions*, vol. 91, pp. 439–444, 2022.

[64] Y. Zhang, Z. Peng, Y. Guan, and L. Wu, "Prognostics of battery cycle life in the early-cycle stage based on hybrid model," *Energy*, vol. 221, p. 119901, 2021.

[65] X. Wu, Q. Ye, and J. Wang, "A hybrid prognostic model applied to sofc prognostics," *International Journal of Hydrogen Energy*, vol. 42, no. 39, pp. 25 008–25 020, 2017.

Article 5

Advancements in Extreme Gradient Boosting for Enhanced Fault Prognosis: A Continuation Study from Fuse Test Bench Analysis

M. G. Alfarizi, J. Vatn and S. Yin, "Advancements in extreme gradient boosting for enhanced fault prognosis: A continuation study from fuse test bench analysis," *2024 IEEE 33rd International Symposium on Industrial Electronics (ISIE)*, Ulsan, South Korea, pp. 1-8, 2024. (to be submitted)

This article is awaiting publication and is therefore not included.

ISBN 978-82-326-7538-8 (printed ver.)
ISBN 978-82-326-7537-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology