Volume 1 · Issue 3
December 2024

International Journal of AI for Materials and Design

CockpitGemini: A personalized design framework for smart vehicle cockpits integrating generative model-based multi-agent systems and human digital twins

ISSN 3041-0746 (Print)
ISSN 3029-2573 (Online)

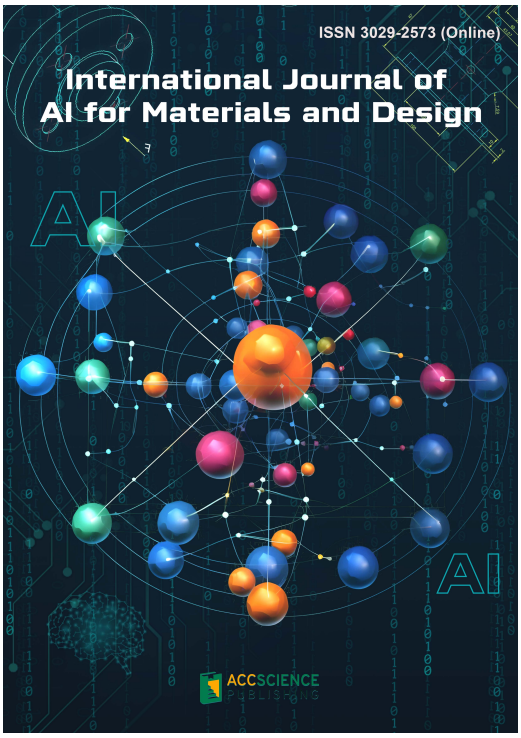
ACCSCIENCE PUBLISHING

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

Print ISSN: 3041-0746
Online ISSN: 3029-2573

The International Journal of AI for Material and Design (IJAMD) is a scholarly publication dedicated to advancing the intersection of artificial intelligence (AI), materials science and design. This peer-reviewed journal provides a platform for researchers, academics, and industry professionals to disseminate cutting-edge research, innovative methodologies, and practical applications that leverage AI techniques to enhance the understanding, development, and optimization of aspects related to materials and design processes. IJAMD seeks to contribute to the advancement of technology, innovation, and sustainability in materials design, engineering disciplines, product manufacturing and process technology.



About the Publisher

AccScience Publishing is a publishing company based in Singapore. We publish a range of high-quality, open-access, peer-reviewed journals and books from a broad spectrum of disciplines.

Contact Us

Managing Editor
ijamd.office@accscience.sg

AccScience Publishing
8 Burn Road, #15-03 Trivex, Singapore 369977.

Volume 1 • Issue 3 • December 2024
ISSN 3041-0746 (print) ISSN 3029-2573 (online)

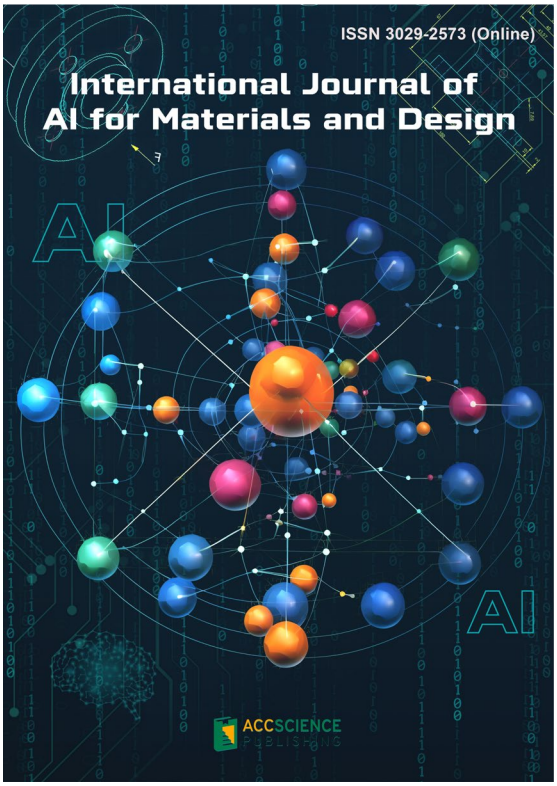
INTERNATIONAL JOURNAL OF AI FOR MATERIALS AND DESIGN

ACCSCIENCE PUBLISHING

Access Science Without Barriers

International Journal of AI for Materials and Design

ISSN: 3041-0746 (print)
ISSN: 3029-2573 (online)



Editorial and Production Credits

Publisher: AccScience Publishing
Managing Editor: Shirley Lu
Production Editor: Sharmila Velapasamy
Article Layout and Typeset: Sinjore Technologies (India)

For all advertising queries, contact ijamd.office@accscience.sg.

Supplementary file

Supplementary files of articles can be obtained at https://accscience.com/journal/IJAMD/1/3.

Disclaimer

AccScience Publishing is not liable to the statements, perspectives, and opinions contained in the publications. The appearance of advertisements in the journal shall not be construed as a warranty, endorsement, or approval of the products or services advertised and/or the safety thereof. AccScience Publishing disclaims responsibility for any injury to persons or property resulting from any ideas or products referred to in the publications or advertisements. AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of AI for Materials and Design

Editorial Board

*Editorial Board Members as of November 29, 2024

ACCSCIENCE PUBLISHING

Volume 1 | Issue 3 | December 2024 | Page 1 – 98 |

International Journal of AI for Materials and Design

https://accscience.com/journal/IJAMD

CONTENTS

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

EDITORIAL

Unlocking the potential of digital twins in heating, cooling, and thermal energy storage technologies

Angelo Maiorino and Fabio Petruzziello*

Department of Industrial Engineering, Università di Salerno, Fisciano, Salerno, Italy

(This article belongs to the *Special Issue: Cooling, Heating, and Thermal Energy Storage Systems: Harnessing the Power of Digital Twins*)

With the intensification of global concerns regarding energy consumption and environmental sustainability, developing innovative strategies to manage and optimize energy systems has become increasingly urgent. Heating and cooling systems are critical in this context because they account for a substantial portion of the global energy consumption and carbon emissions.[1] Effective management of such systems is essential to reduce operational costs, mitigate adverse environmental impacts, and facilitate the transition to sustainable energy practices.[2] A promising solution in this regard is the digital twin technology, a cutting-edge approach offering unprecedented capabilities in complex system simulations, analyses, and real-time management. A digital twin is a virtual replica of a physical system that is continuously updated with data from sensors and Internet of Things (IoT) devices.[3] The twin aims to accurately mimic the behavior of the system under various conditions. In the context of heating, cooling, and thermal energy storage systems, digital twins can drive substantial transformations, enabling improved performance monitoring, predictive maintenance, and energy optimization.[4-7] Hence, digital twins promise considerable improvements in system efficiency, reliability, and sustainability.

Thermal energy storage systems play a pivotal role in heating and cooling applications by storing excess energy for later use, thereby balancing supply and demand and improving the overall system efficiency.[8] Effective interactions among heating, cooling, and thermal energy storage systems are critical for optimal energy utilization. Digital twins offer a unique real-time platform for managing and fine-tuning these interactions.

The special issue "Cooling, Heating, and Thermal Energy Storage Systems: Harnessing the Power of Digital Twins" launched by the *International Journal of AI for Materials and Design* aims to explore and disseminate cutting-edge research and advancements in this rapidly evolving field of digital twin techniques applied to cooling, heating and thermal energy storage systems.

Digital twins rely on accurate numerical modeling and simulations to replicate the behavior of heating, cooling, and thermal energy storage systems. These simulations offer in-depth insights into the complex interactions among thermal dynamics, fluid mechanics, and material properties. Through such high-fidelity simulations, engineers, and researchers can optimize the design phase of cooling, heating and thermal energy storage systems by testing different design configurations and operational strategies to identify the most efficient and sustainable solutions.

Digital twins serve as powerful tools for improving system performance by simulating various operational scenarios. This capability is particularly valuable for optimizing

***Corresponding author:**
Fabio Petruzziello
(fpetruzziello@unisa.it)

Citation: Maiorino A, Petruzziello F. Unlocking the potential of digital twins in heating, cooling, and thermal energy storage technologies. *Int J AI Mater Design*. 2024;1(3):1-3.
doi: 10.36922/ijamd.4532

Received: August 15, 2024

Published Online: October 10, 2024

the performance of heating, cooling, and thermal energy storage systems, thereby maximizing energy efficiency and minimizing operational costs. By leveraging the data-driven insights derived from digital twins, the overall system efficiency can be enhanced, leading to increased sustainable energy usage.

The potential of digital twins extends directly to real-time control and automation. By leveraging the continuous data updates provided by physical systems to inform immediate decision-making processes, the operations of heating, cooling, and thermal energy storage systems can be dynamically adjusted, ensuring optimal performance and efficient energy management.

Another key application of digital twins is in fault detection and predictive maintenance. Early detection of anomalies or performance degradation in heating, cooling, and thermal energy storage systems is essential for maintaining operational efficiency and reliability. By continuously monitoring and comparing actual data with the predictions of a digital twin model, potential issues can be detected before they become severe problems. This enables the timely implementation of maintenance strategies that extend the lifespan of these systems and improve energy efficiency.

Offering accurate and timely data to digital twins is critical for the effective real-time management of heating, cooling, and thermal energy storage systems. Advanced sensors, thermal energy metering, and IoT devices play a crucial role in this regard by providing the real-time data required to keep digital twins updated and functioning optimally.

The ultimate aim of implementing the digital twin technology in heating and cooling systems is to improve energy efficiency and promote sustainability. By optimizing system performance, enabling real-time control, and facilitating predictive maintenance, digital twins can substantially reduce energy consumption and minimize the environmental impact.

By presenting pioneering studies on numerical modeling that contribute to the development of high-fidelity digital twins, particularly focusing on the intricate relations between heating, cooling, and thermal energy storage systems, this special issue aims to emphasize the key role of digital twins in fostering a more sustainable future for heating and cooling systems. This issue seeks to explore the integration of digital twins into control systems to enhance energy management. Moreover, it focuses on the latest advancements in sensor technology and IoT integration as well as the possibility of effectively managing heating, cooling, and thermal energy storage systems. Furthermore, it aims to investigate the potential of digital twins in suggesting predictive maintenance strategies to extend system lifespan and improve energy efficiency.

Researchers, practitioners, and industry experts are encouraged to submit original research articles, case studies, and reviews for publication in this special issue. As we strive toward a more energy-efficient and sustainable world, the importance of the digital twin technology in heating, cooling, and thermal energy storage systems cannot be overstated.

We look forward to your contributions, which will help shape the future of heating and cooling systems and foster innovation in the field of heating and cooling system.

Conflict of interest

Angelo Maiorino and Fabio Petruzziello are the Guest Editors of this special issue.

References

1. Maiorino A, Petruzziello F, Aprea C. Refrigerated transport: State of the art, technical issues, innovations and challenges for sustainability. *Energies*. 2021;14:7237.

   doi: 10.3390/en14217237

2. Balali Y, Chong A, Busch A, O'Keefe S. Energy modelling and control of building heating and cooling systems with data-driven and hybrid models--A review. *Renew Sustain Energy Rev*. 2023;183:113496.

   doi: 10.1016/j.rser.2023.113496

3. Jafari M, Kavousi-Fard AA, Chen T, Karimi M. A review on digital twin technology in smart grid, transportation system and smart city: Challenges and future. *IEEE Access*. 2023;11:17471-17484.

   doi: 10.1109/ACCESS.2023.3241588

4. Vering C, Borges S, Coakley D, Kruetzfeldt H, Mehrfeld P, Müller D. Digital Twin Design with On-line Calibration for HVAC Systems in Buildings. In: *Proceedings of Building Simulation 2021: 17th Conference of IBPSA*; 2021.

   doi: 10.26868/25222708.2021.30117

5. Vering C, Mehrfeld P, Nürenberg M, Coakley D, Lauster MR, Müller D. Unlocking Potentials of Building Energy Systems' Operational Efficiency: Application of Digital Twin Design for HVAC systems. In: 16th *International Conference of the International Building Performance Simulation Association* (*Building Simulation 2019*). Rome, Italy; 2019. p. 1304-1310.

   doi: 10.26868/25222708.2019.210257

6. Hodavand F, Ramaji IJ, Sadeghi N. Digital twin for fault detection and diagnosis of building operations: A systematic review. *Buildings*. 2023;13:1426.

doi: 10.3390/buildings13061426

7. Arowoiya VA, Moehler RC, Fang Y. Digital twin technology for thermal comfort and energy efficiency in buildings: A state-of-the-art and future directions. *Energy Built Environ*. 2024;5:641-656.

doi: 10.1016/j.enbenv.2023.05.004

8. Feng PH, Zhao BC, Wang RZ. Thermophysical heat storage for cooling, heating, and power generation: A review. *Appl Therm Eng*. 2020;166:114728.

doi: 10.1016/j.applthermaleng.2019.114728

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

CockpitGemini: A personalized design framework for smart vehicle cockpits integrating generative model-based multi-agent systems and human digital twins

Mengyang Ren[1,2], Junming Fan[1]*, Chunyang Yu[3], and Pai Zheng[1,2,3,4]*

[1]Department of Industrial and Systems Engineering, Faculty of Engineering, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China

[2]Laboratory for Artificial Intelligence in Design, Hong Kong Science Park, Hong Kong Special Administrative Region, China

[3]Design-AI Lab, China Academy of Art, China

[4]State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan, Hubei Province, China

Abstract

The evolution of smart vehicle cockpits is transitioning from serving as mere driving tools to becoming intimate partners that significantly enhance user experiences through advanced technologies. This research addresses the growing demand for personalized design in smart vehicle cockpits by proposing a framework, CockpitGemini. This framework integrates generative model-based multi-agent systems and human digital twins, enabling tailored designs and services based on user preferences and real-time status. The capabilities of the proposed framework are illustrated through four dimensions: personalized product design, personalized interactive interface design, user state monitoring and personalized regulation, and personalized driving strategy recommendations. A case study on the design of personalized vehicle seats demonstrates the feasibility and usability of the CockpitGemini framework, highlighting its potential to enhance user satisfaction in smart vehicle cockpits.

Keywords: Generative models; Multi-agent systems; Human digital twin; Personalized design; Immersive user experience; Smart vehicle cockpits

***Corresponding authors:**
Junming Fan
(jun-ming.fan@connect.polyu.hk);
Pai Zheng
(pai.zheng@polyu.edu.hk)

Citation: Ren M, Fan J, Yu C, Zheng P. CockpitGemini: A personalized design framework for smart vehicle cockpits integrating generative model-based multi-agent systems and human digital twins. *Int J AI Mater Design*. 2024; 1(3):4-19.
doi: 10.36922/ijamd.4220

Received: July 12, 2024

Accepted: August 14, 2024

Published Online: October 10, 2024

1. Introduction

With the rapid development of digitalization and intelligence technologies, the automotive industry is undergoing a transformative shift toward smart vehicles.[1] Smart vehicle cockpits, as a core component of smart vehicles, serve as the interfaces between the driver and the vehicle's advanced functionalities. They integrate advanced sensing, decision-making, and interaction technologies to provide drivers with a safe, efficient, and comfortable driving experience, evolving from a mere driving tool into an intimate partner[2] that supports a wide range of activities, including entertainment, communication, and productivity. In this context, the role of smart devices cannot be

overstated. Devices such as smartphones, wearables, and Internet of Things (IoT)-enabled gadgets are prevalent in our daily lives, facilitating seamless connectivity and access to information. For smart vehicles, this connectivity is leveraged to create a more integrated and responsive driving environment, allowing the cockpit to adapt to the driver's needs and preferences in real time. Generative artificial intelligence (GAI) has experienced unprecedented growth; with applications like Chat Generative Pre-Trained Transformer (ChatGPT) emerging as the fastest-adopted consumer software in history[3] and can further amplify the capabilities of smart vehicle cockpits. Technologies such as generative models and deep learning algorithms facilitate the processing of vast amounts of data from both the vehicle and the driver. This enables an advanced decision-making process that can lead to the development of personalized design, adaptive user interfaces, personalized driving strategies, and real-time monitoring and regulation of driver states.

Meanwhile, users have higher expectations for personalized and humanized cockpit design and anticipate that the cockpit can provide tailored products and services according to their own preferences and real-time status, so exploring the personalized design of the smart vehicle cockpits is of great significance.[4] Personalized design is crucial for addressing the diverse needs of users, enhancing the user experience (UX), fostering product innovation, and maintaining enterprise competitiveness.[5] The personalized design of smart vehicle cockpits offers drivers unique products, interactive interfaces, environment adjustments, and tailored driving strategies, thereby enhancing driving safety, comfort, and the overall UX. This approach not only satisfies the requirements of various clients but also advances the development of smart vehicle cockpits toward greater user-friendliness and intelligence. Currently, although some luxury vehicles have started to incorporate personalization features, the constraints of traditional vehicle architecture prevent most mainstream models from achieving true personalized design. Existing solutions typically involve basic parameter adjustments and lack a profound understanding of the driver's state and the capacity for flexible response. Furthermore, most current research is confined to individual technical aspects, such as the personalization of human-machine interfaces or driver assistance systems, without a holistic and comprehensive design framework. Integrating advanced AI techniques, such as generative models, multi-agent systems (MAS), and digital twins (DTs), into the personalized design of smart vehicle cockpits remains a significant challenge.

Considering existing constraints, generative model-based MAS and human digital twin (HDT) have demonstrated potential that surpasses traditional design methodologies. Generative models, exemplified by Generative Pre-trained Transformer 4 (GPT-4) released by OpenAI, have initiated a paradigm shift in the field of artificial intelligence (AI). These models enable users to intuitively generate a vast array of high-quality, multimodal content in a short time span,[6] encompassing text, images, video, and interactive 3D content (e.g., avatars, 3D models, and 3D environments[7]) based on users' instructions.[8] MAS are composed of autonomous entities known as agents, which possess inherent learning and decision-making abilities and can collaborate internally to address complex design tasks.[9] Moreover, generative models present a promising way for augmenting the capabilities of agents within MAS.

Meanwhile, HDT refers to the digital representation of human beings in the physical world,[10] which relies on the continuous collection of data through wearables and sensors that capture crucial health metrics, physical activities, personal preferences, and environmental interactions. This extensive data can reflect both physiological features and intrinsic cognitive characteristics of humans in a virtual space, thereby prioritizing physical and mental health in smart product-service systems.[11] Furthermore, HDT can perform activities on behalf of humans in virtual environments, overcoming the physical constraints of the real world and broadening the scope of human activities.[10]

Therefore, this study proposes CockpitGemini, a novel design framework that integrates generative models, MAS, and HDT technologies, enabling highly personalized design for smart vehicle cockpits. The framework can efficiently provide drivers with unique product designs, interactive interfaces, environment adjustments, and driving strategies based on their preferences and real-time status, thereby realizing an overall personalized experience. This integration enhances the personalization of services and products, ultimately leading to a more user-centric approach in the vehicle sector. The major innovation of CockpitGemini can be summarized as follows:

(i) This study presents a novel personalized design framework that combines generative model-based MAS with HDT models to enable the efficient delivery of personalized smart cockpit designs and services based on user preferences and states.
(ii) The four primary functions achieved by the innovative design framework are demonstrated in personalized product design, personalized interactive interface design, user state monitoring and personalized regulation, and personalized driving strategy recommendations.
(iii) An elaborated case study of personalized vehicle seat design is presented to show the feasibility and usability of the proposed personalized design framework.

The rest of the paper is organized as follows: Section 2 reviews the relevant studies. Section 3 outlines the system framework and implementation details of the proposed personalized design framework. In Section 4, a case study of personalized vehicle seat design for smart vehicle cockpits is presented. Limitations and future work are discussed in Section 5, and conclusions are provided in Section 6.

2. Literature review

The main focus of this research is achieving personalized design for smart vehicle cockpits through generative models, MAS, and HDT models. The state-of-the-art research is reviewed from three main aspects: (i) Personalized design methods, (ii) applications of generative models in design, and (iii) HDT-enabled design. Finally, the research identifies existing gaps in the literature.

2.1. Personalized design methods

With advancements in technology and production techniques, competition in an ever-changing and unpredictable global market has compelled manufacturers to transform their production models.[12] They have shifted their focus from products to customers, offering personalized design solutions and enabling customer participation in the product design process. Concurrently, numerous studies on personalized design have been conducted. For instance, Dou *et al.*[5] introduced an interactive genetic algorithm that allows users to actively engage in the personalized product design process, applying it to a personalized design system for automobile wheel hubs. The experimental results indicated that the algorithm enhanced system performance. Moreover, configuration has been recognized as a mainstream approach to achieving personalized design.[13,14] Dong *et al.*[15] proposed a knowledge graph-based configuration approach for mass personalization, considering users' affective and functional demands. Hu *et al.*[16] employed a heuristic process planning method to achieve personalized product configuration design. Ren *et al.*[17] introduced a proactive interaction method to recommend personalized services based on context-aware prediction in the targeted service scenarios. In addition, Zheng *et al.*[18] presented a three-model-based generic framework for UX-oriented product development for mass personalization, demonstrating its feasibility through a case study of a personalized smart wearable product. Zheng *et al.*[19] proposed a universal data-centric approach for developing personalized smart connected products within a cloud-based environment.

The essence of personalized products lies in addressing customer demands and enhancing user satisfaction. Customer-oriented strategies, such as configuration, interactive genetic algorithms for proactive user engagement,[6] and collaborative information tools,[20] are vital in achieving personalized design, particularly in the context of new product development. Nevertheless, current strategies often restrict consumers to limited choices, such as altering product colors or selecting configuration components. These approaches fall short of fully addressing consumers' affective and functional needs. This study aims to empower consumers to engage proactively and immersively in personalized product design through natural language interaction, focusing on both functions and design style aspects.

2.2. Applications of generative models in design

Large language models (LLMs), including ChatGPT, BERT,[21] LLaMA[22,] and other GAI techniques (e.g., DALL·E, Midjourney), belong to the category of generative models. These models can efficiently produce high-quality, multimodal digitalized content at a rapid pace by interpreting human-guided instructions.[6] Consequently, they present significant opportunities for design innovation. For instance, Wang *et al.*[23] introduced a task-decomposed approach that integrates LLMs with the function-behavior-structure (FBS) model to inspire designers during the design conceptualization phase. Jiang *et al.*[24] proposed AutoTRIZ, a design ideation tool powered by LLMs, to facilitate automated design and ideation processes. In addition, Yin *et al.*[25] utilized Midjourney, a text-to-image generative model, to support collaborative design in product-service systems.

However, leveraging a single LLM or multi-foundation model sometimes fails to effectively accomplish complex tasks. Consequently, generative model-based MAS are gaining popularity. These systems consist of multiple agents enhanced by LLMs or generative models that collaborate or compete to handle complicated design tasks more effectively. AutoGen[26] is an emerging multi-agent architecture for building generative applications, supporting customized and conversational agents composed of generative models, human inputs, and tools. This enables the system to assign different roles in processing domain-specific human expertise and integrating robust generative capabilities.[27] For instance, Chen *et al.*[28] combined generative models (i.e., LLM and Stable Diffusion 2.1) with the Who, What, Where, When, Why, How (5W1H) method, FBS model, and Kansei Engineering to generate innovative design solutions. Celen *et al.*[29] introduced I-Design, a personalized LLM-based interior designer that converts user prompts into personalized interior design scenes, further developing them into 3D design scenarios. Xu *et al.*[30] employed LLM-enabled MAS to simulate design sessions among different stakeholders during the collaborative design process.

Generative models have made significant advancements in interpreting human context and creating meaningful content. Concurrently, MAS can integrate generative models with domain expertise to simulate collaboration design processes and autonomously produce design results. Therefore, this paper investigates the application of MAS, enhanced by multiple generative models, in achieving personalized design solutions for smart vehicle cockpits.

2.3. HDT-enabled design

With the increasing popularity of DTs across various domains,[31-33] research on HDT has swiftly emerged, incorporating humans as the foundational physical entities to broaden the conventional scope of DTs.[34] The creation of HDT mainly relies on the dynamic collection of human-related data (e.g., external data, physiological data, behavioral data, social interaction data, and environment data) in the physical world, which is then sensed and transmitted to the digital realm.[35] This data transmission facilitates feedback mechanisms that can influence human physical entities directly or indirectly, such as by aiding decision-making processes.[34]

Furthermore, HDT has garnered much attention across various disciplines. For instance, Okegbile *et al.*[36] explored the opportunities and challenges of HDT for personalized healthcare. Chen *et al.*[37] proposed a mobile AI-generated content (AIGC)-based HDT system architecture and demonstrated its feasibility through case studies on personalized surgery planning and medication. Wang *et al.*[10] conducted a survey reviewing the development and future perspectives of HDT in the industrial sector. Fan *et al.*[38] leveraged a vision-based HDT modeling approach to recognize human status for human-centric manufacturing. Aboulsafa *et al.*[39] presented an educational HDT model for personalized e-learning. He *et al.*[40] summarized the progress from digital human modeling to HDT and further examined HDT from the perspective of human factors.

However, the implementation of HDT remains in its early stages, despite exhibiting key characteristics of DT, such as real-time response, high fidelity, and interoperability.[36] Furthermore, HDT can represent humans at various levels, ranging from individual body parameters to the entire human being.[41] Given its inherent complexity, developing a comprehensive HDT encompassing all aspects of the human being is currently unfeasible.[42] Consequently, HDT should be tailored to specific contexts and scenarios, focusing on a particular aspect of humans, such as specific body components, physiological features, or cognitive properties.[34] In this paper, we develop a physiology-related HDT to extend the users' activity range for design experiences in a virtual space and simulate human biological responses to support their decision-making and design iterations.

2.4. Research gaps

In summary, despite technological advancements and increasing user expectations, there remains a lack of systematic personalized design approaches for proactive user engagement in immersive environments. To address this research gap, we propose CockpitGemini, a generative model-based MAS and HDT-enabled personalized design framework for smart vehicle cockpits in immersive design space.

3. Methods

This section demonstrates the generative model-based MAS and HDT-empowered generic personalized design framework, main functions, and its specific implementation, respectively.

3.1. Generic personalized design framework

The proposed personalized design framework, integrating generative model-based MAS and HDT technologies, is shown in Figure 1. This framework efficiently accomplishes fundamental design tasks, provides preliminary design results, and enables design simulation and iteration within an immersive environment. It primarily consists of four components: the physical human, the digital human, the automatic design process, and the design simulation.

A physical user equipped with a virtual reality (VR) headset can voice input their design requirements to obtain design results through a multi-agent collaborative automatic design process. This process achieves personalized design by assigning specific roles to each agent. In addition, the user must wear various physical sensors, including eye trackers, inertial measurement units (IMUs), temperature sensors, and others, to collect multidimensional individual data, which serves as the basis for constructing the HDT.

The automatic design process is performed by multi-agent internal collaboration, enabled by multi-modal generative models or algorithms such as GPT4, Midjourney, StyleGAN,[43] and Dreamfusion.[44] Each agent is endowed with capabilities for perception, decision-making, and execution, allowing them to fulfill distinct roles and execute specific tasks. For example, the consumer agent functions as an automated user analyzer, extracting user requirement keywords from natural language inputs. The designer agent then utilizes these extracted requirements to engage modality-specific generative models, producing personalized design solutions. The engineer agent evaluates the generated conceptual designs and optimizes the details based on design constraints. Finally, the UX

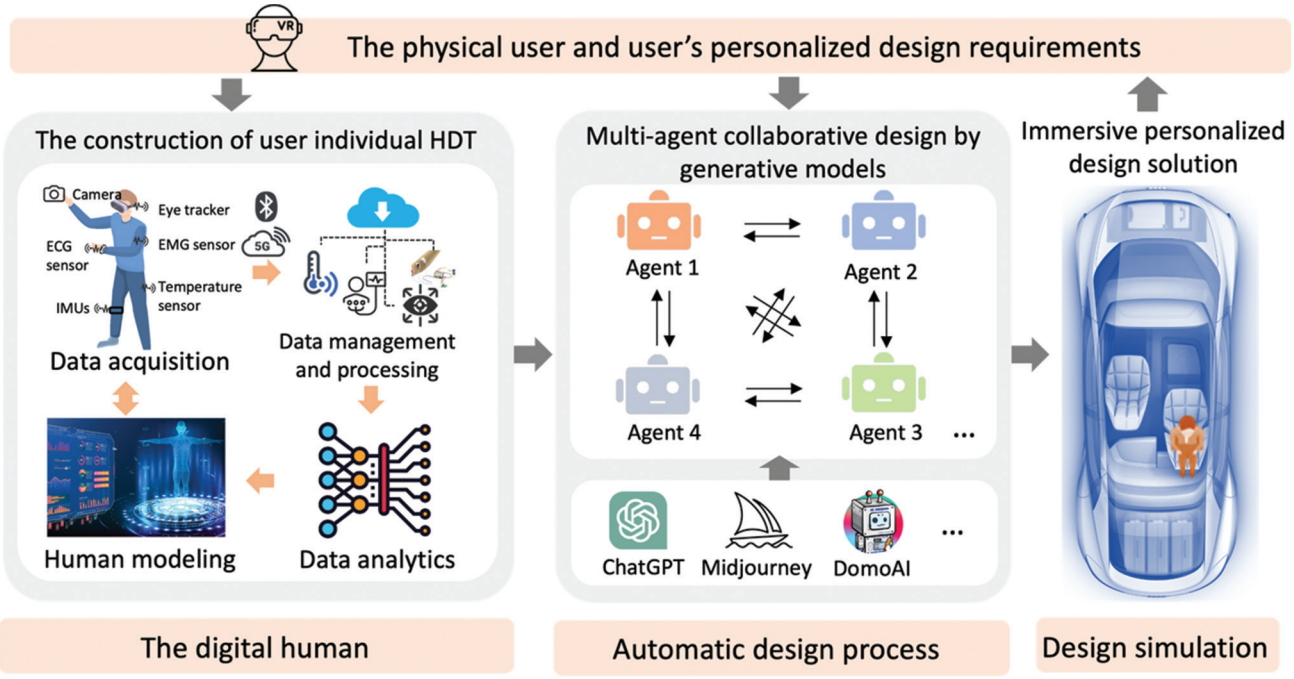


Figure 1. CockpitGemini: The proposed personalized design framework integrating generative model-based multi-agent systems and human digital twin technologies in an immersive environment

agent monitors and analyzes user behavior data from HDT to further discover user preferences, providing a foundation for design optimization and iteration. The user receives the design results from the human-in-the-loop multi-agent collaboration process, visualized in an immersive environment that enhances the sense of realism and improves the overall UX.[45]

HDTs can significantly enhance real-time status monitoring, design verification, performance evaluation, and iterative design within virtual environments, representing a critical component of immersive design. The development of HDT encompasses four primary stages: data acquisition, data management, data analytics, and human modeling. Initially, data acquisition forms the cornerstone of HDT, encompassing four key dimensions: physical data, physiological data, psychological data,[10,34] and biological data.[46] Physical data refers to the user's body features, such as anthropocentric data (e.g., standing height, seating height, etc.) obtained through RGB-D cameras,[47] physical movements captured by optical motion cameras, and scanning systems. Physiological data involves the digital representation of various biological signals measurable from the human body, including eye movement data, electrocardiogram, electromyogram, body temperature, and acceleration. These data provide insights into an individual's physiological state and responses to different stimuli or activities. Psychological data reflects users' mental status, such as emotion, fatigue, and stress, which can be inferred from partial or comprehensive physiological data. For instance, Tanwar *et al.*[48] leveraged hybrid deep learning models to process multi-modal physiological data (heart and electrodermal activity) to predict users' stress states. In addition, biological data refers to the data types at the organ or cell level, primarily applied in smart healthcare and medicine. Given the complexity and variety of data for building HDT, data management plays a crucial role in transmitting, storing, integrating, and managing heterogeneous data within databases. This is often achieved through cloud computing, wireless sensor networks, and database technologies. Subsequently, data pre-processing is performed to eliminate redundant and irrelevant data, involving processes such as data filtering, augmentation, standardization, and feature dimensional reduction.[49] Afterward, data analytics involves multi-dimensional user individual data using advanced algorithms to discover the internal mechanism of the human body and establish an individual HDT that is highly compatible with personal characteristics. Finally, human modeling entails the digital representation of physical counterparts, intuitively reflecting their behaviors and characteristics. The variety of HDT models aligns with the obtained data dimensions, including physical, psychological, and biological models. The physical model, driven by physical and physiological data, can reflect human body appearance, motion, ergonomics, and performing basic activities (e.g., walking, jumping) in a

virtual environment. The psychological facilitates real-time monitoring of the user's mental health status, providing a non-invasive method for continuous psychological well-being tracking. The biological model simulates specific internal mechanisms at the molecular and cellular levels. Given the diversity of data and the accuracy of HDT models, HDT can contribute to personalized design, verification, human-machine interaction experiences, and the applicability of human-centric product design.

3.2. Main functions and specific implementation

The main functions of the proposed personalized design framework for smart vehicle cockpits are illustrated in Figure 2. These functions are categorized into four main areas: (i) Personalized product design; (ii) personalized interactive interface design; (iii) user state monitoring and personalized regulation; and (iv) personalized driving strategy recommendations. The specific implementation details for each aspect are elaborated below.

3.2.1. Personalized product design

Utilizing generative model-based MAS and HDT technologies to assist in personalized cockpit product design can significantly enhance design efficiency and user satisfaction. First, a detailed HDT model, including physical model, psychological, and biological dimensions, is created for each user. This model integrates the user's body parameters, preferences, and mental status, derived from computer vision technology,[50] collected heterogeneous data, and user-provided information. During the automated design process, the MAS comprises a user agent, designer agent, engineer agent, and UX analyst agent. The user agent leverages the HDT model to simulate the user's actual experience in the virtual cockpit, providing detailed design requirements and personal references to the designer and engineer. The designer agent is responsible for generating the design solution and balancing esthetics and functionality. The engineer agent ensures the manufacturability and structural integrity of the design solution. Finally, the UX analyst agent evaluates the UX factors of the design and offers suggestions to the designer, thus facilitating the collaborative design process.

The automatic design process begins with the user agent extracting the user's inputting personalized demands, such as the design style of the vehicle seat. The designer agent then retrieves or generates a preliminary cockpit design based on these specific requirements, employing a retrieval model, generative adversarial networks (GANs), or other generative models to ensure design variety and innovation. The engineer agent then evaluates the manufacturability and structural integrity of the preliminary design. If issues are identified, the engineering agent suggests modifications and provides feedback to the designer agent. The UX analyst agent assesses the UX aspect of the design based on the user's HDT, using simulation and emulation techniques to model the user's actions within the cockpit and assess comfort and convenience. Based on feedback from the UX analyst agent, the designer agent and the engineer agent collaborate to optimize the design. Throughout this process, the user agent continuously provides feedback to ensure that the design remains aligned with the user's needs.

After several iterations, the final cockpit design is generated. Through VR technology, users can experience the final design in an immersive environment and provide final feedback. The final design is then validated by the engineer's agent before moving into the production phase. This approach achieves highly personalized design through multi-agent collaboration and HDT technologies,

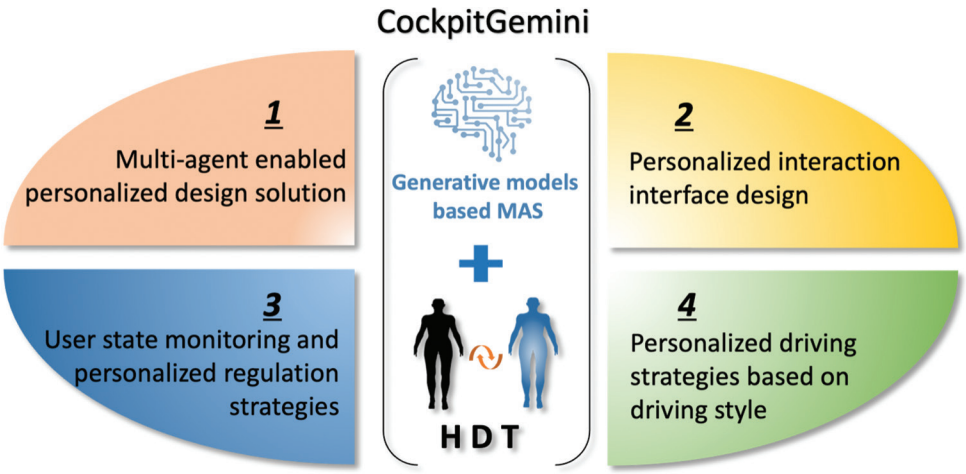


Figure 2. The main functions of generative model-based multi-agent systems and human digital twin based personalized design framework in virtual space

meeting the unique needs of each user, improving design efficiency, shortening the design cycle, and ensuring continuous feedback from the user agent and UX analyst agent to guarantee user satisfaction. The integration of these technologies can significantly enhance the quality and efficiency of personalized cockpit product design.

3.2.2. Personalized interactive interface design

The design of personalized interactive interfaces is crucial for enhancing UX and safety. An individual user HDT model can capture the user's habits and provide support for personalized interactive interface design by understanding specific preferences in the information display. These preferences include visual elements (e.g., font size, color contrast, area of interest), information needs (e.g., real-time speed, navigation paths, and awareness of the surrounding environment and pedestrians), interaction modes (e.g., voice control and touch control), and safety alert preferences (e.g., frequency and form of warning messages).

MASs and HDT technologies play a key role in personalized interaction design, enabling real-time perception and analysis of environmental information around the vehicles as well as the user's state and requirements. The system can comprise three agents: the environment-aware agent, the user-behavior agent, and the interaction agent, as illustrated in Figure 3. The environment-aware agent employs a vision-language model to conduct real-time perception of road conditions, traffic signals, other vehicles, and pedestrians. The user-behavior agent monitors the user's physiological status derived from the individual HDT. The interaction agent dynamically adjusts the interactive interfaces according to the user's preferences and the current environment information to provide real-time feedback and suggestions. This includes adjusting the information hierarchy, displaying real-time information about the vehicle's status, designing various interface styles, and providing warning information to meet the user's needs. By sharing information and collaborating, each agent ensures rapid response and personalized service. For instance, when the environment-aware agent detects an obstacle in front of the car, the user-behavior agent analyzes the user's attention state and reaction ability, and the interaction agent can initiate a timely warning to the user and suggest appropriate measures.

Simulating diverse driving situations (e.g., urban environments, highways, inclement weather conditions) within a virtual driving environment and assessing the performance of various interactive interfaces under these can provide a robust evaluation of the interaction design's applicability and the precision of information transmission in real-world driving contexts.

3.2.3. User state monitoring and personalized regulation

The autonomous decision-making capabilities of the MAS and the precise simulation of the HDT technology

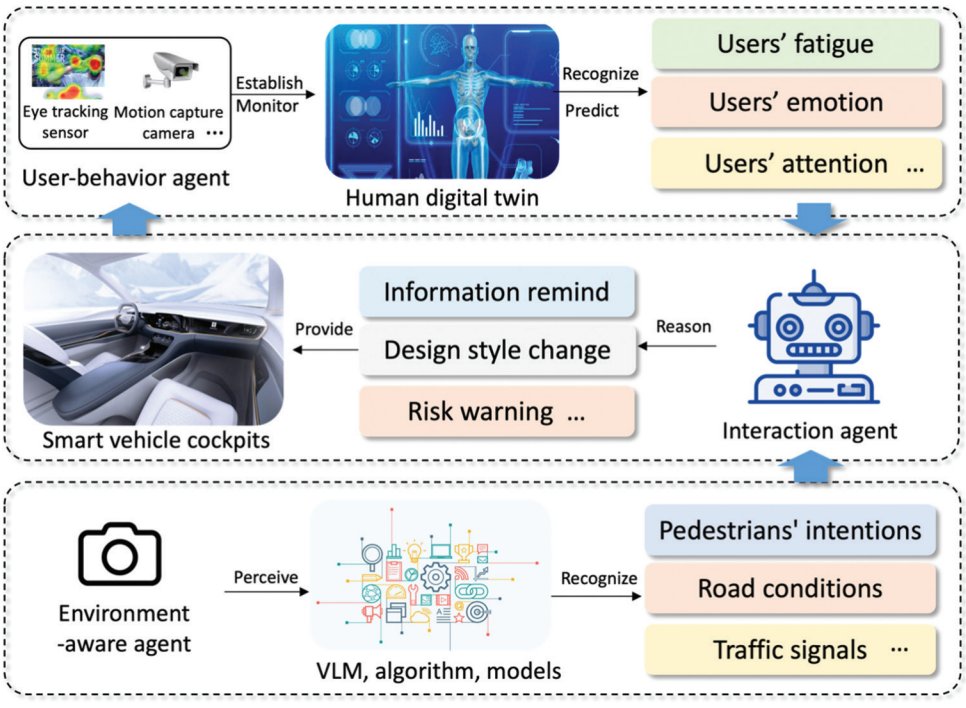


Figure 3. The personalized interactive interface design empowered by the multi-agent system and human digital twin

allow for comprehensive state monitoring (e.g., emotion recognition, fatigue detection, stress monitoring, and health assessment) and enable personalized adjustment to the passenger's condition. This significantly enhances the mobility experience, safety, and efficacy of personalized regulation strategies. The collection of user data through sensors (e.g., wearables, cameras, microphones) supports the construction and maintenance of the HDT model, ensuring continuous updates that accurately reflect users' physiological and mental status in real time.

MASs comprise a perception agent, a decision-making agent, and an execution agent. The perception agent is responsible for monitoring and collecting user status and intensity information. The decision-making agent analyzes this status information and devises regulation strategies. The execution agent is tasked with the actual implementation of conditioning measures, such as adjusting music or providing health reminders and rest advice to drivers or passengers. The HDT detects user states and transmits this information to the perception agent of the MAS, ensuring that the decision-making and execution agents base their decisions and operations on accurate and up-to-date information. For instance, the perception agent detects the frequency of the driver's eye and head movement in real time and determines that the driver may be experiencing fatigue. This status information is then broadcast to all related agents. Based on the recognized fatigue and intensity of the driver, the decision-making agent decides to implement measures such as adjusting the seat angle, playing soothing music, initiating proactive, empathetic speech, and adjusting the vehicle's interior lighting. These tasks are assigned to different execution agents, such as the seat control agent, the entertainment system agent, and the environment control agent. Each execution agent then takes actions to adjust interior temperature and lighting, play soothing music, and adjust the seat angle according to the instructions. The perception agent continuously monitors the driver's status and transmits feedback to the decision-making intelligence, which adjusts the strategy based on this feedback to ensure that the passengers regain their comfort.

3.2.4. Personalized driving strategy recommendations

The integration of the HDT model and the MAS can recommend personalized driving strategies to drivers. Historical driving behavior data, including metrics such as driving speed, acceleration, gas pedal angle, brake pedal frequency and force, and steering angle, are collected to construct an HDT model. These data are then classified into different driving styles (conservative, aggressive, and moderate) using advanced machine learning and deep learning algorithms. In this framework, the MAS comprises a perception agent, a decision-making agent, and an execution agent. The perception agent employs generative models to collect visual information and recognize the surrounding environment. The decision-making agent generates personalized driving strategies based on the driver's identified driving style and real-time environmental conditions. Multi-modal feedback, such as voice reminders and visual cues, is provided to the driver. For example, if the driver's HDT model identifies an aggressive driving style and the perception agent detects that the driver is currently navigating in rainy weather with care and several pedestrians crossing the road ahead, the decision-making agent generates a deceleration command to remind the driver to stay focused. The executive agent then issued a voice alert to the driver, stating, "There are pedestrians ahead; please slow down to 30 km/h and pay attention to safety."

4. Case study

In this section, we leverage the proposed CockpitGemini personalized design methodology to execute a personalized product design and present a case study focusing on the personalized automobile seat design to showcase the viability of the proposed approach within a targeted context. It is demonstrated from the following two aspects: individual HDT construction and generative model-based MAS for personalized seat design, respectively.

4.1. Individual HDT construction

We employed a vision-based HDT modeling approach to capture customers' RGB-D data in real time using an Azure Kinect camera, establish individualized human meshes, and perform ergonomic analysis in a virtual environment. The implementation details of the vision-based HDT modeling approach are depicted in Figure 4. The model comprises an advanced convolution neural network backbone, a body part segmentation branch, and an ergonomics branch. The input data consists of RGB-D information on the human's various gestures. Subsequently, a high-resolution network (HRNet-W32[51]) is utilized to develop the body segmentation and ergonomics branches, fulfilling the perceptual requirements of the HDT modeling process. This results in the creation of a highly personalized human model. The HDT perception model is implemented utilizing the widely adopted PyTorch deep learning framework, with Nvidia RTX 3090 GPU providing enhanced hardware acceleration. The Adam optimizer is employed with a learning rate of 5×10^{-5}, a batch size of 2, and a total of 100 training epochs.

Regarding the part segmentation branch, we introduced a body part attention mechanism to construct an HDT

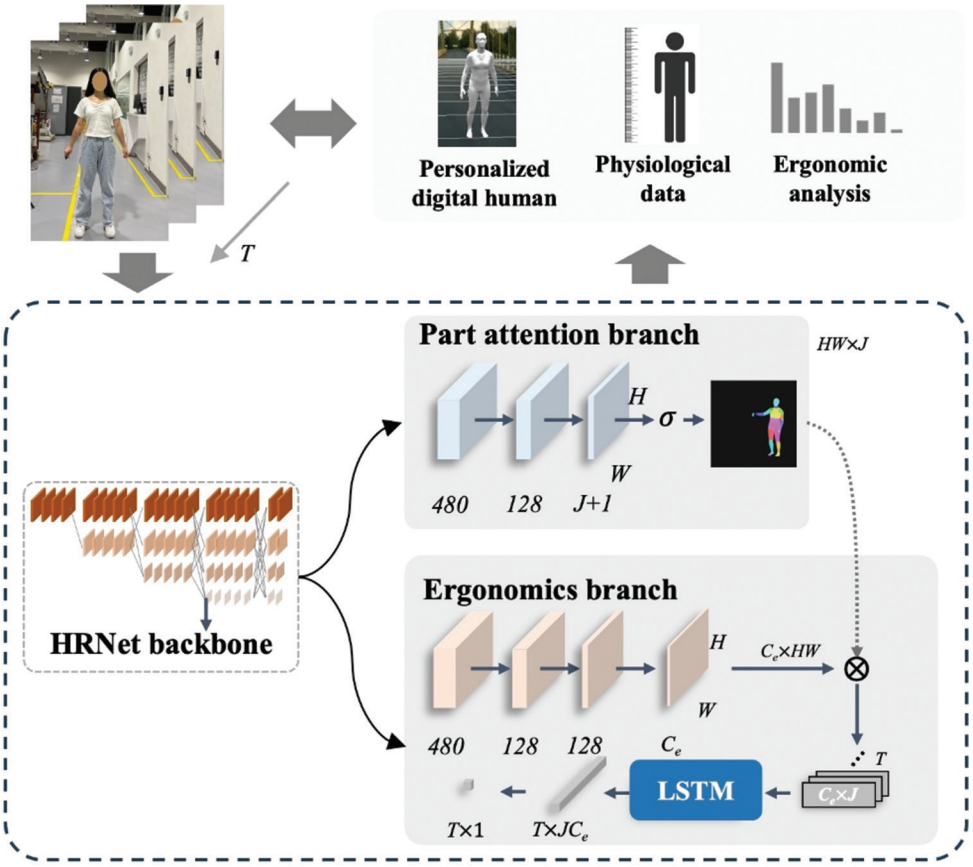


Figure 4. The implementation detail of the vision-based human digital twin modeling approach
Abbreviation: LSTM: Long short-term memory.

model by specifying body positions and highlighting different body parts. Based on the intermediate features F_{int} extracted from the backbone network, the feature map is initially reduced using two 3 × 3 convolutional layers with 480 and 128 channels, respectively. Subsequently, it is compressed by another convolutional layer to $(J+1)\times H\times W$. Here, J = 24 denotes the number of body joints and parts, with the additional channel denoting the background, and $H = W$ =56 indicates the height and width of the feature map. Utilizing the soft-max function σ, the feature map is converted into a partial attention mask $F_{part} \in \mathbb{R}^{J\times H\times W}$, where each pixel value indicates the probability that the pixel value belongs to a specific body part. Finally, the part attention mask is applied to the task-specific feature maps of the following branches to constrain the features to focus on specific body regions. To facilitate subsequent multiplication, F_{part} is reshaped to $J \times H \times W$.

For the ergonomics branch, it is significant to perform automatic ergonomic analyses to develop personalized product design solutions. In this work, we utilized the rapid entire body assessment (REBA)[52] ergonomic assessment tool to evaluate human postures on a scale of 1 – 15. The REBA scores were subsequently processed using an additional neural network within the ergonomics branch. Key components of the HDT model could achieve 3D human mesh reconstruction, anthropometric data measurement, and REBA score prediction.

To quantitatively assess the performance of the HDT perception model, a comparative experiment was conducted, utilizing mean error for the regression of ergonomic scores. Figure 5 displays various examples of results generated by the HDT model, offering a visual representation that aims to provide an intuitive understanding of its performance. As indicated in Table 1, our model outperforms the multi-task learning (MTL) baseline model.[53]

4.2. Generative model-based MAS for personalized seat design

Our MAS originates from the personalized requirements of users and progresses through three specialized agents: the user requirement analysis agent, the designer agent, and the design verification agent. These agents collaboratively derive feasible design solutions that meet the specified

Table 1. Evaluation results of the human digital twin perception model

Method	Ergonomic scores mean error
MTL[53]	0.83
Ours	0.66

Abbreviation: MTL: Multi-task learning.

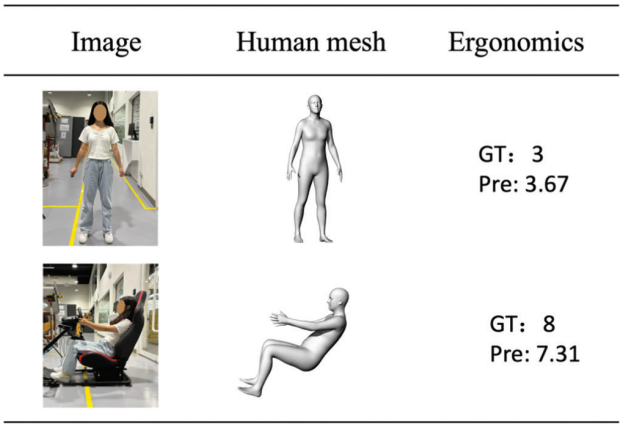


Figure 5. Qualitative examples of the human digital twin perception model
Abbreviations: GT: Ground truth; Pre: Prediction.

criteria of users. The internal collaborative mechanisms of generative model-based MAS for personalized seat design are illustrated in Figure 6.

4.2.1. User requirement analysis agent

The capture and expression of user requirements are crucial components of the product design process. User's emotional needs are often implicit and challenging to fully quantify through traditional research methods. Kansei Engineering provides a methodology to translate users' esthetic demands into design parameters by utilizing Kansei vocabulary to express users' affective responses. When combined with the powerful natural language processing capabilities of LLMs, this approach enables a more systematic and accurate generation of prompts for design inputs, guiding the designer agent to create more user-friendly designs that meet user expectations. Initially, user's voice input detailing their personalized emotional demands. For instance, users might describe their desire for a new automobile seat to "be comfortable to use" or "look fashionable." The user requirements analysis agent must identify Kansei words relevant to the design problem. These words can include adjectives that characterize the user's desired feelings, such as "comfortable," "fashionable," or nouns that represent sources of inspiration, such as "nature." Subsequently, the LLM can reason and output visual design elements such as shapes, colors, and materials based on the identified Kansei vocabulary.

For example, if a user describes their requirement as "I want a vintage vehicle seat," the system needs to deconstruct the concept of "vintage" into its key design elements to create a specific prompt for generating a vintage vehicle seat. The system identifies "vintage" as a key Kansei word and maps it to design elements in terms of shape, color, and materials. The specific prompt output by GPT-4o-mini for a vintage vehicle seat is described as follows:

(i) Shape: Incorporate curved lines and ornate details typical of vintage furniture. The overall form should be reminiscent of classic armchairs or lounge chairs from the early 20th century.
(ii) Color: Use warm, rich tones such as mahogany, deep browns, or muted greens. The color palette should evoke a timeless and elegant feel.
(iii) Materials: Utilize high-quality materials such as leather for the upholstery, with tufted detailing to enhance the vintage look. The frame should be made of polished wood and include brass accents or embellishments to add an extra touch of sophistication.

4.2.2. Designer agent

In this subsection, the TAPS3D model,[54] a state-of-the-art text-to-3D generative model, has been fine-tuned to function as a designer agent for creating personalized vehicle seat design solutions. By leveraging advanced AI technologies, TAPS3D can generate high-quality, textured 3D meshes with arbitrary topologies based on design prompts derived from user requirement analysis agents. This functionality enables the model to address both esthetic and functional specifications provided by customers. The implementation details for the text-to-3D generative model are illustrated in Figure 7, which primarily encompasses three components: pseudo caption generation, 3D textured mesh generator, and 3D textured mesh discriminator.

First, a 3D object dataset is rendered into multi-view images using Blender v2.90.0. The system then leverages the contrastive language-image pre-training (CLIP) model,[55] a pre-trained model for image-text similarity and zero-shot image classification, to establish pseudo-captions. These pseudo captions are subsequently encoded as embeddings by the CLIP model. The caption embeddings, along with two random noise vectors, are fed into the 3D textured mesh generator backbone from GET3D,[55] which controls the 3D shape and texture, respectively. Finally, two discriminators (an RGB image and a silhouette discriminator), based on the StyleGAN[43] architecture, evaluate the authenticity of the generated objects.

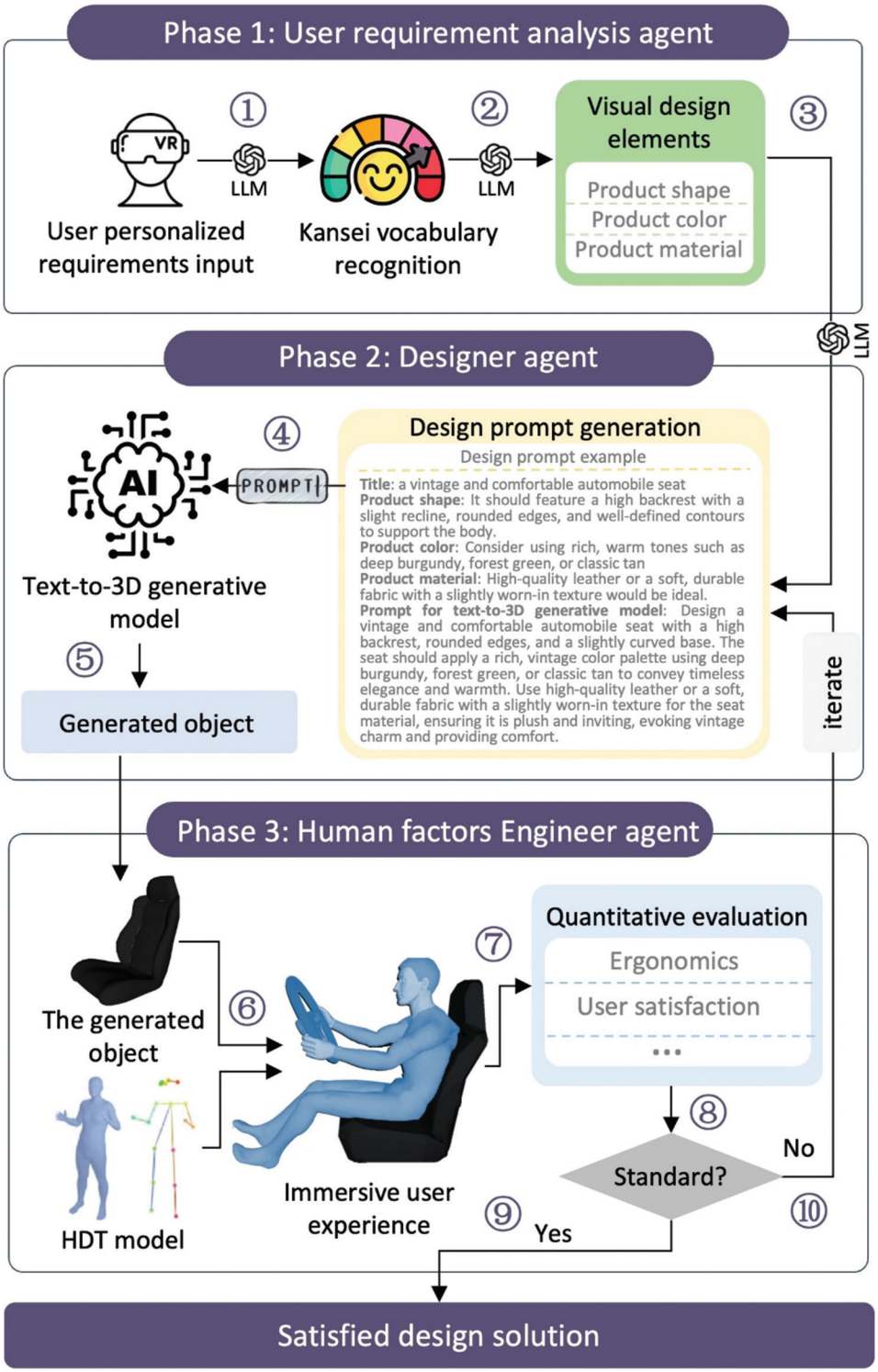


Figure 6. The internal collaborative mechanism of generative model-based multi-agent systems for personalized seat design
Abbreviations: HDT: Human digital twin; LLM: Large language models.

A total of 280 three-dimensional vehicle seat object models were extracted from the Objaverse dataset[56] for the purpose of training a generative model of vehicle seats. Manual data pre-processing was conducted using Blender 3.6.1. The dataset was randomly split into a training set (70%), a validation set (10%), and a test set (20%), with

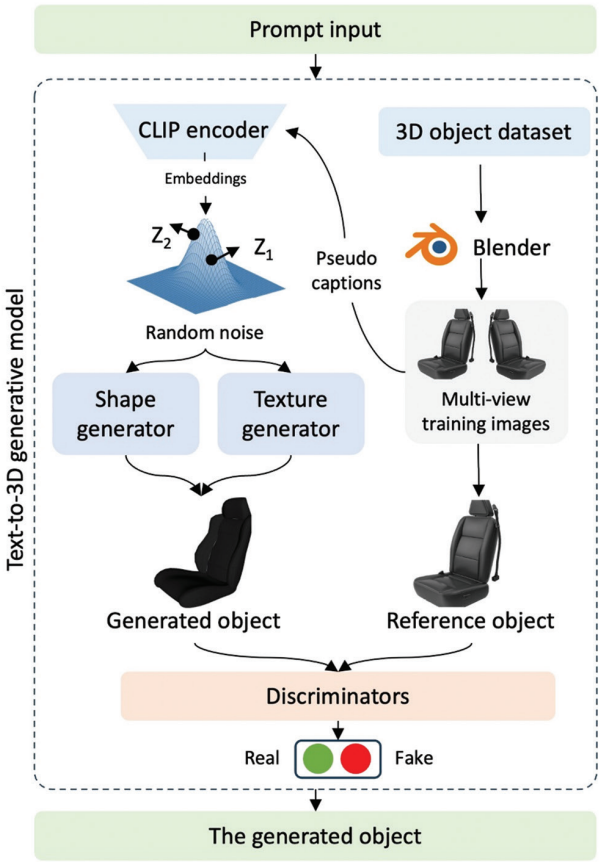


Figure 7. The implementation details of text-to-3D generative model
Abbreviation: CLIP: Contrastive language-image pre-training.

measures taken to ensure the removal of any duplicated test samples from the training set. To augment the diversity of the training dataset, each object was rendered from 100 distinct random perspectives.

4.2.3. Human factors engineer agent

To conduct an automated ergonomic analysis using an HDT model interacting with a generated vehicle seat, the process initiates with the creation of a virtual environment. This entails developing a comprehensive 3D model of the smart vehicle cockpit, encompassing the generated vehicle seat and other relevant components such as the steering wheel, dashboard, and control interfaces. Detailed anthropometric data and REBA scores are obtained from the vision-based HDT, which replicates the user's physical characteristics and behaviors. The digital twin is programmed to perform various activities related to interacting with the vehicle seat, such as sitting down, adjusting the seat position, reaching for controls, and maintaining different postures, ensuring that the movements are realistic and grounded in actual human biomechanics.

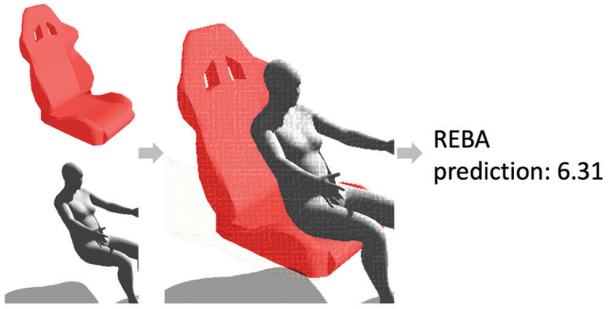


Figure 8. The ergonomic analysis utilizing human digital twin and the generated objects
Abbreviation: REBA: Rapid entire body assessment.

During the interaction simulation, data on the human body segment positions are collected and analyzed using the ergonomics branch to assess the musculoskeletal disorder risk on different body parts. The ergonomic assessment involves evaluating the comfort and accessibility of the generated objects, ensuring that the HDT model can comfortably reach all necessary controls and interfaces, and that all displays are within the line of sight. Ease of use is assessed by simulating the HDT operating various controls and using comfort metrics such as seat pressure distribution and back support. Potential ergonomic risks, such as awkward postures, repetitive movements, and excessive force requirements are identified. An automated feedback loop uses performance metrics to suggest design modifications, which are then used to update the generative model and create new iterations of the vehicle seat design. This iterative process ensures continuous refinement and validation of the design through virtual simulations and real-world testing, optimizing the vehicle seat for user comfort and safety. The ergonomic analysis process utilizing a vision-based HDT approach and the generated vehicle seat from a text-to-3D generative model is shown in Figure 8.

5. Limitations and future work

Despite the promising potential of the CockpitGemini framework, several limitations need to be addressed in future research. First, the technology for generating 3D objects using large generative models is still in its early stages. The current state of 3D generative models does not yet produce outputs that meet product-level quality standards. This limitation impacts the overall effectiveness of the personalized design framework, as the generated 3D models may lack the precision and detail required for practical application in smart vehicle cockpits. Consequently, further advancements in 3D generative modeling are necessary to enhance the fidelity and usability of the generated designs. Second, the

current approach to establishing the HDT model relies predominantly on vision recognition, which provides a limited perspective of the user's state and preferences. This single-channel method fails to incorporate data from other potential sources, such as physiological sensors, voice recognition, and behavioral analysis, which are essential for creating a comprehensive and accurate digital twin. The lack of multi-modal data integration restricts the ability to fully capture and respond to the user's needs, thereby limiting the personalization capabilities of the system. Future research should focus on integrating diverse data channels to build a more holistic and robust HDT model.

6. Conclusion

This research proposes a personalized design framework for smart vehicle cockpits, CockpitGemini, which integrates generative model-based MAS and HDT technologies. The proposed framework addresses the growing demand for personalized vehicle experiences by providing unique designs and services tailored to individual user preferences and states. The framework's capabilities are demonstrated through four key aspects: personalized product design, personalized interactive interface design, user state monitoring and personalized regulation, and personalized driving strategy recommendations. The feasibility and usability of CockpitGemini are further demonstrated through a detailed case study focusing on personalized cockpit element design, specifically a personalized vehicle seat. This study underscores the potential of integrating advanced technologies to enhance UX and satisfaction in smart vehicle cockpits, paving the way for future innovations in personalized design.

Acknowledgments

None.

Funding

This research is partially funded by the Collaborative Project funded by Design-AI Lab, China Academy of Art, China (No.: CAADAI2022A002) and State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology (No.: IMETKF2024010).

Conflict of interest

Pai Zheng serves as the Editorial Board Member of the journal but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, 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.

Author contributions

Conceptualization: Mengyang Ren
Investigation: Mengyang Ren, Junming Fan
Methodology: All authors
Writing – original draft: Mengyang Ren
Writing – review & editing: Junming Fan, Chunyang Yu, Pai Zheng

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data used in this work are available from the corresponding author upon reasonable request.

References

1. Chen L, Li Y, Huang C, *et al*. Milestones in autonomous driving and intelligent vehicles: Survey of surveys. *IEEE Trans Intell Veh*. 2023;8(2):1046-1056.

   doi: 10.1109/TIV.2022.3223131

2. Li W, Wu L, Wang C, *et al*. Intelligent cockpit for intelligent vehicle in metaverse: A case study of empathetic auditory regulation of human emotion. *IEEE Trans Syst Man Cybern Syst*. 2023;53(4):2173-2187.

   doi: 10.1109/TSMC.2022.3229021

3. Stappen L, Dillmann J, Striegel S, Vögel HJ, Flores-Herr N, Schuller BW. Integrating Generative Artificial Intelligence in Intelligent Vehicle Systems. In: *IEEE Conference on Intelligent Transportation Systems, Proceedings*, (*ITSC*). United States: Institute of Electrical and Electronics Engineers Inc.; 2023. p. 5790-5797.

   doi: 10.1109/ITSC57777.2023.10422003

4. Chen H, Gao R, Fan L, *et al*. Scenario-function system for automotive intelligent cockpits: Framework, research progress and perspectives. *IEEE Trans Intell Veh*. 2024;9:4890-4904.

   doi: 10.1109/TIV.2024.3382995

5. Dou R, Lin D, Nan G, Lei S. A method for product personalized design based on prospect theory improved with interval reference. *Comput Ind Eng*. 2018;125:708-719.

   doi: 10.1016/j.cie.2018.04.056

6. Cao Y, Li S, Liu Y, *et al*. A Comprehensive survey of AI-Generated Content (AIGC): A history of generative AI

from GAN to ChatGPT. 2023.

doi: 10.48550/arXiv.2303.04226

7. Qin HX, Hui P. Empowering the Metaverse with Generative AI: Survey and Future Directions. In: *Proceedings-2023 IEEE 43rd International Conference on Distributed Computing Systems Workshops*, (*ICDCSW 2023*). United States: Institute of Electrical and Electronics Engineers Inc.; 2023. p. 85-90.

doi: 10.1109/ICDCSW60045.2023.00022

8. Ren M, Zheng P. Towards smart product-service systems 2.0: A retrospect and prospect. *Adv Eng Inform*. 2024;61:102466.

doi: 10.1016/j.aei.2024.102466

9. Dorri A, Kanhere SS, Jurdak R. Multi-agent systems: A survey. *IEEE Access*. 2018;6:28573-28593.

doi: 10.1109/ACCESS.2018.2831228

10. Wang B, Zhou H, Li X, *et al*. Human digital twin in the context of industry 5.0. *Robot Comput Integr Manuf*. 2024;85:102626.

doi: 10.1016/j.rcim.2023.102626

11. Barricelli BR, Casiraghi E, Gliozzo J, Petrini A, Valtolina S. Human digital twin for fitness management. *IEEE Access*. 2020;8:26637-26664.

doi: 10.1109/ACCESS.2020.2971576

12. Wang Z, Zhang M, Sun H, Zhu G. Effects of standardization and innovation on mass customization: An empirical investigation. *Technovation*. 2016;48-49:79-86.

doi: 10.1016/j.technovation.2016.01.003

13. Myrodia A, Kristjansdottir K, Hvam L. Impact of product configuration systems on product profitability and costing accuracy. *Comput Ind*. 2017;88:12-18.

doi: 10.1016/j.compind.2017.03.001

14. Levandowski CE, Jiao JR, Johannesson H. A two-stage model of adaptable product platform for engineering-to-order configuration design. *J Eng Des*. 2015;26(7-9):220-235.

doi: 10.1080/09544828.2015.1021305

15. Dong L, Ren M, Xiang Z, Zheng P, Cong J, Chen CH. A novel smart product-service system configuration method for mass personalization based on knowledge graph. *J Clean Prod*. 2023;382:135270.

doi: 10.1016/j.jclepro.2022.135270

16. Hu K, Qiu L, Zhang S, Wang Z, Fang N. ICCP: A heuristic process planning method for personalized product configuration design. *Appl Intell*. 2023;53(24):30887-30910.

doi: 10.1007/s10489-023-05186-z

17. Ren M, Dong L, Xia Z, Cong J, Zheng P. A proactive interaction design method for personalized user context prediction in smart-product service system. *Procedia CIRP*. 2023;119:963-968.

doi: 10.1016/j.procir.2023.01.021

18. Zheng P, Yu S, Wang Y, Zhong RY, Xu X. User-experience based product development for mass personalization: A case study. *Procedia CIRP*. 2017;63:2-7.

doi: 10.1016/j.procir.2017.03.122

19. Zheng P, Xu X, Chen CH. A data-driven cyber-physical approach for personalised smart, connected product co-development in a cloud-based environment. *J Intell Manuf*. 2020;31(1):3-18.

doi: 10.1007/s10845-018-1430-y

20. Peng DX, Heim GR, Mallick DN. Collaborative product development: The effect of project complexity on the use of information technology tools and new product development practices. *Prod Oper Manag*. 2014;23(8):1421-1438.

doi: 10.1111/j.1937-5956.2012.01383.x

21. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. 2018.

doi: 10.48550/arXiv.1810.04805

22. Touvron H, Lavril T, Izacard G, *et al*. LLaMA: Open and efficient foundation language models. 2023.

doi: 10.48550/arXiv.2302.13971

23. Wang B, Zuo H, Cai Z, *et al*. A Task-decomposed Ai-aided Approach for Generative Conceptual Design. In: *Proceedings of the ASME Design Engineering Technical Conference*. Vol 6. New York: American Society of Mechanical Engineers (ASME); 2023.

doi: 10.1115/detc2023-109087

24. Jiang S, Luo J. Autotriz: Artificial ideation with triz and large language models. 2024.

doi: 10.48550/arXiv.2403.13002

25. Yin H, Zhang Z, Liu Y. The exploration of integrating the midjourney artificial intelligence generated content tool into design systems to direct designers towards future-oriented innovation. *Systems*. 2023;11(12):566.

doi: 10.3390/systems11120566

26. Wu Q, Bansal G, Zhang J, *et al*. AutoGen: Enabling next-gen llm applications via multi-agent conversation. 2023.

doi: 10.48550/arXiv.2308.08155

27. Hong S, Zhuge M, Chen J, *et al*. MetaGPT: Meta programming for a multi-agent collaborative framework. 2023.

doi: 10.48550/arXiv.2308.00352

28. Chen L, Jing Q, Tsang Y, Wang Q, Sun L, Luo J. DesignFusion: Integrating generative models for conceptual design enrichment. *J Mech Des*. 2024;146(11):111703.

doi: 10.1115/1.4065487

29. Çelen A, Han G, Schindler K, *et al*. I-Design: Personalized

LLM interior designer. 2024.

doi: 10.48550/arXiv.2404.02838

30. Xu S, Wei Y, Zheng P, Zhang J, Yu C. LLM enabled generative collaborative design in a mixed reality environment. *J Manuf Syst*. 2024;74:703-715.

doi: 10.1016/j.jmsy.2024.04.030

31. Chen D, Cheng S, Hu J, Kasoar M, Arcucci R. Explainable global wildfire prediction models using graph neural networks. 2024.

doi: 10.48550/arXiv.2402.07152

32. Gong H, Cheng S, Chen Z, *et al*. An efficient digital twin based on machine learning SVD autoencoder and generalised latent assimilation for nuclear reactor physics. *Ann Nucl Energy*. 2022;179:109431.

doi: 10.1016/j.anucene.2022.109431

33. Zhong C, Cheng S, Kasoar M, Arcucci R. Reduced-order digital twin and latent data assimilation for global wildfire prediction. *Nat Hazards Earth Syst Sci*. 2023;23(5):1755-1768.

doi: 10.5194/nhess-23-1755-2023

34. Lauer-Schmaltz MW, Cash P, Hansen JP, Maier A. Towards the human digital twin: Definition and design--A survey. 2024.

doi: 10.48550/arXiv.2402.07922

35. Lin Y, Chen L, Ali A, *et al*. Human digital twin: A survey. *J Cloud Comp.* 2024;13:131.

doi: 10.1186/S13677-024-00691-Z

36. Okegbile SD, Cai J, Niyato D, Yi C. Human digital twin for personalized healthcare: Vision, architecture and future directions. *IEEE Netw*. 2023;37(2):262-269.

doi: 10.1109/MNET.118.2200071

37. Chen J, Yi C, Du H, *et al*. A revolution of personalized healthcare: Enabling human digital twin with mobile AIGC. *IEEE Netw*. 2024.

doi: 10.1109/MNET.2024.3366560

38. Fan J, Zheng P, Lee CKM. A Vision-based human digital twin modeling approach for adaptive human-robot collaboration. *J Manuf Sci Eng*. 2023;145(12):121002.

doi: 10.1115/1.4062430

39. Aboulsafa EI, Khayat GAE, Elmorsy SA. An Educational Human Digital Twin Proposed Model for Personalized E-Learning. In: *2023 1st IEEE Afro-Mediterranean Conference on Artificial Intelligence, AMCAI 2023 - Proceedings*. United States: Institute of Electrical and Electronics Engineers Inc.; 2023.

doi: 10.1109/AMCAI59331.2023.10431503

40. He Q, Li L, Li D, *et al*. From digital human modeling to human digital twin: Framework and perspectives in human factors. *Chin J Mech Eng*. 2024;37(1):9.

doi: 10.1186/s10033-024-00998-7

41. El Saddik A, Hossain MS, Kantarci B, Editors K. *Connected Health in Smart Cities*. Berlin: Springer Nature; 2020.

doi: 10.1007/978-3-030-27844-1

42. Goodwin GC, Seron MM, Medioli AM, Smith T, King BR, Smart CE. A systematic stochastic design strategy achieving an optimal tradeoff between peak BGL and probability of hypoglycaemic events for individuals having type 1 diabetes mellitus. *Biomed Signal Process Control*. 2020;57:101813.

doi: 10.1016/j.bspc.2019.101813

43. Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks. 2018.

doi: 10.48550/arXiv.1812.04948

44. Poole B, Jain A, Barron JT, Mildenhall B. DreamFusion: Text-to-3D using 2D Diffusion. 2022.

doi: 10.48550/arXiv.2209.14988

45. Gonçalves G, Coelho H, Monteiro P, Melo M, Bessa M. Systematic review of comparative studies of the impact of realism in immersive virtual experiences. *ACM Comput Surv*. 2022;55(6):115.

doi: 10.1145/3533377

46. Barricelli BR, Casiraghi E, Fogli D. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access*. 2019;7:167653-167671.

doi: 10.1109/ACCESS.2019.2953499

47. Manghisi VM, Uva AE, Fiorentino M, Bevilacqua V, Trotta GF, Monno G. Real time RULA assessment using Kinect v2 sensor. *Appl Ergon*. 2017;65:481-491.

doi: 10.1016/j.apergo.2017.02.015

48. Tanwar R, Phukan OC, Singh G, Pal PK, Tiwari S. Attention based hybrid deep learning model for wearable based stress recognition. *Eng Appl Artif Intell*. 2024;127:107391.

doi: 10.1016/j.engappai.2023.107391

49. Debie E, Fernandez Rojas R, Fidock J, *et al*. Multimodal fusion for objective assessment of cognitive workload: A review. *IEEE Trans Cybern*. 2021;51(3):1542-1555.

doi: 10.1109/TCYB.2019.2939399

50. Ramalakshmi K, Srinivasa Raghavan V, Rajagopal S, *et al*. An extensive analysis of artificial intelligence and segmentation methods transforming cancer recognition in medical imaging. *Biomed Phys Eng Express*. 2024;10(4):045046.

doi: 10.1088/2057-1976/ad555b

51. Wang J, Sun K, Cheng T, *et al*. Deep high-resolution representation learning for visual recognition. *IEEE Trans Pattern Anal Mach Intell*. 2021;43:3349-3364.

doi: 10.1109/TPAMI.2020.2983686

52. Hignett S, Ergonomist LM. Rapid Entire Body Assessment (REBA). *Appl Ergon*. 2000;31:201-205.

doi: 10.1016/S0003-6870(99)00039-3

53. Parsa B, Banerjee AG. A Multi-Task Learning Approach for Human Activity Segmentation and Ergonomics Risk Assessment. In: *2021 IEEE Winter Conference on Applications of Computer Vision* (*WACV*); 2021.

    doi: 10.48550/arXiv.2008.03014

54. Wei J, Wang H, Feng J, Lin G, Yap KH. TAPS3D: Text-guided 3D textured shape generation from pseudo supervision. 2023.

    doi: 10.48550/arXiv.2303.13273

55. Radford A, Kim JW, Hallacy C, *et al*. Learning transferable visual models from natural language supervision. 2021.

    doi: 10.48550/arXiv.2103.00020

56. Deitke M, Schwenk D, Salvador J, *et al*. Objaverse: A universe of annotated 3D objects. 2022.

    doi: 10.48550/arXiv.2212.08051

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

Prediction of wall geometry for cold-metal-transfer-based wire-arc additive manufacturing

Robin Kromer* and Eric Lacoste

Univ. Bordeaux, CNRS, Bordeaux INP, I2M, UMR 5295, F-33400, Talence, France

(This article belongs to the *Special Issue: AI Usage in the Analysis of the Additive Manufacturing Process*)

Abstract

Wire-arc additive manufacturing (WAAM) is an advanced technique for fabricating large metal components through layer-by-layer material deposition using arc welding methods. This study focused on optimizing the WAAM process by employing machine learning models to predict and control bead geometries, specifically bead height (BH) and bead width (BW), while ensuring consistent height increments in multibead walls. Based on CMT technology in cold metal transfer experiments, linear regression models achieved high accuracy in predicting BH and BW. Analysis of variance results highlighted the considerable influence of voltage (*V*) and travel speed (*TS*) on bead geometries. For multibead wall characteristics, polynomial regression models incorporating non-linear terms, such as travel speed (TS^2) and dwell time (Dt^2), were developed to predict height (*H*) and waviness (W). Various optimization metrics were employed to balance the trade-offs between *H* and *W* for identifying optimal welding conditions that achieved the target *H* while minimizing *W*. A notable innovation of this research is the optimization of dwell time (*Dt*) for each layer to achieve a linear incremental *H* profile, minimizing *W* and ensuring consistent layer quality.

Keywords: Machine learning; Dwell time; Bead geometry; Process modeling; Wire-arc additive manufacturing

***Corresponding author:**
Robin Kromer
(robin.kromer@u-bordeaux.fr)

Citation: Kromer R, Lacoste E. Prediction of wall geometry for cold-metal-transfer-based wire-arc additive manufacturing. *Int J AI Mater Design*. 2024;1(3): 20-32.
doi: 10.36922/ijamd.4285

Received: July 19, 2024

Accepted: September 2, 2024

Published Online: October 10, 2024

1. Introduction

Wire-arc additive manufacturing (WAAM) is a process that creates large metal parts by depositing the associated material in successive layers based on arc metal transfer. This technique offers several advantages, including optimized raw part designs, reduced material waste, and high production efficiency.[1] WAAM employs various welding techniques to melt metal wires, including gas metal arc welding (GMAW), gas tungsten arc welding, and plasma arc welding. Among these, GMAW is particularly preferred for its high material deposition rate.[2] A variation of GMAW, called cold metal transfer (CMT), features a controlled dip transfer mode, making it popular in WAAM owing to its low heat input and reduced spatter.[3]

Cong *et al.* examined different arc modes in CMT to fabricate parts with zero porosity,[4] whereas Ali *et al.* examined the influence of arc energy and thermal fields on the mechanical properties and microstructures of hot-work tool steel in CMT-based WAAM.[5] The shape of each weld bead in WAAM depends on the energy input and determines the quality and dimensional precision of the resulting parts. Hence, numerous studies have focused on the influence of process parameters on single-bead geometry. Fu *et al.* investigated the interplay between wire feed speed (WFS) and travel speed (TS), examined its influence on the width (BW) and height (BH) of a single weld bead in bainite steel WAAM, and reported a decrease in BW/BH with increasing WFS across all TS levels.[6] Ayarkwa *et al.* highlighted the importance of the WFS/TS ratio and discovered that a higher ratio results in wider and taller beads during the fabrication of aluminum walls through CMT.[7] Kazanas *et al.* studied the fabrication of inclined steel and aluminum walls using CMT at a constant WFS/TS ratio, discovered the substantial influence of TS on wall quality, and recommended a value between 0.2 and 0.25 m/min for optimal surface smoothness.[8]

The temperature of the workpiece at the beginning of deposition of each new layer, known as the interlayer temperature, is a critical factor in WAAM.[9-11] This temperature influences the microstructures and characteristics of the final components. For instance, in Ti–Al WAAM, increasing the interlayer temperature from 100°C to 500°C decreases the alpha phase content, leading to reduced hardness.[12] If not properly controlled, the interlayer temperature can cause longitudinal cracking and substantial residual stress in the initial layers of Fe–Al parts.[13] In addition, in Ti6Al4V, rising interlayer temperatures influence bead geometry along the build direction.[14] Xiong *et al.*[15] discovered that low interlayer temperatures enhance the surface quality of thin-walled components fabricated using WAAM, whereas another study demonstrated that maintaining the interlayer temperature within a specific low range improves the final quality of steel thin-walled parts.[16] These processes are influenced by material properties, such as the impact of the contact angle on bead geometry. Notably, the development of multibead walls is essential for advancing WAAM technology. Kumar *et al.*[17] attempted to optimize GMAW-based WAAM for creating multilayer beads on steel, focusing on process parameters to improve dimensional precision and mechanical properties. They employed a response surface methodology design for single-layer deposits, considering variables such as *TS*, voltage (*V*), current, and gas flow rate. Their findings revealed that *TS* critically influenced BW and BH, altering them by 52.29% and 43%, respectively. By applying a desirability function, they identified optimal parameters and successfully fabricated a multilayer structure. Mai *et al.*[18] explored the fabrication of 308L stainless steel parts using WAAM, by combining experimental design and optimization through analysis of variance (ANOVA). The optimized parameters resulted in remarkable mechanical properties, highlighting the importance of optimization. Chaudhari *et al.*[19] analyzed the influence of WAAM process parameters on bead geometry for single-layer deposits, focusing on the trends of BW and BH under varying *TS*, WFS, and *V*. Using a Box–Behnken design, they identified WFS as the most influential factor for BW and BH, followed by *V* and *TS*. Their optimized settings (TS = 141 mm/min, WFS = 5.50 m/min, and *V* = 19 V) led to the successful fabrication of a multilayer structure. Natryan *et al.*[20] used the Taguchi method to examine the effects of TS, welding current, and filler diameter on the quality of welded joints. By employing an orthogonal array design and performing statistical analyses, they identified an optimal parameter combination that minimized defects and improved bead geometry and weld penetration. Vora *et al.*[21] optimized bead shape for GMAW-based WAAM using a Box–Behnken design for bead-on-plate tests. They applied ANOVA to analyze regression equations and employed a teaching-learning-based optimization method to determine the best parameters, achieving a minimum BW of 4.73 mm and a maximum BH of 7.81 mm. The optimized parameters enabled the fabrication of a multilayer structure without layer disbonding. Kumar *et al.*[22] employed a genetic algorithm to identify optimal process parameters for WAAM, achieving near-net-shaped deposition with fewer layers. The genetic algorithm effectively optimized these parameters to yield the desired outcomes. Liberini *et al.*[23] focused on selecting optimal process parameters for WAAM through a multiobjective optimization approach, successfully identifying the best values for BW, BH, porosity, and deposition rate. Wang *et al.*[24] employed a multiwire indirect-arc-directed energy deposition method, discovering that WFS, current, and wire angle substantially impacted indirect-arc-directed energy deposition. This method achieved favorable microstructures and mechanical properties compared with conventional methods. Mishra *et al.*[25] optimized the topology and deposition direction in WAAM using a mathematical model and a combination of genetic algorithm- and gradient-based optimization techniques. This approach improved part quality and reduced manufacturing durations compared with traditional methods.

To achieve high surface quality and precise dimensions in WAAM, ensuring predictable and controllable weld beads for each layer is essential.[26] Developing models that can accurately forecast weld bead geometry based

on various process parameters is critical in this regard. Xiong *et al.*[27] developed models that linked the process parameters of GMAW-based WAAM (WFS, TS, *V*, and nozzle-to-plate distance [NdP]) to bead geometry (BW and BH) using artificial neural networks (ANNs) and second-order regression analysis. For pulsed-gas-tungsten-arc-welding-based and plasma-arc-welding-based WAAM, an analytical model was developed to predict layer height and wall width based on process parameters, highlighting the considerable influence of interlayer temperature on wall geometry.[28] ANNs have also been employed to correlate process variables with bead geometry in shielded metal arc welding,[29] whereas linear regression has been utilized to predict bead geometry in pulsed GMAW.[21] However, limited research has been conducted on the inclusion of interlayer temperature into models for predicting bead geometry in CMT-based WAAM. Disregarding interlayer temperature can lead to inaccurate geometry predictions in multilayer WAAM owing to heat accumulation. The dwell time (*Dt*; cooling time in the WAAM process) between the deposition of successive layers is critical for achieving the appropriate temperature. For instance, during the manufacturing of a multilayer steel alloy (ER100) component, the 15th and 26th layers were found to be required for the interlayer temperature to stabilize.[30] Different deposition strategies yield varying types of microstructures, highlighting the need to improve dwell time.[31] Spencer *et al.*[32] proposed cooling each deposited layer to a low temperature, such as 120°C, before adding the next layer to improve surface quality. However, this method substantially decreases deposition efficiency as the cooling time surpasses the build time. This inefficiency is exacerbated for larger parts owing to extended cooling durations.

This study developed a machine learning method for predicting BH and BW. These BH and BW predictions were then utilized to obtain the desired bead geometry ratio for multibead wall manufacturing. Experimental setups were assembled to analyze multibead production, particularly focusing on incremental height (*H*) and waviness (*W*) measurements. Machine learning techniques were employed to train and test models for multibead wall manufacturing optimization based on experimental datasets. To enhance prediction accuracy, the algorithm leveraged comprehensive datasets derived from diverse experimental scenarios, encompassing various bead strategies and process settings. The integration of machine learning models enables the provision of geometric information to manufacturers of WAAM parts. Furthermore, the paper includes a discussion to elucidate the underlying assumptions of the study.

2. Methodology

2.1. Experimental setup

Figure 1A displays a schematic of the experimental setup. A CMT welding power source (Fronius TPS320i) was connected to a three-axis laboratory bench, allowing numerical control over its movement and CMT synchronization. The gantry, with the substrate mounted on it, followed the deposition path. A 316L stainless steel consumable wire electrode with a diameter of 0.8 mm was utilized. Shielding was provided by a gas mixture containing 85% argon and 15% CO_2, flowing at a rate of 18 L/min. Single- and multi-weld beads were deposited onto a 304L substrate measuring 280 mm in length, 60 mm in width, and 10 mm in *H*. The substrate was mounted onto a support to avoid strain development. During deposition, the torch was held perpendicular to the substrate, maintaining a nozzle-to-plate distance, denoted as *D*, of 10 mm.

As illustrated in Figure 1, the bead and wall lengths were set to 120 mm. Each wall comprised 15 layers of two beads arranged in a zigzag pattern. The start and end points of each bead, influenced by arc ignition and extinguishment, were excluded from measurements to avoid inconsistencies. To ensure measurement consistency, only the central 60 mm section was used for data acquisition, given that this section of the bead geometry is stable. The average values of the acquired data were recorded as the final measurement results.

A laser line profiler (Gocator 2430, LMI Technologies) was employed to record profiles at three locations for bead geometry evaluation. The wall height was measured using a mechanical dial gauge at three positions on each layer. The incremental height was also measured for each layer. A 3D smart line profile sensor was utilized to capture the surface of the wall (Figure 2). Notably, this sensor enables the calculation of the peak-to-valley height of the surface, referred to as *W*, which is a crucial parameter for assessing the surface quality of components fabricated using WAAM. *W* is defined as the more widely spaced component of surface texture, encompassing periodic surface deviations larger than those of surface roughness. By measuring the vertical distance between the highest peak and the lowest valley within a specific evaluation length, the total *W* height (Wt) can be determined. The profiles are processed using robust Gaussian filters, as specified in ISO 16610-21, to separate the *W* component originating from roughness. The filtered profiles facilitate accurate calculations of Wt, in accordance with the definitions and procedures outlined in ISO 4287 and ASME B46.1 standards.

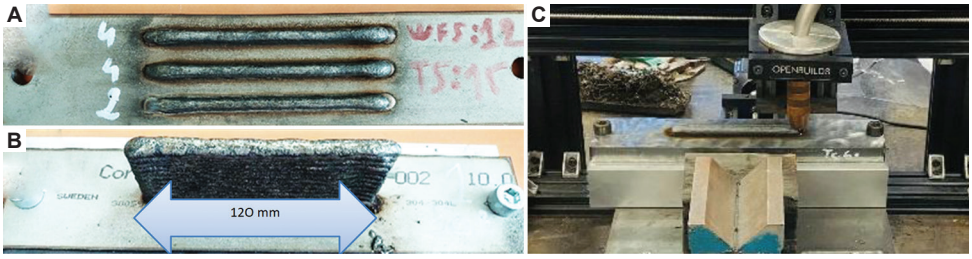


Figure 1. Samples and WAAM deposition setup. (A) Bead samples for BH and BW measurements. (B) Multibead samples. (C) Setup for sample production. Abbreviations: BH: Bead height; BW: Bead width.

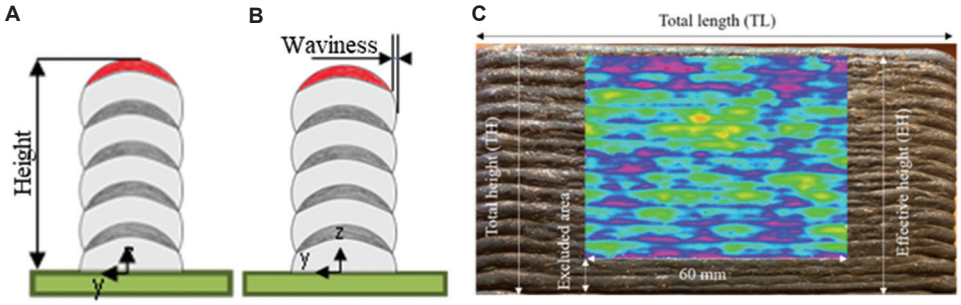


Figure 2. Wall modeling and analysis. (A) Height measurement. (B) Peak-to-valley measurement. (C) Example of laser line profile measurement.

2.2. Input and output selection

In a CMT welding system, the welding *V* and current parameters are automatically adjusted based on the WFS selected by the user to ensure stable and continuous metal transfer. Factors such as wire material, diameter, welding mode, and type of shielding gas are also considered. For this study, three independent process parameters were considered inputs (Table 1). Notably, the selection of TS and *V* for bead geometry as well as TS and dwell time (Dt) for wall production is fundamental owing to their substantial influence on the weld bead and wall characteristics in WAAM. TS affects the deposition rate: a slower TS results in a taller and wider bead owing to the deposition of more material at one spot, whereas a faster TS produces a shorter and narrower bead. *V* controls the arc length and heat input: higher voltages generate more heat and a broader bead, while lower voltages produce a narrower bead. A constant TS ensures uniform layer deposition, influencing the incremental height of each layer. Dt, reflecting the period between the deposition of successive layers, is crucial for thermal management. Adequate Dt allows the deposited layers to cool, preventing excessive heat buildup that can lead to distortions and poor interlayer bonding. Appropriate Dt minimizes *W* and maintains structural integrity. Thus, the strategic selection of TS and *V* for bead geometry, as well as TS and dwell time for wall production, enables precise control over the WAAM process, optimizing both the geometric accuracy and mechanical properties of the fabricated components. The geometric characteristics of the weld bead (BW and BH), along with height (*H*) and *W*, were selected as outputs.

Table 1. Experimental conditions for the CMT-based WAAM process

Input factors	Values/levels
Voltage (*V*)	22, 24, and 26
Travel speed, *TS* (mm/s)	5, 10, and 15
Dwell time, *Dt* (s)	5, 60, and 120
Distance, *D* (mm)	10
Gas flow rate (*L*/min)	18
Weld bead length (mm)	120

Abbreviations: CMT: Cold metal transfer; WAAM: Wire-arc additive manufacturing.

The study employed a full factorial design (for *V* and TS) to systematically explore the effects of critical welding parameters on BH and BW (Table 2). However, even when using advanced measurement techniques, slight inaccuracies could arise owing to inherent equipment limitations and variations along the tracks. A 2% error margin was selected to account for slight variations across three zones of the beads and walls. This experimental strategy covered all possible combinations of factors at different levels, facilitating a comprehensive analysis of both primary effects and interactions. The following Python libraries for data analysis and optimization were utilized: Pandas and NumPy for data manipulation and

Table 2. Results of bead and wall geometries (2% errors)

V (V)	TS (mm/min)	BH (mm)	BW (mm)	TS (mm/min)	Dt (s)	H (mm)	W (mm)
24	10	4.61	6.85	10	60	61.5	1.38
24	5	6.46	7.82	5	60	63	2.3
22	15	3.71	5.05	15	5	52	2.4
24	15	3.31	6.25	15	60	56.8	1.32
26	15	3.71	8.55	15	120	58	1.11
24	5	7.2	8.01	5	60	62.8	2.2
24	10	4,78	6.73	10	60	61	1.36
26	10	4.51	7.91	10	120	62	1.4
26	5	5.47	9.2	5	120	64	2.11
22	5	6.52	6.25	5	5	59	2.98
24	10	4.9	6.69	10	60	60.8	1.41
26	10	4.51	8	10	120	61	1.44
24	15	3.73	6.07	15	60	57.2	1.26
22	10	4.51	5.98	10	5	57.4	1.88
22	10	6.13	5.78	10	5	57.2	1.84

Abbreviations: V: Voltage; TS: Travel speed; BH: Bead height; BW: Bead width; TS: Travel speed; Dt: dwell time; H: Height; W: Waviness.

Statsmodels for statistical modeling. Ordinary least squares regression was used for linear modeling. Scipy's minimize function was used for optimization tasks, whereas Matplotlib was employed to create visualizations for result interpretation. The collected data were used to develop predictive models that helped establish the relationships between welding parameters and response variables. Linear regression models were used to establish a direct relationship between the predictors (*V* and TS) and the response variables (BH and BW). Model performance was evaluated using metrics such as the mean squared error (MSE) and coefficient of determination (R^2) to ensure accuracy and reliability. Subsequently, ANOVA was employed to statistically assess the significance of each factor and their interactions. Notably, ANOVA compares the variance explained by the model with the variance within the data to determine whether changes in welding parameters substantially impact BH and BW. Factors with low *P*-values (typically $P < 0.05$) indicate a strong influence on the response variables, as detailed in a previous paper.[33]

Through this analysis, we identified optimal welding conditions using various metrics, such as efficiency (BH/BW ratio), harmonic mean, exponential score, normalized difference, and cost function. Specifically, we targeted a BH/BW value close to one. Each metric offered a different perspective on balancing the trade-off between maximizing BH and minimizing BW. The geometric quality of the wall was also modeled using different TS and dwell time values. Table 2 summarizes the measurement results of height and *W*. Another full factorial design, involving polynomial regression modeling and ANOVA, was adopted to acquire in-depth insights regarding height and *W* for wall geometry prediction. Incremental height was measured for each bead following the design of experiments, using a distance sensor after the deposition of each layer. This approach was aimed at collecting data to guide the WAAM process, ultimately enhancing both quality and performance. The height and *W* of each layer were measured (data can be obtained from the authors upon request). Furthermore, all bead-on-plate samples were visually inspected and found to be adequately fused and free of porosity or similar defects, thereby establishing the workable parameter range for our study.

3. Results

3.1. BH and BW with a linear predictions

The welding parameters selected for the ANOVA are listed in Table 2 alongside the results obtained for bead and wall geometries. The analysis of BH and BW predictions using linear regression models provided valuable insights into the relationship between the welding parameters – *V* and TS – and the resulting weld bead characteristics. In the linear regression models, *V* and TS were directly related to BH and BW through a straightforward linear equation. The development and assessment of these models highlighted important trends and statistical relationships. For BH, the linear model exhibited an MSE of 0.043 and an R^2 value of 0.95, indicating that the model could explain approximately 95% of the variance in BH. This suggests a strong linear relationship between BH and the input parameters *V* and

TS, including their interaction. Further supporting this linear model, the ANOVA results for BH indicated that neither *V* nor *TS* alone was statistically significant at the 0.05 level, with *P*-values of 0.046 and 0.005, respectively. The interaction between *V* and *TS* also lacked statistical significance (*P*-value = 0.106). These results imply that the additive linear effects of *V* and *TS* are sufficient for modeling BH, highlighting the robustness of the linear regression model for this response variable.

For BW, similar to BH, the linear model yielded an MSE of 0.061 and an R^2 value of 0.88, indicating that the linear model explained approximately 88% of the variance in BW. This suggests a strong linear dependency of *V* and TS. The ANOVA results for BW revealed that *V* and TS are significant factors (*P*-value = 0.004 and 0.002, respectively). This highlights the predominant influence of *V* on BW and suggests that a simple linear model is sufficient for capturing the relationship between the input parameters and BW. Figure 3 illustrates the two linear models, with the actual and predicted results. The linear model can be used easily to enhance WAAM productivity.

The aforementioned analysis highlights the importance of selecting appropriate models based on the response variable. Insights from ANOVA highlight the importance of specific parameters and their interactions, providing a strong foundation for refining the welding process. The regression analysis yields distinct best-fit equations for BH and BW, illustrating the relationships between these dependent variables and the independent parameters *V* and TS. For BH, the optimal model is linear and is expressed as $y = 11.74 - 0.17V - 0.28TS$. This equation indicates that an increase in *V* or *TS* reduces BH. Such a linear relationship simplifies the control over BH, allowing straightforward adjustments to *V* and *TS* to achieve the desired height. Similarly, the best-fit model for BW is also linear and is expressed as $y = -7.55 + 0.66V - 0.13TS$. This equation demonstrates that *V* and *TS* affect BW. The simplicity and statistical significance of this model indicate that these factors are critical in determining BH and BW, and their effects are effectively captured by the linear terms without requiring higher-order polynomial terms.

Evaluating the optimization of the welding process using various metrics offers detailed insights into the optimal welding conditions. The full factorial design predicts BH and BW values across different combinations of *V* and *TS*. For instance, when $V = 22$ and $TS = 5$, the predicted BH is 6.67 mm, whereas the predicted BW is 6.35 mm. Hence, by adjusting the values of *V* and *TS*, a specific ratio can be targeted. Several metrics are employed to assess these predictions, including efficiency, harmonic mean, exponential score, normalized difference, and cost function. Figure 4 illustrates various metrics for the BH and BW predictions of the linear regression model. A combination of these parameters, helps to interpret predictive values, focusing on how different V and *TS* levels impact BH and BW. These metrics offer varied perspectives on the trade-off between BH and BW. For instance, the efficiency metric, which favors higher BH and lower BW, yields an optimal score of 1.05. Similarly, the harmonic mean and exponential score metrics emphasize the balance between achieving a high BH and an acceptable BW, with scores of 6.54 and 1.49, respectively. Normalizing BH and BW values is essential for a fair comparison and interpretation across different scales. The normalized values for BH (BH_norm) and BW (BW_norm) are 1.00 and 0.66, respectively. These normalized values are used

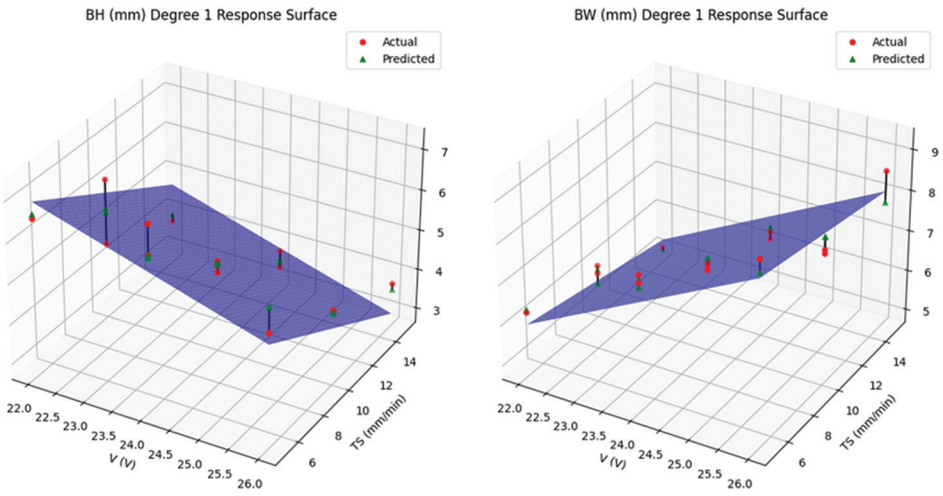


Figure 3. Linear model representations of BH and BW
Abbreviations: BH: Bead height; BW: Bead width.

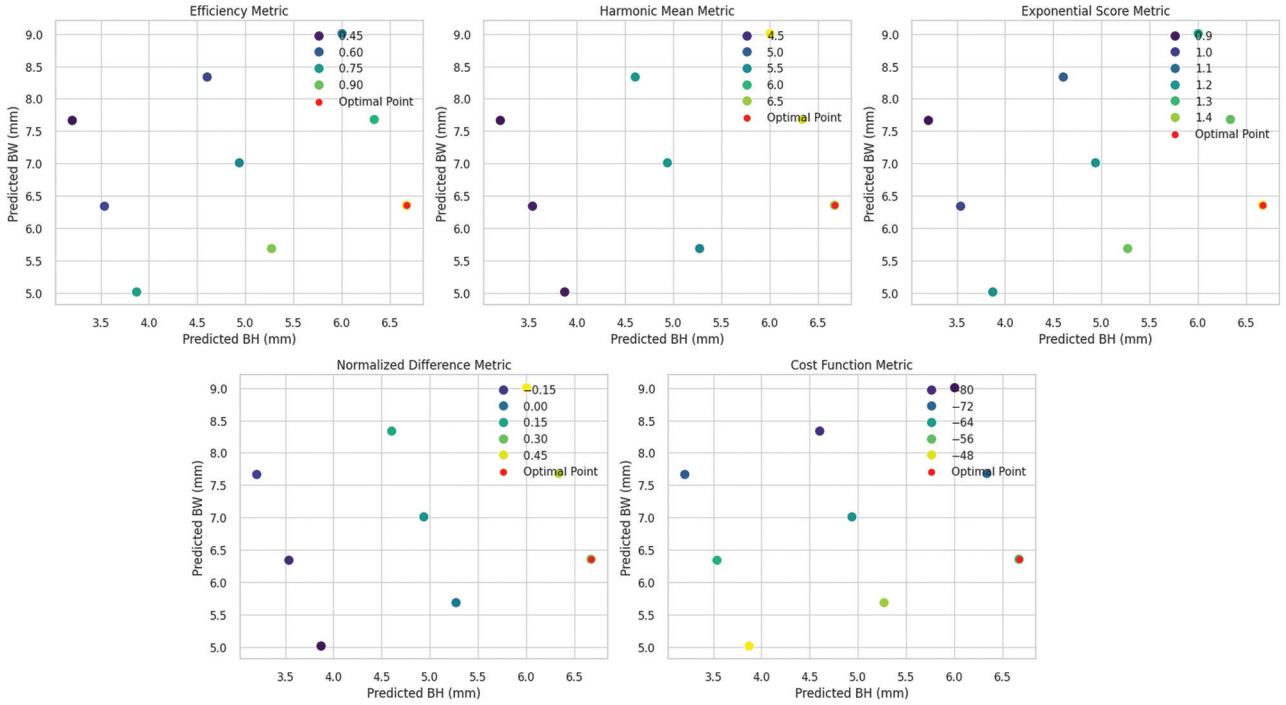


Figure 4. Different metrics for BH/BW ratio predictions
Abbreviations: BH: Bead height; BW: Bead width.

in computing the normalized difference metric, which yields a value of 0.33, indicating a slight preference for BH. In addition, the cost function metric integrates multiple factors to reflect the overall cost or desirability of the welding conditions, identifying the optimal point with the minimum cost function value. This highlights the practical applicability of the predicted conditions for a single bead.

3.2. Multibead wall analysis

In contrast to the linear model, the polynomial models demonstrate superior performance for the multibead wall characteristics, namely height and *W*. The polynomial regression model for height exhibits a substantial level of explanatory power, with an R^2 value of 0.83. The adjusted R^2 value of 0.765, while slightly lower, still indicates a strong model fit, although with some reduction in robustness compared with the linear regressions. The model's overall significance is underscored by an F-statistic of 12.38 and a highly significant *P*-value of 0.001, demonstrating the collective importance of the considered predictors. Notably, *TS* is not significant ($P = 0.393$), suggesting that it does not individually make meaningful contributions to the model. Conversely, the quadratic term TS^2 exhibits marginal significance ($P = 0.049$), indicating a potential non-linear relationship with height. Similarly, *Dt* is marginally significant ($P = 0.053$), suggesting some influence on height, whereas the quadratic term Dt^2 is not significant ($P = 0.441$), implying a limited impact on height. The polynomial regression model for *W* boasts an R^2 value of 0.80, with an adjusted R^2 value of 0.730, affirming a solid fit. The model's F-statistic of 10.46 and significant *P*-value of 0.002 emphasize its overall validity. Both *TS* and TS^2 are highly significant ($P < 0.01$), reflecting their substantial roles in determining *W*. *Dt* exhibits marginal significance ($P = 0.054$), indicating a potential effect on *W*, whereas Dt^2 is not significant ($P = 0.353$), suggesting minimal impact. Figure 5 illustrates the polynomial model representations of BH and BW, with the actual and predicted values. This visualization provides valuable insights for engineers involved in WAAM production.

The equation of the optimal model for predicting height is polynomial in nature and is expressed as $y = 58.04 + 0.63Dt - 0.51TS^2$. This equation indicates that TS^2 and *Dt* considerably influence height. The negative coefficient for TS^2 suggests that as *TS* increases, height initially increases and then decreases, exhibiting a parabolic relationship. The equation of the optimal model for predicting *W* is also polynomial but more complex: $y = 4.77 - 0.48TS - 0.01Dt + 0.02TS^2$. This model indicates that *W* is influenced by the linear and non-linear interactions of *TS* and *Dt*. The positive coefficient for TS^2 implies that the relationship between these variables and *W* is curvilinear. The parabolic nature of the *TS* term suggests the existence of optimal values for height and *W*, beyond which further increments

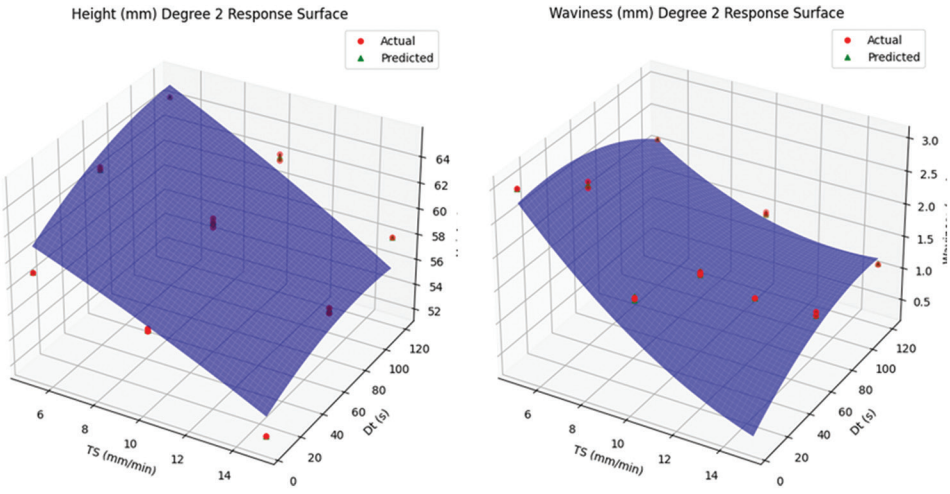


Figure 5. Polynomial model representation for BH and BW
Abbreviations: BH: Bead height; BW: Bead width.

in *TS* may be detrimental. Understanding these physical relationships aids in fine-tuning the process parameters to achieve desired outcomes in WAAM, ensuring that the deposition height is maximized while minimizing surface *W* for better structural integrity and finish of the final product.

The analysis of optimal welding conditions using various metrics enables a comprehensive evaluation of the trade-offs involved in achieving the desired balance between the height and *W* of the weld bead (Figure 3). Similar to the single-bead analysis, this evaluation also utilizes the following metrics: efficiency, harmonic mean, exponential score, normalized difference, and cost function, each offering a unique perspective on the optimal welding parameters. The efficiency metric identifies optimal conditions at a *TS* of 15 mm/min and *Dt* of 120 s, with a height of 58 mm and *W* of 1.11 mm. The corresponding efficiency value of 52.25 indicates a strong balance between the two objectives, maximizing height while minimizing *W*. This optimal condition also performs well across the other metrics, demonstrating consistency in the evaluation approach. The harmonic mean metric, which balances the trade-off between height and *W*, confirms the same optimal condition (*TS* = 15 mm/min, *Dt* = 120 s). The corresponding harmonic mean value of 2.20 supports the robustness of this condition across different evaluation criteria. This agreement among the metrics highlights the stability of the identified optimal condition. Similarly, the exponential score metric, which emphasizes exponential relationships in the trade-off, also confirms the optimal condition at a *TS* of 15 mm/min and *Dt* of 120 s, with an exponential score of 10.96. This reinforces the efficiency of the selected parameters in achieving desirable welding outcomes. The normalized difference metric, which assesses the normalized differences between the actual and desired values, yields a slightly different optimal point at a *TS* of 5 mm/min and *Dt* of 120 s, resulting in a height of 64 mm and *W* of 2.11mm. Although the efficiency value corresponding to this condition is lower at 30.3, the normalized difference of 0.41 indicates a favorable balance for this specific metric. Finally, the cost function metric, which integrates multiple factors into a single cost-oriented measure, identifies the optimal condition at a *TS* of 10 mm/min and *Dt* of 60 s, resulting in a height of 61.5 mm and *W* of 1.38 mm. The corresponding cost function value is 47.7. Given that the cost function metric integrates multiple factors such as *TS*, *Dt*, height, and *W* into a single measure, the above value of 47.7 reflects the economic feasibility of these factors. This value indicates that the selected conditions optimally balance manufacturing efficiency, part quality, and overall production costs.

The results of these analyses are visually presented in Figure 6, which illustrates the performance of different welding conditions under each metric. Figure 6 presents the wall geometry predictions of the polynomial regression models. The depicted data includes variations in *TS* and *Dt*. While this figure includes optimization metrics, its primary focus is on displaying smooth curves or surfaces that represent the underlying trends, which may not correspond directly to the discrete experimental data points detailed in the results table. The plots clearly depict the optimal points, with markers indicating various metric scores. This visual aid helps in understanding the impact of each welding parameter on the overall performance. Based on the preceding analysis, the optimization process identified the following optimal conditions for the welding

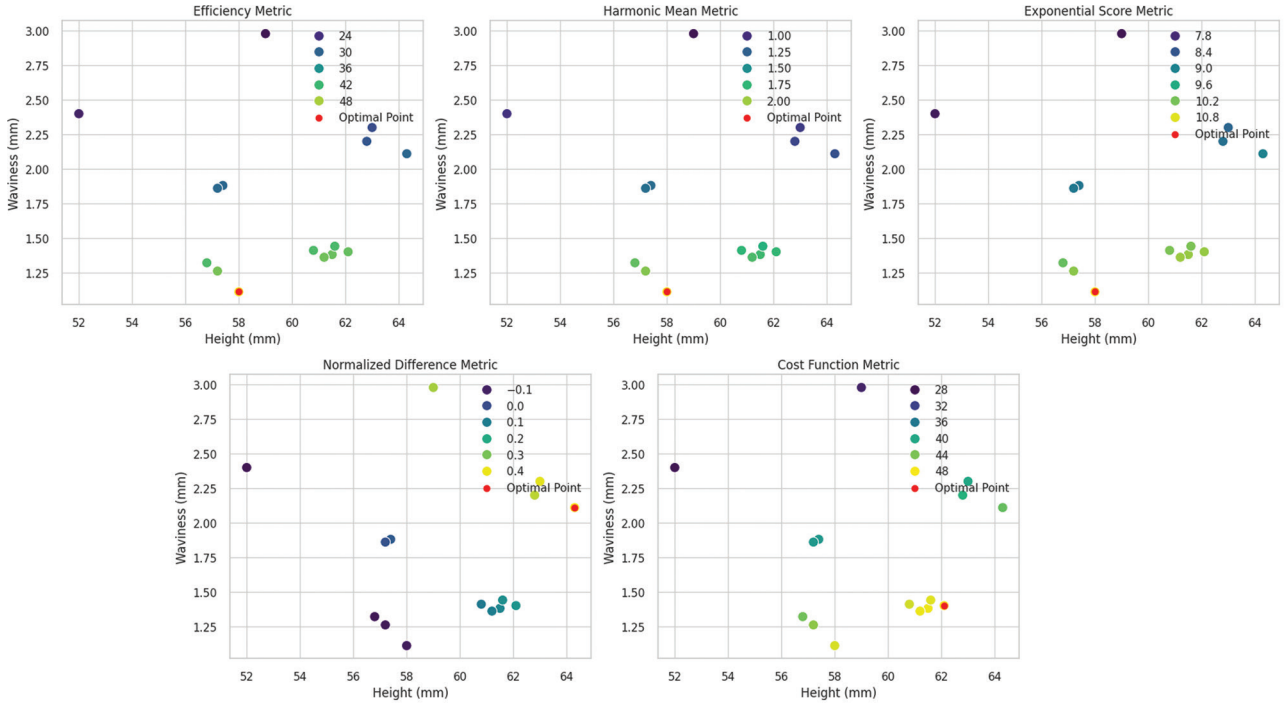


Figure 6. Metrics for height and waviness analysis

process: *TS* = 15 mm/s and *Dt* = 100 s, resulting in an optimal height of 43.85 and an optimal *W* of 1.85 (*V* = 20 V and BH/BW = 1).

3.3. *Dt* optimization

Following the aforementioned analysis, a response surface can be created to identify the optimal *Dt* value based on the measured incremental heights. The underlying goal remains the same: to maximize height and minimize *W*. The process begins with data preparation, which involves loading incremental height data, extracting relevant features and targets, and splitting the data into training and testing sets. Standardization of features and the creation of polynomial features ensure improved model performance. Model training utilizes an XGBoost regressor for height prediction owing to its robustness and capability to tackle complex data relationships, while a gradient-boosting regressor is employed for *W* prediction. These models are selected to improve the accuracy of predictions. The optimization objective function is designed to minimize the difference between the predicted incremental heights and a target incremental height for each layer, incorporating penalties for *W* and ensuring smooth transitions in *Dt* values between layers. This approach guarantees the model achieves the desired incremental height while maintaining smooth transitions and minimal *W*. The optimization process iterates through each layer, employing the "minimize" function from the scipy.optimize library to determine the optimal *Dt* value, which is subsequently stored and plotted.

The response surface plot (Figure 7) graphically represents the variation in predicted incremental height relative to *Dt* and layer number. The X-axis represents the range of *Dt* values from 0 to 120 s, whereas the Y-axis represents the layer number, ranging from 1 to 19. The color scale denotes predicted incremental height values, with lighter colors indicating higher incremental heights. The plot illustrates that in the initial layers (1–5), incremental height increases rapidly with rising *Dt*, particularly for *Dt* values between 0 and 60 s. In the middle layers (6–14), incremental height stabilizes, as evidenced by the broader color bands, whereas in the upper layers (15–19), incremental height decreases slightly or remains constant with increasing *Dt*. This visualization helps identify an optimal range of *Dt* values for achieving desired incremental heights, suggesting that *Dt* values around 60–80 s are optimal for most layers. Beyond a certain *Dt* (approximately 100 s), incremental height becomes more uniform across layers. This response surface plot can be used to optimize the WAAM process by selecting appropriate *Dt* values for each layer, ensuring uniform build quality, minimizing defects, and ultimately improving the reliability and performance of the AM process.

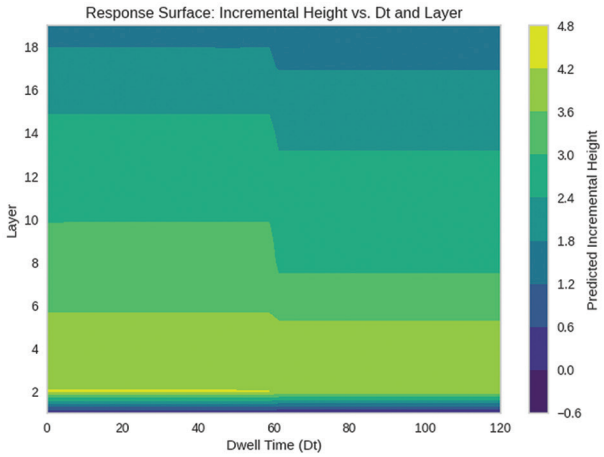


Figure 7. Response surface plot of *Dt* depending on the number of layers and predicted incremental height

4. Discussion

The integration of machine-learning models enhances the predictive capability of current approaches for multilayer bead geometry. *Dt* emerges as a critical parameter influencing microstructure and dimensional stability, as highlighted by Turgut.[34] This study provides empirical evidence demonstrating that varying interlayer dwell times can substantially influence the final quality of parts fabricated using WAAM. These considerations are particularly critical for controlling heat input, which is a pivotal factor in maintaining the structural integrity and dimensional accuracy of manufactured components. In the same study by Turgu, three samples were fabricated through continuous deposition with interlayer dwell times of 60 s and 120 s. Findings revealed that the interlayer dwell time effectively controlled temperature fields, which in turn influenced the microstructural and mechanical properties of the parts. For instance, an increased dwell time resulted in greater hardness and yield strength, highlighting the importance of thermal management in the WAAM process. Based on the microstructural analysis results of our models, microstructures can be predicted. These outcomes emphasize the importance of optimizing interlayer dwell time to achieve the desired material properties.

Furthermore, the analysis extends to multibead wall production, focusing on height and *W*. The optimal model for predicting wall height incorporates *TS*, *Dt*, and their interactions, indicating a parabolic relationship wherein height initially increases with these parameters before subsequently decreasing. Figure 8 illustrates two wall surfaces: one manufactured using optimal parameters and another manufactured using conventional parameters. This indicates the possibility of predicting *W* and surface

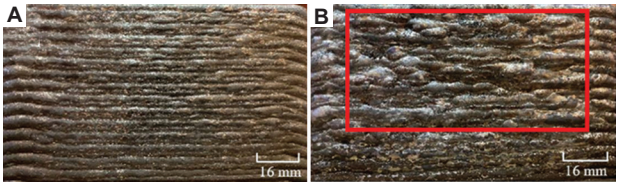


Figure 8. Two walls manufactured with optimal (A) and conventional (B) parameters

machining depth. The main information is to have enough matter for machining to get the right geometry and also a good health matter.

Moreover, the distance between beads was also optimized based on the BW. Previous studies[1,35] suggest that utilizing three or four beads does not substantially influence the predictions of height and *W* for a specific interlayer dwell time. However, heat accumulation varies with cumulative energy inputs, indicating the importance of integrating geometry into the prediction models. These findings align with the results of previous studies emphasizing the importance of thermal management and process parameter optimization in WAAM for producing high-quality metal components.

Finally, in a related study, Hu *et al.*[36] investigated the prediction of welding parameters for various layer heights in robotic WAAM. Their model accurately predicted the required parameters, enhancing adaptability and precision. Wang and Xue[37] conducted WAAM experiments on 316L stainless steel, maintaining a constant deposition rate while varying arc modes. Their results demonstrated that SpeedArc and SpeedPulse manufacturing processes were stable and efficient, revealing correlations between arc mode, microstructure, and mechanical properties. Wahsh *et al.*[38] focused on selecting parameters for the WAAM process, emphasizing the importance of identifying optimal settings to achieve the desired outcomes. Their study offered comprehensive guidelines for enhancing process efficiency and part quality. Kumar *et al.*[17] performed a parametric study and characterization of steel structures fabricated using WAAM. They identified key parameters influencing mechanical properties and dimensional accuracy, also providing a detailed analysis of various settings. Patel and Savsani[39] utilized a multiobjective improved teaching–learning-based optimization algorithm to optimize multiple objectives simultaneously. This algorithm demonstrated substantial improvements in WAAM process optimization. Collectively, these studies highlight the critical role of parameter optimization in improving the WAAM process, leading to enhanced mechanical properties, dimensional accuracy, and overall part quality. However, the interlayer dwell time is rarely investigated, despite being a key factor in part production.

To address this, knowledge formalization using artificial intelligence to correlate geometry and WAAM parameters is essential.

5. Conclusion

This study underscores the critical importance of precise parameter control in WAAM, particularly when utilizing the CMT process. Focusing on key welding parameters – *V* and *TS* – this study demonstrated the influence of these factors on weld bead geometries, specifically BH and BW. Linear regression models exhibited strong predictive capabilities, evidenced by robust performance metrics, indicating a direct linear relationship between welding parameters and bead geometries.

The ANOVA results further validated the aforementioned models, underscoring the statistical significance of *V* in determining BW and highlighting the more nuanced role of TS. The study also employed various optimization metrics, such as efficiency, harmonic mean, exponential score, normalized difference, and cost function, to provide a comprehensive understanding of the trade-offs involved in optimizing BH and BW. These metrics effectively identified optimal welding conditions, demonstrating their utility in fine-tuning the WAAM process.

For multibead wall production, polynomial regression models outperformed linear models, particularly in predicting wall height and *W*, with significant non-linear terms indicating their substantial influence. This nuanced understanding is crucial for optimizing multilayer builds, ensuring structural integrity, and minimizing defects. Our evaluation using various metrics consistently identified optimal conditions, achieving the desired balance between height and *W*.

In addition, *Dt* values were optimized to achieve a linear incremental height profile while minimizing *W* in WAAM. A response surface visualization illustrated the variation in incremental height with *Dt* and layer number, facilitating the selection of optimal *Dt* values for consistent, high-quality builds. Overall, this study provides a robust framework for optimizing WAAM processes through precise parameter control and advanced predictive modeling. The integration of machine learning models and statistical analysis techniques offers powerful tools for manufacturers, enabling them to achieve desired part geometries with greater efficiency and reliability. Our findings enhance the viability and competitiveness of WAAM in large-scale manufacturing applications involving extensive parameters.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare no competing interests.

Author contributions

Conceptualization: Robin Kromer
Investigation: Robin Kromer
Methodology: Robin Kromer
Writing – original draft: All authors
Writing – review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data are available from the corresponding author upon reasonable request.

References

1. Ding D, Pan Z, Cuiuri D, Li H. Wire-feed additive manufacturing of metal components: Technologies, developments and future interests. *Int J Adv Manuf Technol*. 2015;81:465-481.

   doi: 10.1007/s00170-015-7077-3

2. Le VT, Paris H. On the use of gas-metal-arc-welding additive manufacturing for repurposing of low-carbon steel components: microstructures and mechanical properties. *Weld World*. 2021;65:157-166.

   doi: 10.1007/s40194-020-01005-y

3. Frostevarg J, Kaplan AFH, Lamas J. Comparison of CMT with other arc modes for laser-arc hybrid welding of steel. *Weld World*. 2014;58:649-660.

   doi: 10.1007/s40194-014-0146-7

4. Cong B, Ding J, Williams S. Effect of arc mode in cold metal transfer process on porosity of additively manufactured Al-6.3%Cu alloy. *Int J Adv Manuf Technol*. 2015;76:1593-1606.

   doi: 10.1007/s00170-014-6346-x

5. Ali Y, Henckell P, Hildebrand J, Reimann J, Bergmann JP, Barnikol-Oettler S. Wire arc additive manufacturing of hot work tool steel with CMT process. *J Mater Process Technol*. 2019;269:109-116

   doi: 10.1016/j.jmatprotec.2019.01.034

6. Youheng F, Guilan W, Haiou Z, Liye L. Optimization of

surface appearance for wire and arc additive manufacturing of Bainite steel. *Int J Adv Manuf Technol*. 2017;91:301-313.

doi: 10.1007/s00170-016-9621-1

7. Ayarkwa KF, Williams S, Ding J. Investigation of pulse advance cold metal transfer on aluminium wire arc additive manufacturing. *Int J Rapid Manuf*. 2015;5:44.

doi: 10.1504/ijrapidm.2015.073547

8. Kazanas P, Deherkar P, Almeida P, Lockett H, Williams S. Fabrication of geometrical features using wire and arc additive manufacture. *Proc Inst Mech Eng Part B J Eng Manuf*. 2012;226:1042-1051.

doi: 10.1177/0954405412437126

9. Zhao W, Wei Y, Long J, Chen J, Liu R, Ou W. Modeling and simulation of heat transfer, fluid flow and geometry morphology in GMAW-based wire arc additive manufacturing. *Weld World*. 2021;65:1571-1590.

doi: 10.1007/s40194-021-01123-1

10. Vázquez L, Rodríguez N, Rodríguez I, Alberdi E, Alvarez P. Influence of interpass cooling conditions on microstructure and tensile properties of Ti-6Al-4V parts manufactured by WAAM. *Weld World*. 2020;64:1377-1388.

doi: 10.1007/s40194-020- 00921-3

11. Reisgen U, Sharma R, Mann S, Oster L. Increasing the manufacturing efficiency of WAAM by advanced cooling strategies. *Weld World*. 2020;64:1409-1416.

doi: 10.1007/s40194-020-00930-2

12. Ma Y, Cuiuri D, Shen C, Li H, Pan Z. Effect of interpass temperature on in-situ alloying and additive manufacturing of titanium aluminides using gas tungsten arc welding. *Addit Manuf*. 2015;8:71-77.

doi: 10.1016/j.addma.2015.08.001

13. Shen C, Pan Z, Cuiuri D, Ding D, Li H. Influences of deposition current and interpass temperature to the Fe3Al-based iron aluminide fabricated using wire-arc additive manufacturing process. *Int J Adv Manuf Technol*. 2017;88:2009-2018.

doi: 10.1007/s00170-016-8935-3

14. Wu B, Ding D, Pan Z, *et al.* Effects of heat accumulation on the arc characteristics and metal transfer behavior in Wire Arc Additive Manufacturing of Ti6Al4V. *J Mater Process Technol*. 2017;250:304-312.

doi: 10.1016/j.jmatprotec.2017.07.037

15. Xiong J, Li Y, Li R, Yin Z. Influences of process parameters on surface roughness of multi-layer single-pass thin- walled parts in GMAW-based additive manufacturing. *J Mater Process Technol*. 2018;252:128-136.

doi: 10.1016/j.jmatprotec.2017.09.020

16. Yang D, Wang G, Zhang G. Thermal analysis for single-pass multi-layer GMAW based additive manufacturing using infrared thermography. *J Mater Process Technol*. 2017;244:215-224.

doi: 10.1016/j.jmatprotec.2017.01.024

17. Kumar V, Mandal A, Das AK, Kumar S. Parametric study and characterization of wire arc additive manufactured steel structures. *Int J Adv Manuf Technol*. 2021;115:1723-1733.

doi: 10.1007/s00170-021-07261-6

18. Mai DS, Doan TK, Paris H. Wire and arc additive manufacturing of 308L stainless steel components: Optimization of processing parameters and material properties. *Eng Sci Technol Int J*. 2021;24:1015-1026.

19. Chaudhari R, Parmar H, Vora J, Patel VK. Parametric study and investigations of bead geometries of GMAW-based wire-arc additive manufacturing of 316L stainless steels. *Metals*. 2022;12:1232.

doi: 10.3390/met12071232

20. Natrayan L, Anand R, Kumar SS. Optimization of process parameters in TIG welding of AISI 4140 stainless steel using Taguchi technique. *Mater Today Proc*. 2021;37:1550-1553.

doi: 10.1016/j.matpr.2020.07.150

21. Vora J, Parikh N, Chaudhari R, *et al.* Optimization of bead morphology for GMAW-based wire-arc additive manufacturing of 2.25 Cr-1.0 Mo steel using metal-cored wires. *Appl Sci*. 2022;12:5060.

doi: 10.3390/app12105060

22. Kumar A, Maji K. Selection of process parameters for near-net shape deposition in wire arc additive manufacturing by genetic algorithm. *J Mater Eng Perform*. 2020;29:3334-3352.

doi: 10.1007/s11665-020-04847-1

23. Liberini M, Astarita A, Campatelli G, *et al.* Selection of optimal process parameters for wire arc additive manufacturing. *Procedia Cirp.* 2017;62:470-474.

doi: 10.1016/j.procir.2016.06.124

24. Wang L, Wu T, Wang D, *et al.* A novel heterogeneous multi-wire indirect arc directed energy deposition for *in-situ* synthesis Al-Zn-Mg-Cu alloy: Process, microstructure and mechanical properties. *Addit Manuf*. 2023;72:103639.

doi: 10.1016/j.addma.2023.103639

25. Mishra V, Ayas C, Langelaar M, Van Keulen F. Simultaneous topology and deposition direction optimization for Wire and Arc Additive Manufacturing. *Manuf Lett*. 2022;31:45-51.

doi: 10.1016/j.mfglet.2021.05.011

26. Xiong J, Zhang G, Hu J, Li Y. Forecasting process parameters for GMAW-based rapid manufacturing using closed-loop iteration based on neural network. *Int J Adv Manuf Technol*. 2013;69:743-751.

doi: 10.1007/s00170- 013-5038-2

27. Xiong J, Zhang G, Hu J, Wu L. Bead geometry prediction for robotic GMAW-based rapid manufacturing through a neural network and a second-order regression analysis. *J Intell Manuf*. 2014;25:157-163.

doi: 10.1007/s10845- 012-0682-1

28. Ríos S, Colegrove PA, Martina F, Williams SW. Analytical process model for wire + arc additive manufacturing. *Addit Manuf*. 2018;21:651-657.

doi: 10.1016/j.addma.2018.04.003

29. Nagesh DS, Datta GL. Prediction of weld bead geometry and penetration in shielded metal-arc welding using artificial neural networks. *J Mater Process Technol*. 2002;123:303-312.

doi: 10.1016/S0924-0136(02)00101-2

30. Rao PS, Gupta OP, Murty SSN, Rao ABK. Effect of process parameters and mathematical model for the prediction of bead geometry in pulsed GMA welding. *Int J Adv Manuf Technol*. 2009;45:496-505.

doi: 10.1007/s00170-009-1991-1

31. Bourlet C, Zimmer-Chevret S, Pesci R, Bigot R, Robineau A, Scandella F. Microstructure and mechanical properties of high strength steel deposits obtained by Wire-Arc Additive Manufacturing. *J Mater Process Technol*. 2020;285:116759.

doi: 10.1016/j.jmatprotec.2020.116759

32. Spencer JD, Dickens PM, Wykes CM. Rapid prototyping of metal parts by three-dimensional welding. *Proc Inst Mech Eng Part B J Eng Manuf*. 1998;212:175-182.

doi: 10.1243/0954405981515590

33. Sthle L, Wold S. Analysis of variance (ANOVA), chemometrics and intelligent laboratory systems. 1989;6:259-272.

34. Turgut B, Gürol U, Onler R. Effect of interlayer dwell time on output quality in wire arc additive manufacturing of low carbon low alloy steel components. *Int J Adv Manuf Technol*. 2023;126:5277-5288.

doi: 10.1007/s00170-023-11481-3

35. Suryakumara S, Karunakaran KP, Bernard A, Chandrasekhar U, Raghavender N, Sharma D. Weld bead modeling and process optimization in Hybrid Layered Manufacturing. *Comput Aided Des*. 2011;43:331-344.

doi: 10.1016/j.cad.2011.01.006

36. Hu Z, Qin X, Li Y, Ni M. Welding parameters prediction for arbitrary layer height in robotic wire and arc additive manufacturing. *J Mech Sci Technol*. 2020;34:1683-1695.

doi: 10.1007/s12206-020-0331-0

37. Wang L, Xue J, Wang Q. Correlation between arc mode, microstructure, and mechanical properties during wire arc additive manufacturing of 316L stainless steel. *Mater Sci Eng A*. 2019;751:183-190.

doi: 10.1016/j.msea.2019.02.078

38. Wahsh L, ElShater A, Mansour A, *et al.* Parameter selection for wire arc additive manufacturing (WAAM) process. *Mater Sci Technol*. 2018;1:78-85.

doi: 10.7449/2018mst/2018/mst_2018_78_85

39. Patel VK, Savsani VJ. A multi-objective improved teaching-learning based optimization algorithm (MO-ITLBO). *Inf Sci*. 2016;357:182-200.

doi: 10.1016/j.ins.2014.05.049

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

Layer porosity in powder-bed fusion prediction using regression machine learning models and time-series features

Vivek Mahato[1,2,3,4], Suman Chatterjee[1,2,4]*, Anesu Nyabadza[1,2,4], Annalina Caputo[1,3,4], and Dermot Brabazon[1,2,4]*

[1]I-Form Advanced Manufacturing Research Centre, Dublin City University, Dublin, Ireland

[2]School of Mechanical and Manufacturing Engineering, Dublin City University, Dublin, Ireland

[3]School of Computing, Dublin City University, Dublin, Ireland

[4]Advanced Processing Technology Research Centre, Dublin City University, Dublin, Ireland

(This article belongs to the *Special Issue: AI Usage in the Analysis of the Additive Manufacturing Process*)

Abstract

Additive manufacturing (AM) using laser powder-bed fusion (L-PBF) has become a common industrial process for high-end component production. The uptake of the process has been accelerated through the broad acceptance of the L-PBF process toward achieving high-quality parts with complex geometry. However, the L-PBF process faces challenges from the process's sensitivity to the process build parameters, which, when incorrectly set, can cause defects such as porosity, which in turn have a detrimental effect on the produced part properties. On the other hand, the AM processing equipment generates a vast amount of data captured through *in situ* sensors such as pyrometers and imaging cameras. Having such an abundance of process data facilitates the employment of advanced machine learning (ML) tools to understand and extract patterns and information about the underlying AM process and gain "predictive control." Driven by this idea, we aimed to employ ML tools over pyrometer time-series data from an L-PBF process to predict the porosity percentage of layers of an AM-built part. Sensor data are naturally modeled by time series; however, most ML algorithms work with tabular data (*i.e.*, one single vector describes a feature). In the work presented here, feature engineering tools were used to transform the time-series data into informative features. These features were fed into the tabular ML algorithms for evaluation, broadening the selection of ML algorithms available in the literature. It was hypothesized that the time-series summary features would capture the interaction of melt-pool temperature with resulting porosity, from which the resulting models could better predict porosity occurrence. The dataset contains layer porosity values in the range of 0.00175 – 7.160%, to which we divide the data into "low" and "high" porous layers using a splitting threshold value of 1%. From evaluating these algorithms, it was concluded that classifying "low" versus "high" porosity layers is relatively easier than predicting the layer's porosity percentage.

Keywords: Additive manufacturing; Powder-bed fusion laser-beam; Machine learning; TS-Fresh

***Corresponding authors:**
Dermot Brabazon (dermot.brabazon@dcu.ie)
Suman Chatterjee (suman.chatterjee@dcu.ie)

Citation: Mahato V, Chatterjee S, Nyabadza A, Caputo A, Brabazon D. Layer porosity in powder-bed fusion prediction using regression machine learning models and time-series features. *Int J AI Mater Design.* 2024;1(3): 33-49. doi: 10.36922/ijamd.4812

Received: September 10, 2024

Accepted: November 13, 2024

Published Online: December 16, 2024

1. Introduction

Additive manufacturing (AM) is the process of constructing products by joining materials layers upon layers facilitated by a three-dimensional (3D) computer-aided design (CAD), which differs from traditional subtractive manufacturing processes where products are fabricated by removing excess material from a larger block.[1] Depending on the industry, AM equipment has the flexibility of using a wide array of materials such as glass, metals (steel, titanium, and gold), and thermoplastics (polycarbonate and polylactic acid).[2] AM can decrease energy and material wastage, deliver greater design flexibility, and prolong product life. AM encompasses various processes for building 3D objects layer by layer using different materials. These processes include laser powder-bed fusion (L-PBF), direct energy powder metal deposition, and wire arc AM.[3,4] L-PBF is a metal AM technique that uses a laser beam to selectively melt and fuse metal powder to create dense 3D metal parts.[3,5] It offers advantages such as design freedom and is widely used to manufacture complex metal parts.[6]

L-PBF for AM faces several challenges. The quality and reliability of parts produced are crucial aspects that need monitoring and improvement.[7] The non-uniform melt pool size is a common issue, affecting the printed parts' strength and dimensional accuracy.[8-10]

If the major challenge in AM of improved reproducibility of product quality could be solved, the process could be automated.[11] A recent study evaluated the process capability and baseline variations between multiple L-PBF machine configurations, finding a systematic error in build accuracy across all features.[12] Another study assessed the feasibility of using L-PBF for remanufacturing and identified challenges such as misalignment in restoration, keyhole defects, and gas pores in the boundary area between original and restored parts.[13] A round-robin test investigated the reliability of additive-manufactured specimens and found that optical microstructure inspection was beneficial in determining true porosity and the effects of powder scatter.[13] An inter-laboratory test on reproducing porous samples for sound absorption using different AM technologies highlighted discrepancies due to shape and surface imperfections induced by the manufacturing process.[14] The need for reproducibility and benchmarking in AM is also emphasized in surgical robotics, where a lack of systematic approaches and accurate experimental descriptions hinders progress.[15]

Machine learning (ML) approaches have been increasingly used to predict additively manufactured products' quality and mechanical behavior. These approaches offer advantages over traditional statistical methods in handling complex and non-linear patterns in manufacturing process data. Several studies have demonstrated the effectiveness of ML algorithms such as neural networks, support vector machines, and decision trees (DTs) in predicting the mechanical characteristics and behavior of 3D-printed materials.[16,17] Table 1 presents an overview of the literature on ML approaches for predicting defects in the L-PBF process.

ML algorithms that predict the quality of the built product can significantly enhance the sustainability of the manufacturing process in real time as they reduce material and energy wastage to construct good quality products. When an ML system predicts an anomalous layer (like high porosity) is laid, it can alert the operator; the layer could be re-melted to remove the pores to improve or maintain the part's quality and characteristics.

Previous studies have applied various ML approaches to detect irregularities in additively fabricated parts produced by L-PBF processes. However, limited research has focused on determining the porosity of these parts. Typically, porosity is assessed using post-processing techniques such as X-ray, computed tomography (CT) scans, and synchrotron-based micro-CT. Notably, there is a gap in the literature regarding the use of *in situ* approaches for porosity detection.

This paper presents the development of a new data analysis algorithm for the detection of defects which may occur naturally during the L-PBF production process. In this process, a laser beam scans the part CAD design on the powder bed, which melts and fuses the powder to the underlying solid material. Next, the "build chamber actuator" (Figure 1A) automatically lowers by a specified layer thickness, and a new layer of powder is laid. The process continues until all the designed layers are fused, which completes the part manufacturing process. Furthermore, to prevent any oxidation during fusion, the build chamber is supplied with a continuous flow of inert gas (like argon).[30] Although the L-PBF process shows a lot of potential, it is susceptible to defects occurring within the built part related to the set process parameters (such as layer thickness, laser power, and scan speed). Common problems arising from the L-PBF process are the occurrence of pores, voids, and lack of fusion owing to under- or over-melting.[31] Due to the process's sensitivity to the build parameters, much research has been done on *in situ* monitoring of the L-PBF process.[32]

AM processes such as L-PBF generate a wealth of data,[33] often in the form of raw time series, and are challenging to process owing to their size. For example, *in situ* pyrometers are utilized to gauge the reflected optical emissions emanating from the melt-pool region which can be translated into melt pool temperature data (as depicted in Figure 1B, details are discussed in section 2). Connecting the acquired data to the product's quality is a formidable challenge.[34] ML can be used to analyze these raw time-series data to predict quantifiable outcomes, such as porosity, surface roughness, and fatigue life.[1,31] Previous studies[1,31] have shown that working with

Table 1. Overview of proposed machine learning algorithms to predict defects during the L-PBF process

No.	Insights of research	References
1	• The paper presents a deep learning approach for defect detection. • It focuses on porosity and melt pool geometry segmentation. • Utilizes encoder-decoder networks with optimization techniques for accuracy. • Addresses challenges in noisy microscopy data for multi-layer structures. • Demonstrates superior performance in identifying multiple features simultaneously.	19
2	• The paper proposes a neural network for *in situ* defect prediction. • It focuses on microporosity localization in L-PBF. • Utilizes within hatch stripe sensory data for improved accuracy. • Achieves classification accuracy of 73.13% for porosity detection. • Demonstrates a significant improvement in detecting small porosities • Aims to enhance process control and defect mitigation.	20
3	• The paper addresses porosity detection in AM. • ML methods are compared for porosity classification. • DCNNs outperform traditional methods in accuracy. • DCNN achieved 95% accuracy.	21
4	• The paper develops a model for predicting keyhole porosity. • It uses a closed-form analytical approach without numerical calculations. • The model predicts porosity based on molten pool characteristics. • The proposed model shows good predictive accuracy and computational efficiency.	22
5	• The paper presents a physics-informed ML model for porosity analysis. • It addresses the limitations of machine-dependent porosity prediction models. • The model interprets machine settings into physical effects. • It predicts porosity levels using "pass," "flag," and "fail" categories. • The model achieved a prediction error of 10 – 26%.	23
6	• The paper presents a ML approach for defect detection. • It integrates fuzzy logic and self-organizing maps for analysis. • The model predicts a lack of fusion and keyhole defects. • Experimental validation shows strong performance across various parameters. • Customizable fuzzy rules enhance defect detection accuracy.	24

(Cont'd...)

Table 1. ***(Continued)***

No.	Insights of research	References
7	• Deep learning predicts porosity in L-PBF additively manufactured parts. • ANN model trained with X-ray CT images for accuracy. • Synthetic CT data enhances model performance and predictions. • The study addresses porosity challenges in AM.	25
8	• The paper predicts micropore defects in L-PBF using thermal imaging. • ML models analyze *in situ* thermographic data for predictions. • Key features include time for achieving melting threshold and maximum radiance.	26
9	• The paper presents DSPMs. • DSPMs quantify porosity in metal additive manufacturing. • Synchrotron-based micro-computed tomography identifies defect trends. • Ti-6Al-4 V test blocks were fabricated using varied parameters. • Keyhole and lack-of-fusion defects were analyzed and mitigated. • Processing parameters significantly affect defect formation in L-PBF materials.	27
10	• The paper studies porosity in L-PBF. • X-ray tomography reveals pore formation mechanisms and characteristics. • Pore size and shape vary with process parameters. • Increased power leads to more keyhole mode porosity. • Lack of fusion occurs with poor hatch overlap. • Optimized parameters improve part density and quality. • Insights assist in quality control and process improvement.	28
11	• The study investigates AlSi10Mg alloy properties via L-PBF. • Sub-optimal parameters affect density and microstructure. • Defect orientation impacts mechanical properties significantly. • Findings are relevant for various L-PBF-fabricated AlSi10Mg alloys.	29
12	• The paper investigates pore formation in L-PBF AM. • Six pore formation mechanisms were identified during the L-PBF process. • Pores significantly affect mechanical performance and fatigue life. • Understanding mechanisms aids in developing pore mitigation strategies.	30

(Cont'd...)

Table 1. *(Continued)*

No.	Insights of research	References
13	• The paper develops a predictive model for porosity in additive manufacturing. • Focuses on high-energy-density regimes in powder-bed fusion. • Discusses bubble formation and trapping during solidification. • Aims to control and reduce porosity in manufacturing processes. • Highlights the significance of porosity on fatigue performance.	31

Abbreviations: AM: Additive manufacturing; CT: Computed tomography; DCNN: Deep convolutional neural network; DSPMs: Defect structure process maps; L-PBF: Laser powder-bed fusion; ML: Machine learning.

raw time-series data is essential, as it retains the information regarding the underlying process, which is sequential by nature, that could be lost when using summary statistics (such as data mean and variance values). Classification or regression based ML methods can be used to analyse time series data.[39-46] The extent of data available for training models can be increased via the combined efforts of increased experimentation as well as ML data boosting methods.[47-53].

In this research, the problem of detecting pores during the manufacturing process is examined via two different tasks: predicting an AM block's layer to be of low or high porosity (a binary classification task) and predicting the porosity percentage of each of the layers (a regression task). The capability to successfully predict such events can be translated into detecting faulty layers as they occur, allowing the process to be stopped or corrected, resulting in significant waste reduction. The temperature time series is one of the main contributing factors to pore formation which is captured via an in-line pyrometer. The resulting time-series data can capture the complex temporal interaction between the melt-pool temperature and the porosity of that zone. For this purpose, we have created a dataset of naturally occurring pores (from a build of 10 mm × 10 mm × 10 mm nickel-titanium [NiTi] metal blocks) for developing and analyzing the prediction and classification steps. The dataset reports the melting pool's temperature as the laser traces the part on the powder-bed bit-by-bit and layer-by-layer. Following a dataset analysis, we devised a pipeline to balance the data using data augmentation. Based on the dataset, we employed a package of strategies that take advantage of the sequential nature of these time-series data to extract summary features that traditional ML models can effectively optimize.

The following section defines our system for capturing data and describes the used data augmentation technique. Next, we provide brief descriptions of our feature engineering and summarizing tool and all the ML algorithms employed (using Python 3.11) in Section 3. In Section 4, we summarize our evaluation pipeline, the two ML tasks (classification and regression) operated on the AM data, and the evaluation metrics. The results are then presented in Section 5, with conclusions and future work discussed at the end of the paper.

2. Materials and methods

2.1. Experimental setup

Rectangular objects were fabricated using the L-PBF process in the present research work. The specimens were produced using the AconityMINI 3D laser powder-bed fusion-laser system (Aconity, Germany), which incorporates a 200-Watt 1068-nanometer Ytterbium (Yb) fiber laser (Figure 2). The laser system can attain up to 2000 mm/s velocities and possesses a minimum laser beam diameter of 32 mm.[54] The "Skywriting" functionality was activated to fabricate the rectangular samples $10 \times 10 \times 10$ mm^3 (length, width, height), and each sample had identical dimensions to maintain the uniformity of the study. The samples were produced on a NiTi substrate with a diameter of 125 mm. The L-PBF-fabricated samples may contain various anomalies in the manufactured parts, such as internal porosity, hot cracking, or the formation of material aggregation on the component's surface.[55] The present study uses ML algorithms to identify and predict the internal porosity in the built-up samples. To oversee the current manufacturing procedure, two *in situ* pyrometers from Kleiber Infrared GmbH were used to record the melt pool temperature at 100 kHz at each x, y, z location.

The focus of this investigation is to predict the porosity of an AM block's layer based on its temperature time-series. Based on the porosity distribution manifested by the dataset, the plan is to target this problem as both classification and regression tasks. Classification predicts whether the layer is of low or high porosity (*i.e.*, one binary classification). Subsequently, the regression models, one for low porosity (0 – 1%) and another for high porosity (>1%) samples predict the layer's porosity percentage, a non-negative real number. The threshold value (of 1%) was selected based on a visible split in porosity distribution. The details of the aforementioned ML tasks and their evaluation metrics and strategies are presented in the following subsections.

2.2. Exploratory data analysis (NiTi×3B data)

Employing AconityMINI, we constructed a batch of 10 mm × 10 mm × 10 mm NiTi metal blocks with varying layer thicknesses (30, 60, and 90 microns). The samples were fabricated according to a 2^3 design of experiments, resulting in eight sample fabrication permutations using the process parameters of laser power (160 W and 200 W), scanning

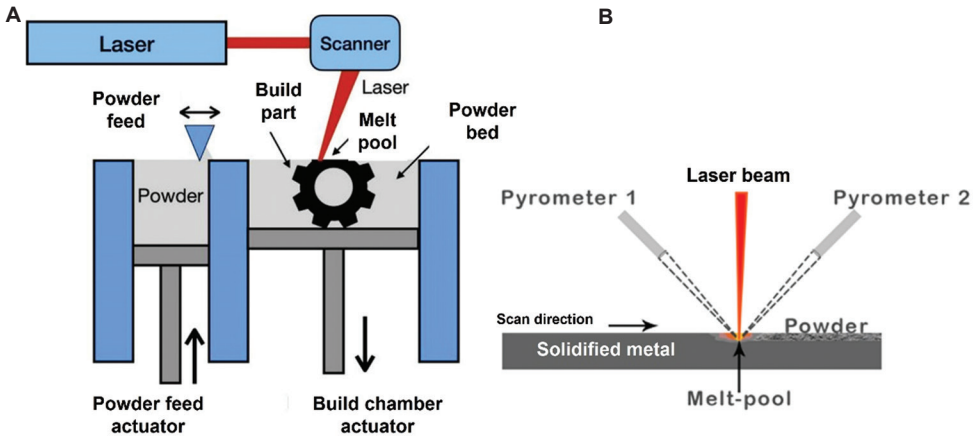


Figure 1. Schematic of (A) layout of AconityMINI's laser powder-bed fusion manufacturing process and (B) laser beam melting the powder and leaving behind a solidified metal trace with the two *in situ* pyrometers capturing the melt-pool temperature[1].

Figure 2. AconityMINI 3D printer utilized for the presented study to produce the sample and record the melt pool temperatures[1].

speed (750 mm/s and 1100 mm/s), laser spot size (60 µm and 80 µm) and with hatch spacing fixed at 80 µm. The pyrometers detect the temperature-correlated infrared emission in the range of 1500 – 1700 nm, denoting the melt-pool temperature. The light is divided into two tracks through optical filters and transmitted through optical fiber cables to the pyrometers. Similar to the study by Mahato *et al.*,[1] the scanner and the pyrometers are configured to cover x and y values (for each layer) in the range of 0 – 32,768 bit covering a physical area of 400 × 400 mm, which results in a calibration value of 81.92 bit/mm. In addition, the frequency of the sensors is set to 100 kHz. Figure 3A displays the pyrometer data of a single layer from a block. The porosity of the block was measured using a µ-CT scanner. The scanner captures a series of two dimensional (2D) X-ray images (Figure 3B) which are then reconstructed and processed into a 3D model. The percentage of porosity was calculated for each of these 2D images.

The 2D slices recorded in the µ-CT scan each cover a block's layer thickness of 8.337 µm, which is lower than the three AM layer thicknesses used during the building of the block (30, 60, and 90 µm). Therefore, there are multiple µ-CT 2D image slices in every pyrometer layer (i.e., images per layer = layer thickness/8.337). Hence, concerning the layer thickness of the AM built block, we take the average porosity of the image slices belonging to a particular block's layer. Figure 4 illustrates the combined porosity distribution of the layers in the three blocks. The plot illustrates a significantly higher count of layers (around 82% of the dataset) with low porosity (0 – 0.7%), while a small percentage of layers (around 18% of the dataset) has high porosity (2 – 7%).

Further discussion on the datasets and the ML tasks employing it is presented in the evaluation section (section 4).

2.3. Synthetic minority oversampling technique (SMOTE)

Like many anomaly detection datasets, the NiTi blocks dataset (NiTi×3B) suffers from a substantial imbalance in the data, where the occurrence of one event (*i.e.*, porosity) is much lower than the number of events when no porosity occurs. Therefore, the majority of the samples belong to the "ideal" category, and very few represent anomalous samples. Hence, one main challenge of employing ML over these imbalanced datasets is that most algorithms overlook the minority samples. However, it is the performance over minority samples the most critical or of interest.

One strategy widely utilized in literature to tackle problems related to imbalanced datasets is data augmentation. This technique increases the number of data samples by adding new synthetic data or modified copies of existing samples in the dataset. SMOTE[35] is a data augmentation technique that oversamples minority classes by synthesizing them from existing samples.

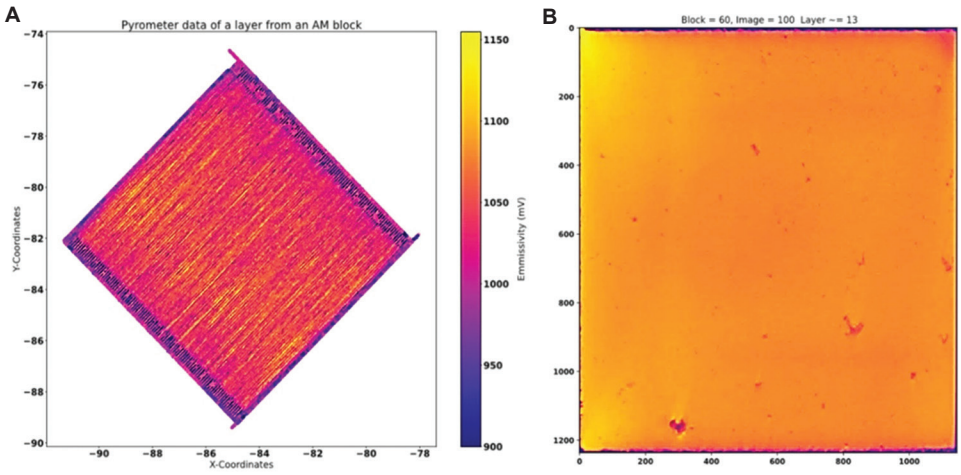


Figure 3. Illustration of the data from a layer of an additive manufacturing-built NiTi block with 60-micron thickness. (A) Pyrometer data and (B) slice of μ-computed tomography scan representing the layer.

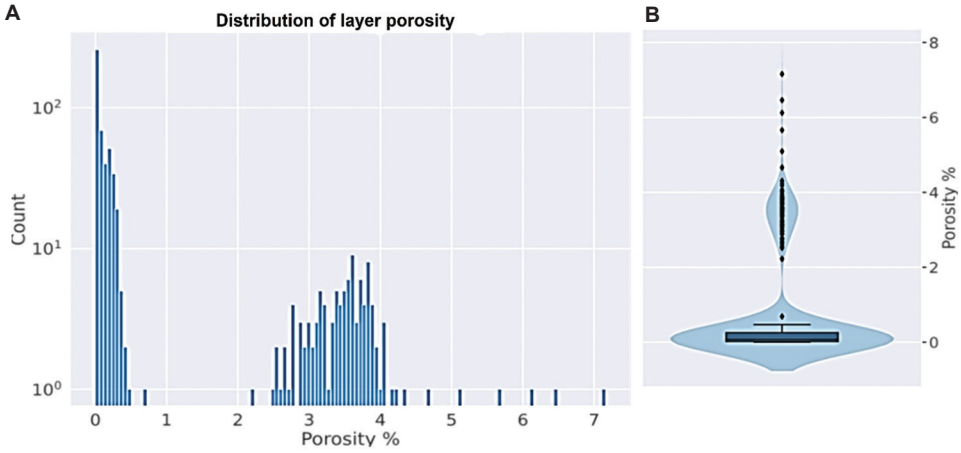


Figure 4. Distribution of layer porosity through the dataset: (A) Histogram and (B) violin and boxplot.

First, SMOTE randomly selects one sample (s_r) from the minority class in dataset S (NiTi×3B in our case) and then finds the k nearest neighbors (NNs) of the selected sample s_r in the feature space to create a set N. Next, it randomly selects a neighbor n_r (where $n_r \in N$) and connects s_r and n_r to form a line segment (i.e. $\overline{s\_rn\_r}$) in the feature space. A new synthetic sample is created by selecting a random point (*i.e.*, a vector in the feature space) along this line segment. The process continues until the dataset is reasonably balanced and the resultant modified dataset is D (Algorithm 1).

SMOTE's reasonable strategy for synthetic data generation, which directly involves minority samples, makes the newly constructed data relatively close to the minority samples in the feature space, making the approach effective.

2.3.1. SMOTE-edited NN (ENN)

One drawback of SMOTE is that the algorithm overlooks the majority class during data augmentation. According to Chawla *et al.*,[35] combining the random undersampling of the majority class and SMOTE (to oversample the minority class) offers better performance than the simple undersampling approach.

SMOTE uses the NN strategy to synthesize new samples and maintain closer proximity to the samples belonging to the same class in the feature space (*i.e.*, samples sharing a class are closer to one another than samples of a different class). Like most NN algorithms, which are susceptible to noise, SMOTE also suffers from creating many noisy data points in the feature space. Regarding handling such a scenario, ENN[36] is a popular algorithm that finds and removes ambiguous and noisy samples from the dataset.

When used with SMOTE, ENN is an undersampling method for extensive data cleaning over the oversampled dataset to achieve much cleaner clustered samples. For each sample d_i in the SMOTE oversampled dataset D, ENN finds d_i's NNs set N of size k. If the d_i's class y_i and the majority

Algorithm 1:
SMOTE

Data: S: Original Dataset
Result: D: Modified Dataset
/* θ_k=user-defined k nearest neighbor to consider (by default 3) */
$k \leftarrow \theta_k$

/* iterate the algorithm until the desired class balance in dataset S is achieved */
while *notBalanced* (S) do

/* get a random sample from the minority class in set S */
$s_r \leftarrow random\ (S)$

/* get s_r's k nearest neighbors in S */$N \leftarrow$nearestNeighbors (s_r)

/* get a random neighbor in set N */$n_r \leftarrow random\ (N)$

/* create a sample at a random location along the line segment connecting s_r and n_r in the feature-space */
$s \leftarrow createSample\ (s_r, n_r)$

/* add the minority class sample to set S */
$S \leftarrow S \cup \{s\}$
end

$D \leftarrow S$

Algorithm 2:

Edited Nearest Neighbor

Data: D: Modified dataset after SMOTE
Result: T: Reduced set
/* θ_k = user-defined k nearest neighbor to consider (by default 3) */
$k \leftarrow \theta_k$
while notProportionate (D) do
/* iterate through all samples in D */for $d_i \in D$ do

/* get d_i's k nearest neighbors in D */
$N \leftarrow nearestNeighbors\ (d_i, D, \text{k})$

/* get d_i's class */
$y_i = getClass\ (d_i)$

/* get majority class in the set N */
$\hat{y}_i = getMajorityClass\ (N)$

/* if y_i and $\hat{y}_i$ do not match, remove sample d_i from D */
if $y_i \neq \hat{y}_i$ then
$D \leftarrow D \setminus \{di\}$
end
end
end
$T \leftarrow D$

class in set N $\hat{y}_1$ do not match, then it removes d_i from D. ENN repeats the process until a desired proportion of each class is fulfilled. The resultant reduced dataset after the procedure is T (Algorithm 2).

2.3.2. SMOTE for regression

SMOTE is a sampling strategy to handle datasets with imbalanced class distribution for classification tasks. SMOTE for Regression (SMOTER)[37] is an adaptation of SMOTE for regression tasks where the objective is to predict rare extreme values, which are more critical herein. Like SMOTE, SMOTER oversamples the observations with rare extreme values and undersamples the remaining normal observations, balancing the distribution of the values. A user-defined threshold value determines the distinction between rare and normal values. The algorithm also accepts parameters that govern the percentages of desired under- and over-sampling and the k-NNs to consider generating the new synthetic samples.

Traditional ML algorithms, such as DTs and Naïve Bayes, work on datasets with feature-vector representation (*i.e.*, a single vector describing a feature). A time series is a series of data points taken sequentially in time. When it comes to features described by time series, models like DTs fails to work. These models treat each data point in the series independently and hence miss the opportunity of capturing the series's sequential information. ML algorithms working with time series can take two different approaches. Either special ML time-series models are employed (like k-NN with Dynamic Time Warping, Time-Series-Forest, and BOSS), or the time series itself is transformed into a set of features that describe the sequence on which the traditional ML models are then optimized. Deep Learning (DL) does exhibit high performance for such tasks and provides a flexible framework to model sequential data; however, in some contexts, the lack of big data currently available in AM restricts DL's application. Hence, this paper investigates applying the traditional ML algorithm to time-series data without losing sequential information. For this purpose, we explored by using summary tools capable of extracting several features characterizing the time series. The summary vectors can then be employed in a standard ML algorithm for classification and regression tasks.

In the following subsections, we describe the summary tool (TS-Fresh package) and the ML models employed in this study.

2.3.3. Time series feature extraction on the basis of scalable hypothesis tests

The wide availability of cheaper sensors allowed their usage in various domains such as manufacturing, health, and

automotive industries. This has resulted in the availability of big temporally annotated datasets, generally known as time series. Like many others, these time-series datasets may also suffer from poor data quality, noise, missing values and redundancy. Therefore, it is essential that proper feature engineering[1] is conducted on the time-series data, which considers its sequential nature before providing the data to the ML models.

Time series feature extraction on the basis of Scalable Hypothesis tests, also known as TS-Fresh,[38] is a Python package that automatically extracts and selects relevant features from the given time-series dataset. TS-Fresh consists of more than 63 time-series characterization strategies, which aids it to compute more than 1500 informative time-series summary features. Having too many features, especially if some of them are irrelevant, may impair ML models' performance. Hence, TS-Fresh provides a fast feature selection algorithm based on statistical hypothesis tests. This algorithm is configured to automatically remove redundant features based on their type and the intended ML task. Additionally, the hyperparameters were also tuned to optimize model performance. Grid SearchCV from the sklearn3 library was employed for each model and the optimal model based on lowest root-mean-square error (RMSE) value was considered.

2.4. ML algorithms

Herein several ML models were employed for the regression tasks including Linear Regression (LR), Support Vector Regression (SVR), k-NN, DTs, Random Forest (RF), Gradient Boosting (GB) Trees, and Extremely Randomized Trees (XT), were primarily compared for regression tasks. The models were evaluated using accuracy and RMSE.

For model evaluation, RMSE is a crucial metric for regression problems because it provides an absolute measure of prediction error, directly reflecting the deviation of predicted values from actual values in the original units. This makes RMSE more suitable for this regression tasks than classification metrics such as recall, precision, and F1-score, which focus on percentage-based evaluations for categorical predictions. Furthermore, RMSE allows for direct comparison with other ML models from the literature, as it provides a consistent and absolute measure across studies. Accuracy is also used in evaluation of both classification and regression, offering a holistic view of model performance by assessing how well the predictions align with actual outcomes, while RMSE remains the most appropriate and comparable metric for evaluating regression models. Lastly, an evaluation matrix, using previously unseen data, called absolute error matrix, was also employed during model validation.

The original dataset consists of 586 samples (479 low and 107 high), which pass through SMOTE-ENN's data augmentation algorithm. The resultant modified dataset consists of 931 samples (462 low and 469 high). Figure 5 illustrates the class distribution in the modified dataset's train and test splits. It is worth mentioning that prior experiments demonstrated that both undersampling and the use of imbalanced data resulted in lower accuracy and higher RMSE values. Consequently, oversampling was found to be the optimal approach for improving the performance of models when analyzing pyrometer data. Given the significant split between low and high porosity layers (illustrated in Figure 4), it is considered relatively straightforward, as this distinction is also reflected in the underlying pyrometer data.

Figure 6 illustrates the pipeline we employed in the experiments for a rigorous evaluation. First, we read the dataset through an initial preprocessing stage, where we divided the data into "low" and "high" porous layers using a splitting threshold value of 1%. Layers with <1% porosity belong to the "low" category, and the rest fall into the "high" category. We scaled the target variable (*i.e.*, porosity percentage) in both datasets to range from 0 to 1. Next, the datasets are passed to the TS-Fresh algorithm to generate or extract informative features (around 698) from the time-series data. Furthermore, to address the imbalance of the target class or its uneven distribution, we modified the dataset using SMOTE-ENN (for classification) or SMOTER (for regression).

An ML model is as good as the features in the data supplied to it. By removing irrelevant features, we enhanced the prediction power of the ML algorithm while speeding up its execution and decreasing its memory and computation costs.[56] We employed Recursive Feature

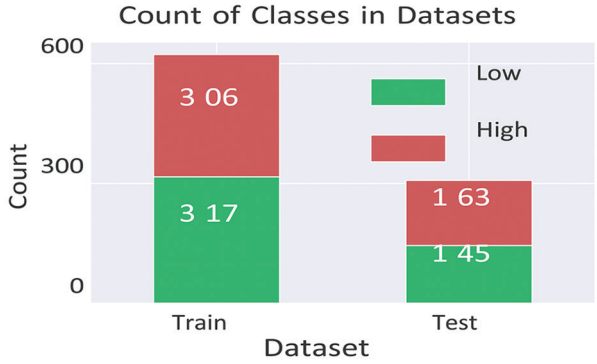


Figure 5. Illustration of class distribution in train and test datasets

[1] Feature engineering is the pre-processing step that transforms raw data into informative features that machine learning algorithms can utilize.

Elimination with Cross-Validation (RFECV) from the Scikit-learn library on the post-SMOTE-ENN or SMOTER dataset to achieve this. The idea is to eliminate irrelevant features for the intended ML task that TS-Fresh might have introduced. Furthermore, we split the final dataset obtained after the RFECV stage into a train and test dataset of size 67% and 33% of the total observations, respectively. We held off the test dataset until the final evaluation.

We employed the training data and evaluated all the models using 10-fold cross-validation. For the cross-validation mechanism (repeated Stratified K-Fold with the number of splits and repeats was employed). Based on the cross-validated scores obtained by these models, we selected the best performing model ($model_{best}$). In the next stage, $model_{best}$ went through intensive parameter tuning (cross-validated on the training set) to discover the best set of parameters to enhance its performance. We gauged the true performance of the $model_{best}$ by supplying it with the best parameters obtained in the last stage, training it over the training data, and evaluating its performance over the held-out test data. This evaluation pipeline's robustness guarantees that the ML model's final evaluation scores reflect its performance in unseen real-world environment or data.

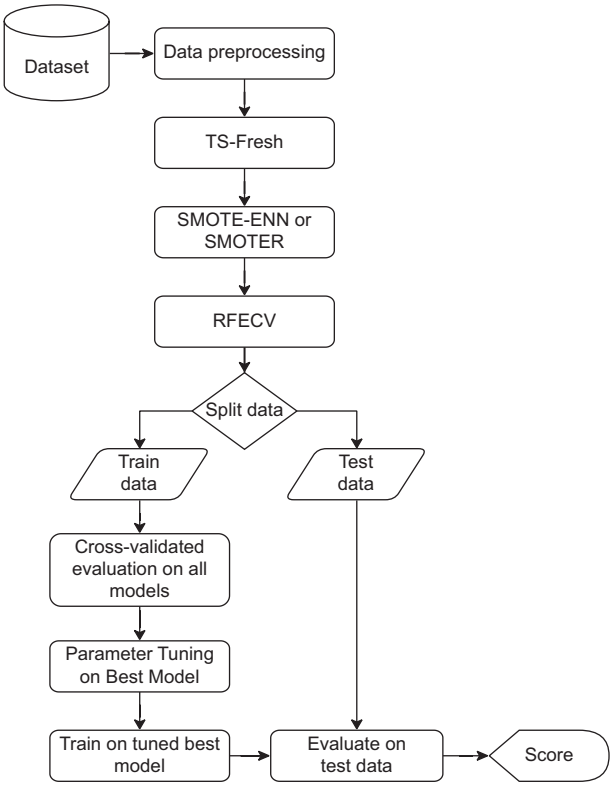


Figure 6. Evaluation pipeline employed in this study
Abbreviations: ENN: Edited Nearest Neighbor; RFECV: Recursive Feature Elimination with Cross-Validation; SMOTE: Synthetic minority oversampling technique; SMOTER: SMOTE for Regression.

The regression task is defined as predicting the porosity percentage of a given layer. We targeted this ML task with two different regression models, motivated by the significant divergence in porosity versus layer counts distributions observed in Figure 4. Therefore, the two regression tasks we created are associated with the two different datasets resulting from splitting the original dataset into "low" and "high" porosity layers (Figure 7).

The rationale behind this choice is that distinguishing between low and high porosity is easy. A simple classification algorithm should be able to achieve that. However, when predicting the actual porosity percentage of the layer, we need ad-hoc models capable of capturing the subtle differences in the temperature fluctuation during the formation of pores. Therefore, to avoid regression models inclining to either low- or high-porosity samples owing to the frequency of occurrence of either one of them, we targeted the problem with two models: one for low-porosity layers and another for high-porosity layers.

Given a query q, the suited regression model tries to predict the exact porosity percentage (*i.e.*, the target or dependent variable). Predicting the porosity percentage of

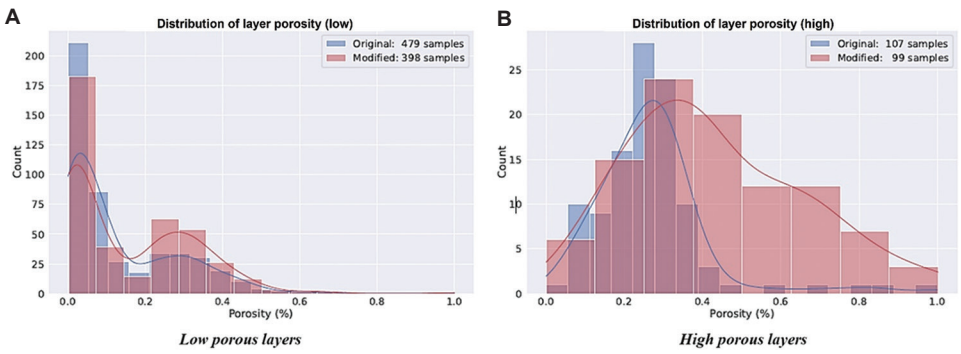


Figure 7. Original and modified (using SMOTER) distribution of layer porosity across samples. (A) Low-porosity layers and (B) high-porosity layers. To balance the distribution, SMOTER augments more high-porosity samples, which occurred less often.

a layer is critical in the manufacturing process since a slight variance in the porosity at this level could affect the overall quality of the built product.

As shown in Figure 7, the significant split in the distribution (x-axis) observed in Figure 3 is absent once we split the dataset. Hence, there is no clear distinction between the two distributions once we separate the dataset into "low" and "high" subsets. Therefore, predicting the porosity level should be more challenging than classification.

3. Results and discussion

3.1. Regression

All the regression models were evaluated using the RMSE separately on the "low" and the "high" porosity training datasets. The models' performances are shown in Figure 8. In Figure 8A shows the performance of the regression models on the "low" porosity dataset. Linear Regression and ensemble models (which are RF, XT, GB, Adaptive Boosting [AdaBoost], and eXtreme Gradient Boosting [XGBoost]) attained similar RMSE values, whereas the XT model achieved the lowest RMSE score of 0.04, signifying better performance. Regarding the performance on the "high" porosity dataset, both XT and AdaBoost are tied with the best RMSE score of 0.12. To keep our evaluations consistent and comparable, we selected XT regressor for both datasets and further investigated this model in the hyper-parameter optimization stage.

The best hyper-parameters for the XT regressor model from Grid SearchCV on the "low" dataset are: *n_estimators* = 700, *max_depth* = 76, *min samples_leaf* = 2, and *min_samples_split* = 2. The best hyper-parameters discovered for the model on the "high" dataset are: *n_estimators* = 841, *max_depth* = 138, *min_samples_leaf* = 2, and *min_samples_split* = 3.

As the final step, the two XT regressors supplied with the hyper-parameters discovered in the previous step separately on the "low" and "high" datasets were trained. Then, the models evaluate their performance on the unseen held-out test dataset. Figure 9 illustrates the absolute error and RMSE scores of the two models. The black dotted line represents the identity line, and the red line denotes the line of best fit.

Figure 10A shows the performance of the XT regressor on the "low" porosity dataset. The model achieves a low RMSE score of 0.0367, echoed in the absolute n error plot, as the best-fit line fairly aligns with the identity line. Likewise, the model has also achieved a relatively low RMSE score of 0.108 on the "high" porosity dataset (Figure 10B). However, the best-fit line has a slightly different slope compared to the "low" dataset result, showcasing a slightly weaker performance.

Table 2 provides a summary of the 10 regression models, including the general speed of each model based on the given dataset size, assuming GridSearch CV is used for hyperparameter tuning, which is considered part of the model training. The models displayed low RMSE values comparable to those reported by Ho *et al.*[57] who achieved RMSE values of around 5% when using DL models including CNN and transfer learning models such as VGG-16. These DL models generally take much longer to train, highlighting a strong advantage of these ML models. Liu *et al.*[21] also reported the use of ML in predicting porosity, achieving significantly higher RMSE values between 10% and 26%, which are much higher compared to those achieved in this study.

The execution time of Python scripts can be measured using the *time* function to provide an indication of model performance. LR is the fastest in terms of real-time processing due to the simplicity of solving a single linear equation and the absence of hyperparameter tuning. However, the preference for models like XT over LR, despite similar RMSE values (0.05 and 0.04) and slower speed, stems from their ability to capture non-linear patterns in the data. LR is limited to linear relationships,

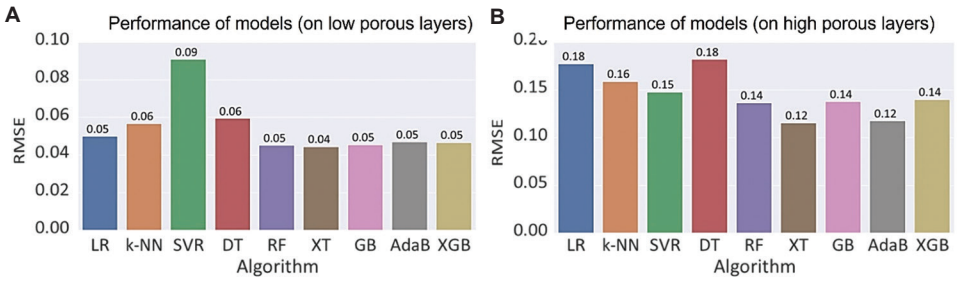


Figure 8. Performance of all the models on training dataset (lower RMSE is better): (A) Low-porosity layers and (B) high-porosity layers. Abbreviations: AdaB: Adaptive Boosting; DT: Decision Tree; GB: Gradient Boosting; k-NN: k-Nearest Neighbors; LR: Linear Regression; RF: Random Forest; RMSE: Root-mean-square error; SVR: Support Vector Regression; XGB: eXtreme Gradient Boosting; XT: Extremely Randomized Trees.

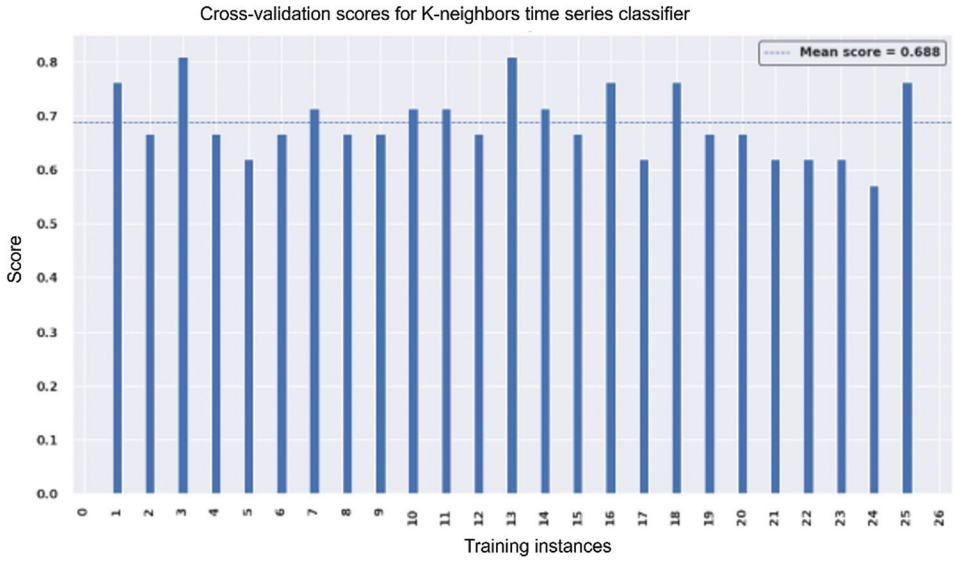


Figure 9. Accuracy score of k-Nearest Neighbors model trained with best set of parameters across cross-validation folds over training data. Here, the mean cross-validated accuracy is around 69%.

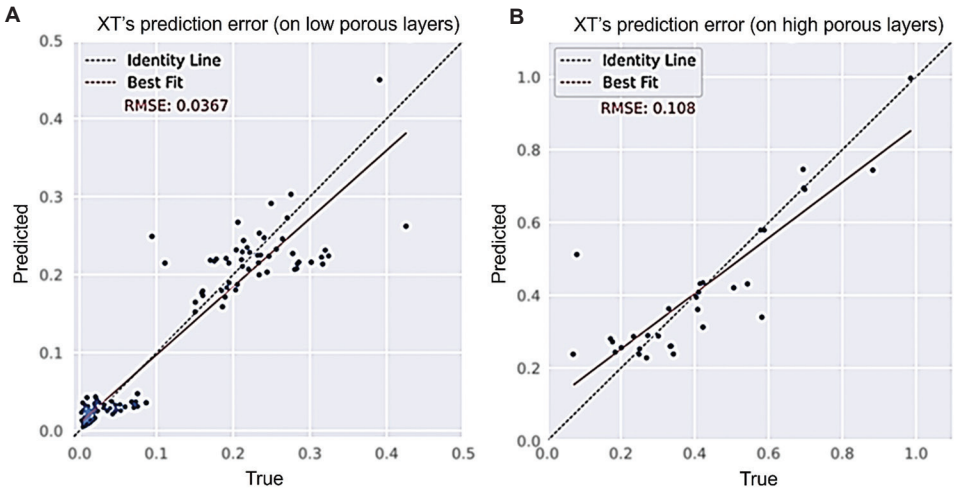


Figure 10. Absolute error plots showcasing the model's performance on test data: (A) Low-porosity layers and (B) high-porosity layers. Abbreviation: XT: Extremely randomized trees.

whereas XT can model more complex, non-linear trends, making it a better choice for problems where there is a risk of underfitting or the data exhibits more intricate patterns that linear models fail to capture. XT also mitigates the risk of overfitting by using multiple DTs and random splits, providing a more robust model.

3.2. XGBoost and k-NN for classification

Although the main aim of this study is the prediction of the porosity percentage (regression), it is also valuable to quickly classify samples as either low or high porosity for applications where rapid decision-making is required. In such cases, items with low porosity but still within usable limits can proceed to further analysis, such as applying the regression models discussed in section 3.1 to predict the exact amount of porosity. This dual approach helps streamline processes where quick classification is essential before more detailed assessments are conducted.

XGBoost and k-NN were employed for classification tasks, each bringing unique strengths to handling our large dataset. XGBoost is particularly well-suited for this scenario due to its speed, scalability, and ability to capture non-linear relationships, which makes it effective when working with large datasets. Its gradient-boosting

mechanism allows it to handle complex patterns in the data efficiently, while also offering adaptability through hyperparameter tuning, making it a popular choice in ML literature for large-scale classification tasks. Conversely, k-NN, though less efficient with large datasets due to the need to compute distances between all points, was used for its simplicity and interpretability. It provides a straightforward approach to classification by relying on proximity in feature space. By employing both models, we ensure that XGBoost handles complex, non-linear classifications efficiently, while k-NN provides a valuable comparison through its simpler, distance-based approach.

For obtaining the best set of parameters of the k-NN and XGBoost model, Grid SearchCV from the sklearn3 library was employed. For the cross-validation mechanism, Repeated Stratified K-Fold with the number of splits and repeats was employed. The system is only exposed to the training data for hyperparameter optimization. Figure 9 illustrates the model accuracies across the training instances (*i.e.*, iterations). Here, it can be observed that the best parameters render a mean accuracy of the k-NN model around 69% on training data. The k-NN model was trained on training data supplied with the best parameters obtained in the previous step and evaluated it on the held-out test data. As a result, the model predicts the classes of the held-out samples with an accuracy of 65%. Undergoing a cross-validated hyper-parameter tuning process for XGBoost on training data, the best set of parameters obtained are as follows: learning rate = 0.01, estimators = 400, gamma = 5, max depth = 5, min child weight = 10, subsample = 1.0, and colsample-bytree = 0.8. The XGBoost model attained an accuracy, recall, precision, and F1 score of ~100% after 10-fold cross-validation as demonstrated by the confusion matrix and classification report in Figure 11. The XGBoost model proved to be a significantly better classifier of porosity (high vs. low) compared to the k-NN model and is, therefore, the recommended choice for analyzing pyrometer data.

Table 2. Summary of the advantages and disadvantages of models used for the predictive quality assessment of porosity within metal additive manufacturing

Model	Training speed	Real-time classification speed	RMSE (low porosity)
LR	Very Fast	Very Fast	0.05
k-NN	Slow	Slow	0.06
SVR	Medium	Medium	0.09
DT	Fast	Very Fast	0.06
RF	Slow	Fast	0.05
XT	Medium	Fast	0.04
GB	Slow	Medium	0.04
AdaBoost	Slow	Medium	0.05
XGBoost	Fast	Fast	0.05

Abbreviations: AdaBoost: Adaptive Boosting; DT: Decision Tree; GB: Gradient Boosting; k-NN: k-Nearest Neighbors; LR: Linear Regression; RF: Random Forest; SVR: Support Vector Regression; XGBoost: eXtreme Gradient Boosting; XT: Extremely Randomized Trees.

In this study, pyrometer data was found to be valuable for predicting porosity percentage during AM, addressing a major issue in printed parts. Porosity can lead to reduced density and poor mechanical properties, but with the assistance of ML models, porosity can be detected in real-time, allowing for process parameters to be adjusted accordingly. The two *in situ* pyrometers, provided by KLEIBER Infrared GmbH, measured reflected light from the melt-pool area, detecting heat emissions in the 1500 – 1700 nm range, which represent the melt pool temperature. The data captured was mapped in x and y coordinates for each layer, with a resolution calibrated

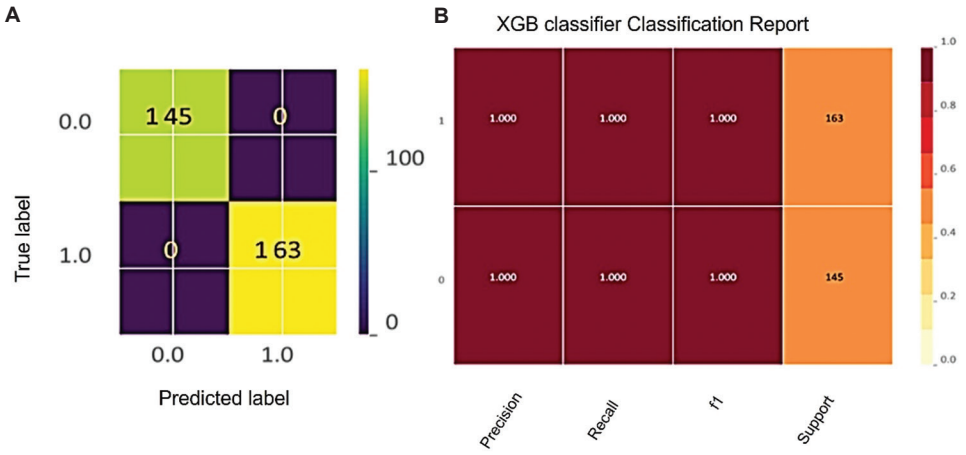


Figure 11. Performance of eXtreme Gradient Boosting Classifier on validation data: (A) Confusion matrix and (B) classification report.

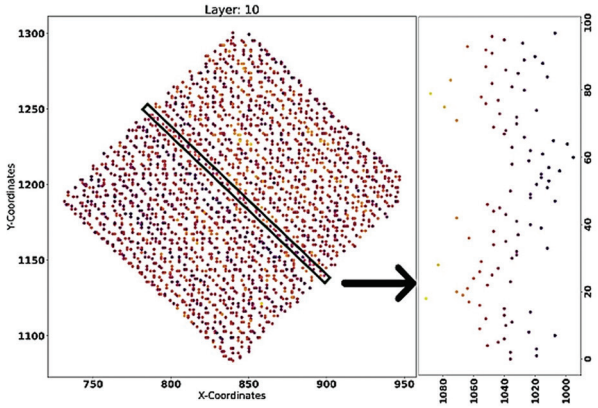


Figure 12. Pyrometer data from a layer of an additive-manufactured block with natural pores, and a single raster scan.

at 81.92 bit/mm across a 400 × 400 mm area. As shown in Figure 12, each layer's pyrometer data is divided into raster scans, and the data around visible pores is truncated to reduce noise and emphasize the underlying patterns. Raster scans passing through visible pores are labeled as porous, while others are classified as non-porous. The connection between temperature fluctuations and porosity is well established in the literature,[57,58] supporting the use of pyrometer data for this predictive modeling. The prediction of porosity in AM is strongly linked to the underlying physical mechanisms of melt pool behavior, as captured through *in situ* pyrometer data. Temperature fluctuations within the melt pool play a crucial role in porosity formation, with rapid heating and cooling cycles often resulting in the entrapment of gas or incomplete fusion of metal powders. These temperature changes can cause melt pool instability, leading to defects like pores due to insufficient material flow and uneven solidification. Studies have shown that higher porosity levels tend to occur when melt pool temperatures are inconsistent, as rapid solidification can trap gas within the material.[57,68] By monitoring these temperature fluctuations in real-time, ML models such as XGBoost can predict porosity and enable proactive adjustments to process parameters, mitigating defects and improving part quality. This approach allows for effective control of porosity, addressing a critical issue in the mechanical properties and structural integrity of AM parts.

4. Future works

To address the gaps identified in this study and further improve the predictive performance of ML models in AM, several targeted directions are proposed. First, increasing the dataset size is a priority. We plan to expand the dataset by at least 100%, focusing particularly on rare occurrences such as high porosity. Collecting more data in these cases is crucial for enhancing model robustness, especially for models that struggle with imbalanced datasets.

In terms of model performance evaluation, future works will also include a detailed analysis of the impact of dataset size and class distribution on model accuracy and RMSE. Given the potentially significant influence of imbalanced data on performance, this analysis will help better understand the scalability and generalizability of the models under different conditions. Data augmentation, which may include techniques such as generative adversarial networks or data synthesis algorithms to generate realistic data, can be explored as a viable alternative to conducting additional experiments. This approach is particularly beneficial given the expensive and time-consuming nature of processing NiTi via AM, offering a cost-effective way to both balance the data and increase dataset size without the need for extensive physical experiments. Additionally, emerging state-of-the-art algorithms such as DrCIF and MINIROCKET, which can directly work with raw time-series data, will be explored for their potential in porosity prediction. These models could offer improvements in both accuracy and speed by reducing the need for feature engineering. We aim to benchmark these algorithms against traditional models to quantify improvements in RMSE, with a target of reducing RMSE by at least 10% for high-porosity cases.

Moreover, while directly comparing the effects of hyperparameters across ten diverse models is challenging due to the variation in hyperparameters (e.g., SVM uses parameters such as C, gamma, and kernel type, while RF requires tuning of the number of trees, minimum samples per split, and splitting criterion), a future study will focus on recommending models based on the ease of hyperparameter tuning. Some models, such as LR, require no hyperparameter tuning, while others, such as neural networks, involve complex hyperparameter settings. Such study will involve identifying the key hyperparameters for each model (*e.g.*, focusing on 3 – 4 most significant hyperparameters) and using techniques such as factor analysis and principal component analysis to evaluate how many hyperparameters must be adjusted from default settings to achieve a specific performance threshold (*e.g.*, 90% accuracy). We will also factor in model training time to provide a comprehensive comparison. This will enable more informed decision-making when selecting models based on both performance and ease of tuning, filling a current gap in the literature. Herein, automated feature selection was employed, in the future, an analysis of how certain physical features directly contribute to porosity prediction can be conducted, complementing the automated approach provided by TS-Fresh.

By following these quantitative goals and exploring new algorithms, we expect to significantly improve the performance of the ML models used for porosity prediction in AM. These directions not only address the reviewer's concerns about the vagueness of the original future work section but also provide clear, measurable objectives aimed at improving both accuracy and RMSE while considering practical aspects such as hyperparameter tuning and computational efficiency.

5. Conclusion

This study addresses a knowledge gap by developing and implementing an *in situ* method for porosity prediction in additively manufactured parts. We approached the problem in two steps: first, classifying each layer as either low or high porosity, and second, predicting the exact porosity level using two separate regression models. Herein, the ability of 10 ML models to predict the porosity of NiTi additively manufactured parts was compared in terms of RMSE, accuracy, and speed. The dataset was collected from the Aconite MINI AM machine which consisted of pyrometer data. The porosity of additively manufactured parts was predicted using our experimental findings, highlighting several important conclusions and suggesting future directions for research. We found that pyrometer data effectively captures relevant information about the underlying AM process, which can be valuable for analysis, though it requires careful preprocessing to reduce noise and clean the data. By using an appropriate ML model, such as an XT regressor, we achieved strong predictive performance for these tasks. Specifically, classifying a layer as low (<1%) or high (≥1%) porosity proved to be a relatively straightforward classification task, whereas predicting the exact porosity percentage was more challenging.

In our experiments, the XT regressor achieved an RMSE of 0.0367 on the "low" porosity dataset (479 observations) but a slightly higher RMSE of 0.108 on the "high" porosity dataset (107 observations), which had fewer data points. When we undersampled the "low" dataset to match the size of the "high" dataset, the model's performance dropped, resulting in an RMSE of 0.165. These findings underscore the importance of having large and balanced datasets, especially for rare occurrences such as high porosity, to achieve strong performance in regression tasks.

For additional assessment of the ML algorithms' performance in AM, more data on rare occurrences (*i.e.*, high porosity) can be explored in the future. Further analyses on the effects of the dataset size and class distribution on the performance of the ML models should also be examined.

Acknowledgements

None.

Funding

This publication reflects research supported by a research grant from Science Foundation Ireland (SFI) under grant number 16/RC/3872, 21/RC/10295_P2, and co-funded by the European Regional Development Fund.

Conflict of interest

Dermot Brabazon is an Editorial Board Member of this journal and Guest Editor of this special issue, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: Vivek Mahato, Annalina Caputo, Dermot Brabazon

Formal analysis: Vivek Mahato

Investigation: Vivek Mahato, Annalina Caputo

Methodology: Vivek Mahato, Annalina Caputo, Dermot Brabazon

Writing–original draft: Vivek Mahato

Writing–review & editing: Suman Chatterjee, Anesu Nyabadza, Annalina Caputo, Dermot Brabazon

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data is available from the corresponding author upon reasonable request.

References

1. Mahato V, Obeidi MA, Brabazon D, Cunningham P. Detecting voids in 3D printing using melt pool time series data. *J Intell Manuf.* 2022;33:842-852.

   doi: 10.1007/s10845-020-01694-8

2. Herzog T, Brandt M, Trinchi A, Sola A, Molotnikov A. Process monitoring and machine learning for defect detection in laser-based metal additive manufacturing. *J Intell Manuf.* 2024;35(4):1407-1437.

   doi: 10.1007/s10845-023-02119-y

3. Gradl P, Tinker DC, Park A, *et al*. Robust metal additive manufacturing process selection and development for aerospace components. *J Mater Eng Perform*. 2022;31(8):6013-6044.

doi: 10.1007/s11665-022-06850-0

4. Mahale RS, Shamanth V, Hemanth K, *et al*. Processes and applications of metal additive manufacturing. *Mater Today Proc*. 2022;54:228-233.

doi: 10.1016/j.matpr.2021.08.298

5. Hehr A, Norfolk M, Kominsky D, Boulanger A, Davis M, Boulware P. Smart build-plate for metal additive manufacturing processes. *Sensors* (*Basel*). 2020;20(2):360.

doi: 10.3390/s20020360

6. Huang WB, Zhang LW, Li WL, *et al*. Various types and applications of additive manufacturing. *DEStech Trans Comput Sci Eng*. 2019:377-381.

doi: 10.12783/dtcse/ammso2019/30160

7. Papy K, Jean-Marc S, Alexey S, Andras B. Additive manufacturing feasibility of WC-17Co cermet parts by laser powder bed fusion. *Procedia CIRP*. 2022;111:153-157.

doi: 10.1016/j.procir.2022.08.049

8. Dejene ND, Lemu HG. Current status and challenges of powder bed fusion-based metal additive manufacturing: Literature review. *Metals*. 2023;13(2):424.

doi: 10.3390/met13020424

9. Hussain SZ, Kausar Z, Koreshi ZU, Shah MF, Abdullah A, Farooq MU. Linear active disturbance rejection control for a laser powder bed fusion additive manufacturing process. *Electronics*. 2023;12(2):471.

doi: 10.3390/electronics12020471

10. Pan T, Li L, Zhang X, *et al*. Investigation of significant factors on deformation with powder bed fusion system. *Proc Inst Mech Eng B J Eng Manuf*. 2021;235(5):902-911.

doi: 10.1177/0954405420970088

11. Ford S, Despeisse M. Additive manufacturing and sustainability: An exploratory study of the advantages and challenges. *J Clean Prod*. 2016;137:1573-1587.

doi: 10.1016/j.jclepro.2016.04.150

12. Gradl PR, Tinker DC, Ivester J, Skinner SW, Teasley T, Bili JL. Geometric feature reproducibility for laser powder bed fusion (L-PBF) additive manufacturing with Inconel 718. *Addit Manuf*. 2021;47:102305.

doi: 10.1016/j.addma.2021.102305

13. Kayacan MY, Özsoy K, Duman B, Yilmaz N, Kayacan MC. A study on elimination of failures resulting from layering and internal stresses in Powder Bed Fusion (PBF) additive manufacturing. *Mater Manuf Processes*. 2019;34(13):1467-1475.

doi: 10.1080/10426914.2019.1655151

14. Zieliński TG, Opiela KC, Pawłowski P, *et al*. Reproducibility of sound-absorbing periodic porous materials using additive manufacturing technologies: Round robin study. *Addit Manuf*. 2020;36:101564.

doi: 10.1016/j.addma.2020.101564

15. Faragasso A, Bonsignorio F. Reproducibility challenges in robotic surgery. *Front Robot AI*. 2023;10:1127972.

doi: 10.3389/frobt.2023.1127972

16. Li X, Zhang M, Zhou M, *et al*. Qualify assessment for extrusion-based additive manufacturing with 3D scan and machine learning. *J Manuf Processes*. 2023;90:274-285.

doi: 10.1016/j.jmapro.2023.01.025

17. Gu Z, Mani Krishna KV, Parsazadeh M, *et al*. Deep learning-based melt pool and porosity detection in components fabricated by laser powder bed fusion. *Prog Addit Manuf*. 2024:1-18.

doi: 10.1007/s40964-024-00603-2

18. Atwya M, Panoutsos G. *In-situ* porosity prediction in metal powder bed fusion additive manufacturing using spectral emissions: A prior-guided machine learning approach. *J Intell Manuf*. 2024;35(6):2719-2742.

doi: 10.1007/s10845-023-02170-9

19. Satterlee N, Torresani E, Olevsky E, Kang JS. Comparison of machine learning methods for automatic classification of porosities in powder-based additive manufactured metal parts. *Int J Adv Manuf Technol*. 2022;120(9):6761-6776.

doi: 10.1007/s00170-022-09141-z

20. Wang W, Ning J, Liang SY. Analytical prediction of keyhole porosity in laser powder bed fusion. *Int J Adv Manuf Technol*. 2022;119(11):6995-7002.

doi: 10.1007/s00170-021-08276-9

21. Liu R, Liu S, Zhang X. A physics-informed machine learning model for porosity analysis in laser powder bed fusion additive manufacturing. *Int J Adv Manuf Technol*. 2021;113(7):1943-1958.

doi: 10.1007/s00170-021-06640-3

22. Ero O, Taherkhani K, Hemmati Y, Toyserkani E. An integrated fuzzy logic and machine learning platform for porosity detection using optical tomography imaging during laser powder bed fusion. *Int J Extrem Manuf*. 2024;6(6):065601.

doi: 10.1088/2631-7990/ad65cd

23. Mohammed AS, Almutahhar M, Sattar K, Alhajeri A, Nazir A, Ali U. Deep learning based porosity prediction for additively manufactured laser powder-bed fusion parts. *J Mater Res Technol*. 2023;27:7330-7335.

doi: 10.1016/j.jmrt.2023.11.130

24. Estalaki SM, Lough CS, Landers RG, Kinzel EC, Luo T. Predicting defects in laser powder bed fusion using in-situ thermal imaging data and machine learning. *Addit Manuf.* 2022;58:103008.

doi: 10.1016/j.addma.2022.103008

25. Gordon JV, Narra SP, Cunningham RW, *et al*. Defect structure process maps for laser powder bed fusion additive manufacturing. *Addit Manuf.* 2020;36:101552.

doi: 10.1016/j.addma.2020.101552

26. Du Plessis A. Effects of process parameters on porosity in laser powder bed fusion revealed by X-ray tomography. *Addit Manuf*. 2019;30:100871.

doi: 10.1016/j.addma.2019.100871

27. Kan WH, Nadot Y, Foley M, Ridosz L, Proust G, Cairney JM. Factors that affect the properties of additively-manufactured AlSi10Mg: Porosity versus microstructure. *Addit Manuf.* 2019;29:100805.

doi: 10.1016/j.addma.2019.100805

28. Hojjatzadeh SMH, Parab ND, Guo Q, *et al*. Direct observation of pore formation mechanisms during LPBF additive manufacturing process and high energy density laser welding. *Int J Mach Tools Manuf*. 2020;153:103555.

doi: 10.1016/j.ijmachtools.2020.103555

29. Vastola G, Pei QX, Zhang YW. Predictive model for porosity in powder-bed fusion additive manufacturing at high beam energy regime. *Addit Manuf.* 2018;22:817-822.

doi: 10.1016/j.addma.2018.05.042

30. Sun S, Brandt M, Easton MJL. Powder bed fusion processes: An overview. *Laser Additive Manufacturing*. Sawston, UK: Woodhead Publishing; 2017. p. 55-77.

31. Mahato V, Obeidi MA, Brabazon D, Cunningham P. An evaluation of classification methods for 3d printing time-series data. *IFAC-PapersOnLine*. 2020;53(2):8211-8216.

doi: 10.1016/j.ifacol.2020.12.1992

32. Grasso M, Colosimo BM. Process defects and *in situ* monitoring methods in metal powder bed fusion: A review. *Meas Sci Technol.* 2017;28(4):044005.

doi: 10.1088/1361-6501/aa5c4f

33. Li Y, Gu D. Thermal behavior during selective laser melting of commercially pure titanium powder: Numerical simulation and experimental study. *Addit Manuf.* 2014;1:99-109.

doi: 10.1016/j.addma.2014.09.001

34. Yadav P, Rigo O, Arvieu C, Le Guen E, Lacoste E. *In situ* monitoring systems of the SLM process: On the need to develop machine learning models for data processing. *Crystals*. 2020;10(6):524.

doi: 10.3390/cryst10060524

35. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321-357.

doi: 10.1613/jair.953

36. Wilson DL. Asymptotic properties of nearest neighbor rules using edited data. *IEEE Trans Syst Man Cybern*. 1972;2(3):408-421.

doi: 10.1109/TSMC.1972.4309137

37. Torgo L, Ribeiro RP, Pfahringer B, Branco P. Smote for regression. In: *Portuguese Conference on Artificial Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 378-389.

38. Christ M, Braun N, Neuffer J, Kempa-Liehr AW. Time series feature extraction on basis of scalable hypothesis tests (tsfresh-a python package). *Neurocomputing*. 2018;307:72-77.

doi: 10.1016/j.neucom.2018.03.067

39. Peng CYJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. *J Educ Res*. 2002;96(1):3-14.

doi: 10.1080/00220670209598786

40. Mahato V, Cunningham P. A Case-study on the Impact of Dynamic Time Warping in Time Series Regression. In: *3rd ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*. 2018.

doi: 10.48550/arXiv.2010.05270

41. Mahato V, Obeidi MA, Brabazon D, Cunningham P. An evaluation of classification methods for 3d printing time-series data. *IFAC-PapersOnLine*. 2020;53(2):8211-8216.

doi: 10.1016/j.ifacol.2020.12.1992

42. Badiane M, Cunningham P. An empirical evaluation of kernels for time series. *Artif Intell Rev*. 2022;55(3):1803-1820.

doi: 10.1007/s10462-021-10050-y

43. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern*. 1991;21(3):660-674.

doi: 10.1109/21.97458

44. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.

doi: 10.1023/A:1010950718922

45. Biau G, Scornet E. A random forest guided tour. *Test*. 2016;25:197-227.

doi: 10.1007/s11749-016-0481-7

46. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63:3-42.

doi: 10.1007/s10994-006-6226-1

47. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci*. 1997;55(1):119-139.

doi: 10.1007/3-540-59119-2_166

48. Wang L, Sugiyama M, Yang C, Zhou ZH, Feng J. On the Margin Explanation of Boosting Algorithms. In: *21st Annual Conference on Learning Theory - COLT 2008*. Helsinki, Finland; 2008. p. 479-490.

49. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat*. 2001;1:1189-1232.

doi: 10.1214/aos/1013203451

50. Chen T, Guestrin C. Xgboost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*. 2016. p. 785-794.

doi: 10.1145/2939672.2939785

51. Ogunleye A, Wang QG. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans Comput Biol Bioinform*. 2019;17(6):2131-2140.

doi: 10.1109/TCBB.2019.2911071

52. Gumus M, Kiran MS. Crude Oil Price Forecasting Using XGBoost. In: *2017 International Conference on Computer Science and Engineering* (*UBMK*). IEEE; 2017. p. 1100-1103.

53. Deng J, Xu Y, Zuo Z, Hou Z, Chen S. Bead geometry prediction for multi-layer and multi-bead wire and arc additive manufacturing based. In: *Transactions on Intelligent Welding Manufacturing*. Vol. 2. Singapore: Springer; 2019. p. 125-135.

54. Monu MC, Ekoi EJ, Hughes C, Kumar S, Brabazon D. Resultant physical properties of as-built nitinol processed at specific volumetric energy densities and correlation with *in-situ* melt pool temperatures. *J Mater Res Technol*. 2022;21:2757-2777.

doi: 10.1016/j.jmrt.2022.10.073

55. McCann R, Obeidi MA, Hughes C, *et al*. *In-situ* sensing, process monitoring and machine control in laser powder bed fusion: A review. *Addit Manuf*. 2021;45:102058.

doi: 10.1016/j.addma.2021.102058

56. Blum AL, Langley P. Selection of relevant features and examples in machine learning. *Artif Intell*. 1997;97(1-2):245-271.

doi: 10.1016/S0004-3702(97)00063-5

57. Ho S, Zhang W, Young W, *et al*. DLAM: Deep learning based real-time porosity prediction for additive manufacturing using thermal images of the melt pool. *IEEE Access*. 2021;9:115100-115114.

doi: 10.1109/ACCESS.2021.3105362

58. Dilip JJS, Zhang S, Teng C, *et al*. Influence of processing parameters on the evolution of melt pool, porosity, and microstructures in Ti-6Al-4V alloy parts fabricated by selective laser melting. *Prog Addit Manuf*. 2017;2:157-167.

doi: 10.1007/s40964-017-0030-2

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

REVIEW ARTICLE

Role of phase change materials and digital twin technology in thermal energy storage system: A review

Mohammad Waseem[1]*, Mumtaz Ahmad[1], G. Sree Lakshmi[2†], Areti M.S.V. Sushma[3], Sanjay Paul[4†], and Mohammad Afazal[5]

[1]Department of Mechanical Engineering Section, Faculty of Engineering and Technology, Jamia Millia Islamia, University Polytechnic, New Delhi, Delhi, India

[2]Department of Electrical and Electronics Engineering, CVR College of Engineering, Hyderabad, Telangana, India

[3]Department of Electrical and Electronics Engineering, SRKR Engineering College, West Godavari, Andhra Pradesh, India

[4]Department of Mechanical Engineering, Faculty of Mechanical Engineering, Rajshahi University of Engineering and Technology, Rajshahi, Bangladesh

[5]Centre for Biomedical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India

(This article belongs to the *Special Issue: Cooling, Heating, and Thermal Energy Storage Systems: Harnessing the Power of Digital Twins*)

Abstract

The exponential growth in energy consumption and demand, along with the depletion of natural resources, is exerting a catastrophic impact on global ecosystems. Recent advances in research and development have focused on the distribution of renewable energy sources and the reduction of traditional energy usage as strategies to address pressing environmental concerns, such as climate change and global warming. Moreover, there is an urgent need for appropriate technologies that can enhance the thermal performance of buildings, given the rapid increase in global cooling and heating demands. This study examines the role of phase change materials (PCMs) and digital twin (DT) technology in thermal energy storage (TES), drawing on an analysis of 89 research articles sourced from multiple databases and references. The findings demonstrate that TES systems optimized through meticulous selection of PCMs can effectively meet thermal comfort requirements. Integrating DT technology with building systems allows for the analysis of cooling effects and optimization of energy demand through DT models of smart buildings. The present study provides a comprehensive overview of the different PCMs used in cooling applications and explores the implementation of DT technologies within building systems. In addition, practical applications of DT technologies for TES systems are presented, providing insights into their potential for enhancing energy efficiency in building systems.

Keywords: Phase change material; Thermal energy storage; Digital twin; Intelligent building

†These authors contributed equally to this work.

***Corresponding author:**
Mohammad Waseem
(waseem159088@st.jmi.ac.in)

Citation: Waseem M, Ahmad M, Lakshmi GS, Sushma AMSV, Paul S, Afazal M. Role of phase change materials and digital twin technology in thermal energy storage system: A review. *Int J AI Mater Design.* 2024;1(3):50-65.
doi: 10.36922/ijamd.4696

Received: August 29, 2024

Accepted: October 28, 2024

Published Online: November 27, 2024

1. Introduction

Rapid population and economic growth have led to a dramatic rise in global energy consumption, which in turn has had devastating effects on the planet's ecosystems.[1-5] The

building industry is a major contributor to carbon dioxide (CO_2) emissions and the largest consumer of energy worldwide, accounting for more than 33.33% of total energy consumption.[2,6] In addition, heating, ventilation, and air conditioning (HVAC) systems are responsible for 50 – 60% of a building's energy usage.[7,8] Governments and communities aim to reduce energy consumption without compromising thermal comfort in buildings, regardless of weather conditions.[9,10] One significant challenge in this area is the development of new technologies to promote energy conservation and efficiency. Thermal energy storage (TES) systems, when installed in buildings, can reduce peak energy demand and enhance heating and cooling efficiency.[11] The building industry has shown considerable interest in phase change materials (PCMs) due to their high energy storage density and the ability to store thermal energy through a permanent phase transition.[12] Applications of PCMs, a comprehensive TES technology, have gained significant attention for improving cooling efficiency while reducing overall building energy consumption, making them a popular topic among engineers and architects.[13]

Heat storage can be achieved through sensible, latent, or thermochemical means, which are the three primary methods used to store thermal energy.[14] Among the various energy storage technologies, TES stands out as highly environmentally beneficial.[15] PCM-based TES systems show significant potential to absorb or release substantial energy during phase transitions under nearly isothermal conditions, making them a promising avenue for research.[16] There are primarily three types of PCMs, classified according to their charging and discharging processes: organic, inorganic, and eutectic. Salt hydrates are an example of an inorganic PCM, while organic PCMs include substances, such as paraffin waxes, fatty acids, and fatty acid esters. Eutectic PCMs, which are polymeric compounds, include eutectic salts and solutions.[1,17] Another way to categorize PCMs is based on the type of phase transition they undergo: solid-liquid, solid-solid, and solid-gas.[18] Among these, solid-liquid PCMs are generally the best choice for energy storage due to their low cost, high latent heat, and other desirable properties.[19,20] However, despite these advantages, limitations such as super-cooling and low latent heat can restrict their widespread use.[21]

Several reviews have focused on PCMs and their potential applications for cooling in buildings. For instance, Oropeza-Perez *et al.*[22] reviewed active and passive cooling in homes, while Monghasemi *et al.*[23] explored research on integrating solar chimneys with HVAC systems in buildings. Khan *et al.*[24] provided a review of PCM applications in solar-powered refrigeration systems. Additional PCM research topics include climate control and heating (PCM-TES),[25] investigations under specific environmental conditions,[26,27] measurement methodologies, encapsulation techniques, and nano-PCMs.[28-30]

A critical issue in TES and heat insulation within buildings remains a major challenge.[31] For example, Jiang *et al.*[32] developed a novel self-reinforced composite PCM through hot-melt extrusion, designed for heat storage.[32] Wang *et al.*[33] constructed a unique packed-bed structure for storing heat, capable of withstanding medium to high temperatures.[33] In the context of smart buildings, Antoniadou-Plytaria *et al.*[34] proposed an energy management system for microgrids based on batteries. Their simulations demonstrated that the framework could reliably predict operational costs and enhance battery performance.[34] Ren *et al.*[35] compared two solar-powered systems: System P, which used solar thermal collectors (to transform sunlight into heat) and photovoltaic panels (to convert sunlight into electricity), and System Q, which employed photovoltaic thermal collectors to generate both thermal and electrical energy from solar power.[35] In addition, Lin *et al.*[36] introduced a thermoelectrically-powered wireless sensor network for environmental monitoring in building envelopes. Their design, which features self-powered capabilities, thermally optimized components, and milliwatt power control, could offer significant benefits for low-power, inexpensive environmental monitoring.[36]

Yu *et al.*[37] explored a novel approach to harnessing solar energy by integrating two form-stable PCMs within a humid environment containing dissolved CO_2. The distinct electrical resistivities and phase transition temperatures of these PCMs enable energy harvesting independent of conventional thermoelectric devices. The process utilizes the Seebeck effect, where the temperature differential between the PCMs generates electricity, offering an innovative method for sustainable energy production using PCMs and environmental CO_2.

1.1. Role of digital twin (DT) technology in PCM-TES systems

Lv *et al.*[38] discovered that DT technology significantly improves the efficiency and performance of PCM-based TES systems. By creating a virtual replica of the physical system, DT enables real-time monitoring, simulation, and optimization of PCM behavior during energy storage and release. The virtual models are continuously updated to accurately reflect changes in the physical system, facilitating precise predictions of temperature distributions, phase transitions, and energy transfer dynamics. In PCM-TES systems, DT technology plays a vital role in enhancing system operations by providing accurate forecasts of energy demand and storage capacity. For example, DT can simulate various environmental conditions, such as

fluctuating temperatures and changing energy demands, allowing for proactive adjustments that enhance energy efficiency. Furthermore, DT improves control strategies for heating and cooling applications, particularly within building energy management systems. By integrating real-time data, DT systems enable predictive maintenance, identify system inefficiencies, and improve overall operational reliability.

1.2. Contributions

The main contributions of this article are outlined below:

- A total of 89 research articles, including case studies, review papers, experimental analyses, and letters, were reviewed in the current study, sourced from databases such as Science Direct, IEEE, MDPI, Wiley, and Springer, among others.
- The present study describes TES as a versatile form of renewable energy, utilizing materials that absorb heat when heated and release it when cooled through sensible heat, latent heat, and thermochemical energy.
- A review of different PCMs is provided, focusing on their classifications and selection criteria for regulating temperatures within a narrow range by releasing latent heat during phase transitions.
- Practical applications of various PCMs for cooling in TES buildings are identified and discussed.
- The role of DT technology in smart buildings is explored, along with real-world examples of how DT is integrated into thermal storage systems to enhance the intelligence and functionality of buildings.

1.3. Scope and organization of the present article

The reviewed studies highlight the potential of PCM-TES systems as an innovative building technology capable of significantly enhancing thermal efficiency and sustainability in the future. It is evident that further research is needed to identify the optimal PCMs for different climate zones. The aim of this article is to provide a concise overview of the role of PCMs and DT technology, their properties, and integration techniques, as well as to examine their cooling applications in buildings through appropriate classification. The article is structured as follows: Section 1 offers a brief description of TES systems, PCMs, and DT technology; Section 2 outlines the methods of TES, including sensible heat storage (SHS), latent heat storage (LHS), and thermochemical energy storage systems; Section 3 describes the classification, selection criteria, and practical applications of PCMs in buildings; Section 4 summarizes the integration of DT technology in smart buildings, along with real-world applications. A flow diagram summarizing the structure of the current study is presented in Figure 1.

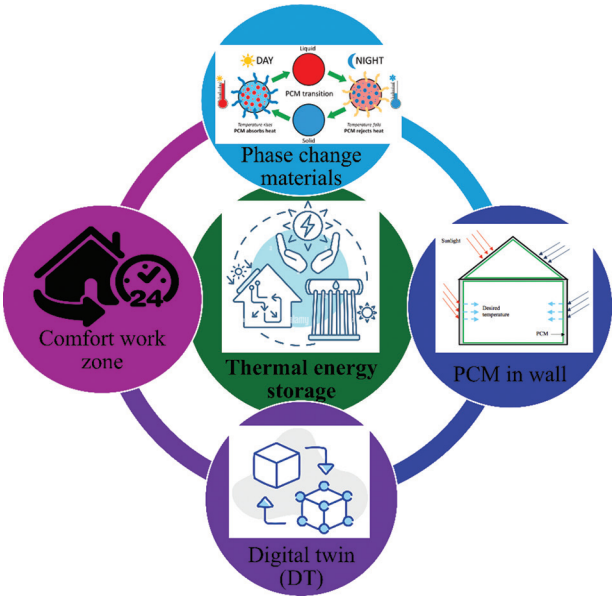


Figure 1. Structural flow diagram of the review
Abbreviation: PCM: Phase change material.

2. TES

Energy storage systems work by transforming energy into a form that can be stored and made available when needed. One type of renewable energy storage system is the TES system, which uses materials that absorb heat when heated and release it when cooled. The three primary forms of TES are depicted in Figure 2.

TES is widely used in buildings and modern energy cycles, where it involves storing excess energy – typically surplus energy from renewable sources or waste heat – to be used later for heating, cooling, or power generation purposes.[39] To store thermal energy, TES systems use heat- or cold-storage media. The stored energy can then be utilized to generate electricity for temperature regulation systems. Heat can be stored through three main mechanisms: sensible heat, latent heat (heat released or absorbed during phase transitions), and thermochemical energy (heat released or absorbed during chemical reactions).[40]

Lin *et al.*[41] emphasized that PCMs are essential for improving the efficiency of TES systems due to their ability to absorb and release large amounts of latent heat during phase transitions, typically between the solid and liquid states. In TES systems, PCMs are utilized to store thermal energy during charging cycles, where heat is absorbed, and the material transitions to a liquid state, and to release this stored energy during discharging cycles, when the material returns to its solid state. This process helps stabilize temperature fluctuations and promotes efficient energy storage over extended periods, making PCMs particularly

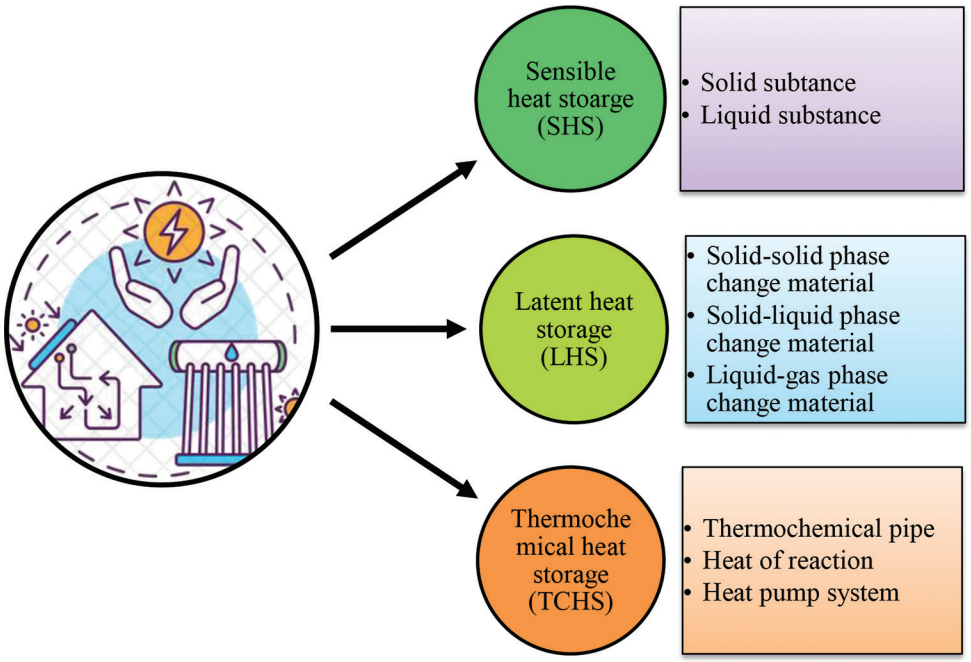


Figure 2. Three distinct classifications of thermal energy storage systems

suitable for applications in buildings, renewable energy systems, and industrial heat recovery.

Bland *et al.*[42] found that the primary advantage of PCMs in TES systems is their high energy density, which allows for substantial energy storage in a compact volume. Furthermore, PCMs offer an almost iso TES process, which is particularly beneficial for maintaining stable temperatures in applications such as HVAC systems in buildings or solar thermal power plants. The selection of efficient materials for PCM is critical to enhancing the performance of TES systems, as factors such as the material's melting point, thermal conductivity, and energy storage capacity directly influence the system's effectiveness. In renewable energy applications, such as solar thermal plants, PCMs are vital in addressing the intermittent nature of energy supply, as noted by Masood *et al.*[43] They store excess heat during peak production times and release it when energy generation is low. The role of PCMs in enhancing the flexibility and reliability of TES systems is crucial for advancing sustainable energy solutions. TES systems can be classified into three categories, as discussed below.

2.1. SHS

In SHS systems, energy storage is achieved by utilizing the heat capacity that a material gains as its temperature increases. The energy storage capacity of an SHS system is determined by three factors: the specific heat capacity of the material, the quantity of the material, and the temperature change gradient. The temperature of the storage material can be increased by radiation, convection, or conduction.[44] This method involves freezing or heating a solid or liquid storage medium to store heat.[45] The temperature of an energy storage substance can be altered using the SHS approach, which can involve both solid and liquid storage materials.[46] A commonly used medium for SHS is water due to its low cost. The SHS process is based on the heat capacity of the storage material, which is related to the temperature differential during the charge and discharge phases.[45]

The sensible heat of a system can be determined by its mass (m), specific heat (C), and temperature differential (ΔT), all of which are key components of the thermodynamic system affecting temperature. Mathematically, the sensible heat ($Q_{sensible}$) of a thermodynamic system is given by Equation I:

$$Q_{sensible}=mC\Delta T \quad \text{(I)}$$

Water is the most common medium used for this type of heat storage and release. However, various other solid materials, such as sandy soil, pebbles, rock, stone, brick, aluminum, and cast iron, can also be used for SHS.

2.2. LHS

LHS systems are based on a material's ability to absorb or release heat during a phase transition – such as from solid to liquid, liquid to gas, or vice-versa. The appeal of this quasi-static process lies in the high energy density stored per unit mass. The LHS capacity of the system, when using a PCM, is given by the following equation:

$$Q = \int_{T_1}^{T_{pc}} C_{solid} + \Delta h + \int_{T_{pc}}^{T_2} C_{liquid} \quad \text{(II)}$$

where T_{pc} is the temperature of the phase change, Δh is the latent heat of fusion, and C_{solid} and C_{liquid} are the specific heat capacities of the PCM in the solid and liquid phases, respectively.

A PCM is utilized in LHS systems, with the temperature at which the phase transition occurs referred to as the phase-change temperature.[45,47] In addition to offering high energy density, LHS systems also face challenges related to low thermal conductivity and material degradation over time. PCMs are commonly employed in LHS systems because of their ability to absorb and release large amounts of energy as they change phase, which alters their physical state.[48] In the LHS process, heat is stored as the PCM undergoes a phase transition at a nearly constant temperature, with the amount of heat being proportional to the material's latent heat of fusion. Due to their high energy capacity and ability to maintain a stable temperature throughout the phase transition, PCMs are widely considered ideal materials for LHS systems.[49]

2.3. Thermochemical energy storage

Thermochemical energy storage systems absorb and release energy during reversible reactions, which involve the breaking and reforming of chemical bonds. Chemical reactions that release heat form the basis of thermochemical energy storage systems. The heat involved in the calcium hydroxide cycle is given by Equation III as follows:

$$CaO + H_2O \rightleftharpoons Ca(OH)_2 + \text{Heat} \quad \text{(III)}$$

3. PCMs

PCMs are substances that can regulate temperatures within a narrow range by releasing latent heat during phase transitions.[50] As the surrounding temperature increases to the point where the PCM melts, an endothermic process begins, breaking the chemical bonds and transforming the PCM from a solid to a liquid while absorbing energy. This process is referred to as the PCM charging phenomenon. The next step involves the temperature decreasing until the PCM solidifies, at which point the bonds regenerate, and the PCM returns to its solid state by releasing heat in an exothermic process. Therefore, PCMs are defined as thermal reservoirs.[51] When charged and discharged, the volume of PCMs changes by a negligible amount – approximately 10% of their original volume.[52] Figure 3 provides a visual representation of this process.

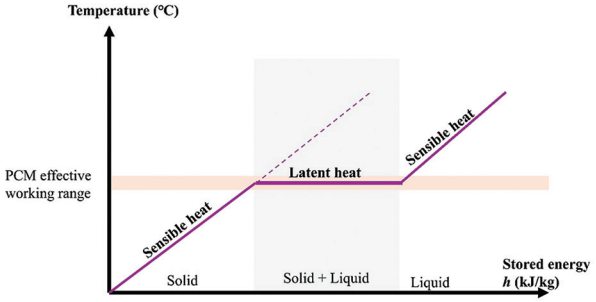


Figure 3. Sensible and latent heat behavior of phase change materials with respect to temperature
Abbreviation: PCM: Phase change material.

3.1. Classification of PCMs

PCMs can be primarily classified into three types based on their charging and discharging processes within the solid-liquid transition category: organic, inorganic, and eutectic PCMs.[50] Each of these categories contains additional sub-categories. Organic PCMs include both paraffin and non-paraffin materials. Non-paraffin materials are further subdivided into lipids, sugar esters, and carbohydrates. Inorganic PCMs, which undergo phase transitions without requiring super-cooling, retain their heat of fusion throughout cycling.[53] Depending on their composition, inorganic PCMs can be metals, molten salts, or salt-hydrates. Eutectic PCMs, in turn, can be synthesized by combining organic and inorganic components.[54] Figure 4 provides a concise overview of the different PCM classes.

Organic PCMs are primarily composed of carbon-hydrogen bonds. These materials are non-corrosive, self-nucleating, and typically have consistent melting points. While organic PCMs offer numerous advantages, they also have several limitations that can reduce their effectiveness.[55] They do not experience phase separation or degradation, meaning they can undergo continuous melting and freezing cycles. Key advantages of organic PCMs include their chemical stability, lack of super-cooling, non-corrosiveness, and recyclability. However, they are combustible and generally have low conductivity.[56] One widely used organic heat exchange material is paraffin wax,[57] which is composed primarily of alkanes. The formula of paraffin is C_nH_{2n+2}, where "n" indicates the number of carbon atoms in the alkane chain.

In contrast, inorganic PCMs are non-combustible, have a higher LHS capacity, and exhibit more distinct phase transitions. However, their main drawbacks include inherent corrosiveness, phase segregation, and super-cooling.[58] Among organic PCMs, salt hydrates are particularly promising due to their higher LHS capacity

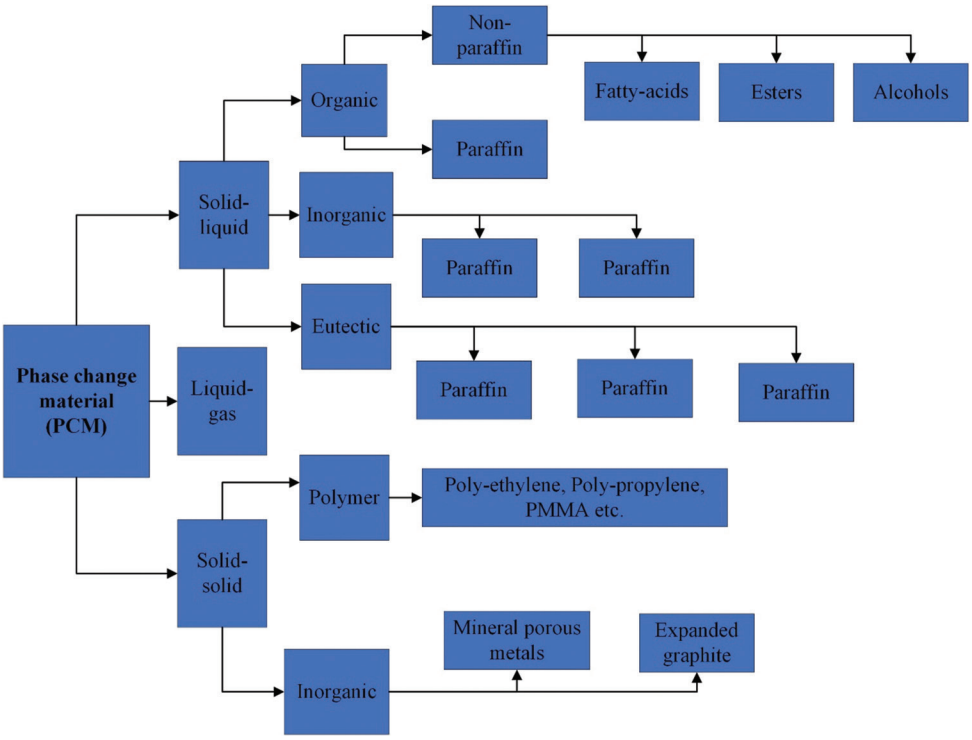


Figure 4. Different categories of phase change materials

– approximately twice that of organic compounds with a high fusion temperature.[59] Salt hydrates also have a thermal conductivity of about 0.5 W/m-K, are non-combustible, readily available, inexpensive, and have minimal environmental impact.[60] Most importantly, they can be recycled.

A newer type of PCM, called eutectic, consists of a combination of elements that form crystals as they melt and freeze. Eutectic alloys contain both organic and inorganic elements, with the most common being inorganic salt hydrates. These materials exhibit homogeneous melting and freezing properties, avoiding phase segregation.[61] The advantages and disadvantages of organic and inorganic PCMs are summarized in Table 1.

3.2. Selection criteria of PCM

Materials can only be used as PCMs if they meet specific requirements related to thermal, physical, chemical, and mechanical properties. Ecological and economic considerations also play a crucial role in the selection process for thermal storage applications. From a thermal perspective, the phase transition temperature of a PCM must fall within the target operating temperature range to ensure efficient heat absorption and release during operation.[63] Another essential property for effective TES is high thermal conductivity, which facilitates the efficient absorption and release of thermal energy. In addition, the PCM must have a high latent heat of fusion relative to its volume to achieve adequate energy storage density.[64] Finally, the phase transition must occur uniformly to ensure that the process is repeatable, stable, and durable over time.[15] The selection criteria and key properties of PCM materials are illustrated in Figure 5.

In terms of physical properties, the vapor pressure at operating temperatures must be limited, and the material should undergo minimal volume change during the phase transition. The material's density is also an important factor to consider.[64] Given that different temperatures are required for melting and solidifying, an insufficient nucleation rate can lead to super-cooling in the liquid state, which significantly alters the kinetics of the charging and discharging cycles. Therefore, both high nucleation rates and enhanced crystal formation are necessary for optimal performance.[65] In addition, the material must be able to withstand repeated freezing and melting cycles without deterioration. It must also be environmentally safe when used.[66] From a financial perspective, PCMs should be inexpensive, abundant, and readily available.

3.3. Practical applications of PCMs in TES for buildings

PCMs have a wide range of potential applications, including heating, cooling, and hybrid systems. The principle behind

Table 1. Pros and cons of organic and inorganic PCMs

PCM	Pros	Cons	References
Organic	• Wide temperature range for phase transitions. • Withstands heat without deterioration. • Does not react chemically. • Safe for use around metals. • No phase segregation during homogeneous melting. • High latent heat. • Strong nucleation ability. • Minimal volume variation. • Recyclable plastic.	• Combustible. • Low density. • Poor thermal conductivity.	62
Inorganic	• Excellent heat conduction. • Efficient use of enthalpy. • Minimal volume change during phase transition. • Non-flammable.	• Phase separation during incongruent melting. • Poor nucleation. • Prone to degradation. • Potentially hazardous. • Not all building materials are compatible. • Higher cost.	62

Abbreviation: PCM: Phase change material.

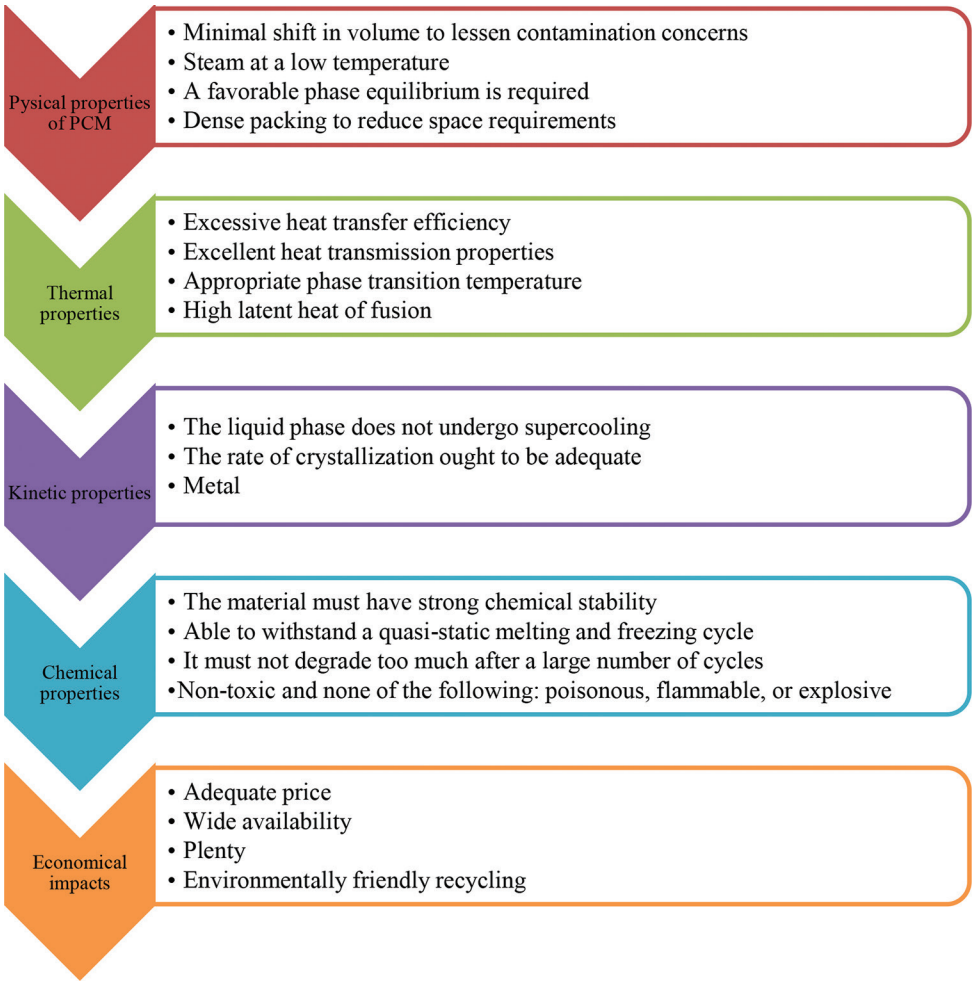


Figure 5. Selection criteria and properties of PCMs
Abbreviation: PCM: Phase change material.

PCM-based cooling applications is as follows: when the temperature outside the designated thermal zone exceeds the pre-determined comfort level, PCMs activate to absorb and store the excess heat. Table 2 provides examples of practical PCM cooling applications in buildings across different countries.[67-73]

An investigation was conducted by Ning *et al.*[67] in the coastal area of Yantai, China, focused on a bedroom enclosure equipped with PCM boards. The study aimed to understand how the thickness and transition temperature of the PCM affect thermal performance. It was demonstrated that raising the PCM melting point helped decrease the variance in return air temperature. In another study, Wang *et al.*[68] monitored the thermal performance of a 30 × 30 × 3 cm micro PCM honeycomb wallboard on a daily basis. Their findings showed that the thermal insulation of the micro PCM honeycomb wallboard was 4 h in case 1, 4.7 h in case 2, and 4.7 h in case 3. These results indicate that the thermal performance of the wallboard is influenced by the interior environment. In field experiments, Lee *et al.*[69] exposed two similar model exhibit spaces to varying meteorological conditions. They found that incorporating cellulose insulation into a paraffin-based PCM mixture did not negatively affect the PCM's latent heat. The results demonstrated a peak heat flow reduction of 26.6% when all four walls were together.

4. DT technology

DT technology, which creates a digital model of a real-world process, person, location, system, or device, enables intelligent buildings to establish a data bridge and facilitate communication between the physical and digital realms. Using highly realistic, interactive virtual models, DT technology can simulate the physical world, including the states and behaviors of its inhabitants.[74] It can even predict or recreate the actions of these inhabitants, either in advance or in real-time.[74]

In TES systems, the design and data processing procedure, with a focus on DT technology, follows a systematic methodology. This process incorporates real-time monitoring, simulations, and data-driven optimization to enhance the efficiency and reliability of TES. Figure 6 illustrates the step-by-step design process procedure, highlighting the role of DT in TES.

4.1. Integrating DT technology into smart buildings

To integrate DT technology into real-world building applications, the first step is to create a DT model of the building. By adding two additional dimensions – data and connection – to the traditional three-dimensional model, this study develops DT models with five dimensions.[75] Figure 7 illustrates the DT intelligent building model, derived from the DT system in five dimensions. In this

Table 2. Practical applications of PCMs in thermal energy storage systems for buildings

PCM category	Type of analysis	Applied country	Remarks	References
External bedroom envelope equipped with PCM	Numerical	China (for cooling)	PCM effectively reduces the impact of external thermal environment changes	Ning *et al.*[67]
Microencapsulated PCM based on paraffin	Experimental	Taiwan (for cooling)	Protection from heat in cases 1, 2, and 3: 4 h, 4.7 h, and 4.7 h, respectively	Wang *et al.*[68]
Filling wall voids with a mixture of paraffin PCM and cellulose insulation	Experimental	USA (for cooling)	Average cost savings: 3 cents/m^2. Peak heat flux reduction: 26.6% per hour	Lee *et al.*[69]
Shape-stabilized PCM wallboards (PCMW)	Numerical and experimental	China (for cooling)	PCMW enhances thermal comfort, reduces indoor temperature fluctuations, and increases energy efficiency	Yao *et al.*[70]
Composite wallboard made of gypsum with shape-stabilized nano-PCM	Numerical and experimental	USA (for cooling)	No decrease in heat gains when set points are 22°C and 23.3°C	Biswas *et al.*[71]
PCM-based bio-composite wall	Numerical	France (for cooling)	Heavy walls with PCM perform best in reducing stratification effects compared to lightweight walls	Kharbouch *et al.* [72]
Enhanced PCM container modified for use as a solar chimney	Experimental	Lab environment	Three solar heat fluxes were tested: 700, 600, and 500 W/m^2. The lowest output temperature was at 700 W/m^2, and the thermal efficiency was highest at 500 W/m^2	Li *et al.*[73]

Abbreviations: PCM: Phase change material; USA: United States of America.

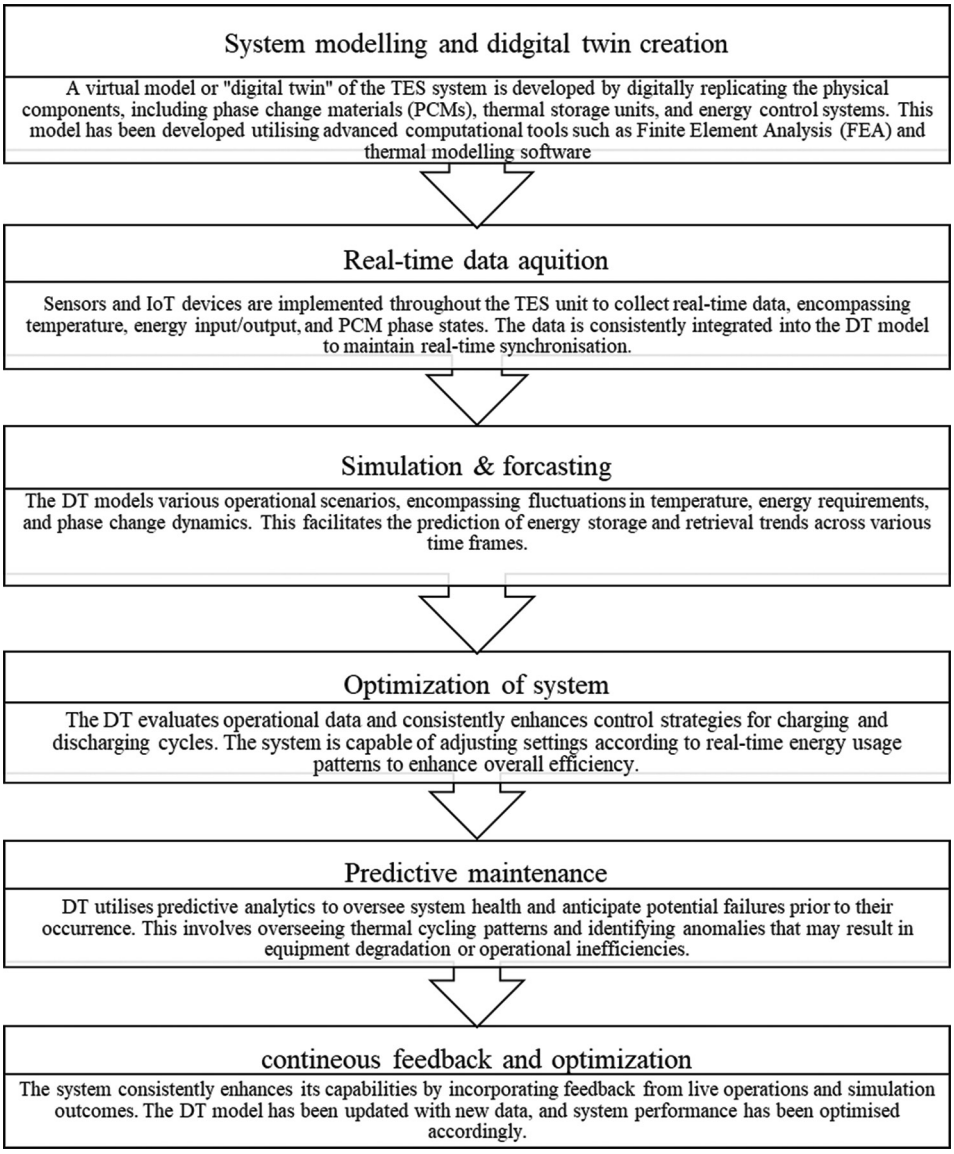


Figure 6. Design or data processing procedure of DT technology roles in the overall TES unit
Abbreviations: DT: Digital twin; IoT: Internet of Things; PCM: Phase change material; TES: Thermal energy storage.

setup, the wall functions as an energy storage device in combination with wind turbines, solar panels, and battery backup systems. The latter two devices are utilized exclusively for backup purposes.[76]

The development of the DT smart building concept is shown in Figure 5. Before proceeding, the virtual entity component must be defined, and the corresponding physical entity must be clearly identified.[77] The next step is to create a digital representation of the virtual entity, decouple its physical layout, and use this model to assess the performance of simulations. The third step involves gathering design-related information – specifically, identifying the data to be collected when the smart building is operational and determining the tools required for this purpose.[78] Next, a link should be established between the real world and the building's digital representation, ensuring that the protocol is understood by all stakeholders. Finally, it is essential to verify that the smart building is functioning correctly and meets all necessary standards.[79]

An intelligent building model is created and then tested by the user, who provides detailed recommendations for improvement. In addition, the user can input further criteria into the system, including models for visual and physical elements, and suggest improvements to

conventional scheduling techniques. Developers will incorporate customer feedback into iterative software updates until the software reaches a satisfactory level. Only then will it be ready for full-scale implementation, testing, and maintenance.

The phase change wall (PCW) structure serves as the foundation for TES within the building. In response to temperature fluctuations, a PCM undergoes a transformation in its physical properties, which either increases or decreases its heat absorption capacity.[80] Most PCMs used in construction materials today are stored in solid-liquid mixtures. Liu *et al.*[81] developed form-stabilized PCMs with appropriate latent heat enthalpies.[81] These materials demonstrated excellent heat control and shape retention properties.[81] Within this indoor construction framework, the PCM is integrated into the building structure. Figure 8 depicts the intelligent building's PCW material and thermal network modeled using DT technology.

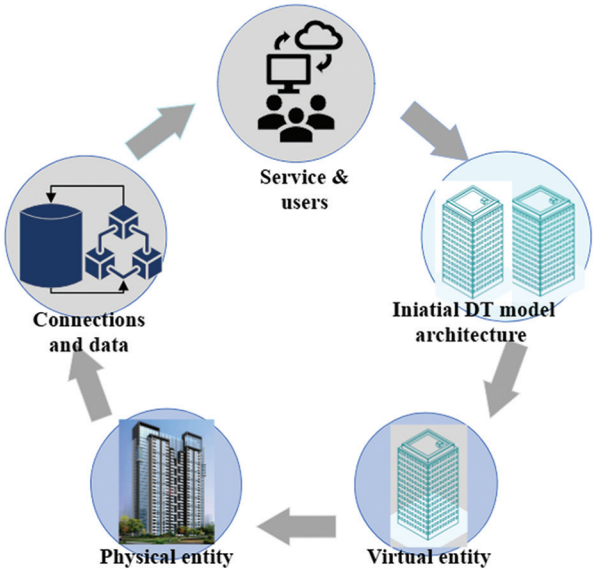


Figure 7. Energy conversion through DT technology
Abbreviation: DT: Digital twin.

4.2. Real-world applications of DT technology in TES

The growing global demand for energy is a major concern, particularly with the increasing number of people living in urban areas. The depletion of energy resources is contributing significantly to environmental issues, such as climate change, the expansion of the ozone hole, and global warming. In a study by South Korean researchers Seo *et al.*,[82] the energy consumption of a university classroom building was reduced by 60% after switching to LED lighting, using a turn-off technique. However, they also found that LED lights tend to consume more energy than expected and require adjustments. The study did not address the potential financial costs of LED bulbs, or the level of satisfaction occupants might have with the lighting's performance. DT technology has been the subject of several studies aimed at improving building energy management, with the goal of enhancing thermal comfort and energy efficiency.[83] Table 3 provides real-world applications of DT technology in intelligent buildings.

Qian *et al.*[84] conducted a year-long study of thermal sensation votes using a DT platform to learn about indoor and semi-open space environmental conditions. Their findings revealed that occupants felt hot and uncomfortable throughout the day and night during the summer. In another study focused on human-centered monitoring in intelligent buildings, Deng *et al.*[85] utilized digital identity technology to gain insights into indoor thermal comfort. The study classified two types of identification – digital and non-digital –into two sets. According to the responses, participants with digital identities reported improved

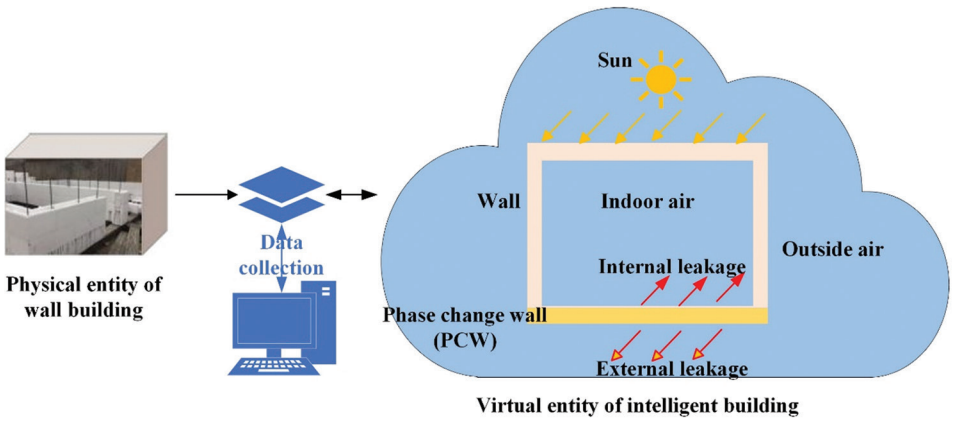


Figure 8. Energy conversion through digital twin technology

Table 3. Real-world applications of DT technology in smart buildings

Applications	Techniques employed	Key findings	References
Case study	Cloud-based IoT platform and DT	Humidity and temperature tracking within a structure	Nurumova *et al.*[86]
Case study	DT and BIM	Energy-saving technologies and strategies for existing building	Seo *et al.*[82]
Experimental	DT and BIM	Remote visual management for reduced energy consumption	Tan *et al.*[83]
Case study	DT platform	Monitoring environmental factors and understanding how consumers' comfort and pain levels change throughout the year	Qian *et al.*[84]
Experimental	DT and NN	Centralized database for building occupant data and environmental monitoring	Gnecco *et al.*[87]
Case study	Human-centric monitoring using DT and BIM	Accurate depiction of relative humidity and temperature	Deng *et al.*[85]
Experimental	NN and DT	Estimating occupiers' individual thermal comfort	Abdelrahman *et al.*[88]
Case study	DT and Graph Neural Network	Improve subjective data on thermal comfort by considering spatial proximity	Abdelrahman *et al.*[89]

Abbreviations: BIM: Building Information Modeling; DT: Digital twin; IoT: Internet of Things; NN: Neural network.

thermal comfort. However, the research was limited to just two aspects of thermal comfort: relative humidity and room temperature.

5. Conclusion and future perspectives

This study highlights TES systems, including sensible heat, latent heat, and thermochemical energy, as essential solutions for promoting environmental sustainability. These systems are crucial for addressing the depletion of energy resources, mitigating climate change, and reducing greenhouse gas emissions. A concise overview of PCMs and DT technology, including their properties and selection criteria, is provided to support efforts aimed at reducing energy consumption in buildings. Various case studies, along with numerical and experimental analyses, are examined to explore practical cooling applications and the integration of PCM-DT technology in the development of intelligent buildings.

One important area that requires further investigation is the economic analysis of PCM systems. Despite its significance in commercializing and improving these systems, these aspects remain under-discussed in some publications. Future research should focus on measuring thermal comfort indices in both indoor and outdoor environments in real-time, possibly through the use of computer software and other digital instruments. This approach could enhance heat monitoring and offer more accurate results than existing formulaic methods.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Mohammad Waseem, Mumtaz Ahmad, Mohammad Afazal

Supervision: Mohammad Waseem, Mumtaz Ahmad

Writing – original draft: Mohammad Waseem, G. Sree Lakshmi, Areti M. S. V. Sushma

Writing – review & editing: Mohammad Waseem, Mumtaz Ahmad, Sanajay Pual

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Du K, Calautit J, Wang Z, Wu Y, Liu H. A review of the applications of phase change materials in cooling, heating and power generation in different temperature ranges. *Appl Energy*. 2018;220:242-273.

   doi: 10.1016/j.apenergy.2018.03.005

2. Husein M, Chung IY. Optimal design and financial feasibility of a university campus microgrid considering renewable energy incentives. *Appl Energy.* 2018;225:273-289.

   doi: 10.1016/j.apenergy.2018.05.036

3. Peker M, Kocaman AS, Kara BY. Benefits of transmission

switching and energy storage in power systems with high renewable energy penetration. *Appl Energy*. 2018;228:1182-1197.

doi: 10.1016/j.apenergy.2018.07.008

4. Liu J, Mei C, Wang H, Shao W, Xiang C. Powering an island system by renewable energy-a feasibility analysis in the Maldives. *Appl Energy*. 2018;227:18-27.

doi: 10.1016/j.apenergy.2017.10.019

5. Bloess A, Schill WP, Zerrahn A. Power-to-heat for renewable energy integration: A review of technologies, modeling approaches, and flexibility potentials. *Appl Energy*. 2018;212:1611-1626.

doi: 10.1016/j.apenergy.2017.12.073

6. Devaux P, Farid MM. Benefits of PCM underfloor heating with PCM wallboards for space heating in winter. *Appl Energy*. 2017;191:593-602.

doi: 10.1016/j.apenergy.2017.01.060

7. Akeiber H, Nejat P, Majid MZA, *et al*. A review on Phase Change Material (PCM) for sustainable passive cooling in building envelopes. *Renew Sustain Energy Rev*. 2016;60:1470-1497.

doi: 10.1016/j.rser.2016.03.036

8. Young BA, Falzone G, Wei Z, Sant G, Pilon L. Reduced-scale experiments to evaluate performance of composite building envelopes containing phase change materials. *Constr Build Mater*. 2018;162:584-595.

doi: 10.1016/j.conbuildmat.2017.11.160

9. Liu C, Zhou Y, Li D, Meng F, Zheng Y, Liu X. Numerical analysis on thermal performance of a PCM-filled double glazing roof. *Energy Build*. 2016;125:267-275.

doi: 10.1016/j.enbuild.2016.05.002

10. De Gracia A, Cabeza LF. Phase change materials and thermal energy storage for buildings. *Energy Build*. 2015;103:414-419.

doi: 10.1016/j.enbuild.2015.06.007

11. Olivieri L, Tenorio JA, Revuelta D, Navarro L, Cabeza LF. Developing a PCM-enhanced mortar for thermally active precast walls. *Constr Build Mater*. 2018;181:638-649.

doi: 10.1016/j.conbuildmat.2018.06.013

12. Maccarini A, Hultmark G, Bergsøe NC, Afshari A. Free cooling potential of a PCM-based heat exchanger coupled with a novel HVAC system for simultaneous heating and cooling of buildings. *Sustain Cities Soc*. 2018;42:384-395.

doi: 10.1016/j.scs.2018.06.016

13. Boussaba L, Foufa A, Makhlouf S, Lefebvre G, Royon L. Elaboration and properties of a composite bio-based PCM for an application in building envelopes. *Constr Build Mater*. 2018;185:156-165.

doi: 10.1016/j.conbuildmat.2018.07.098

14. Kalnæs SE, Jelle BP. Phase change materials and products for building applications: A state-of-the-art review and future research opportunities. *Energy Build*. 2015;94:???.

doi: 10.1016/j.enbuild.2015.02.023

15. Cabeza LF, Castell A, Barreneche C, De Gracia A, Fernández AI. Materials used as PCM in thermal energy storage in buildings: A review. *Renew Sustain Energy Rev*. 2011;15(3):1675-1695.

doi: 10.1016/j.rser.2010.11.018

16. Sharma A, Tyagi VV, Chen CR, Buddhi D. Review on thermal energy storage with phase change materials and applications. *Renewable and Sustainable Energy Reviews*. 2009;13(2):318-345.

doi: 10.1016/j.rser.2007.10.005

17. Waqas A, Ji J, Ali M, Alvi JZ. Effectiveness of the phase change material-based thermal energy storage integrated with the conventional cooling systems of the buildings - A review. *Proc Inst Mech Eng A J Power Energy*. 2018;232(6):735-766.

doi: 10.1177/0957650917754033

18. Chen C, Liu W, Wang H, Zhu L. Synthesis and characterization of novel solid-solid phase change materials with a polyurethaneurea copolymer structure for thermal energy storage. *RSC Adv*. 2016;6(105):102997-103005.

doi: 10.1039/C6RA23141A

19. Ferrer G, Solé A, Barreneche C, Martorell I, Cabeza LF. Review on the methodology used in thermal stability characterization of phase change materials. *Renew Sustain Energy Rev*. 2015;50:665-685.

doi: 10.1016/j.rser.2015.04.187

20. Chen K, Yu X, Tian C, Wang J. Preparation and characterization of form-stable paraffin/polyurethane composites as phase change materials for thermal energy storage. *Energy Convers Manag*. 2014;77:13-21.

doi: 10.1016/j.enconman.2013.09.015

21. Faraj K, Khaled M, Faraj J, Hachem F, Castelain C. Phase change material thermal energy storage systems for cooling applications in buildings: A review. *Renew Sustain Energy Rev*. 2020;119:109579.

doi: 10.1016/j.rser.2019.109579

22. Oropeza-Perez I, Stergaard PA. Active and passive cooling methods for dwellings: A review. *Renew Sustain Energy Rev*. 2018;82:531-544.

doi: 10.1016/j.rser.2017.09.059

23. Monghasemi N, Vadiee A. A review of solar chimney integrated systems for space heating and cooling application. *Renew Sustain Energy Rev*. 2018;81:2714-2730.

doi: 10.1016/j.rser.2017.06.078

24. Khan MMA, Saidur R, Al-Sulaiman FA. A review for phase

change materials (PCMs) in solar absorption refrigeration systems. *Renew Sustain Energy Rev*. 2017;76:105-137.

doi: 10.1016/j.rser.2017.03.070

25. Iten M, Liu S, Shukla A. Experimental validation of an air-PCM storage unit comparing the effective heat capacity and enthalpy methods through CFD simulations. *Energy*. 2018;155:495-503.

doi: 10.1016/j.energy.2018.04.128

26. Muresan AA, Attia S. Energy efficiency in the Romanian residential building stock: A literature review. *Renew Sustain Energy Rev*. 2017;74:349-363.

doi: 10.1016/j.rser.2017.02.022

27. Xie Y, Gilmour MS, Yuan Y, Jin H, Wu H. A review on house design with energy saving system in the UK. *Renew Sustain Energy Rev*. 2017;71:29-52.

doi: 10.1016/j.rser.2017.01.004

28. Nematpour Keshteli A, Sheikholeslami M. Nanoparticle enhanced PCM applications for intensification of thermal performance in building: A review. *J Mol Liq*. 2019;274:516-533.

doi: 10.1016/j.molliq.2018.10.151

29. Milián YE, Gutiérrez A, Grágeda M, Ushak S. A review on encapsulation techniques for inorganic phase change materials and the influence on their thermophysical properties. *Renew Sustain Energy Rev*. 2017;73:983-999.

doi: 10.1016/j.rser.2017.01.159

30. Yu K, Liu Y, Yang Y. Review on form-stable inorganic hydrated salt phase change materials: Preparation, characterization and effect on the thermophysical properties. *Appl Energy*. 2021;292:116845.

doi: 10.1016/j.apenergy.2021.116845

31. Xiong Y, Song C, Ren J, *et al*. Sludge-incinerated ash based shape-stable phase change composites for heavy metal fixation and building thermal energy storage. *Process Saf Environ Prot*. 2022;162:346-356.

doi: 10.1016/j.psep.2022.04.004

32. Jiang Z, Navarro Rivero ME, Liu X, She X, Xuan Y, Ding Y. A novel composite phase change material for medium temperature thermal energy storage manufactured with a scalable continuous hot-melt extrusion method. *Appl Energy*. 2021;303:117591.

doi: 10.1016/j.apenergy.2021.117591

33. Wang W, He X, Shuai Y, Qiu J, Hou Y, Pan Q. Experimental study on thermal performance of a novel medium-high temperature packed-bed latent heat storage system containing binary nitrate. *Appl Energy*. 2022;309:118433.

doi: 10.1016/j.apenergy.2021.118433

34. Antoniadou-Plytaria K, Steen D, Tuan LA, Carlson O, Fotouhi Ghazvini MA. Market-based energy management model of a building microgrid considering battery degradation. *IEEE Trans Smart Grid*. 2021;12(2):1794-1804.

doi: 10.1109/TSG.2020.3037120

35. Ren F, Wei Z, Zhai X. Multi-objective optimization and evaluation of hybrid CCHP systems for different building types. *Energy*. 2021;215:119096.

doi: 10.1016/j.energy.2020.119096

36. Lin Q, Chen YC, Chen F, DeGanyar T, Yin H. Design and experiments of a thermoelectric-powered wireless sensor network platform for smart building envelope. *Appl Energy*. 2022;305:1794-1804.

doi: 10.1016/j.apenergy.2021.117791

37. Yu C, Konlan J, Li G. Energy harvesting and electricity production through dissolved carbon dioxide by connecting two form-stable phase change materials. *J Mater Chem A*. 2024;12:7943-7955.

doi: 10.1039/d3ta06766a

38. Lv Z, Cheng C, Lv H. Digital twins for secure thermal energy storage in building. *Appl Energy.* 2023;338:120907.

doi: 10.1016/j.apenergy.2023.120907

39. Vazquez S, Lukic SM, Galvan E, Franquelo LG, Carrasco JM. Energy storage systems for transport and grid applications. *IEEE Trans Ind Electron*. 2010;57(12):3881-3895.

doi: 10.1109/TIE.2010.2076414

40. International Eletrotechnical Comission. Electrical Energy Storage White Paper. Vol 2; 2011.

41. Lin Y, Jia Y, Alva G, Fang G. Review on thermal conductivity enhancement, thermal properties and applications of phase change materials in thermal energy storage. *Renew Sustain Energy Rev*. 2018;82:2730-2742.

doi: 10.1016/j.rser.2017.10.002

42. Bland A, Khzouz M, Statheros T, Gkanas EI. PCMs for residential building applications: A short review focused on disadvantages and proposals for future development. *Buildings*. 2017;7(3):78.

doi: 10.3390/buildings7030078

43. Masood U, Haggag M, Hassan A, Laghari M. A Review of phase change materials as a heat storage medium for cooling applications in the built environment. *Buildings*. 2023;13(7):1595.

doi: 10.3390/buildings13071595

44. Sharaf M, Huzayyin AS, Yousef MS. Performance enhancement of photovoltaic cells using phase change material (PCM) in winter. *Alexandria Eng J*. 2022;61(6):4229-4239.

doi: 10.1016/j.aej.2021.09.044

45. Lilley D, Menon AK, Kaur S, Lubner S, Prasher RS. Phase

change materials for thermal energy storage: A perspective on linking phonon physics to performance. *J Appl Phys*. 2021;130(22):220903.

doi: 10.1063/5.0069342

46. Bhagat K, Saha SK. Numerical analysis of latent heat thermal energy storage using encapsulated phase change material for solar thermal power plant. *Renew Energy.* 2016;95:323-336.

doi: 10.1016/j.renene.2016.04.018

47. Anisur MR, Kibria MA, Mahfuz MH, Metselaar IHSC, Saidur R. Latent Heat Thermal Storage (LHTS) for energy sustainability. In: *Green Energy and Technology*. New Delhi: Springer; 2015. p. 201.

doi: 10.1007/978-81-322-2337-5_10

48. Liang L, Diao Y, Kang Y, Zhao Y, Wei X, Chen C. Characteristic of latent heat thermal energy storage strengthened by flat micro heat pipe array-copper foam composite structure. *Huagong Xuebao/CIESC J*. 2018;69:34-42.

doi: 10.11949/j.issn.0438-1157.20180778

49. Tyagi V V., Kaushik SC, Tyagi SK, Akiyama T. Development of phase change materials based microencapsulated technology for buildings: A review. *Renew Sustain Energy Rev*. 2011;15(2):1373-1391.

doi: 10.1016/j.rser.2010.10.006

50. Jelle BP, Kalnæs SE. Phase change materials for application in energy-efficient buildings. *In: Cost-Effective Energy Efficient Building Retrofitting: Materials, Technologies, Optimization and Case Studies*. United Kingdom: Woodhead Publishing; 2017.

doi: 10.1016/B978-0-08-101128-7.00003-4

51. Wang Q, Wu R, Wu Y, Zhao CY. Parametric analysis of using PCM walls for heating loads reduction. *Energy Build*. 2018;172:328-336.

doi: 10.1016/j.enbuild.2018.05.012

52. Amaral C, Vicente R, Marques PAAP, Barros-Timmons A. Phase change materials and carbon nanostructures for thermal energy storage: A literature review. *Renew Sustain Energy Rev*. 2017;79:1212-1228.

doi: 10.1016/j.rser.2017.05.093

53. Memon SA. Phase change materials integrated in building walls: A state of the art review. *Renew Sustain Energy Rev*. 2014;31:870-906.

doi: 10.1016/j.rser.2013.12.042

54. Zhai XQ, Wang XL, Wang T, Wang RZ. A review on phase change cold storage in air-conditioning system: Materials and applications. *Renew Sustain Energy Rev*. 2013;22:108-120.

doi: 10.1016/j.rser.2013.02.013

55. Solangi NH, Mubarak NM, Karri RR, *et al*. MXene-based phase change materials for solar thermal energy storage. *Energy Convers Manag*. 2022;273:116432.

doi: 10.1016/j.enconman.2022.116432

56. Pandey AK, Hossain MS, Tyagi VV, Abd Rahim N, Selvaraj JAL, Sari A. Novel approaches and recent developments on potential applications of phase change materials in solar energy. *Renew Sustain Energy Rev*. 2018;82:281-323.

doi: 10.1016/j.rser.2017.09.043

57. Waqas A, Ud Din Z. Phase Change Material (PCM) storage for free cooling of buildings - A review. *Renew Sustain Energy Rev*. 2013;18:607-625.

doi: 10.1016/j.rser.2012.10.034

58. Chandel SS, Agarwal T. Review of current state of research on energy storage, toxicity, health hazards and commercialization of phase changing materials. *Renew Sustain Energy Rev*. 2017;67:581-596.

doi: 10.1016/j.rser.2016.09.070

59. Luo J, Zou D, Wang Y, Wang S, Huang L. Battery thermal management systems (BTMs) based on Phase Change Material (PCM): A comprehensive review. *Chem Eng J*. 2022;430:132741.

doi: 10.1016/j.cej.2021.132741

60. Drissi S, Ling TC, Mo KH. Thermal efficiency and durability performances of paraffinic phase change materials with enhanced thermal conductivity-a review. *Thermochim Acta*. 2019;673:198-210.

doi: 10.1016/j.tca.2019.01.020

61. Luo X, Zhu YS, Zhong YQ, Qin Y. An effective and reliable power data transmission scheme based on smart antenna WLAN technology. *J Clean Energy Technol*. 2013;1:327-330.

doi: 10.7763/jocet.2013.v1.74

62. Podara CV., Kartsonakis IA, Charitidis CA. Towards phase change materials for thermal energy storage: Classification, improvements and applications in the building sector. *Appl Sci*. 2021;11(4):1490.

doi: 10.3390/app11041490

63. Leong KY, Abdul Rahman MR, Gurunathan BA. Nano-enhanced phase change materials: A review of thermo-physical properties, applications and challenges. *J Energy Storage*. 2019;21:18-31.

doi: 10.1016/j.est.2018.11.008

64. Da Cunha SRL, De Aguiar JLB. Phase change materials and energy efficiency of buildings: A review of knowledge. *J Energy Storage*. 2020;27:101083.

doi: 10.1016/j.est.2019.101083

65. Nie B, Palacios A, Zou B, Liu J, Zhang T, YunrenLi. Corrigendum to "Review on phase change materials for cold thermal energy storage applications" [Renew. Sustain.

Energy Rev. 134 (2020) 110340/RSER-D-20-00822]. *Renew Sustain Energy Rev*. 2021;139:110642.

doi: 10.1016/j.rser.2020.110642

66. Mehling H, Brütting M, Haussmann T. PCM products and their fields of application-an overview of the state in 2020/2021. *J Energy Storage*. 2022;51:104354.

doi: 10.1016/j.est.2022.104354

67. Ning M, Jingyu H, Dongmei P, Shengchun L, Mengjie S. Investigations on thermal environment in residential buildings with PCM embedded in external wall. *Energy Procedia*. 2017;142:1888-1895.

doi: 10.1016/j.egypro.2017.12.387

68. Wang SM, Matiašovský P, Mihálka P, Lai CM. Experimental investigation of the daily thermal performance of a mPCM honeycomb wallboard. *Energy Build.* 2018;159:419-425.

doi: 10.1016/j.enbuild.2017.10.080

69. Lee KO, Medina MA, Sun X, Jin X. Thermal performance of Phase Change Materials (PCM)-enhanced cellulose insulation in passive solar residential building walls. *Solar Energy*. 2018;163:113-121.

doi: 10.1016/j.solener.2018.01.086

70. Yao C, Kong X, Li Y, Du Y, Qi C. Numerical and experimental research of cold storage for a novel expanded perlite-based shape-stabilized phase change material wallboard used in building. *Energy Convers Manag*. 2018;155:20-31.

doi: 10.1016/j.enconman.2017.10.052

71. Biswas K, Lu J, Soroushian P, Shrestha S. Combined experimental and numerical evaluation of a prototype nano-PCM enhanced wallboard. *Appl Energy*. 2014;131:517-529.

doi: 10.1016/j.apenergy.2014.02.047

72. Kharbouch Y, Mimet A, El Ganaoui M. Thermal impact study of a bio-based wall coupled with an inner PCM layer. *Energy Procedia*. 2017;139:10-15.

doi: 10.1016/j.egypro.2017.11.165

73. Li Y, Liu S, Lu J. Effects of various parameters of a PCM on thermal performance of a solar chimney. *Appl Therm Eng*. 2017;127:1119-1131.

doi: 10.1016/j.applthermaleng.2017.08.087

74. Agostinelli S, Cumo F, Guidi G, Tomazzoli C. Cyber-physical systems improving building energy management: Digital twin and artificial intelligence. *Energies (Basel)*. 2021;14(8):2338.

doi: 10.3390/en14082338

75. Zou Y, Li R, Zhang X, Song J. Five-dimensional model research of complex product assembly driven by digital twin. *Int J Wirel Mob Comput*. 2021;21(3):198-206.

doi: 10.1504/IJWMC.2021.120883

76. Lu X, Huang J, Wong WY, Qu JP. A novel bio-based polyurethane/wood powder composite as shape-stable phase change material with high relative enthalpy efficiency for solar thermal energy storage. *Solar Energy Mater Solar Cells*. 2019;200:109987.

doi: 10.1016/j.solmat.2019.109987

77. Huang J, Su J, Weng M, *et al*. An innovative phase change composite with high thermal conductivity and sensitive light response rate for thermal energy storage. *Solar Energy Mater Solar Cells*. 2022;245:111872.

doi: 10.1016/j.solmat.2022.111872

78. Rudra Murthy BV, Gumtapure V. Thermo-physical analysis of natural shellac wax as novel bio-phase change material for thermal energy storage applications. *J Energy Storage*. 2020;29:101390.

doi: 10.1016/j.est.2020.101390

79. Lu X, Huang J, Kang B, Yuan T, Qu J ping. Bio-based poly (lactic acid)/high-density polyethylene blends as shape-stabilized phase change material for thermal energy storage applications. *Solar Energy Mater Solar Cells*. 2019;192:170-178.

doi: 10.1016/j.solmat.2018.12.036

80. Sam MN, Caggiano A, Mankel C, Koenders E. A comparative study on the thermal energy storage performance of bio-based and paraffin-based PCMs using DSC procedures. *Materials*. 2020;13(7):1705.

doi: 10.3390/ma13071705

81. Liu L, Fan X, Zhang Y, *et al*. Novel bio-based phase change materials with high enthalpy for thermal energy storage. *Appl Energy*. 2020;268:114979.

doi: 10.1016/j.apenergy.2020.114979

82. Seo H, Yun WS. Digital twin-based assessment framework for energy savings in university classroom lighting. *Buildings*. 2022;12(5):544.

doi: 10.3390/buildings12050544

83. Tan Y, Chen P, Shou W, Sadick AM. Digital Twin-driven approach to improving energy efficiency of indoor lighting based on computer vision and dynamic BIM. *Energy Build*. 2022;270:112271.

doi: 10.1016/j.enbuild.2022.112271

84. Qian Y, Leng J, Chun Q, Wang H, Zhou K. A year-long field investigation on the spatio-temporal variations of occupant's thermal comfort in Chinese traditional courtyard dwellings. *Build Environ*. 2023;228:109836.

doi: 10.1016/j.buildenv.2022.109836

85. Deng M, Wang X, Li D, Menassa CC. Digital ID framework for human-centric monitoring and control of smart buildings. *Build Simul*. 2022;15(10):1709-1728.

doi: 10.1007/s12273-022-0902-3

86. Nurumova K, Ramaji I, Kermanshachi S. Leveraging Digital Twin for Enhancing Occupants Comfort: A Case Study. *In: Computing in Civil Engineering 2021-Selected Papers from the ASCE International Conference on Computing in Civil Engineering*; 2021.

    doi: 10.1061/9780784483893.052

87. Gnecco VM, Vittori F, Pisello AL. Digital twins for decoding human-building interaction in multi-domain test-rooms for environmental comfort and energy saving via graph representation. *Energy Build*. 2023;279:112652.

    doi: 10.1016/j.enbuild.2022.112652

88. Abdelrahman MM, Chong A, Miller C. Personal thermal comfort models using digital twins: Preference prediction with BIM-extracted spatial-temporal proximity data from Build2Vec. *Build Environ*. 2022;207:108532.

    doi: 10.1016/J.BUILDENV.2021.108532

89. Abdelrahman MM, Miller C. Targeting occupant feedback using digital twins: Adaptive spatial-temporal thermal preference sampling to optimize personal comfort models. *Build Environ*. 2022;218:109090.

    doi: 10.1016/j.buildenv.2022.109090

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

REVIEW ARTICLE

Examining the critical aspects of gas turbine blade failures caused by erosion using computational models: A comprehensive review

Surajit Mondal and Shankha Shubhra Goswami*

Department of Mechanical Engineering, Abacus Institute of Engineering and Management, Hooghly, West Bengal, India

Abstract

Gas turbines play a key role in industries such as aeronautics, maritime, and power generation, but high operating temperatures in these settings would expose critical components, particularly turbine blades, to significant wear. Among the multiple failure modes, erosion induced by high-velocity particle impacts is identified as one of the primary causes of turbine blade degradation. This review provides a comprehensive analysis of erosion-induced failure mechanisms in gas turbine blades, emphasizing recent advancements in understanding erosion processes, material degradation, and mitigation strategies. Key areas covered include the impacts of solid particle erosion, the effectiveness of erosion-resistant coatings, and the role of superalloys designed for high-temperature resilience. Furthermore, this review explores how computational models, such as machine learning, computational fluid dynamics, and finite element analysis, contribute to predicting erosion patterns and designing improved turbine components. By integrating experimental findings and computational methods, this review aims to inform future research and guide the development of advanced materials and protective strategies to enhance turbine blade durability in demanding environments.

Keywords: Gas turbine; Failures; Blade erosion; Turbine blade failures

***Corresponding author:**
Shankha Shubhra Goswami
(ssg.mech.official@gmail.com)

Citation: Mondal S, Goswami SS. Examining the critical aspects of gas turbine blade failures caused by erosion using computational models: A comprehensive review. *Int J AI Mater Design.* 2024;1(3): 66-98.
doi: 10.36922/ijamd.5188

Received: October 17, 2024
Accepted: November 28, 2024
Published Online: December 23, 2024

1. Introduction

Gas turbine engines (GTEs) represent a pivotal means for generating mechanical power through the controlled combustion of fuel. This technology has been adopted by many industries including aviation, maritime, rail transportation, electricity generation, and various industrial applications.[1] The wide-scale adoption of GTEs is primarily attributed to their compact form, reduced weight, and a commendably high power-to-weight ratio. Integral to their functioning are key components such as the compressor, combustor, and turbine, each playing a crucial role in the energy conversion process. Complementary components, such as the intercooler, superheater, and recuperator, have been integrated to enhance overall operational efficiency.[2,3] The operation of a GTE adheres to the ideal Brayton cycle, characterized by two isobaric heat interaction processes and two reversible adiabatic work interaction processes. The pursuit of an optimal GTE necessitates an emphasis on the efficiency of all constituent parts, ensuring seamless coordination and maximal energy conversion rates.

The operational efficiency of a system is intricately tied to the thermal environment, where an observed trend in literature indicates a notable correlation between increased maximum temperatures and heightened efficiency, while lower minimum temperatures contribute to this effect. Notably, the lower limit of attainable temperatures is constrained by ambient conditions, creating a direct and crucial relationship between overall efficiency and the maximum achievable temperature.[4,5] This relationship is particularly pronounced at the turbine inlet, where the gas temperature reaches its peak, rendering the turbine blades exceptionally vulnerable to significant thermal stresses. Consequently, current research endeavors are ardently focused on the development of advanced superalloys capable of withstanding and operating effectively under these elevated temperature regimes.[3,4,6,7] Such endeavors are pivotal in ensuring sustained performance and durability under demanding thermal conditions.

The focus of this article lies not in the investigation of thermal stress impacts but rather in the comprehensive analysis of the effects stemming from solid particulate matter entrained within the airflow. As these particles traverse the diffuser and compressor sections, an augmented air velocity ensues, carrying the solid particulates alongside.[2,3,8] Consequently, the impingement of these particles, combined with the elevated temperature and velocity, induces a pronounced erosion effect, posing a significant threat to the structural integrity of the blade material. This challenge turns the spotlight to the necessity in developing erosion-resistant superalloys, with researchers actively refining these advanced materials.[4,5,9,10] Moreover, to fortify the durability of the blades, a multifaceted approach involving the integration of specialized coatings, applied using diverse methodologies, is being pursued, aiming to bolster the protective capabilities of the blade substrate alloys. Such comprehensive strategies aim to ensure the sustained operational efficacy and longevity of the turbine blades within demanding environments characterized by solid particulate-laden airflow.

Solid particle erosion (SPE) denotes the progressive degradation of solid matter caused by the relentless bombardment of minuscule solid elements, posing a significant challenge in industrial settings, notably within the domain of gas turbine blade applications.[6,11] This erosive phenomenon has garnered substantial attention due to its detrimental impact on the operational efficiency and structural integrity of critical components. In this comprehensive review, the current landscape of research pertaining to SPE is meticulously surveyed, encompassing recent breakthroughs, methodologies, and underlying SPE-induced material degradation.[12,13] By examining the cutting-edge insights and advancements in this field, this review aims to consolidate a comprehensive perspective on the evolving strategies and mitigation techniques necessary to combat the persistent threat posed by SPE within industrial equipment and gas turbine systems. The article provides an in-depth look at erosion-induced damage in gas turbine blades, identifying failure mechanisms and influencing factors that significantly impact the performance and longevity of these components. The primary mechanisms of failure due to erosion are SPE, material fatigue, and structural deformation, each influenced by a set of operational and environmental factors.[14,15] Table 1 depicts the different failure mechanisms in gas turbine blades.

1.1. Primary failure mechanisms

The high-velocity impact of solid particles carried by the turbine's airflow is the leading cause of surface degradation in turbine blades.[16] When these particles strike the blade surfaces, particularly the leading and trailing edges, they cause localized material removal.[7] The high-temperature environment amplifies this effect, as particles with elevated kinetic energy penetrate and erode the surface, leading to cumulative surface damage. SPE causes severe impact in areas exposed to massive turbulent flows, such as the blade's leading edge, where particle impacts are frequently unpredictable and direct.

The repeated impacts of solid particles create small pits, scratches, and micro-cracks on the blade surface, which serve as stress concentrators. These stress concentrators are weak points that contribute to material fatigue, which leads to gradual expansion into larger cracks over time.[17,18] Under cyclic mechanical and thermal loading conditions typical of gas turbine operation, these cracks propagate further, weakening the blade structure and leading to failure.[8] Studies referenced in the article, including those by Khushbash *et al.*,[9] demonstrate that SPE-initiated microstructural damage reduces the blade's load-bearing capacity over time, eventually leading to catastrophic failure if left unaddressed.

High operational temperatures not only increase susceptibility to erosion but also cause time-dependent deformation or "creep" in the material. When erosion strips away protective coatings, it exposes the base material to oxidizing conditions and creep deformation, further reducing the structural integrity of the blade.[9,10,19] This effect is particularly noticeable in high-temperature environments where oxidation and thermal cycling accelerate material degradation, ultimately leading to blade rupture or failure. Past researches also highlight how erosion exposes deeper layers of material to heat, which

exacerbates creep deformation and fatigue.[11,20] Factors affecting erosion rates are tabulated and described in Table 2.

1.2. Influencing factors on erosion damage

Particle velocity, size, and density are significant determinants of erosion severity. Higher particle velocities lead to more forceful impacts, resulting in more extensive material degradation.[8,9,21,22] Fang *et al.*,[11] conducted experiments that demonstrated how particles with velocities between 100 and 500 m/s and diameters ranging from 1 to 5 mm caused considerable erosion, highlighting the critical role of particle characteristics in erosion dynamics. The angle at which particles collide with the blade surface also influences erosion outcomes. High impact angles (close to 90°) allow for maximum kinetic energy transfer, leading to greater material removal. Lower angles, by contrast, cause surface deformation rather than significant material loss, resulting in a smoother erosion pattern but not less overall wear over time.[10,11] Researchers observed that high-angle impacts cause more severe erosion on the blade leading edge, where impacts are direct and forceful.

The blade material's hardness, toughness, and coating type are central to its resistance against erosion. Harder materials generally exhibit higher erosion resistance, as they are able to better absorb and dissipate the energy from particle impacts.[12,23,24] In addition, porous coatings, such as those tested by other authors, are more susceptible to erosion because the voids within the coating material weaken its structural resistance, allowing particles to chip or erode sections of the surface more easily.[25] Advanced coatings such as thermal barrier coatings (TBCs) are often applied to mitigate this erosion; however, once these coatings wear away, the underlying material becomes exposed to the full force of particle impacts.[6,9,10,26] Turbine blades are constantly exposed to extreme temperatures, which play a critical role in the rate and extent of erosion-induced damage. High temperatures promote oxidation and thermal fatigue, which further erode the blade surface over time.[10,11] The thermal stresses involved weaken the blade material and accelerate erosion, especially when combined with the impact of high-speed particles. Past researches showed that erosion rates increase with temperature and gas velocity, and higher operating temperatures contribute to material softening, which exacerbates the effects of erosion.[12,27,28] This is particularly critical in high-temperature regions of the turbine where protective coatings may degrade over time, exposing the base material to accelerated wear. Moreover, the synergistic effect of thermal fatigue and particle impingement can significantly reduce the operational lifespan of turbine blades.[29] Table 3 presents the influence of particle characteristics on erosion.

Table 1. Failure mechanisms in gas turbine blades

Failure mechanism	Description	Contributing factors	Impact on blade performance
Solid particle erosion	Material loss caused by high-velocity particle impacts	Particle velocity, size, angle of impact, and material properties	Surface damage, reduced aerodynamic efficiency
Fatigue	Propagation of stress-induced cracks due to repeated loading cycles	Stress concentrators, thermal cycling, and material properties	Structural weakness and eventual blade failure
Oxidation and corrosion	Decline of material properties due to chemical reactions	High temperatures, exposure to oxidizing gases, and erosion	Accelerated material loss and reduced durability
Creep	Time-dependent deformation under constant stress in high-temperature conditions	Elevated temperatures, erosion-induced thinning of material	Reduced load-bearing capacity and rupture

Table 2. Factors affecting erosion rates

Factor	Description	Effect on erosion	Mitigation strategies
Particle properties	Velocity, size, density, and hardness of particles	Increases material loss with higher values of each parameter	Use harder materials and optimized coatings
Operational temperature	Gas and material temperatures during turbine operation	Accelerates oxidation, thermal fatigue, and material softening	Apply heat-resistant coatings (e.g., TBCs)
Blade geometry	Shape and surface profile of blades	Irregular geometries increase turbulence and particle impingement	Optimize blade design to minimize flow separation
Environmental conditions	Presence of oxidizing and corrosive gases	Accelerates chemical degradation and erosion synergistically	Use anti-corrosion coatings and maintain controlled environments

Abbreviation: TBCs: Thermal barrier coatings.

1.3. Mitigation strategies

The use of coatings and erosion-resistant materials is essential in mitigating the effects of erosion. TBCs, which consist of multiple layers including bond coats and thermal barriers, help to protect the underlying substrate from extreme temperatures and particle impacts. These coatings reduce the thermal and mechanical stresses exerted on the blade.[13,30] However, when TBCs begin to wear away under sustained particle impacts, the underlying material becomes vulnerable, accelerating erosion and leading to more rapid structural degradation.

Modifications to the geometry of turbine blades can also help minimize erosion. Redesigning the leading edges to reduce the incidence of particle impacts or altering the airflow pattern can reduce the rate of erosion, as highlighted by simulations and studies employing computational fluid dynamics (CFD) and finite element analysis (FEA) to predict erosion-prone zones and stress points.[14,31] Reducing turbulence and optimizing airflow in these regions helps lower erosion intensity, thereby prolonging blade life.

In summary, erosion-induced failures in gas turbine blades result from a complex interplay of mechanical and environmental factors, including particle impact characteristics, material properties, and operational conditions.[13,14,32] By understanding these mechanisms and the factors influencing erosion, researchers can develop more resilient materials, effective coatings, and optimized blade designs to reduce erosion's impact, thereby extending the operational life of gas turbine blades. These advancements also enable more efficient maintenance strategies, reducing downtime and operational costs.[33,34] Erosion mechanisms across turbine blade zones are listed and described in Table 4.

2. Method for predicting and mitigating erosion

Recent advancements in analytical methods have revolutionized the field of material degradation, specifically in addressing challenges such as erosion-induced failures in gas turbine blades.[10,12,14] Machine learning (ML) and other advanced computational techniques such as CFD and FEA are now being increasingly employed to enhance predictive capabilities and provide more efficient mitigation strategies for erosion.[35] Table 5 delineates the methods for predicting and mitigating erosion.

2.1. ML applications

In recent years, ML has emerged as a powerful tool in the field of material science and mechanical engineering, particularly in predictive maintenance and failure analysis.[15,36] The application of ML in gas turbine blade erosion studies has opened new avenues for improving reliability, extending operational life, and optimizing maintenance schedules. The ability of ML to analyze large datasets and extract complex patterns has revolutionized the way engineers approach predictive modeling for gas turbine blade erosion.[37] ML algorithms are becoming invaluable in identifying patterns

Table 3. Influence of particle characteristics on erosion

Parameter	Range tested	Effect on erosion	Key findings
Particle velocity	100 – 500 m/s	Higher velocities increase kinetic energy, causing greater material removal.	Erosion severity is directly proportional to particle velocity.[5,6,11]
Particle size	1 – 5 mm	Larger particles cause deeper pits and more extensive surface damage.	Larger diameters exacerbate erosion rates significantly.[3,4,6]
Impact angle	20° – 90°	Steeper angles (close to 90°) transfer more energy, causing higher erosion.	Leading edges of blades are most vulnerable due to direct impacts.[8,9,12,13]

Table 4. Erosion mechanisms across turbine blade zones

Blade zone	Erosion mechanism	Key factors	Impact on blade performance
Leading edge	High-frequency particle impingement	High impact angles, high velocity, turbulence at the boundary layer	Severe material removal, loss of aerodynamic efficiency
Trailing edge	Flow separation-induced erosion	Low impact angles, erratic particle trajectories, reduced turbulence dissipation	Localized erosion, structural fatigue, and eventual cracking
Blade surface (mid-body)	Uniform particle impacts with low turbulence.	Lower velocities and impact angles compared to edges	Gradual thinning, minor structural degradation
Blade tips	Combined mechanical and thermal erosion	High gas velocities, elevated temperatures, and centrifugal forces	Accelerated wear, material fatigue, and potential failure

Table 5. Methods for predicting and mitigating erosion

Technique	Purpose	Methodology	Outcome
CFD	Predicts particle trajectories and erosion	Simulates gas flow, turbulence, and particle impacts	Identifies high-risk zones on turbine blades
FEA	Assesses structural response to erosion	Simulates stress and deformation under particle loads.	Provides insights into fatigue and crack propagation.
ML	Enhances predictive maintenance.	Uses operational data to predict erosion patterns	Optimizes maintenance schedules and material selection (e.g., YSZ coatings).
Experimental testing	Validates computational models	Simulates real-world erosion conditions	Provides empirical data for material performance (e.g., TBC erosion rates)

Abbreviations: CFD: Computational fluid dynamics; FEA: Finite element analysis; ML: Machine learning; TBC: Thermal barrier coating; YSZ: Yttria-stabilized zirconia.

in complex datasets generated during material wear and erosion studies. For instance, ML can be employed to predict the failure modes of gas turbine blades based on historical operational data, considering variables such as particle velocity, impact angles, and material properties.[16,17,38] A supervised learning approach can categorize the severity of erosion based on these factors, allowing for the development of targeted interventions and more accurate maintenance scheduling. A recent research has demonstrated that integrating ML with existing computational models significantly enhances the prediction accuracy of erosion rates, particularly in high-velocity environments like those found in gas turbine systems.[18]

2.1.1. Predictive maintenance and anomaly detection

One of the most impactful contributions of ML is its role in predictive maintenance. Various operational parameters, such as temperature, pressure, particle velocity, and erosion rates, are monitored in real-time through sensors installed in turbine engines.[39,40] Traditional maintenance relies on fixed intervals, but this approach often leads to either over-maintenance or unexpected failures.[18] ML models, particularly supervised learning algorithms, are increasingly used to predict blade failure based on historical data, thus optimizing maintenance schedules.[41]

Blinov *et al.*[19] demonstrated that deep learning models, such as long short-term memory (LSTM) networks, are highly effective in processing time-series data from turbines to detect gradual degradation due to erosion. Their study showed that LSTMs could predict the remaining useful life of turbine blades with higher accuracy than traditional statistical methods. Similarly, Mishra and Kumar[20] employed support vector machines (SVM) and random forest (RF) classifiers to classify erosion severity based on operational data, such as particle velocities and impingement angles, providing early warnings for maintenance. Table 6 presents a comparative analysis of coating materials for erosion resistance.

2.1.2. Optimizing erosion resistance through material selection

ML has also proven valuable in optimizing the selection of erosion-resistant materials. By analyzing large datasets containing material properties – such as hardness, tensile strength, and thermal resistance – ML models can identify correlations between material composition and performance under erosive conditions.[21,42,43] Rani and Agrawal[22] applied decision-tree models to predict the erosion resistance of newly developed superalloys and coatings, allowing researchers to focus their experiments on materials with the highest predicted success rates in field trials.

Further advancing the field, Bonu and Barshilia[23] explored the use of reinforcement learning (RL) to dynamically adjust material compositions during operational simulations. Their research demonstrated that RL systems could suggest real-time adjustments to material properties, improving erosion resistance without the need for prolonged testing. This approach offers significant potential for accelerating the development of more robust materials for gas turbine applications. Table 7 presents some of the experimental techniques used for erosion studies.

2.1.3. Integration of ML with CFD

Another important application of ML in turbine erosion studies is its integration with CFD. While CFD simulations provide detailed insights into airflow and particle behavior, they are computationally expensive.[24,25,44] ML models trained on CFD-generated data can serve as surrogate models, making predictions about erosion patterns much faster and without the computational cost of running new simulations.

Yenugula *et al.*[26] applied convolutional neural networks (CNNs) to CFD datasets, successfully predicting erosion-prone zones on turbine blades based on airflow and particle trajectories. Their research showed that CNNs

Table 6. Comparative analysis of coating materials for erosion resistance

Coating material	Properties	Advantages	Limitations	Applications
YSZ	High thermal insulation and hardness	Excellent erosion and heat resistance	Reduced effectiveness with high porosity	TBCs
MCrAlY	Oxidation and hot corrosion resistance	Strong adhesion layer for ceramic coatings	Less effective against direct particle erosion	Bond coats in TBC systems
Alumina-based coatings	High wear resistance	Effective against abrasion in moderate conditions	Reduced resistance at elevated temperatures	Erosion protection at lower temperatures
Composite ceramic matrix	Combines ceramic and metallic matrices	Exceptional erosion resistance in harsh environments	High manufacturing costs	Advanced turbine applications

Abbreviations: TBC: Thermal barrier coating; YSZ: Yttria-stabilized zirconia.

Table 7. Experimental techniques used for erosion studies

Experimental technique	Purpose	Materials/Conditions tested	Key findings
High-velocity gas tunnels	Simulate erosion under real-world conditions	Zirconia-based TBCs at temperatures > 980°C[16,19,20]	Demonstrated erosion rate dependence on velocity and temperature
SEM	Analyzes surface damage morphology	Ceramic coatings, superalloys (e.g., Nimonic-105)	Revealed crack initiation points and material microstructure degradation
Taguchi design of experiments	Identify factors influencing erosion rates	Plasma-sprayed YSZ coatings with varying impact angles	Highlighted particle velocity as the most significant factor affecting erosion
Particle impingement testing	Measures erosion rates under controlled conditions	Alumina and titanium-based coatings	Established relationship between particle size and material loss

Abbreviations: SEM: Scanning electron microscopy; YSZ: Yttria-stabilized zirconia.

could analyze complex geometries and turbulence patterns in a fraction of the time required by traditional CFD simulations, enabling engineers to test multiple design configurations rapidly before finalizing a blade's geometry.

2.1.4. Unsupervised learning for feature discovery

In addition to supervised learning, unsupervised ML algorithms have proven useful in discovering hidden patterns in large, unlabeled datasets generated from turbine operations.[27,45,46] Techniques such as k-means clustering and principal component analysis have been employed to identify latent factors that contribute to erosion, such as variations in particle composition or interactions between environmental conditions.

Chowdhury *et al.*[28] used k-means clustering to analyze turbine operational data and discovered previously unrecognized correlations between particle composition and increased erosion rates. Their findings led to the development of new protective coatings better suited to resist the specific erosive forces encountered in real-world turbine environments.

2.1.5. Hybrid models and the future of ML in erosion studies

As ML continues to evolve, hybrid models combining multiple techniques are being developed for more robust predictions and failure analysis. Combining ML with physics-based models allows researchers to incorporate domain-specific knowledge into the learning process, creating models that are both accurate and computationally efficient.[29,47,48] Talebi *et al.*[30] demonstrated the effectiveness of hybrid models that fuse ML with fluid dynamics simulations, predicting erosion rates with higher precision under complex conditions.

Moreover, the integration of ML with real-time data from Internet of Things devices embedded in turbine engines is paving the way for continuous, real-time monitoring systems.[49,50] This approach, as highlighted by Chen *et al.*,[31] enables proactive responses to erosion risks, reducing the need for reactive maintenance and improving the overall efficiency and operational lifespan of turbines.

2.1.6. Future prospects

The future of ML in gas turbine blade erosion studies is promising. As more data become available and algorithms become more sophisticated, the accuracy and precision of failure predictions will continue to improve.[32,33,51] ML, when combined with traditional material science approaches and computational simulations, will enable the design of more resilient gas turbine systems and a reduction in operational costs across various industries.[52]

2.2. CFD and FEA

Advancements in CFD and FEA continue to provide detailed insights into the behavior of solid particles within GTEs. These methods simulate the interaction of fluid flow with turbine blades, offering a deeper understanding of how erosion patterns develop under varying operational conditions.[34,53] When combined with ML, these simulations can be optimized to predict critical erosion-prone zones, providing more effective erosion control measures. Incorporating such methods helps bridge the gap between theoretical understanding and practical solutions. The application of CFD and FEA has become indispensable in understanding the complex fluid-structure interactions, material degradation, and failure modes associated with gas turbine blade erosion.[35,36,54] Both techniques allow researchers to simulate real-world operational conditions in a virtual environment, thus offering deeper insights into the mechanisms of erosion and enabling more precise design optimization.[37,55,56] Recent advancements in CFD and FEA, when combined with ML, have significantly enhanced the ability to predict erosion-prone areas, optimize blade geometries, and select appropriate materials and coatings for gas turbine blades.

2.2.1. CFD for erosion prediction and flow analysis

CFDs play a critical role in simulating the flow of gases and particles within GTEs, which is essential for understanding the dynamics of erosion.[57] By modeling the airflow patterns and particle trajectories, CFD provides a detailed visualization of how solid particles interact with turbine blades at various velocities, temperatures, and angles of impingement.[38,39,58] The resulting simulations help identify regions of high erosion risk, enabling engineers to design blades that minimize erosion damage.

In their study, Peng *et al.*[40] employed CFD to simulate the behavior of solid particles in the gas flow of a turbine. They found that particles with higher velocities tend to follow more erratic trajectories, which significantly increases erosion near the leading edge of turbine blades. Their research demonstrated that regions experiencing the highest flow velocity and turbulence were the most vulnerable to particle impingement. Alqallaf and Teixeira[41] expanded on this by integrating CFD with Eulerian–Lagrangian models to predict erosion rates more accurately, considering both the fluid flow and the particle interaction with the blade surface. Their results highlighted that, while particle size and velocity are important factors, the angle of particle impact is critical in determining erosion severity.

Further advancements in CFD, as shown by Goswami *et al.*,[42] involve the use of large eddy simulations (LES) and direct numerical simulations (DNS) to capture small-scale turbulence and particle interactions in high-detail. LES, in particular, has been shown to provide more accurate results compared to traditional Reynolds–Averaged Navier–Stokes (RANS) models, especially in simulating highly turbulent flows around turbine blades. In gas turbines, flow conditions are often turbulent. Models such as LES and RANS are used to model turbulence. LES resolves the large-scale turbulent structures while modeling the smaller scales, which is crucial for accurately predicting particle impacts in highly turbulent flow regions. Rivaz *et al.*[43] applied LES to model gas-particle interactions in a turbine operating at high temperatures, revealing the intricate erosion patterns caused by turbulent eddies near the blade's trailing edge. Their research demonstrated that more accurate CFD models can significantly improve the prediction of erosion-prone zones and inform better blade design. Table 8 presents computational models for erosion prediction and their different aspects.

2.2.2. FEA for structural deformation and erosion effects

While CFD primarily focuses on the fluid dynamics aspect, FEA is essential for understanding the structural response of turbine blades to erosion-induced damage. FEA allows researchers to simulate the mechanical stresses, strains, and material deformation that occur in turbine blades under operational loads.[44,45,59] Erosion, particularly when combined with high thermal and mechanical stresses, can lead to significant structural degradation, and FEA provides a framework for assessing the extent of this damage.

Wei *et al.*[46] employed FEA to simulate the impact of repeated particle impingement on turbine blade surfaces, showing that erosion leads to micro-cracking and pitting, which weakens the material over time. Their FEA model incorporated both thermal and mechanical loads, providing a more comprehensive understanding of how erosion accelerates fatigue failure in turbine blades. Their study concluded that the combination of erosion and thermal cycling creates stress concentrations, which propagate cracks and lead to premature blade failure.

Similarly, Mortazavi *et al.*[47] used FEA to analyze the effects of particle size, impact velocity, and angle on the deformation of turbine blades. Their study found that larger particles caused deeper pits and more extensive material removal, while higher velocities increased the plastic deformation of the blade surface. The integration of FEA with CFD, as performed by Prashar and Vasudev,[48] provided even more precise results. By coupling fluid flow data from CFD with the structural analysis capabilities of FEA, they were able to model not only the particle

Table 8. Computational models for erosion prediction

Model	Application	Strengths	Limitations
CFD	Predicts particle trajectories and flow patterns	Provides detailed insights into high-velocity environments	Computationally expensive, limited by turbulence models
FEA	Simulates structural response to particle impacts	Models stress distribution and crack propagation	Requires accurate input data from CFD
Eulerian–Lagrangian Framework	Tracks particle motion and impingement locations	Combines fluid dynamics and particle interactions	Accuracy depends on mesh quality and particle assumptions
ML	Predicts erosion patterns based on prior data	Allows rapid predictions, reduces reliance on computational resources	Requires extensive and accurate training datasets

Abbreviations: CFD: Computational fluid dynamics; FEA: Finite element analysis; ML: Machine learning.

trajectories but also the resulting stresses on the blade surface, leading to a more accurate prediction of material fatigue and failure due to erosion.

2.2.3. Integrated CFD and FEA approach

The combination of CFD and FEA provides a more holistic approach to understanding gas turbine blade erosion. CFD offers insights into the flow dynamics and particle behavior, while FEA helps analyze the material's response to erosion-induced stresses.[49,60] This integration is particularly beneficial when investigating how erosion weakens the structural integrity of blades over time, as well as how changes in blade geometry affect erosion patterns.

Rezamand *et al.*[50] applied an integrated CFD-FEA framework to study the long-term effects of particle impingement on turbine blades. Their approach involved using CFD to simulate particle flows and impact angles, followed by FEA to assess how these impacts translated into material degradation. Their study revealed that blade regions with complex geometries, such as the leading and trailing edges, were the most vulnerable to erosion and subsequent structural failure. By combining CFD and FEA, they were able to propose design modifications that reduced turbulence and particle impact on critical blade areas, thus extending the operational life of the blades.

This integrated approach was further advanced by Li *et al.*,[51] who combined multi-scale modeling techniques with CFD and FEA to assess how micro-scale erosion events (*e.g.*, pitting and surface roughness) accumulate over time to affect macro-scale blade performance. Their research demonstrated that small-scale surface irregularities caused by erosion can lead to increased drag and reduced aerodynamic efficiency, which ultimately affects the overall performance of the turbine. Their integrated model allowed for more accurate life-cycle predictions for gas turbine blades, highlighting the need for more erosion-resistant materials and coatings.

2.2.4. Computational challenges and future directions

Despite the significant advancements in both CFD and FEA, several challenges remain. CFD simulations, especially those using LES and DNS, are computationally expensive and require substantial processing power.[61] FEA models that account for complex material behaviors, such as creep and fatigue under high temperatures, also demand considerable computational resources.[52,53,62] Moreover, fully coupling CFD and FEA for real-time analysis is still an area of ongoing research due to the difficulty of synchronizing fluid dynamics and material responses in a single simulation.[63]

However, the integration of ML into CFD and FEA has the potential to overcome these challenges. Olabi *et al.*[53] explored the use of ML-based surrogate models to approximate CFD and FEA simulations. By training neural networks (NNs) on existing CFD and FEA data, they were able to generate accurate predictions of fluid-structure interactions at a fraction of the computational cost. This approach opens the door to real-time erosion monitoring and adaptive turbine blade designs that respond dynamically to changing operational conditions.

In the future, advancements in parallel computing and cloud-based simulations could make high-fidelity CFD and FEA models more accessible. Han *et al.*[54] predicted that the use of multi-physics simulation platforms – which combine CFD, FEA, thermal analysis, and ML – will become increasingly common in turbine design and maintenance. Such platforms could allow for continuous monitoring and optimization of turbine blades, ensuring higher efficiency and longer operational life. Table 9 summarizes the key research contributions that are exclusively highlighted in this article.

3. Theoretical analysis

The theoretical framework of this research offers a structured lens to examine erosion-induced failures in

Table 9. Key research contributions highlighted in this article

Researcher (s)	Focus area	Methodology used	Key findings
Hamed *et al.*[55]	Erosion mechanisms in high-velocity environments	CFD and experimental testing	Identified critical erosion-prone zones on turbine blades
Taherkhani *et al.*[56]	Effect of particle size and velocity on erosion rates	FEM simulations and experimental studies	Showed particle velocity as the dominant factor for erosion
Shin and Hamed[57]	TBC erosion under high temperatures and velocities	Taguchi method and SEM analysis	Highlighted porosity's impact on erosion resistance
Branco *et al.*[58]	Erosion of alumina-based coatings	Particle impingement testing	Porosity reduces erosion resistance by weakening material
Zhu *et al.*[59]	Thermal gradient effects on TBC erosion	Experimental tests and SEM imaging	Demonstrated spalling and fatigue under combined conditions

Abbreviations: CFD: Computational fluid dynamics; FEM: Finite element method; SEM: Scanning electron microscopy; TBC: Thermal barrier coating.

gas turbine blades by integrating theories from multiple disciplines, including fluid dynamics, material science, and mechanical engineering.[60,61] The framework outlines how various theories come together to explain the failure mechanisms of turbine blades, focusing on how erosion compromises the structural integrity of the blades and affects the overall operational efficiency. It also draws on computational methods to model and predict failures, helping to design more resilient blades and optimize maintenance schedules.[64,65] In addition, it incorporates insights from material science to identify critical thresholds for wear and erosion resistance. These multidisciplinary approaches ensure a comprehensive understanding of blade failure dynamics. The three main theoretical pillars are discussed in the following sub-sections.

3.1. Erosion mechanism theories

The first pillar of the theoretical framework focuses on understanding how erosion mechanisms affect gas turbine blades, particularly how solid particles suspended in airflow impinge on the blade surfaces, leading to material removal and surface degradation.[62,66] These mechanisms are vital for explaining the localized wear and tear observed on turbine blades during high-stress operational environments.

Particle impact theory explains the interaction between solid particles and the surface of turbine blades. Particles suspended in the high-velocity gas stream collide with the turbine blades, leading to localized surface damage.[63,67] The rate and severity of erosion depend on the kinetic energy of the particles, which is a function of their velocity, size, mass, and the hardness of both the particle and the impacted material. Kumar *et al.*[64] explored how high-velocity particles impact surfaces at various angles, contributing to material loss through a process known as SPE. They found that high-impact angles, particularly those closer to 90 degrees, cause more severe erosion, as the energy transfer during these collisions is maximized.

Erosion rate can be mathematically modeled to predict the amount of material loss over time. One of the early models, Finnie's equation,[65,68] posits that the erosion rate is directly proportional to the kinetic energy of the particles. It also emphasizes the importance of the impact angle, where shallow angles tend to result in surface plowing (material deformation without removal), and higher angles cause actual material loss. These models are essential for understanding where and how fast erosion will occur on turbine blades, especially in high-speed regions such as the leading edge of the blades.

The boundary layer theory[66] offers insights into how fluid flow behaves near the blade surface. It explains the development of turbulence close to the surface, which affects how particles move and impinge on the blade. In high-velocity gas flows, the turbulent boundary layer can cause particles to move erratically, leading to more frequent and unpredictable particle impacts.[67,69] This theory helps explain why certain areas of a turbine blade (*e.g.*, leading edges and blade tips) experience more severe erosion compared to smoother, laminar-flow regions.

Flow separation and particle rebound involves that the interaction between fluid flow and blade surfaces often leads to flow separation, especially at sharp edges or curved surfaces of the blades. When this occurs, particles rebound off surfaces in chaotic patterns, increasing the likelihood of high-energy impacts. Quintanar-Gago *et al.*[68] showed that flow separation near the trailing edges can cause erosion patterns that are more concentrated in certain zones, further exacerbating blade wear. Table 10 outlines the influence of coating porosity on erosion resistance.

This pillar is critical for understanding the physical forces behind erosion, allowing engineers to design blade shapes and materials that can better resist particle impacts and prolong blade life. In the study of erosion-induced damage in gas turbine blades, predicting the erosion

Table 10. Influence of coating porosity on erosion resistance

Porosity level (%)	Coating material	Erosion resistance	Failure mode	Key observations
<10%	Dense ceramic coatings	Very high: minimal particle penetration	Micro-cracking and delamination	Effective at high temperatures but prone to brittleness
10 – 15%	YSZ	Moderate: balances insulation and toughness.	Surface pitting and spalling	Optimal balance for TBC performance[57]
>15%	Porous ceramic coatings	Low: High particle ingress and reduced strength	Accelerated erosion and chipping	High porosity reduces mechanical integrity[58]

Abbreviations: TBC: Thermal barrier coating; YSZ: Yttria-stabilized zirconia.

rate and the impact of particle collisions on the material surface is essential.[69,70] Here, we consider models such as Finnie's erosion rate equation shown by Equation I, as well as more advanced particle impact models derived from fluid dynamics. Finnie's equation provides a basic model for estimating the erosion rate of a material based on particle velocity and impact angle. The erosion rate "E" is expressed, as shown in Equation I.

$$E = C\left(\frac{\rho_p v^2}{2}\right) sin^2\left(\theta\right) \quad \text{(I)}$$

where E is the erosion rate (material loss per unit area); C is a constant dependent on material properties and particle characteristics; ρ_p is the particle density; v is the velocity of the impacting particle; and θ is the angle of impact of the particle.

This model shows that impact angle plays a significant role in erosion severity, with steep angles (close to 90 degrees) causing the highest erosion due to maximal energy transfer.[71] The $sin^2(\theta)$ term captures the influence of angle on erosion. Once particle impacts are determined, empirical models such as Finnie's erosion rate equation or more complex erosion models[48,49] are used to calculate the erosion rate at each impact site. For instance, Finnie's equation given by Equation I provides a direct method for estimating material removal due to each particle impact.

In more complex models, the erosion rate is also a function of particle size and material hardness. For example, Anand and Parammasivam[72] developed an empirical model that accounts for particle size "d," hardness "H," and velocity "v," as given by Equation 2.

$$W = k\frac{v^n \cdot d^m}{H} \quad \text{(II)}$$

where E is the erosion rate; k is an empirically determined constant; v is the particle velocity; d is the particle diameter; H is the hardness of the material; and n and m are exponents typically determined through experimentation.

In gas turbines, particles with larger diameters and higher velocities cause greater erosion, while materials with higher hardness (*e.g.*, superalloys or coated surfaces) exhibit better erosion resistance.

3.2. Material degradation and fatigue theory

The second theoretical pillar explores how erosion-induced material degradation leads to fatigue failure in gas turbine blades.[73] While the previous pillar focused on how particles cause surface damage, this section focuses on how that damage propagates into more severe material failures over time, considering both mechanical and thermal stressors.

Material fatigue refers to the progressive weakening of a material due to repeated stress cycles, which eventually leads to crack formation and structural failure. Alqallaf and Teixeira[74] demonstrated that in gas turbine blades, erosion not only causes surface damage but also accelerates fatigue by creating stress concentrators – localized areas where stress is intensified due to surface irregularities such as pits, scratches, and micro-cracks. These stress concentrators initiate cracks that grow over time, ultimately leading to catastrophic failure of the blade. The Paris–Erdogan law[75] is frequently used to model crack growth, which relates the crack propagation rate to the applied stress intensity. Fatigue failure theory helps in predicting how long a turbine blade can operate under erosive conditions before failure is likely to occur.

In high-temperature environments, turbine blades experience creep, a slow, time-dependent deformation that occurs when materials are exposed to constant stress at elevated temperatures. Erosion exacerbates creep by thinning the blade material, reducing its ability to resist the mechanical stresses that it experiences during operation. Rajabinezhad *et al.*[76] highlighted that as erosion removes protective coatings and material from the surface, it exposes deeper layers to higher temperatures, accelerating creep deformation. This leads to reduced load-bearing capacity and increases the likelihood of blade rupture.

To mitigate the effects of erosion and thermal stresses, advanced surface coating theories are often applied to turbine blades. These coatings act as a protective barrier between the blade material and the harsh operational environment. Shin and Hamed[77] investigated the effectiveness of TBCs in reducing erosion and oxidation. TBCs are typically made of ceramics, which are highly resistant to heat and erosion. However, as erosion wears away these coatings, the underlying blade material becomes exposed, leading to accelerated degradation. The diffusion theory[69,70] explains how heat and oxygen diffuse through these coatings over time, contributing to their eventual breakdown.

This section explains the long-term effects of erosion on turbine blades. As the blades experience repeated cycles of erosion, high temperatures, and mechanical loads, the damage accumulates, eventually leading to failure.[21,39,45,57,71,72] This understanding is crucial for developing maintenance strategies that address the cumulative effects of erosion and fatigue.

The Paris–Erdogan law[70] for crack growth, combined with erosion models, can describe how erosion-induced fatigue leads to material failure. The Paris–Erdogan law[70,71] is used to model the propagation of cracks in materials subjected to cyclic loading, which is typical in gas turbines due to operational cycles.

The rate of crack growth per cycle, $\frac{da}{dN}$, is given by Equation III:

$$\frac{da}{dN} = C \cdot \Delta K^m \tag{III}$$

where a is the crack length; N is the number of loading cycles; C and m are material-dependent constants; and ΔK is the stress intensity factor range (difference between maximum and minimum stress intensities during a cycle).[78] The stress intensity factor range, ΔK, is given by Equation IV:

$$\Delta K = K_{max} - K_{min} = \Delta\sigma \cdot \sqrt{\pi a} \tag{IV}$$

where $\Delta\sigma$ is the stress range (difference between maximum and minimum stress in a cycle); and a is the crack length. This equation links erosion-induced cracks, which serve as stress concentrators, with the fatigue life of the blade. Repeated erosion weakens the blade's surface, increasing the rate of crack propagation under cyclic thermal and mechanical loads, leading to failure.

3.3. Computational modeling and simulation theories

The third pillar focuses on computational methods – specifically CFD and FEA – to model and predict erosion patterns and structural failure in turbine blades.[16,19] These simulations are invaluable for understanding how fluid dynamics and material behavior interact under real-world conditions.

CFD models simulate the complex interactions between gas flows and turbine blades, particularly focusing on how high-velocity flows and solid particles contribute to erosion. Hamed *et al.*[74] applied the Navier–Stokes equations[32,34,38] to predict airflow patterns around turbine blades. By incorporating Eulerian–Lagrangian models,[14,19] they were able to simulate particle trajectories and the subsequent impacts on the blade surface. Fluid-structure interaction technique allows for precise identification of erosion hotspots, particularly on the leading edges of blades where particles tend to accumulate and cause the most damage.

While CFD focuses on fluid dynamics, FEA models the structural response of the blade material to the forces generated by particle impacts. Taherkhani *et al.*[56] used FEA to simulate how repeated particle collisions create stress concentrations and lead to material deformation and fatigue. von Mises stress theory[56,57,60] was employed to understand how stress is distributed across the blade surface. By combining CFD and FEA, Branco *et al.*[58] created an integrated model that allowed them to simulate not only where particles would hit the blade but also how these impacts would lead to long-term structural damage.

A growing trend in computational simulations is the use of multi-scale modeling, which links microscale erosion events, such as pitting and surface roughness, to macroscale turbine blade performance. Zhu *et al.*[59] demonstrated that micro-scale damage caused by erosion can lead to increased drag, reduced aerodynamic efficiency, and eventually large-scale failure. These models integrate CFD data on particle impacts with FEA data on material behavior to create a comprehensive simulation that spans multiple scales of analysis.

Due to the computational cost of running detailed CFD and FEA simulations, researchers are increasingly using ML to develop surrogate models that approximate these simulations. Shin and Hamed[77] trained NNs on CFD and FEA data to predict erosion patterns and material stress in real time. Surrogate models and ML approach significantly reduces the time and computational power required for full-scale simulations, enabling more frequent and detailed analyses of turbine blade performance under erosive conditions.

The computational modeling and simulation theories section emphasizes the importance of combining fluid dynamics and structural analysis to understand how

gas turbine blades respond to erosive forces.[42,43] This integrated approach helps engineers predict failure points, optimize blade designs, and develop more effective maintenance strategies. In the context of CFD and FEA simulations, fluid dynamics equations govern the airflow and particle trajectories, while structural mechanics equations predict deformation and stress under erosion impacts. CFD is critical in gas turbine analysis because it simulates the complex fluid flow and particle behavior in the high-temperature, high-velocity environment inside turbines.[12,15,27,40] The purpose of CFD in erosion analysis is to understand how the gas flow and solid particles interact with the turbine blades, contributing to surface degradation.

The Navier–Stokes equations describe the motion of fluid (air) around turbine blades, particularly for predicting high-speed airflows and particle paths. The Navier–Stokes equations govern fluid flow, describing the conservation of momentum, mass, and energy in the gas phase.[71] These equations are solved to compute the velocity field, pressure distribution, and temperature gradients in the airflow around the turbine blades. By solving the Navier–Stokes equations, CFD helps predict where the flow separation and vortex formation occur, which are known to concentrate particle impacts. The velocity field around the blades, especially near the leading edges and trailing edges, indicates where the highest particle velocities will lead to severe erosion.[69,73,74] The general form of the Navier–Stokes equations for an incompressible fluid is given by Equation V:

$$\rho\left(\frac{\partial u}{\partial t}+u\cdot\nabla u\right)=-\nabla p+\mu\nabla^2 u+f \tag{V}$$

where ρ is the fluid density; u is the velocity field of the fluid; t is time; p is the pressure; μ is the dynamic viscosity; and f represents external forces (*e.g.*, gravity). In CFD simulations, this equation is solved to predict velocity fields and particle trajectories, which determine where particles will strike the turbine blades and the impact velocities. In our theoretical framework, CFD simulations are critical for identifying erosion-prone regions on turbine blades.[72-74] By solving the Navier–Stokes equations and tracking particle impacts, CFD helps predict surface wear patterns and hotspots where erosion is most likely to occur. This information is then passed to FEA for further structural analysis.

In the Eulerian–Lagrangian framework, the gas phase (airflow) is treated as a continuous medium (Eulerian approach), while particles are tracked individually (Lagrangian approach). The particles carried by the gas flow are modeled using the Eulerian–Lagrangian approach, where the Eulerian frame is used to track the gas flow, and the Lagrangian frame tracks individual particles.[75,76] The particles' velocities and trajectories are governed by the equation of motion given by Equation VI. The equation of motion for a particle under the influence of aerodynamic forces is given by Equation VI:

$$m_p\frac{dv_p}{dt}=F_d+F_g+F_l \tag{VI}$$

where m_p is the mass of the particle; v_p is the velocity of the particle; F_d is the drag force; F_g is the gravitational force; and F_l is the lift force. By solving **Equation VI**, CFD simulations predict where and with what energy particles impact the blade surface. This information is passed to FEA models to analyze the structural response. FEA complements CFD by analyzing how the turbine blades respond to the mechanical and thermal stresses caused by particle impacts.[23,24,31] FEA breaks down the turbine blade into smaller elements, allowing for a detailed examination of stress distributions and material deformations.

FEA is used to predict how the blade structure responds to the stresses induced by particle impacts. FEA solves for the displacement field of the structure under load, using the relationship between stress and strain, as shown by Equation VII, derived from Hooke's Law for linear elastic materials.[39,43,47] The governing equation in FEA for linear elastic materials is derived from Hooke's law and the equilibrium of forces, as shown by Equation VII:

$$\sigma_{ij,j}+f_i=0 \tag{VII}$$

where σ_{ij} is the stress tensor; and f_i is the body force.

FEA divides the blade into a mesh of small elements and solves this equation at each node of the mesh to calculate deformations and stresses under operational loads and particle impacts.[78] Using von Mises stress theory, the failure criterion can be written as Equation VIII:

$$\sigma_v=\sqrt{\frac{1}{2}\left[\left(\sigma_1-\sigma_2\right)^2+\left(\sigma_2-\sigma_3\right)^2+\left(\sigma_3-\sigma_1\right)^2\right]} \tag{VIII}$$

where σ^1, σ^2, and σ^3 are the principal stresses at a given point.

To predict failure, the von Mises stress is used as a criterion for yielding in ductile materials. The material is considered to fail when the von Mises stress exceeds the yield strength of the material. FEA simulations use this to predict when and where blade failure will occur as a result of erosion-induced stresses.[8,9] By coupling CFD and FEA, engineers can predict erosion patterns and stress

distributions, guiding the design of more erosion-resistant blades. This coupling allows for iterative analysis, where structural deformation feedback refines fluid dynamics models, ensuring a more accurate representation of real-world conditions and improving blade performance under harsh environments. Table 11 presents and describes some of the advanced coating systems for turbine blades.

3.3.1. FEA application in erosion-induced fatigue

By applying the impact forces calculated from CFD, FEA simulates how erosion-induced material loss affects the stress distribution across the blade surface. For example, Kedir *et al.*[78] showed that as material is eroded away, the remaining blade material experiences higher localized stresses, leading to plastic deformation and crack formation. FEA models simulate cyclic loading that gas turbine blades undergo during operation. By combining this with the Paris–Erdogan crack growth law, FEA predicts the remaining useful life of the blade by calculating how quickly cracks will propagate from erosion-initiated defects. In high-temperature environments, FEA can also model creep deformation, which occurs alongside erosion. Prashar *et al.*[32] demonstrated that FEA can simulate the creep strain accumulation in areas where erosion has reduced the blade thickness, predicting when and where the blade will fail due to a combination of thermal stresses, erosion, and mechanical fatigue. In our framework, FEA simulations are essential for understanding how erosion-induced damage translates into structural failure. Once CFD has identified erosion-prone areas, FEA can analyze how material loss affects stress distribution and predict the crack initiation and growth due to cyclic loading, fatigue, and creep.

3.3.2. Case example: Predicting erosion patterns using ML surrogate models

High-velocity particle impacts on turbine blades lead to complex erosion patterns that require analysis with extensive CFD and FEA simulations. A single simulation cycle, integrating CFD for airflow and particle trajectory analysis with FEA for structural response, may take several hours or days on high-performance computing systems. To train an ML surrogate model, CFD and FEA simulations are performed on a limited number of turbine blade configurations under varying conditions (*e.g.*, particle velocity, impact angle, and material properties). The outputs, such as erosion rate distributions and stress levels, are recorded for each scenario. An ML algorithm, such as a Gaussian Process Regressor (GPR) or a deep NN (DNN), is trained on the dataset to approximate the relationship between input parameters and outputs. For example, given inputs such as particle velocity (v), angle of impact (θ), and material hardness (H), the model predicts erosion rates (E) and high-stress zones using Equation IX:

$$E = f_{ML}(v, \theta, H) + \epsilon \quad \text{(IX)}$$

where f_{ML} is the surrogate model function approximated by the ML algorithm; and ϵ is the residual error between simulated and predicted values. The trained model is validated using a separate test set to ensure its predictive accuracy. Metrics such as Mean Squared Error or R-squared (R2R^2R2) evaluate the surrogate model's performance. Example validation results include:

- R^2 score: 0.98 (indicating high correlation between predicted and simulated values).
- Computational time reduction: Surrogate models reduced prediction time from 8 h per simulation to <1 s.

The surrogate model is applied to predict erosion patterns for new operational conditions without running full CFD or FEA simulations. For instance:

- Input: Particle velocity = 400 m/s; impact angle = 70°; material hardness = 500 MPa.
- Output: Predicted erosion rate distribution on the blade surface, with high-risk zones highlighted.

Surrogate models eliminate the need for repeated high-fidelity simulations, reducing the computational expense by several orders of magnitude. Engineers can quickly test multiple design modifications (*e.g.*, blade geometries and materials) to identify configurations with minimal erosion

Table 11. Advanced coating systems for turbine blades

Coating system	Composition	Application technique	Advantages	Limitations
TBCs	Ceramic topcoat with MCrAlY bond coat	Plasma spraying or electron beam PVD	High thermal resistance and insulation	Degrades under sustained erosion and spallation
Anti-corrosion coatings	Metal oxide or nitride layers	Thermal spraying or dip coating	Protects against oxidation and chemical degradation	Limited erosion resistance without additional layers
Environmental barrier coatings	Silicon-based ceramics	Slurry or chemical vapor deposition	Protects against oxidation and hot corrosion	High cost and limited thermal fatigue resistance
Wear-resistant coatings	Carbides (e.g., tungsten carbide)	HVOF spraying	High hardness and wear resistance	Susceptible to thermal cracking at high temperatures

Abbreviations: HVOF: High-velocity oxygen fuel; PVD: Physical vapor deposition; TBCs: Thermal barrier coatings.

risk, significantly accelerating the design optimization process. Well-trained surrogate models achieve high accuracy when validated against full CFD-FEA simulations, making them reliable for real-world applications. The use of ML-based surrogate models exemplifies the synergy between advanced computational methods and data-driven approaches. By significantly reducing computational costs and providing rapid predictions, these models empower researchers and engineers to efficiently analyze erosion patterns and optimize turbine blade designs under diverse operating conditions. This integration demonstrates the practical impact of ML in addressing complex engineering challenges.

3.4. Theoretical implications for the study

This theoretical framework outlines a comprehensive understanding of erosion-induced failures in gas turbine blades by integrating theories on erosion mechanics, material degradation, and computational modeling. The framework provides a solid foundation for addressing the following research questions:

(1) R1: How do erosion mechanisms in gas turbines contribute to localized surface damage?
(2) R2: How does erosion-induced fatigue lead to blade failure over time?
(3) R3: How can CFD and FEA simulations predict failure points and improve blade design?

These theoretical pillars not only support the exploration of erosion but also guide the development of practical solutions, including new materials, coatings, and computational tools for predicting and mitigating erosion in gas turbines. This framework forms the basis for future research aimed at enhancing the durability and efficiency of turbine blades, ultimately contributing to more reliable gas turbine systems.

3.4.1. Contribution to localized damage

The erosion mechanisms in gas turbines primarily involve SPE, where hard particles, carried by the high-velocity airflow, impact the turbine blade surfaces.[52,57,61] These particle impacts progressively remove material, causing localized surface damage that varies depending on factors such as particle velocity, size, impact angle, and the material properties of the blade.

Localized erosion is primarily observed on the leading edges, trailing edges, and high-velocity regions of turbine blades. These areas experience more direct impacts from particles, especially where airflow accelerates due to the blade geometry. According to Wang *et al.*,[33] particles that impinge on the surface at steep angles (near 90°) cause the most significant erosion, as the kinetic energy transferred to the blade is maximized. Conversely, lower angles result in surface plowing or smearing, which deforms but does not remove material. This explains why the leading edge of a blade is particularly vulnerable – it experiences frequent, high-angle impacts.

In addition, boundary layer theory helps explain how turbulence near the blade surface contributes to uneven particle impacts. Stokes and Presby[71] found that flow separation near the trailing edge causes particles to rebound erratically, creating localized hotspots where erosion is more severe. In these regions, the turbulent flow combined with high particle velocities leads to repeated impacts, further intensifying the erosion.

Zainuddin *et al.*[52] also noted that the material composition of the blade plays a significant role in the degree of surface damage. Blades made of harder materials or those coated with erosion-resistant coatings, such as TBCs, experience less erosion because these materials can better absorb and dissipate the energy from particle impacts.

Thus, erosion mechanisms lead to non-uniform material removal, predominantly on blade surfaces exposed to high-velocity gas flow and frequent particle impacts. This damage is concentrated on specific parts of the blade, contributing to localized surface wear that degrades the overall performance of the turbine. Over time, this uneven material loss can alter the blade's aerodynamic profile, increasing turbulence and drag while reducing efficiency. In addition, the resulting stress concentrations from erosion-induced surface irregularities accelerate crack formation, further compromising the blade's structural integrity.

3.4.2. Mechanism of fatigue failure

Erosion-induced fatigue occurs as a result of repetitive particle impacts that weaken the blade surface, creating stress concentrators (*e.g.*, pits, scratches, and microcracks), which gradually propagate into larger cracks.[49,52] Over time, these stress concentrators become critical points where fatigue cracks initiate, especially under the cyclic mechanical and thermal loads experienced during gas turbine operation.

The Paris–Erdogan law[70,72,77] provides a basis for understanding how small cracks initiated by erosion grow over time. It establishes a relationship between the stress intensity factor and the rate of crack growth. When particles strike the blade surface, they create microstructural damage that intensifies stress concentrations, accelerating crack propagation under the influence of thermal cycling (due to repeated heating and cooling) and mechanical stress (due to rotational forces).

Ahsan *et al.*[49] examined how erosion-induced pits and cracks become the starting points for fatigue fractures. Their research demonstrated that erosion damages the protective surface layer of the turbine blade, making it more susceptible to oxidation and thermal fatigue. The repeated removal of material leads to localized thinning, which reduces the blade's structural integrity and load-bearing capacity. As cracks propagate, they lead to brittle fracture or ductile tearing, depending on the material's temperature and properties at the time of failure.

In addition, Błachnio *et al.*[36] emphasized the role of creep deformation in conjunction with erosion. At high temperatures, the blades are subjected to creep, a time-dependent deformation. Erosion accelerates this process by removing material and exposing deeper layers to the same high temperatures, leading to faster creep deformation. Over time, the combined effects of erosion, creep, and fatigue result in catastrophic blade failure.

In summary, erosion creates surface defects that evolve into larger cracks under the combined influence of thermal and mechanical stresses, leading to progressive fatigue failure of the turbine blades over time.

3.4.3. Role of CFD in predicting failure points

CFD and FEA are powerful tools used to simulate the interactions between gas flows, solid particles, and turbine blade structures.[4,9,10,11] These simulations allow researchers to model the fluid dynamics of gas turbines and the structural response of blades to impacts, predicting failure points and optimizing blade design for better erosion resistance.

CFD simulations are used to model airflow patterns and particle trajectories within GTEs, providing detailed insights into how particles behave in high-velocity gas streams. Bera *et al.*[39] employed CFD models based on Navier–Stokes equations to simulate the velocity fields and turbulence levels around turbine blades. They identified erosion-prone zones, such as the leading edge and trailing edge, where particle impact is most severe. CFD tracks particles through the gas flow, determining their trajectories, velocities, and impact angles with the blade surface. This information is essential for predicting erosion rates at different locations on the blades. Sun *et al.*[62] also showed that particles striking the leading edge at high velocity and near-normal angles cause significant material loss due to high-energy impacts. These simulations also reveal flow separation and vortex formation, which contribute to localized erosion.

By incorporating Eulerian–Lagrangian models, CFD allows for the simulation of solid particle impacts on the blade surface, predicting where the highest erosion rates will occur. Kishore *et al.*[69] used CFD to simulate particle collisions at various angles and velocities, identifying high-risk areas where erosion would likely lead to cracks and structural damage.

3.4.4. Role of FEA in structural analysis

While CFD focuses on fluid dynamics, FEA is used to simulate how the blade material reacts to repeated mechanical and thermal stresses caused by particle impacts. Khan and Sasikumar[60] applied FEA models to study how erosion-induced material loss affects stress distribution across the blade. Using von Mises stress theory, they predicted the failure points where cracks would initiate due to stress concentration caused by erosion pits.

Moreover, FEA can simulate the fatigue life of a blade by predicting how cracks grow over time under cyclic loads. Sabri *et al.*[24] combined CFD and FEA to create a fluid-structure interaction model, allowing them to simulate both the flow-induced erosion and the structural deformation of the blades. Their integrated model provided a detailed view of how erosion progresses over time, leading to blade failure.

3.4.5. Improving blade design

These simulations are invaluable for optimizing blade geometry and selecting materials to improve erosion resistance. Hoksbergen *et al.*[38] demonstrated how CFD and FEA simulations can be used to test different blade designs, identifying configurations that reduce turbulence and particle impact. For example, redesigning the leading edge of a blade to minimize flow separation or applying erosion-resistant coatings in high-impact areas can extend the operational life of turbine blades. Surrogate models and ML approaches, such as those explored by Yang *et al.*,[35] allow engineers to run simulations more efficiently by approximating CFD and FEA results. This enables real-time predictions of erosion patterns and helps refine blade designs without the need for time-consuming full-scale simulations.

The leading edge of a turbine blade is one of the most erosion-prone areas due to its exposure to high-velocity particle impacts. Modifying its geometry can significantly reduce the intensity of these impacts and improve the airflow over the blade. A rounded leading edge is particularly effective because it minimizes stagnation pressure, reducing the energy transferred from particles to the blade surface on impact. This smoother curvature also facilitates a more gradual transition of airflow over the blade, decreasing the velocity gradient and reducing turbulence intensity. Another common design approach is the elliptical leading edge, which optimizes aerodynamic

performance by creating a laminar boundary layer that helps reduce particle deposition and subsequent erosion. The elliptical shape provides a balance between lowering drag and maintaining airflow smoothness.

The trailing edge, another critical zone, is prone to localized flow separation caused by the abrupt deceleration of airflow. Streamlining the trailing edge can mitigate this issue. A tapered trailing edge design ensures a smoother exit flow by reducing the abrupt changes in velocity that cause vortices to form. These vortices are often laden with particles that can repeatedly collide with the blade surface, exacerbating erosion. Another innovative approach involves the use of chevron patterns or serrations at the trailing edge. These features disrupt large vortices into smaller, less erosive structures, improving pressure recovery, and minimizing turbulence. The result is reduced particle impingement in this region, extending the blade's operational life.

Surface contouring can also play a key role in managing boundary layer behavior and controlling the interaction between fluid flow and solid particles. For example, a concave adjustment along the blade's midsection can stabilize the boundary layer, reducing turbulence and particle collisions. This stabilization encourages smoother streamlines, preventing the reattachment of turbulent flows that can cause concentrated erosion. Similarly, micro-grooved surfaces – small, precisely engineered grooves along the blade – help channel airflow in a controlled manner. These grooves not only reduce drag but also divert particles away from high-stress zones, further minimizing erosion.

Other geometrical changes, such as twisting the blade profile or adjusting the lean of the blade tips, help optimize particle trajectories and airflow distribution. A twisted blade profile ensures that the flow remains more uniform along the blade's length, preventing the formation of localized zones of high velocity where particles can concentrate. This even distribution of flow reduces the severity of erosion by spreading it more evenly across the surface. Similarly, backward-leaning blade tips are designed to align the flow more effectively with the blade geometry, reducing the formation of eddies at the tips, which are known to cause erosion-prone turbulence.

CFD simulations have been instrumental in validating the benefits of these geometrical modifications. For instance, studies have shown that a smoother leading edge can reduce turbulence intensity by up to 20%, significantly decreasing the velocity of particle impacts. Similarly, serrated trailing edges have been found to reduce wake turbulence by 15%, which helps maintain the integrity of erosion-resistant coatings. Micro-grooved surfaces have been shown to reduce drag coefficients by 10 – 12%, indirectly lowering erosion rates by altering particle trajectories.[43,49,51]

In summary, geometrical modifications to turbine blades, including adjustments to the leading and trailing edges, surface contouring, and blade profile optimization, are effective strategies for mitigating erosion. These changes not only enhance the blade's resistance to erosion but also improve aerodynamic performance by optimizing fluid flow. By implementing such modifications, engineers can design more durable blades, ultimately extending their operational lifespan and reducing maintenance costs. CFD and FEA simulations are critical for identifying erosion-prone areas and predicting structural failures. These tools help engineers design more durable turbine blades by optimizing geometries and selecting materials that can better withstand erosive forces.

3.4.6. ML integration

ML plays a critical role in enhancing the prediction of erosion rates, optimizing material selection, and enabling proactive maintenance strategies in gas turbine systems. This section elaborates on the methodology, mathematical frameworks, and applications of ML in the study.

The first step in ML integration involves preparing a robust dataset. Input parameters such as particle velocity (v), impact angle (θ), particle size (d), material hardness (H), and operational temperature (T) are identified as key features influencing erosion. These features are used to predict output labels, which represent erosion severity metrics such as material loss per unit area (E) or erosion rate. The dataset is generated from a combination of CFD and FEA simulations, historical operational data, and experimental results. Splitting the dataset into training, validation, and testing sets ensures the model's reliability and generalizability. The study employs supervised learning algorithms to predict erosion rates. Models such as SVMs or RF classify erosion severity (E_{low}, $E_{moderate}$, E_{high}) or predict continuous values of erosion rate (E). For instance, regression models map relationships between input features and erosion rate using Equation X:

$$E = f(v, \theta, d, H, T) + \epsilon \quad \text{(X)}$$

where f is the function approximated by the ML model; and ϵ is the error term. In addition, feed forward NNs with activation functions like ReLU ($g(x) = \max(0,x)$) are used to capture non-linear relationships between features and erosion severity. For spatially dependent data such as particle impact distribution maps, CNNs are utilized. CNNs analyze input images or grid data to predict erosion-prone zones. The loss function for CNN training, typically the Mean Squared Error, is calculated using Equation XI:

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(E_{true}^{(i)} - E_{pred}^{(i)}\right)^2 \quad \text{(XI)}$$

where E_{true} and E_{pred} represent the actual and predicted erosion values, respectively. CNNs efficiently process complex geometries and predict erosion distribution with high accuracy. The ML models are trained by minimizing the loss function $L(\theta)$ using optimization techniques such as stochastic gradient descent. The training data are used to adjust the model parameters, while the validation set monitors overfitting. After training, the test set evaluates the model's performance, ensuring its accuracy in predicting erosion rates under unseen conditions. The trained ML models enable real-time prediction of erosion rates and severity. These predictions help to identify high-risk zones on turbine blades, allowing engineers to develop targeted interventions. For example, ML can predict areas requiring enhanced coatings or geometric optimization to minimize erosion effects. In addition to erosion prediction, decision-tree-based models and reinforcement learning are applied to optimize material selection. These models analyze material properties such as hardness (H), tensile strength (TS), and thermal resistance (T_{max}) to identify materials with the highest erosion resistance. By narrowing down material options through ML, researchers can focus on testing and deploying the most promising candidates.

This structured integration of ML provides an efficient, data-driven approach to understanding and mitigating erosion in gas turbine blades, paving the way for enhanced durability and performance.

3.5. Coupling CFD and FEA for fluid-structure interaction

The most powerful approach in gas turbine analysis is the coupling of CFD and FEA to simulate fluid-structure interaction. This combined method allows engineers to model both the fluid flow and particle dynamics and the structural response of the blade.[12,16,18] This coupled CFD-FEA approach offers a comprehensive way to predict erosion and failure points, optimizing blade geometry and materials. The following are the key steps in CFD-FEA coupling:

- CFD simulation: The Navier–Stokes equations are solved to obtain the airflow and particle impact forces on the blade surface. Solve the Navier-Stokes equations to predict airflow around the turbine blade and use the Eulerian–Lagrangian approach to track particle trajectories.
- Transfer of impact forces: The particle impact forces from CFD are transferred to the FEA model as external loads. Determine particle velocities, angles, and impact locations.
- FEA simulation: The FEA model calculates the stress distribution, deformation, and failure points based on the applied particle forces and material properties. Use the impact forces from CFD to simulate the structural response of the blade, solving for stresses and deformations.
- Iterative coupling: In more advanced simulations, the deformed geometry from FEA is fed back into CFD to update the flow field and particle trajectories, creating an iterative loop that provides a more accurate representation of erosion progression and structural failure. Apply von Mises stress criteria and fatigue models (Paris Law) to predict crack growth and blade failure.

Molinari and Ortiz[44] applied multi-scale models to simulate how micro-scale erosion affects the macro-scale performance of gas turbine blades. By combining CFD and FEA, they were able to track how surface roughness from erosion influences the fluid dynamics (*e.g.*, increased drag) and how this further exacerbates structural deformation. Their approach also captured the cascading effects of erosion, where small-scale material removal leads to altered flow patterns, increasing turbulence and intensifying localized wear. By coupling CFD and FEA, the theoretical framework provides a comprehensive approach to predicting erosion damage and structural failure. This integration allows for a detailed analysis of the erosion-fatigue interaction and the long-term degradation of turbine blades, offering actionable insights for improving blade design, material selection, and maintenance schedules. Furthermore, the study highlighted how localized stress concentrations from erosion propagate cracks, emphasizing the need for more robust materials and coatings in high-impact zones. The findings also showed that modifications to blade geometry, such as smoothing sharp edges or optimizing surface profiles, can significantly reduce stress amplification and delay failure. These findings underline the critical importance of coupling high-fidelity simulations with experimental validation to ensure practical applicability in real-world turbine operations. Moreover, this multi-scale approach bridges the gap between theoretical models and industrial applications, enabling engineers to proactively address failure risks. The study further demonstrated that integrating these models into predictive maintenance systems can enhance operational reliability by identifying critical wear zones before catastrophic failures occur. The entire CFD-FEA for fluid-structure interaction is depicted using a flow diagram (Figure 1).

3.6. Surrogate models and ML for real-time prediction

As CFD and FEA simulations can be computationally expensive, ML is increasingly used to create surrogate

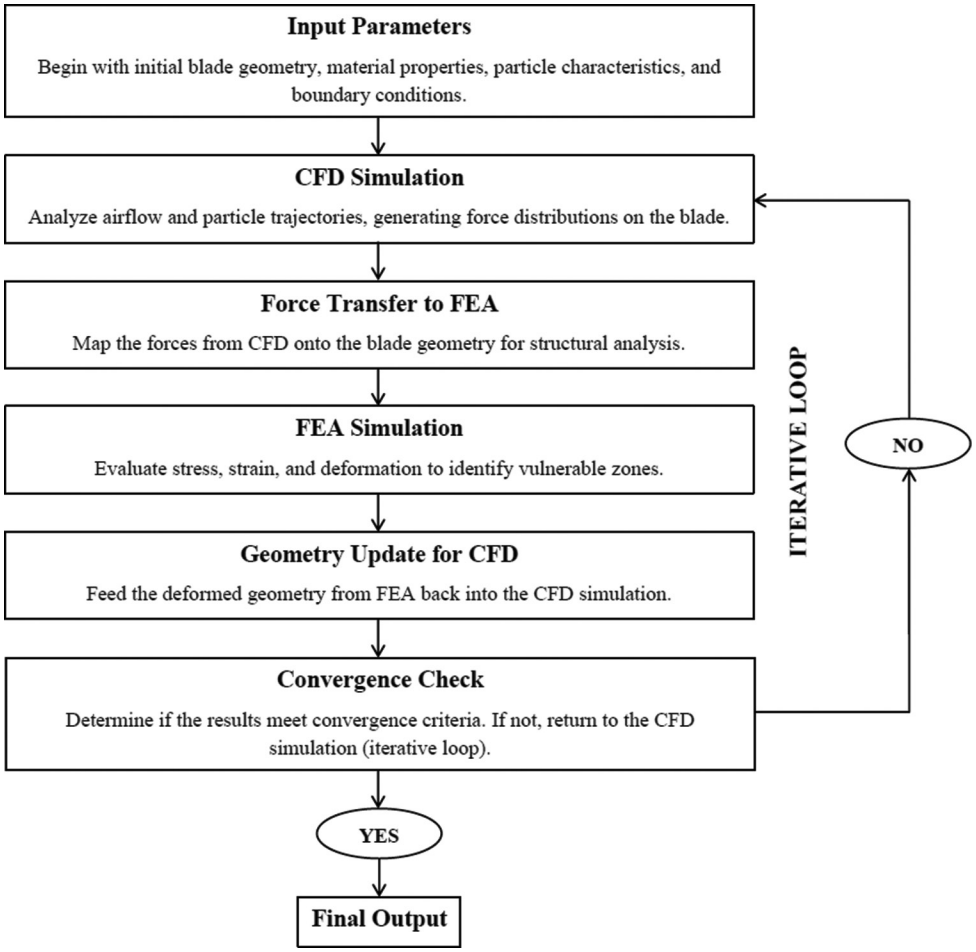


Figure 1. Flow diagram of the entire CFD-FEA process
Abbreviations: CFD: Computational fluid dynamics; FEA: Finite element analysis.

models that approximate the results of full-scale simulations. These models significantly reduce computational time while maintaining accuracy.[33,34,37] By training NNs on large datasets generated from CFD and FEA simulations, surrogate models can predict erosion patterns and structural responses in near real-time. Poursaeidi *et al.*[61] used CNNs to predict particle impact locations and erosion rates based on input parameters such as flow velocity, particle size, and blade geometry. These models can predict erosion hot spots without requiring the full CFD solution, drastically speeding up the design optimization process. In our framework, ML-based surrogate models can be used to complement CFD and FEA simulations, providing a fast and efficient tool for predicting erosion damage and structural degradation in real time.[55,56] This capability is critical for adaptive design and real-time monitoring of turbine blades, improving the overall reliability of gas turbine systems.

The Paris Law is a widely used model in material fatigue theory to describe the rate at which a crack grows under cyclic loading. In the context of gas turbine blades, repeated particle impacts create localized stress concentrators, such as microcracks and pits, which propagate over time due to the cyclic nature of operational stresses. This crack growth eventually compromises the structural integrity of the blade, leading to failure. The Paris Law provides a mathematical framework to predict this progressive damage as presented in Equation III.

In turbine blades, the repeated impacts from high-velocity particles act as cyclic loads that initiate and propagate cracks. These impacts create small pits and scratches on the blade surface, which serve as nucleation sites for cracks. As the turbine operates, thermal and mechanical stresses combine with these impacts to cyclically load the blade material, causing the cracks to grow incrementally with each cycle. The Paris Law allows engineers to estimate the rate of crack growth by correlating it with the stress intensity range (ΔK), which depends on the material properties and loading conditions.

By integrating the equation over the number of cycles, the total crack growth can be predicted using Equation XII:

$$a_f = a_i + \int_{N_i}^{N_f} \frac{C.(\Delta K)^m}{dN} \quad \text{(XII)}$$

where a_i is the initial crack length; a_f is the final crack length before failure; and N_i and N_f are the initial and final number of cycles, respectively. Using the Paris Law, engineers can achieve the following:

- Identification of high-risk zones: Crack growth predictions help pinpoint regions on the blade most susceptible to failure, such as leading edges and trailing edges.
- Material selection: Materials with lower values of C and m exhibit slower crack growth rates, making them more resistant to fatigue-induced failure.
- Proactive maintenance: The model enables predictions of the remaining useful life of a blade, allowing for timely maintenance or replacement before catastrophic failure occurs.
- Design improvements: The insights gained from crack growth analysis inform design modifications, such as adding erosion-resistant coatings or altering blade geometry to reduce stress concentrations.

The Paris Law bridges the gap between material fatigue theory and practical applications in turbine blade design. Material fatigue occurs due to the progressive weakening of a material under cyclic loading, which is analogous to the repeated particle impacts experienced by turbine blades. By quantifying the relationship between cyclic stresses and crack growth rates, the Paris Law highlights the importance of understanding and mitigating fatigue effects to enhance the durability and safety of turbine components. The inclusion of the Paris Law in Section 3.6 underscores its critical role in predicting and mitigating blade failure caused by repeated particle impacts. By providing a quantitative framework for crack growth analysis, the Paris Law enables engineers to develop fatigue-resistant designs, select appropriate materials, and implement proactive maintenance strategies, ultimately extending the operational life of turbine blades.

4. Experimental studies

Experimental investigations play a crucial role in advancing our understanding of erosion-induced failures in gas turbine blades. While computational models and theoretical frameworks provide valuable insights into the mechanisms of erosion, real-world testing is essential to validate these models, assess material performance, and identify key factors influencing erosion in operational environments.[25,29,36] In the context of gas turbines, experimental studies typically focus on simulating high-temperature, high-velocity gas flows, replicating particle impacts, and measuring the subsequent material degradation. These experiments offer empirical data on erosion rates, surface damage, and structural integrity, which help refine computational predictions and guide the development of more erosion-resistant materials and protective coatings.[41,58,78] This section explores the significant contributions of experimental research, detailing methodologies, findings, and their implications for turbine blade durability and performance.

Hamed *et al.*[55] conducted an extensive computational research initiative, which was dedicated to the comprehensive analysis of surface degradation within turbine vanes and blades, specifically attributable to the impingement of solid particulate matter. Leveraging a computational framework, Hamed *et al.*[55] were able to predict, based on meticulously computed particle trajectories, the emergence of a localized region characterized by heightened erosion levels along the vane's leading edge. Furthermore, the research findings elucidated a discernible trend wherein the erosion intensity progressively intensified along the pressure surface, culminating in a notable elevation of erosion effects as the trajectory approached the trailing edge. This insightful investigation[55] significantly contributes to the nuanced understanding of the intricate erosion mechanisms prevalent within high-velocity and solid particle-laden environments, thereby providing valuable insights crucial for the development of robust protective measures and advanced material solutions for turbine components.

Taherkhani *et al.*[56] conducted an extensive investigation employing the finite element method to elucidate the influence of both particle velocity and particle diameter on the erosion rate experienced by a smooth surface. The research findings illustrated by Taherkhani *et al.*[56] clearly demonstrate the substantial impact of these parameters on the erosion phenomenon. Notably, Taherkhani *et al.*'s[56] study encompassed a comprehensive exploration of particle velocities spanning the range of 100 – 500 m/s, alongside particle diameters ranging from 1 to 5 mm. Through the implementation of a numerical approach, a direct correlation between particle velocity, diameter, and the resultant erosion rate was established. Furthermore, the researchers delved into the intricate dynamics associated with suspended solid particles within gas turbine flows, offering critical insights into the consequential material degradation of blade surfaces induced by the impingement and interaction of these particles. Their research serves as a pivotal contribution to the nuanced understanding of erosion dynamics within complex high-velocity environments, facilitating the development of tailored

mitigation strategies and advanced protective measures for enhancing the resilience and durability of turbine components. Their work also emphasizes the importance of accounting for particle size, velocity, and flow characteristics, providing a comprehensive framework for optimizing material selection and coating techniques. By integrating these factors into predictive models, engineers can anticipate erosion patterns more accurately and implement targeted design modifications to improve the longevity of turbine blades under extreme operating conditions.

Branco *et al.*[58] investigated the effects of SPE on zirconia and alumina-based ceramic coatings under standard room temperature conditions. The experiment involved subjecting the coatings to a stream of alumina particles, each averaging 50 μm in size, at a velocity of 70 m/s and an impingement angle of 90°. His research, as presented in Figure 2, revealed a significant correlation between the erosion rate and the porosity of the coatings. The researchers further deduced that the porosity influenced the erosion rate in multiple ways. First, the presence of porosity resulted in a reduction in the material's resistance to plastic deformation or chipping, especially at the void's extremity where mechanical support was lacking. Second, the concave surface within a void, not shielded by any surrounding void edge, was particularly susceptible to impacts from particles at angles exceeding the average surface-to-impact angle.[75] Finally, the existence of pores weakened the overall strength of the material, acting as stress concentrators and contributing to the increased erosion rates observed.

Zhu *et al.*[58] outlined the process behind the creation of TBC for turbine blades, taking into account the impact of erosion during operations at elevated temperatures within rotorcraft propulsion systems, as shown in Figure 3. Through comprehensive experiments simulating engine erosion and thermal gradient conditions, they verified the efficacy of these advanced TBC systems, demonstrating notable improvements in performance, as depicted in Figure 4. Moreover, a predictive modeling technique for estimating the lifespan of these turbine blade coatings was introduced, emphasizing the crucial aspect of mitigating erosion effects in high-temperature environments.

Shin and Hamed[77] performed comprehensive experiments to characterize the erosion resistance of plasma-sprayed 7 wt% YSZ TBCs applied on MCrAlY bond coats and IN718 substrates, typically utilized in GTE blade materials. Utilizing 26 μm particles, they conducted experimental runs at varying gas temperatures and velocities, considering different impingement angles based on the Taguchi method. This approach was adopted to develop an effective experimental design aimed at identifying the key test parameters influencing TBC erosion rates. Moreover, through the application of Taguchi's L9 orthogonal array and the analysis of variance, the researchers investigated the impact of these parameters on the erosion rate of the TBCs. They observed that increased gas temperature, gas velocity, and impingement angles, as guided by the Taguchi design of experiments (DOE), led to elevated erosion rates. Notably, the study revealed that 77.07% of the erosion was attributed to particle impact velocity, 14.15% to gas temperature, and 3.56% to impingement angles. By analyzing the SEM surface images, the researchers detected the presence of unmelted and partially melted ceramic powders in the pre-tested samples, which were received from the devices. However, these were notably absent in the post-tested samples, indicating the transformation and alterations induced by the erosion process.[72,77]

Shin and Hamed[57] conducted an experimental investigation to analyze the impact of robust port microstructures on the erosion resistance of air plasma-sprayed 7 wt % YSZ TBCs at elevated temperatures. Two sets of YSZ TBCs with distinct microstructures,

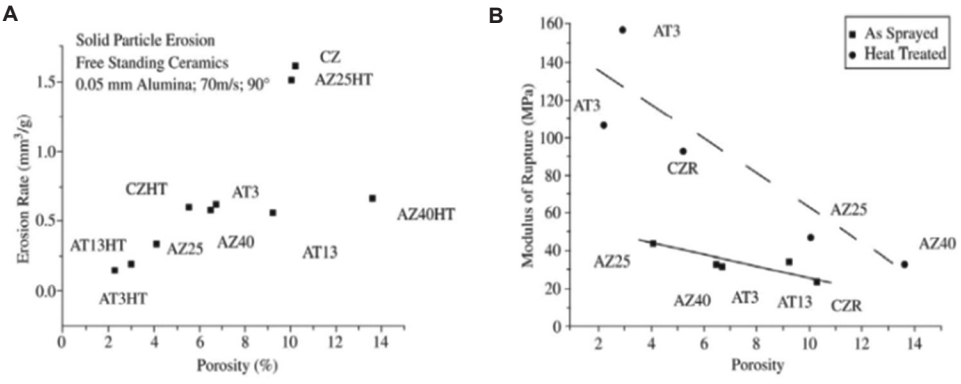


Figure 2. Erosion rate and fracture strength were analyzed in relation to porosity in various alumina and titania-zirconia based materials, both in their initial form and after heat treatment. (A) Porosity versus erosion rate. (B) Porosity versus modulus of rupture. Source: Branco *et al.*[58]

characterized by porosities of 12.9 ± 0.5% and 19.5 ± 1.2%, were subjected to rigorous testing in a high-temperature erosion tunnel designed to simulate the operational conditions of contemporary GTEs. The experimental trials encompassed a temperature range spanning from 537°C to 980°C, with gas velocities ranging between 122 m/s and 305 m/s, and impingement angles spanning 20° – 90°. During the testing process, two types of powders were utilized, nominal 27 µm white fused alumina and 15 µm A3, commonly known as test dust. The study concluded that erosion rates increased in correspondence with gas velocities across all three impingement angles, *that is*, 20°, 60°, and 90°. Furthermore, it was established that the experimental variations adhered to an empirical relation, emphasizing the importance of the relationship between the erosion rates and the tested parameters.

Kedir *et al.*[78] conducted comprehensive tests on various ceramic matrix composites (CMCs), employing different erodent particles at varying velocities ranging between 200 m/s to 300 m/s under ambient temperature

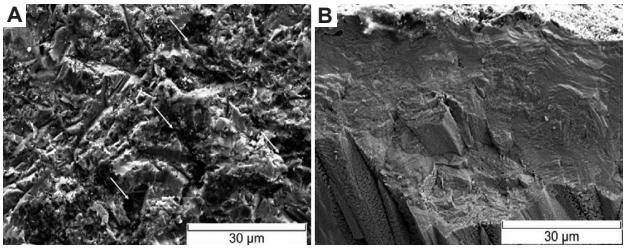


Figure 3. SEM images showing surface damage morphologies of durable, low conductivity thermal barrier coatings during erosion and impact tests. (A) Limited plasticity erosion surface with small fatigue spalling areas (indicated by arrows). (B) Extensive plasticity impact surface with widespread spalling. Source: Zhu *et al.*[59]
Abbreviation: SEM: Scanning electron microscopy.

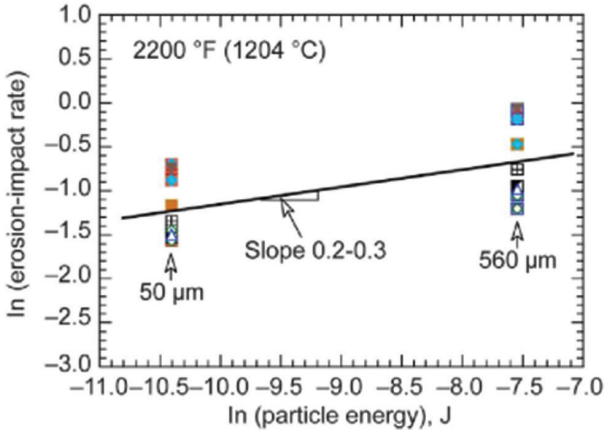


Figure 4. The correlation between erosion and impact rates was examined for 50 and 560 µm erodent particles in specific coating systems under testing at 2200 F (1204°C). Source: Zhu *et al.*[59]

conditions. The study encompassed nine distinct factors thoroughly examined and linked to their inclusive abrasion performance and the root cause failure of turbine blades. Kedir *et al.*[78] determined the minimal solidity using a mass-volume system, while the matrix hardness was assessed through Vicker's hardness of the matrix-rich region. Furthermore, the research emphasizes the necessity of a systematic understanding of both material properties and the operational environment. This approach is crucial in the development and assembling of abrasion-resilient CMC engine materials and machineries, ultimately contributing to the improved reliability and strength of CMC hardware. It also highlights the need for rigorous testing protocols to validate material performance under realistic operating conditions, ensuring long-term durability and efficiency.

Rajabinezhad *et al.*[76] examined the primary factor contributing to the failure of gas turbine blades composed of the Nimonic-105 nickel-based superalloy. Their investigation focused on two specific blades, one fractured at the root and the other at the airfoil, delving into the intricate details to identify the underlying causes of these failures. Images displaying the fracture surface of the root-failed blade, captured using a stereomicroscope and SEM, are depicted in Figure 5. They meticulously analyzed the material properties and structural integrity of the blades to gain a comprehensive understanding of the failure mechanisms at play. Figure 6 illustrates the fractured surface of a gas turbine blade, which has become detached from the aerofoil. The aerofoil is the primary section of the blade responsible for directing airflow and converting thermal energy into mechanical power. In this image, the blade's root or attachment point to the turbine hub has failed, which typically indicates a significant mechanical stress or fatigue fracture.[76] The fracture surface provides critical insights into the failure mechanism, including whether the failure was driven by fatigue, corrosion, or

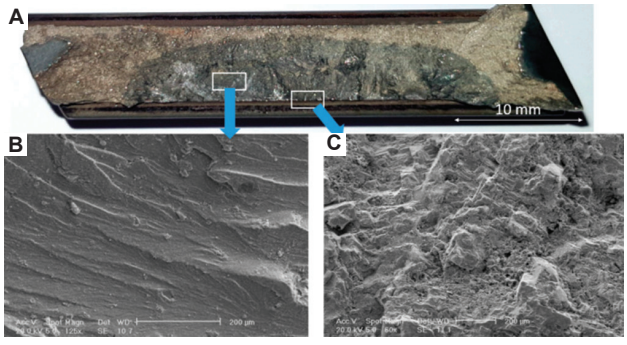


Figure 5. Micrographs of the fractured surface of the root-failed blade. (A) Stereomicroscope image. (B and C) SEM micrographs. Source: Rajabinezhad *et al.*[76]
Abbreviation: SEM: Scanning electron microscopy.

erosion-induced material weakening. The detachment of the blade from the aerofoil is a serious issue that compromises the turbine's functionality and can lead to catastrophic failure if undetected during operation. Figure 7 presents SEM images of the fractured blade surface after it has been separated from the aerofoil. SEM is used to capture highly detailed, high-resolution images of the fracture surface, allowing for an in-depth analysis of the microstructural features of the failure. These images can reveal microcracks, grain boundary separations, and pitting that might not be visible to the naked eye. By analyzing the fracture morphology under SEM, researchers can better understand the underlying causes of failure, such as erosion-induced fatigue or corrosion damage.[76] The SEM images provide critical evidence of how erosion and thermal fatigue contribute to the degradation and eventual failure of gas turbine blades over time.

Their findings revealed the presence of a characteristic dendritic structure, with dendrites growing from the edges toward the center. No surface irregularities, buckling, or deformations were identified, except for a minor dent observed on the edge of one of the

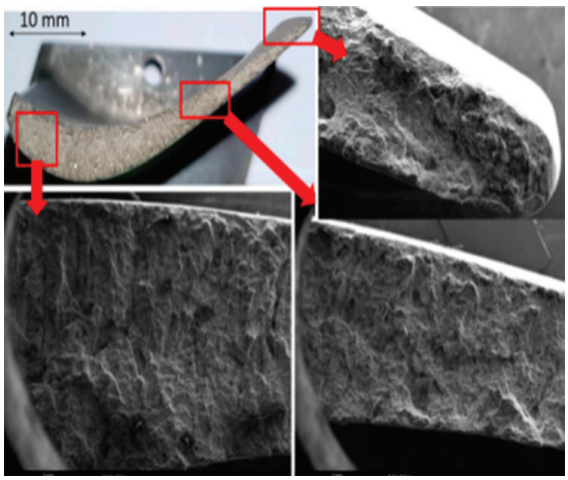


Figure 6. Fractured surface of a blade detached from the aerofoil. Source: Rajabinezhad *et al.*[76]

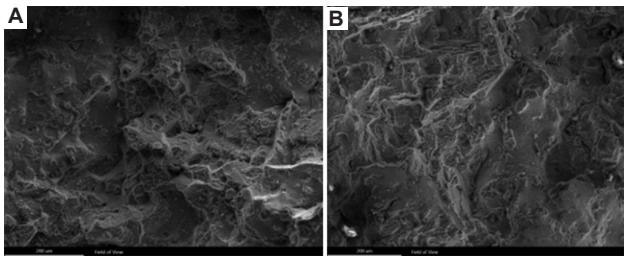


Figure 7. (A and B) SEM images of the fractured blade surface separated from the aerofoil. Source: Rajabinezhad *et al.*[76]
Abbreviation: SEM: Scanning electron microscopy.

fractured blades. Corrosion and erosion emerged as the predominant failure mechanisms, as deduced from their comprehensive analysis. Through meticulous investigation, they concluded that the initial failure occurred at the root, induced by corrosion-fatigue, followed by subsequent breakage. This sequence of events led them to assert that the corrosion-fatigue mechanism played a pivotal role in the eventual failure of the blades.[76] They also emphasized the need for further research into the specific environmental and operational conditions that may exacerbate the corrosion and erosion processes in similar turbine blades. Table 12 presents a detailed comparative analysis of erosion studies by Hamed *et al.*,[55] Taherkhani *et al.*,[56] and Branco *et al.*[58] with erosion pattern comparisons and material properties.

Table 12 presents a detailed comparison of the erosion patterns observed in the studies by Hamed *et al.*,[55] Taherkhani *et al.*,[56] and Branco *et al.*[58] as well as insights into material properties and coatings. It highlights the influence of key factors such as particle size, velocity, material hardness, and coating porosity on erosion behavior, with specific findings from each study. The comparisons emphasize that harder materials and lower porosity coatings consistently offer superior protection, while high particle velocities and larger particles lead to more significant erosion damage. These insights can guide engineers in selecting materials and coatings to improve turbine blade durability.

4.1. Practical erosion mitigation strategies

This section explicitly connects the experimental insights on factors such as porosity, particle impact angles, and material resilience with practical mitigation strategies. By doing so, this section provides clear, actionable recommendations for engineers to improve turbine blade durability through specific coating techniques, material selection, and design modifications.

4.1.1. Porosity in coatings

Porosity is a critical factor influencing the performance of coatings in resisting erosion. Coatings with high porosity allow for easier penetration of particles, which accelerates material wear and degradation. Experimental findings indicate that coatings with lower porosity exhibit better resistance to particle impacts and can extend the lifespan of turbine blades under erosive conditions. Engineers can choose plasma-sprayed coatings or vacuum deposition techniques, which typically yield coatings with reduced porosity. Specifically, YSZ and ceramic coatings with fine microstructures provide excellent resistance to erosion, especially when applied in multi-layer configurations. Surface densification treatments, such as hot isostatic

Table 12. Comparative analysis of erosion studies by different researchers

Study	Focus area	Methodology	Key findings	Erosion pattern comparisons	Material properties and coatings
Hamed *et al.*[55]	Particle velocity and impact angle	CFD simulations for particle trajectories, combined with FEA for structural analysis	High particle velocities result in severe erosion, particularly at high impact angles (60° – 90°).	Erosion is most severe at high impact angles (60° – 90°), with significant material loss observed at the leading edges.	Focus on material hardness and toughness for resisting erosion at high velocities. No explicit coating discussion.
Taherkhani *et al.*[56]	Particle size and material hardness	High-velocity impact tests and finite element modeling for stress analysis on different materials	Larger particles cause deeper pits; harder materials exhibit lower erosion rates. Particle size is dominant.	Larger particles (1 – 5 mm) result in deeper pits and more severe erosion compared to smaller particles.	Harder materials, like tungsten carbide, show significantly reduced erosion rates, highlighting the importance of material hardness.
Branco *et al.*[58]	Coating porosity, temperature, and erosion	Erosion tests on coated versus uncoated materials at varying temperatures and particle velocities	Low porosity coatings offer better protection; elevated temperatures accelerate material degradation.	Coatings with lower porosity reduce material wear; high temperatures exacerbate material degradation in coatings.	Coatings with lower porosity, such as ceramic coatings, provide better protection against particle impact and temperature.
Comparative trends	Erosion patterns based on particle size, velocity, and material properties	Combination of high-velocity tests, CFD simulations, and coating evaluations	All studies agree on the influence of particle velocity, size, and material properties on erosion severity.	Larger particles and higher velocities result in increased erosion depth; smaller particles cause smoother erosion.	Material hardness and coating porosity play a critical role in reducing erosion; coatings significantly improve erosion resistance.
Material properties	Impact of hardness and surface treatments	All studies assess material hardness and resilience against erosion, with emphasis on coatings.	Harder materials, like tungsten carbide, and coatings, like ceramics, are more erosion-resistant.	Harder materials and coatings resist deeper erosion, while softer materials show increased wear and pit formation.	Hardness plays a key role in resistance to erosion; ceramic and tungsten carbide coatings provide optimal protection in high-velocity conditions.
Coatings	Performance of various coatings under erosive conditions	Focus on coating resilience, porosity, and temperature effects	Ceramic coatings and metallic bond coats provide effective protection, though coating porosity affects performance.	Porosity is critical in determining coating effectiveness; non-porous coatings perform better under high-velocity impacts.	Ceramic coatings (e.g., YSZ) and MCrAlY bond coatings provide superior performance, but porosity must be controlled to prevent wear.

Abbreviations: CFD: Computational fluid dynamics; FEA: Finite element analysis; YSZ: Yttria-stabilized zirconia.

pressing (HIP), can further reduce porosity and improve coating adherence, enhancing both mechanical and thermal resistance.

4.1.2. Particle impact angles

The angle at which particles collide with turbine blades plays a significant role in the rate and severity of erosion. Steep impact angles (close to 90 degrees) result in higher erosion rates due to the direct transfer of kinetic energy to the blade surface. Conversely, shallow impact angles reduce the intensity of the impacts, resulting in less material removal. To minimize the effect of particle impacts at steep angles, blade geometry modifications such as rounded or elliptical leading edges can help reduce the severity of impacts. In addition, turbine blade orientation and flow optimization can be engineered to redirect the trajectory of particles to hit the blade at more favorable angles. A tapered blade design can further reduce the frequency of high-impact angles, thereby mitigating erosion.

4.1.3. Material resilience and hardness

Material resilience, including properties such as hardness, tensile strength, and fatigue resistance, directly influences a material's ability to withstand repeated particle impacts. Experimental studies have shown that materials with higher hardness resist wear better, as they are less susceptible to the formation of pits and cracks.[48,49] However, materials that are too brittle may fracture under high-impact loads, which is counterproductive. Engineers can select superalloys such as INCONEL or Waspaloy, which combine high hardness with good toughness to resist both erosion and thermal cycling. In addition, incorporating composite materials such as tungsten carbide coatings or ceramic-metal hybrids can offer superior resistance to high-velocity impacts, particularly when particles are hard. These materials are often paired with bond coats to enhance adhesion and resistance to thermal stresses. Moreover, gradient coatings that transition from hard to more ductile layers can help absorb impact energy while maintaining surface integrity.

4.1.4. Thermal and mechanical fatigue resistance

Thermal cycling and mechanical loading contribute to material fatigue, especially in areas with high particle impingement and elevated temperatures. Oxidation and creep damage further exacerbate erosion by degrading the material's strength and resistance to wear. Experimental data suggest that coatings and materials that provide both thermal barrier properties and high resistance to oxidation can significantly reduce erosion-induced damage. The application of TBCs, such as YSZ, offers excellent protection against high-temperature oxidation and thermal fatigue. These coatings not only insulate the underlying material from extreme heat but also improve overall blade longevity. MCrAlY bond coatings, typically used in combination with TBCs, enhance adhesion and resistance to oxidation. For applications subject to extreme temperature variations, selecting CMCs can provide superior resistance to both erosion and high-temperature fatigue.

4.1.5. Multi-layer coating systems

Multi-layer coatings, where a bond coat is applied before the top ceramic or metallic coating, have proven effective in enhancing erosion resistance. The bond coat serves as a buffer to accommodate thermal expansion differences between the base material and the top coating, preventing spallation under thermal cycling or particle impacts. A multi-layer approach using materials such as MCrAlY (bond coat) followed by ceramic or metallic topcoats (*e.g.*, TBCs or tungsten carbide) significantly improves blade durability. This system not only enhances thermal protection but also resists the mechanical wear caused by high-velocity particles. Engineers should also consider coating thickness: thicker coatings provide greater protection, but excessive thickness can introduce stresses that may cause premature failure. Optimal thickness must be balanced based on operational conditions.

By directly linking experimental findings to practical mitigation strategies, engineers can make informed decisions on materials, coatings, and design modifications to improve turbine blade durability. Through the careful selection of coating techniques, material properties, and geometrical adjustments, engineers can reduce erosion, enhance material resilience, and extend the operational lifespan of gas turbines under harsh conditions. These strategies help bridge the gap between theoretical erosion studies and real-world turbine performance optimization. Table 13 offers a comprehensive overview of erosion effects under varying conditions and materials.

5. Results and discussion

Erosion-induced gas turbine blade failures have been extensively studied and have been found to significantly impact the efficiency and lifespan of gas turbines. The potential causes and mechanisms of erosion can involve various factors, including the impingement of solid particles, high-velocity fluid flows, and abrasive wear. Researchers have often emphasized the crucial role of material selection, surface coatings, and cooling techniques in mitigating erosion effects and enhancing the durability

Table 13. Comparative summary of erosion effects under different conditions

Figure*	Material/Coating	Experimental parameter	Erosion effects observed	Key insights
Figure 2	Nickel-based superalloy	Elevated temperatures (600 – 1000°C)	Increased oxidation and thermal fatigue exacerbate erosion effects at higher temperatures.	Temperature amplifies erosion through combined thermal and mechanical stresses.
Figure 3	Micro-grooved surface coating	High particle velocity with turbulent flow regime	Reduced material loss compared to flat coatings; grooves divert particles away from high-stress regions.	Surface texturing effectively reduces erosion by modifying particle trajectories.
Figure 4	Serrated trailing edge design	Impact of modified blade geometry	Erosion effects significantly reduced; smoother airflow with less turbulence and particle impingement.	Geometrical modifications reduce turbulence and improve erosion resistance.
Figure 5	Hybrid thermal barrier coating	Layered protection under extreme temperature gradients	Excellent thermal resistance; erosion localized to the top layer, preserving the substrate.	Multi-layer designs enhance both thermal and erosion resistance.
Figure 6	Finely textured blade surface	Micro-texturing effects on particle deflection	Improved particle deflection; significantly reduced material loss compared to untextured surfaces.	Micro-textures effectively reduce erosion by altering particle trajectories.
Figure 7	Blunt trailing edge design	High-velocity turbulent flow impacts	Increased material wear at the trailing edge due to wake turbulence; redesign reduces peak erosion rates.	Wake turbulence can be mitigated with smoother trailing-edge designs.

Note: *This summary is in reference to the figures presented in this article.
Abbreviations: CMC: Ceramic matrix composite; YSZ: Yttria-stabilized zirconia.

of turbine blades.[52,53] These studies typically encompass a comprehensive examination of erosion mechanisms, failure modes, and the specific environmental conditions contributing to erosion-induced damage. Discussions often revolve around the development of advanced materials, including erosion-resistant super alloys, innovative coating technologies such as TBCs, and sophisticated cooling strategies to manage the thermal stresses and protect the blades from erosion-related degradation. Furthermore, discussions in such reviews often delve into the latest advancements in computational modeling techniques, including finite element method and CFD, enabling a deeper understanding of the complex dynamics associated with erosion processes within gas turbine environments.

Examining the critical aspects of gas turbine blade failures caused by erosion represents a significant contribution to the ongoing efforts in the field of gas turbine technology and erosion management. By scrutinizing the critical aspects and intricate nuances of erosion-induced gas turbine blade failures, this review delves into the multifaceted nature of erosive processes, addressing the underlying mechanisms, key contributing factors, and the implications of erosion on the overall operational efficiency and longevity of gas turbines. In addition to the detailed analysis of erosion mechanisms, this comprehensive review offers insights into the diverse strategies and technologies employed to mitigate erosion effects and enhance the resilience of gas turbine blades. These discussions might encompass a broad spectrum of topics, including the development of novel materials with enhanced erosion resistance, the application of advanced coating systems, and the implementation of innovative cooling techniques to manage temperature differentials and reduce thermal stresses that can exacerbate erosion damage.

Moreover, it is probable that the review provides a critical appraisal of the latest advancements in computational modeling techniques, exploring how these tools have revolutionized the understanding of erosion dynamics within gas turbine environments. Such discussions would likely highlight the pivotal role of computational simulations, such as FEA and CFD, in simulating complex erosion scenarios and facilitating the design of robust mitigation strategies tailored to specific operating conditions and turbine configurations. This comprehensive review would serve as a valuable resource for researchers, engineers, and industry professionals, offering in-depth insights and practical guidance for addressing the persistent challenges associated with erosion-induced gas turbine blade failures. The integration of findings from this review into the development of future research directions and practical applications could significantly contribute to the advancement of erosion management strategies, ensuring enhanced performance, durability, and reliability of gas turbine systems in diverse industrial applications. Table 14 outlines the comprehensive and categorized core outcome results derived from the ongoing research, summarizing the failure mechanisms, erosion resistance of materials, influencing factors, analysis techniques, mitigation strategies, and blade vulnerability zones in gas turbine blade erosion.

Table 14. Overview of different types of gas turbine blade erosion

Category	Key points	Details	References
Failure mechanisms	SPE	High-velocity particles remove material, especially at leading and trailing edges.	Hamed *et al.*,[55] Wang *et al.*[33]
	Fatigue	Pits and cracks from erosion act as stress concentrators, accelerating structural degradation.	Ahsan *et al.*,[49] Rajabinezhad *et al.*[76]
	Oxidation and hot corrosion	Exposed material reacts with oxidizing gases, causing chemical degradation.	Rajabinezhad *et al.*[76]
	Creep	Erosion-thinned material deforms under high stress, reducing structural integrity.	Rajabinezhad *et al.*[76]
Erosion resistance of materials	YSZ	High resistance in high-temperature environments, suitable for TBCs.	Shin and Hamed[57]
	Alumina-based coatings	Moderate resistance; effective against abrasion but limited at elevated temperatures.	Branco *et al.*[58]
	MCrAlY bond coats	Moderate resistance; supports ceramic topcoats and improves adhesion.	Shin and Hamed[77]
	Nickel-based superalloys	Moderate to high resistance; good for high-stress environments but relies on coatings for erosion protection.	Rajabinezhad *et al.*[76]

(Cont'd...)

Table 14. ***(Continued)***

Category	Key points	Details	References
Factors influencing erosion rates	Particle velocity	Higher velocities increase material removal due to greater kinetic energy transfer.	Taherkhani *et al.*[56]
	Particle size	Larger particles cause deeper and more extensive erosion damage.	Branco *et al.*[58]
	Impact angle	Steep angles (close to 90°) maximize erosion severity by transferring more energy.	Wang *et al.*[33]
	Operational temperature	High temperatures accelerate erosion by softening materials and increasing oxidation.	Zhu *et al.*[59]
Techniques for erosion analysis	CFD	Simulates gas flow, turbulence, and particle trajectories to predict erosion-prone zones.	Hamed *et al.*,[55] Taherkhani *et al.*[56]
	FEA	Models stress distribution and crack propagation under erosion-induced loads.	Rajabinezhad *et al.*[76]
	SEM	Visualizes surface damage, revealing pits, cracks, and coating degradation.	Shin and Hamed,[57] Branco *et al.*[58]
	Experimental testing	Validates computational models and provides empirical erosion rate data.	Zhu *et al.*[59], Shin and Hamed[57]
Erosion mitigation strategies	Use of TBCs	Apply YSZ-based TBCs with optimized porosity to protect against thermal and mechanical erosion.	Shin and Hamed[57]
	Optimized blade geometry	Redesign leading edges to minimize particle impacts and turbulence.	Hamed *et al.*[55]
	Material hardening	Use superalloys and composite ceramics to enhance resistance against erosion and thermal stress.	Branco *et al.*,[58] Kedir *et al.*[78]
	Predictive maintenance	Integrate IoT and ML for real-time erosion monitoring and proactive maintenance.	Zhu *et al.*,[59] Shin and Hamed[57]
Blade vulnerability zones	Leading edge	Frequent steep-angle impacts cause severe material loss, reducing aerodynamic performance.	Wang *et al.*,[33] Hamed *et al.*[55]
	Trailing edge	Flow separation leads to uneven erosion patterns and localized structural weakening.	Hamed *et al.*[55]
	Mid-blade surface	Uniform erosion due to low-impact angles; results in gradual material thinning.	Branco *et al.*[58]
	Blade tips	Combined thermal and mechanical erosion accelerates wear, increasing the risk of catastrophic failure.	Rajabinezhad *et al.*[76]

Abbreviations: CFD: Computational fluid dynamics; FEA: Finite element analysis; IoT: Internet of things; ML: Machine learning; SEM: Scanning electron microscopy; SPE: Solid particle erosion; TBCs: Thermal barrier coatings; YSZ: Yttria-stabilized zirconia.

6. Conclusions

In this comprehensive review, a thorough analysis of the critical aspects of gas turbine blade failures caused by erosion has been conducted. The findings underscore the multifaceted nature of erosion-induced failures, highlighting the significance of this issue within the realm of gas turbine operations. The current review identifies that erosion, stemming from various factors including particle impingement, high-velocity fluid flow, and environmental conditions, significantly contributes to the deterioration of gas turbine blades, impacting their performance and longevity. Through the examination of various case studies and experimental investigations, it becomes evident that erosion mechanisms entail a complex interplay between material properties, environmental factors, and operational conditions. These insights emphasize the necessity for a holistic approach to blade design, material selection, and maintenance strategies, taking into account the diverse operational environments and potential erosion sources.

Furthermore, the role of advanced materials, coatings, and innovative design modifications emerged as promising avenues to mitigate erosion-induced failures. The incorporation of protective coatings, such as TBCs and erosion-resistant materials, displayed potential in enhancing the erosion resistance of gas turbine blades. In addition, the adoption of aerodynamic enhancements and modifications in blade geometries has shown promise in reducing erosion susceptibility, thereby improving the overall durability and efficiency of gas turbine systems. However, it is imperative to acknowledge the need for

continued research and development efforts to advance our understanding of erosion mechanisms, as well as to facilitate the implementation of effective preventive measures and maintenance strategies. Robust monitoring techniques and predictive modeling tools should be integrated to enable proactive maintenance practices, ensuring the early detection of erosion-related issues and the implementation of timely corrective actions.

This comprehensive review underscores the critical significance of addressing erosion-induced gas turbine blade failures, emphasizing the importance of a multidisciplinary approach that integrates materials science, aerodynamics, and operational considerations. By fostering collaborative efforts between researchers, industry stakeholders, and regulatory bodies, the gas turbine sector can effectively enhance the resilience and reliability of turbine blades, thereby ensuring sustained performance and operational safety in the realm of energy production and aviation. From the extensive literature review and research outlined above, the following key findings and implications can be highlighted:

- The speed at which particles impact the surface is the most significant factor affecting the rate of erosion, with gas temperature and impingement angle following closely in influence.
- Implementing an optimal layer of TBC presents a viable strategy to minimize the erosion of GTE blades.
- The erosion rate is also notably affected by the porosity of the TBC, underscoring the importance of minimizing porosity to mitigate erosion.
- Erosion demonstrates a direct correlation with the size of particles that impinge on the surface, emphasizing the need for measures to control particle size in operational environments.
- The leading edge of the vane emerges as the critical zone for erosion, with erosion on the pressure surface intensifying progressively toward the trailing edge.

These conclusions underscore the nuanced interplay between various parameters affecting erosion in gas turbine blades. The findings emphasize the significance of controlling particle impact velocities, optimizing TBC characteristics, and managing particle size distribution to effectively minimize erosion. Moreover, the critical areas identified for erosion highlight the necessity for targeted protective strategies, especially focusing on the leading edge and pressure surface to ensure prolonged blade life and optimal performance. Further research efforts should concentrate on refining these key factors to develop robust preventive measures and advanced protective coatings, thus enhancing the overall resilience and durability of gas turbine blades in demanding operational environments.

6.1. Theoretical implications

Theoretical implications arising from the examination of critical aspects of gas turbine blade failures caused by erosion are significant in shaping the theoretical framework and understanding of materials science, aerodynamics, and turbine engineering. Several theoretical implications can be derived from this comprehensive review:

- Erosion mechanisms and material science: This study highlights the intricate relationship between erosion mechanisms and material properties, emphasizing the need for a deeper theoretical understanding of material behavior under erosive conditions. Exploring the theoretical foundations of material degradation and erosion processes can facilitate the development of predictive models that account for material responses to varying environmental factors, aiding in the design of more erosion-resistant materials and coatings.
- Aerodynamic considerations and erosion dynamics: This review underscores the intricate link between aerodynamic factors and erosion dynamics, necessitating a more nuanced theoretical approach to comprehend the impact of fluid flow on erosion patterns. This calls for the development of theoretical models that integrate fluid dynamics with erosion mechanisms, enabling a comprehensive understanding of how gas flow characteristics contribute to erosion phenomena in gas turbine blades.
- Multidisciplinary approach to turbine design: Theoretical implications emphasize the significance of a multidisciplinary approach to gas turbine design, incorporating theoretical frameworks from materials science, fluid dynamics, and structural engineering. This integration facilitates the development of comprehensive theoretical frameworks that consider the interplay of various factors, leading to the creation of more robust and efficient gas turbine systems resistant to erosion-induced failures.
- Predictive maintenance and monitoring techniques: The current review highlights the importance of theoretical frameworks in the development of advanced predictive maintenance and monitoring techniques for detecting erosion-related issues in gas turbine blades. Theoretical implications emphasize the need for sophisticated theoretical models that enable the prediction of erosion rates and the identification of critical erosion-prone areas, facilitating the implementation of proactive maintenance strategies to ensure long-term reliability and performance of gas turbine systems.
- Sustainability and durability in energy systems: Theoretical implications emphasize the significance

of incorporating sustainability and durability considerations into the theoretical frameworks governing gas turbine design and operation. This requires the development of theoretical models that assess the long-term environmental impact of erosion-induced failures and the integration of theoretical concepts to promote the use of sustainable materials and practices, fostering the development of more environmentally friendly and durable energy systems.

By incorporating these theoretical implications into future research and development endeavors, this review contributes to the advancement of theoretical frameworks that underpin the design, operation, and maintenance of gas turbine systems, paving the way for the development of more efficient, resilient, and sustainable energy solutions.

6.2. Managerial implications

Managerial implications arising from the comprehensive review of critical aspects of gas turbine blade failures caused by erosion are crucial for guiding strategic decision-making and operational practices within the gas turbine industry. The following managerial implications can be derived from the review:

- Risk management and maintenance strategies: Managers in the gas turbine industry need to develop robust risk management strategies that incorporate the findings of this review to proactively address erosion-induced failures. This involves implementing comprehensive maintenance schedules that target critical erosion-prone areas and integrating advanced monitoring techniques to detect early signs of erosion, thereby minimizing downtime and ensuring optimal turbine performance.
- Investment in research and development: Managers should prioritize investment in research and development initiatives aimed at advancing erosion-resistant materials, innovative coating technologies, and advanced manufacturing processes. By leveraging the insights from the review, managers can direct resources toward the development of cutting-edge solutions that enhance the durability and performance of gas turbine blades, ultimately improving the overall reliability and competitiveness of their products and services.
- Supply chain management and quality control: Ensuring the quality and reliability of materials and components supplied for gas turbine manufacturing is crucial. Managers should establish stringent quality control measures within the supply chain to guarantee the procurement of high-quality materials, coatings, and components that align with the erosion-resistance requirements outlined in this review. This necessitates fostering strong relationships with reliable suppliers and enforcing strict quality assurance protocols to maintain product integrity.
- Regulatory compliance and industry standards: Managers may implement erosion control in gas turbine operations by aligning their practices with regulatory requirements and industry benchmarks. Managers can also demonstrate adherence to best practices, ensuring compliance with safety standards and environmental regulations, thus fostering trust among stakeholders and customers.
- Training and skill development: Enhancing the skills and expertise of the workforce is essential in implementing effective erosion prevention and maintenance practices. Managers should prioritize training programs that educate employees about erosion mechanisms, preventive measures, and the use of advanced monitoring technologies. This enables the workforce to make informed decisions, respond effectively to erosion-related challenges, and contribute to the overall efficiency and safety of gas turbine operations.

By incorporating these managerial implications into their strategic planning and operational frameworks, managers can effectively navigate the challenges posed by erosion-induced gas turbine blade failures, fostering a culture of proactive maintenance, innovation, and compliance, thereby enhancing the overall competitiveness and sustainability of their organizations in the dynamic gas turbine industry.

6.3. Limitations

Several limitations should be considered when examining the critical aspects of gas turbine blade failures caused by erosion. These limitations may have impacted the scope and depth of the comprehensive review, and they include:

- Data availability and accessibility: Limited access to comprehensive and updated data on erosion-induced gas turbine blade failures might have restricted the completeness of this review. Inadequate data availability could have affected the ability to conduct a thorough analysis of various case studies and experimental investigations, potentially leading to a partial understanding of the complex nature of erosion mechanisms and their impacts.
- Research focus and publication bias: The review's findings could have been influenced by a potential bias in the selection of research studies and literature. The emphasis on certain types of erosion mechanisms or specific mitigation strategies within the literature could have skewed the overall perspective, potentially overlooking alternative viewpoints and pertinent

research that might offer additional insights into the topic.
- Technological advancements and evolving research: Rapid advancements in materials science, coating technologies, and erosion mitigation strategies might have rendered some of the reviewed literature outdated or less relevant. The evolving nature of research in the field of gas turbine blade erosion could have limited the review's ability to incorporate the latest developments and emerging trends, thereby presenting an incomplete picture of the current state of knowledge.
- Industry-specific factors and confidentiality: Limited access to proprietary data and industry-specific information within the gas turbine sector might have constrained the review's ability to delve into certain detailed aspects of erosion-induced failures. Confidentiality agreements and proprietary restrictions could have limited the inclusion of real-world case studies and practical insights, potentially affecting the review's applicability to specific industrial contexts.
- Interdisciplinary complexities and overlapping variables: The multidisciplinary nature of the topic, encompassing aspects of materials science, aerodynamics, and engineering, introduces complexities and overlapping variables that might not have been fully addressed in this review. Inadequate exploration of the interconnected nature of these disciplines could have limited the review's ability to provide a comprehensive understanding of the intricate interplay between various factors contributing to gas turbine blade erosion.

Acknowledging these limitations is essential to foster a more nuanced understanding of the constraints and challenges associated with the examination of critical aspects of gas turbine blade failures caused by erosion, thereby encouraging further research efforts and the development of more holistic approaches to address these limitations in future studies.

6.4. Scope of future work

This comprehensive review on the critical aspects of gas turbine blade failures caused by erosion sets the stage for several promising avenues of future research. Building on the insights gained from the review, the following areas offer significant scope for future work and investigation:

- Advanced materials development: Future research can focus on the exploration and development of novel erosion-resistant materials tailored specifically for gas turbine blade applications. This may involve the investigation of new material compositions, innovative coating technologies, and advanced manufacturing processes to enhance the erosion resistance and overall durability of gas turbine blades in harsh operational environments.
- Integrated modeling and simulation studies: There is a need for the development of integrated modeling and simulation frameworks that incorporate fluid dynamics, materials science, and erosion mechanisms involving complex interactions to minimize gas turbine blade erosion. Future research can focus on the refinement of predictive models and computational simulations to assess erosion rates, predict failure modes, and optimize maintenance strategies for enhanced performance and longevity.
- Experimental validation and field testing: Conducting extensive experimental validation and field testing of erosion mitigation strategies and protective coatings is crucial to corroborate the findings of the review. Future research endeavors should emphasize the implementation of large-scale testing facilities and real-world case studies to evaluate the efficacy of various erosion control measures and validate their practical applicability in mitigating gas turbine blade failures caused by erosion.
- Sustainable erosion control strategies: Future work can explore the integration of sustainability principles into erosion control strategies for gas turbine blades. This may involve the investigation of environmentally friendly coating materials, the development of eco-conscious manufacturing processes, and the implementation of erosion prevention techniques that minimize environmental impact while ensuring long-term operational efficiency and safety.
- Data-driven maintenance and prognostics: Future research efforts can concentrate on the integration of data-driven maintenance and prognostic techniques to enable proactive monitoring and early detection of erosion-related issues in gas turbine blades. This includes the development of advanced sensing technologies, data analytics frameworks, and ML algorithms that facilitate real-time condition monitoring, predictive maintenance scheduling, and timely intervention to prevent catastrophic failures and optimize the operational lifespan of gas turbine systems.

By focusing on these areas of future research, scholars and industry practitioners can advance the knowledge frontier on gas turbine blade erosion, fostering the development of innovative solutions, best practices, and sustainable strategies to address the critical challenges posed by erosion-induced failures in gas turbine operations.

Acknowledgments

We would like to express our heartfelt gratitude to all individuals and organizations whose support has been instrumental in completing this research article. We sincerely thank the authors and researchers of the cited works for their valuable contributions, which have served as the foundation for this review. Finally, we are deeply grateful to our families and friends for their unwavering encouragement and understanding throughout this endeavor.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Shankha Shubhra Goswami
Writing – original draft: Surajit Mondal
Writing – review & editing: Shankha Shubhra Goswami

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Rodriguez E, Flores M, Pérez A, *et al.* Erosive wear by silica sand on AISI H13 and 4140 steels. *Wear*. 2009;267(11):2109-2115.

   doi: 10.1016/j.wear.2009.08.009

2. Mourad AHI, Almomani A, Sheikh IA, Elsheikh AH. Failure analysis of gas and wind turbine blades: A review. *Eng Fail Anal*. 2003;146:107107.

   doi: 10.1016/j.engfailanal.2023.107107

3. Alqallaf J, Ali N, Teixeira JA, Addali A. Solid particle erosion behaviour and protective coatings for gas turbine compressor blades-A review. *Process*. 2020;8(8):984.

   doi: 10.3390/pr8080984

4. Singh S, Kharub M, Singh J, Singh J, Jangid V. Brief survey on mechanical failure and preventive mechanism of turbine blades. *Mater Today Proc*. 2012;38:2515-2524.

   doi: 10.1016/j.matpr.2020.07.546

5. Fathyunes L, Mohtadi-Bonab MA. A review on the corrosion and fatigue failure of gas turbines. *Metals*. 2023;13(4):701.

   doi: 10.3390/met13040701

6. Pauzi AA, Ghazali MJW, Zamri WFH, Rajabi A. Wear characteristics of superalloy and hardface coatings in gas turbine applications-A review. *Metals*. 2020;10(9):1171.

   doi: 10.3390/met10091171

7. Mathioudakis K, Alexiou A, Aretakis N, Romesis C. Signatures of compressor and turbine faults in gas turbine performance diagnostics: A review. *Energies*. 2024;17(14):3409.

   doi: 10.3390/en17143409

8. Wang J, Zheng X. Review of geometric uncertainty quantification in gas turbines. *J Eng Gas Turbines Power*. 2020;142(7):070801.

   doi: 10.1115/1.4047179

9. Khushbash S, Hameed A, Mumtaz A, Khan HA, Shahzad A. Investigation of temperature assisted corrosion failure of an aircraft turbine blade. *Eng Fail Anal*. 2024;2024:108909.

   doi: 10.1016/j.engfailanal.2024.108909

10. Wood RJ, Lu P. Leading edge topography of blades-a critical review. *Surf Topogr Metrol Prop.* 2021;9(2):023001.

    doi: 10.1088/2051-672X/abf81f/meta

11. Fang G, Gao X, Song Y. A review on ceramic matrix composites and environmental barrier coatings for aero-engine: Material development and failure analysis. *Coatings*. 2023;13(2):357.

    doi: 10.3390/coatings13020357

12. Alshahrani R, Yenugula M, Algethami H, *et al*. Establishing the fuzzy integrated hybrid MCDM framework to identify the key barriers to implementing artificial intelligence-enabled sustainable cloud system in an IT industry. *Exp Syst Appl*. 2023;238:121732.

    doi: 10.1016/j.eswa.2023.121732

13. Song X, Xing Z, Jia Y, *et al*. Review on the damage and fault diagnosis of wind turbine blades in the germination stage. *Energies*. 2022;15(20):7492.

    doi: 10.3390/en15207492

14. Kablov EN, Muboyadzhyan SA. Erosion-resistant coatings for gas turbine engine compressor blades. *Russ Metall*. 2017;494-504.

    doi: 10.1134/S0036029517060118

15. Ridha WKH, Kashyzadeh KR, Ghorbani S. Common failures in hydraulic Kaplan turbine blades and practical solutions. *Mater*. 2023;16(9):3303.

    doi: 10.3390/ma16093303

16. Shi L, Guo S, Yu P, Zhang X, Xiong J. A review on leading-edge erosion morphology and performance degradation

of aero-engine fan and compressor blades. *Energies*. 2023;16(7):3068.

doi: 10.3390/en16073068

17. Mishnaevsky L Jr. Root causes and mechanisms of failure of wind turbine blades: Overview. *Materials* (*Basel*). 2022;15(9):2959.

doi: 10.3390/ma15092959

18. Hassan MK, Sindi WM, Mohamed A, Backar AH. Failure analysis of gas turbine in first stage turbine blades in an urban power plant. *J Eng Therm Sci*. 2024;4.

doi: 10.21595/jets.2024.23922

19. Blinov VL, Zubkov IS, Bogdanets SV, Komarov OV, Deryabin GA. Studies of erosive wear of the blading in axial compressors of gas turbines. *Therm Eng*. 2023;70(6):430-442.

doi: 10.1134/S0040601523060022

20. Mishra RK, Kumar V. Investigation of turbine blade failure in a turboprop engine: A case study. *J Fail Anal Prev*. 2022;22(2):458-463.

doi: 10.1007/s11668-022-01334-8

21. Pugh K, Nash JW, Reaburn G, Stack MM. On analytical tools for assessing the raindrop erosion of wind turbine blades. *Renew Sustain Energy Rev*. 2021;137:110611.

doi: 10.1016/j.rser.2020.110611

22. Rani P, Agrawal AK. Failure analysis of a low-pressure stage steam turbine blade. *Nondestruct Test Eval*. 2023;38(4):668-682.

doi: 10.1080/10589759.2022.2156503

23. Bonu V, Barshilia HC. High-temperature solid particle erosion of aerospace components: Its mitigation using advanced nanostructured coating technologies. *Coatings*. 2022;12(12):1979.

doi: 10.3390/coatings12121979

24. Sabri K, Gaceb M, Si-Chaib MO. Analysis of a directionally solidified (DS) GTD-111 turbine blade failure. *J Fail Anal Prev*. 2020;20:1162-1174.

doi: 10.1007/s11668-020-00920-y

25. Wang W, Yan Y, Zhou Y, Cui J. Review of advanced effusive cooling for gas turbine blades. *Energies*. 2022;15(22):8568.

doi: 10.3390/en15228568

26. Yenugula M, Goswami SS, Kaliappan S, *et al*. Analyzing the critical parameters for implementing sustainable AI cloud system in an IT industry using AHP-ISM-MICMAC integrated hybrid MCDM model. *Mathematics*. 2023;11(15): 3367.

doi: 10.3390/math11153367

27. Boggarapu V, Gujjala R, Ojha S. A critical review on erosion wear characteristics of polymer matrix composites. *Mater Res Exp*. 2020;7(2):022002.

doi: 10.1088/2053-1591/ab6e7b/meta

28. Chowdhury TS, Mohsin FT, Tonni MM, Mita MNH, Ehsan MM. A critical review on gas turbine cooling performance and failure analysis of turbine blades. *Int J Thermofluids*. 2023;18:100329.

doi: 10.1016/j.ijft.2023.100329

29. Mirhosseini AM, Nazari SA, Pour AM, Haghighi SE, Zareh M. Failure analysis of first stage nozzle in a heavy-duty gas turbine. *Eng Fail Anal*. 2020;109:104303.

doi: 10.1016/j.engfailanal.2019.104303

30. Talebi SS, Tousi AM, Madadi A. The effects of turbine blade erosion on the gas path measurements and performance of micro gas turbines. *AUT J Mech Eng*. 2023;7(4):419-438.

31. Chen K, Seo D, Canteenwalla P. The effect of high-temperature water vapour on degradation and failure of hot section components of gas turbine engines. *Coatings*. 2021;11(9):1061.

doi: 10.3390/coatings11091061

32. Prashar G, Vasudev H, Thakur L. A comprehensive review on the hot corrosion and erosion performance of thermal barrier coatings. *Prot Metals Phys Chem Surf*. 2023;59(3):461-492.

doi: 10.1134/S2070205122060132

33. Wang W, Xue Y, He C, Zhao Y. Review of the typical damage and damage-detection methods of large wind turbine blades. *Energies*. 2022;15(15):5672.

doi: 10.3390/en15155672

34. Sławiński D, Ziółkowski P, Badur J. Thermal failure of a second rotor stage in heavy duty gas turbine. *Eng Fail Anal*. 2020;115:104672.

doi: 10.1016/j.engfailanal.2020.104672

35. Yang P, Yue W, Li J, Bin G, Li C. Review of damage mechanism and protection of aero-engine blades based on impact properties. *Eng Fail Anal*. 2022;140:106570.

doi: 10.1016/j.engfailanal.2022.106570

36. Błachnio J, Spychała J, Zasada D. Analysis of structural changes in a gas turbine blade as a result of high temperature and stress. *Eng Fail Anal*. 2021;127:105554.

doi: 10.1016/j.engfailanal.2021.105554

37. Szczepankowski A, Przysowa R. Thermal degradation of turbine components in a military turbofan. *Eng Fail Anal*. 2022;134:106088.

doi: 10.1016/j.engfailanal.2022.106088

38. Hoksbergen N, Akkerman R, Baran I. The Springer model for lifetime prediction of wind turbine blade leading edge protection systems: A review and sensitivity study. *Materials* (*Basel*). 2022;15(3):1170.

doi: 10.3390/ma15031170

39. Bera P, Lakshmi RV, Pathak SM, Bonu V, Mishnaevsky L Jr., Barshilia HC. Recent progress in the development

and evaluation of rain and solid particle erosion resistant coatings for leading edge protection of wind turbine blades. *Polym Rev*. 2024;64(2):639-689.

doi: 10.1080/15583724.2023.2270050

40. Peng H, Zhang H, Fan Y, Shangguan L, Yang Y. A review of research on wind turbine bearings' failure analysis and fault diagnosis. *Lubricants*. 2022;11(1):14.

doi: 10.3390/lubricants11010014

41. Alqallaf J, Teixeira JA. Quantifying the economic benefits of using erosion protective coatings in a low-pressure compressor (aero-engine): A case study evaluation. *Process*. 2022;10(2):385.

doi: 10.3390/pr10020385

42. Goswami SS, Banerjee B, Das S, Mondal S, Halder R. Risk assessment and management strategies for turbine blade failures in power generation systems. In: *Advances in Additive Manufacturing Technologies*. United States: CRC Press; 2024.

43. Rivaz A, Anijdan SM, Moazami-Goudarzi M. Failure analysis and damage causes of a steam turbine blade of 410 martensitic stainless steel after 165,000 h of working. *Eng Fail Anal*. 2020;113:104557.

doi: 10.1016/j.engfailanal.2020.104557

44. Molinari JF, Ortiz M. A study of solid-particle erosion of metallic targets. *Int J Impact Eng*. 2002;27(4):347-358.

doi: 10.1016/S0734-743X(01)00055-0

45. Rivaz A, Anijdan SM, Moazami-Goudarzi M, Ghohroudi AN, Jafarian HR. Damage causes and failure analysis of a steam turbine blade made of martensitic stainless steel after 72,000 h of working. *Eng Fail Anal*. 2022;131:105801.

doi: 10.1016/j.engfailanal.2021.105801

46. Wei Z, Zhang S, Jafari S, Nikolaidis T. Gas turbine aero-engines real time on-board modelling: A review, research challenges, and exploring the future. *Prog Aerosp Sci*. 2020;121:100693.

doi: 10.1016/j.paerosci.2020.100693

47. Mortazavi SA, Rahi A, Jafari SM. Effect of damping element damage under erosion on vibration behavior of an industrial gas turbine group-blades. *AUT J Mech Eng*. 2024;8(2):181-198.

doi: 10.22060/ajme.2024.23009.6090

48. Prashar G, Vasudev H. Application of artificial neural networks in the prediction of slurry erosion performance: A comprehensive review. *Int J Interact Des Manuf*. 2024:1-19.

doi: 10.1007/s12008-024-02014-7

49. Ahsan S, Lemma TA, Hashmi MB, Liang X. Investigation of operational settings, environmental conditions, and faults on the gas turbine performance. *Meas Sci Technol*. 2024;35(12):125902.

doi: 10.1088/1361-6501/ad678c/meta

50. Rezamand M, Kordestani M, Carriveau R, Ting DSK, Orchard ME, Saif M. Critical wind turbine components prognostics: A comprehensive review. *IEEE Trans Inst Meas*. 2020;69(12):9306-9328.

doi: 10.1109/TIM.2020.3030165

51. Li D, Jiang P, Sun F, Yuan X, Zhang J, Cao X. Water-droplet erosion behavior of high-velocity oxygen-fuel-sprayed coatings for steam turbine blades. *Corros Rev*. 2022;40(1):39-49.

doi: 10.1515/corrrev-2021-0049

52. Zainuddin NS, Zamri WFHW, Omar MZ, Bin Md Din MF. Comprehensive insight into the failure mechanisms, modes, and material selection of steam turbine blades. *J Curr Sci Technol*. 2024;14(3):47-47.

doi: 10.59796/jcst.V14N3.2024.47

53. Olabi AG, Wilberforce T, Elsaid K, *et al*. A review on failure modes of wind turbine components. *Energies*. 2021;14(17):5241.

doi: 10.3390/en14175241

54. Han L, Guo C, Osman FK, Li D, Wang H, Liu Y, Qin D. Effect and mechanism of erosion in Pelton turbine and case studies-A review. *Phys Fluid*. 2024;36(3):031301.

doi: 10.1063/5.0191051

55. Hamed AA, Tabakoff W, Rivir RB, Das K, Arora P. Turbine blade surface deterioration by erosion. *J Turbomach*. 2005;127(3):445-452.

doi: 10.1115/1.1860376

56. Taherkhani B, Anaraki AP, Kadkhodapour J, Farahani NK, Tu H. Erosion due to solid particle impact on the turbine blade: experiment and simulation. *J Fail Anal Prev*. 2019;19:1739-1744.

doi: 10.1007/s11668-019-00775-y

57. Shin D, Hamed A. Influence of micro-structure on erosion resistance of plasma sprayed 7YSZ thermal barrier coating under gas turbine operating conditions. *Wear*. 2018;396-397:34-47.

doi: 10.1016/j.wear.2017.11.005

58. Branco JRT, Gansert R, Sampath R, Berndt CC, Herman H. Solid particle erosion of plasma sprayed ceramic coatings. *Mater Res*. 2004;7(1):147-153.

doi: 10.1590/S1516-14392004000100020

59. Zhu D, Miller RA, Kuczmarski MA. Development and life prediction of erosion resistant turbine low conductivity thermal barrier coatings. In: *65th Annual Forum and Technology Display*; 2009. Available from: https://ntrs.nasa.gov/api/citations/20100011004/downloads/20100011004.pdf [Last accessed on 2024 Aug 25].

60. Khan MS, Sasikumar C. A water droplet erosion-induced

fatigue crack propagation and failure in X20Cr13 martensitic stainless-steel turbines working at low pressure. *Eng Fail Anal*. 2022;139:106491.

doi: 10.1016/j.engfailanal.2022.106491

61. Poursaeidi E, Sigaroodi MJ, Aieneravaie M. Failure investigation of fatigue crack initiation and propagation in compressor blade. *Eng Fail Anal*. 2024;162:108370.

doi: 10.1016/j.engfailanal.2024.108370

62. Sun S, Wang T, Chu F. *In-situ* condition monitoring of wind turbine blades: A critical and systematic review of techniques, challenges, and futures. *Renew Sustain Energy Rev*. 2022;160:112326.

doi: 10.1016/j.rser.2022.112326

63. Sethy NK, Yenugula M, Goswami SS, Bhola A, Behera DK. Selection of ideal IoT based overhead conductor for optimizing the performance of a small hydropower project. *J Nano Electron Phys*. 2023;15(4):04006.

doi: 10.21272/jnep.15(4).04006

64. Kumar P, Singal SK, Gohil PP. A technical review on combined effect of cavitation and silt erosion on Francis turbine. *Renew Sustain Energ Rev*. 2024;190:114096.

doi: 10.1016/j.rser.2023.114096

65. Zhou D, Wei T, Huang D, Li Y, Zhang H. A gas path fault diagnostic model of gas turbines based on changes of blade profiles. *Eng Fail Anal*. 2020;109:104377.

doi: 10.1016/j.engfailanal.2020.104377

66. Lv B, Jin X, Cao J, Xu B, Wang Y, Fang D. Advances in numerical modeling of environmental barrier coating systems for gas turbines. *J Eur Ceram Soc*. 2020;40(9):3363-3379.

doi: 10.1016/j.jeurceramsoc.2020.03.036

67. Bin G, Li C, Li J, Chen A. Erosion-damage-induced vibration response of aero-gas generator rotor system. *Mech Syst Signal Process*. 2023;195:110298.

doi: 10.1016/j.ymssp.2023.110298

68. Quintanar-Gago DA, Nelson PF, Díaz-Sánchez Á, Boldrick MS. Assessment of steam turbine blade failure and damage mechanisms using a Bayesian network. *Reliab Eng Syst Saf*. 2021;207:107329.

doi: 10.1016/j.ress.2020.107329

69. Kishore K, Das S, Mandal H, *et al*. Failure investigation of a blast furnace top gas recovery turbine: Chronology and mechanism. *J Fail Anal Prev*. 2020;20:1376-1387.

doi: 10.1007/s11668-020-00951-5

70. García-Martínez M, Del Hoyo Gordillo JC, González MPV, Muro AP, Caballero BG. Failure study of an aircraft engine high pressure turbine (HPT) first stage blade. *Eng Fail Anal*. 2023;149:107251.

doi: 10.1016/j.engfailanal.2023.107251

71. Stokes JL, Presby MJ. A dynamic testing approach for particulate erosion-corrosion for gas turbine coatings. *J Eng Gas Turbines Power*. 2024;146(11):111006.

doi: 10.1115/1.4065886

72. Anand VGK, Parammasivam KM. Thermal barrier coated surface modifications for gas turbine film cooling: A review. *J Therm Anal Calorim*. 2021;146(2):545-580.

doi: 10.1007/s10973-020-10032-2

73. Faqihi B, Ghaith F. A comprehensive review and evaluation of heat recovery methods from gas turbine exhaust systems. *Int J Thermofluid*. 2023;18:100347.

doi: 10.1016/j.ijft.2023.100347

74. Alqallaf J, Teixeira JA. Numerical study of effects of solid particle erosion on compressor and engine performance. *Result Eng*. 2022;15:100462.

doi: 10.1016/j.rineng.2022.100462

75. Singh J, Chohan JS, Hlail SH, Saxena KK. Future perspective on the flow disturbances in Francis turbine caused by sediment particles and implementation of protection technique. *Mater Today Proc*. 2023.

doi: 10.1016/j.matpr.2023.09.118

76. Rajabinezhad M, Bahrami A, Mousavinia M, Seyedi SJ, Taheri P. Corrosion-fatigue failure of gas-turbine blades in an oil and gas production plant. *Materials* (*Basel*). 2020;13(4):900.

doi: 10.3390/ma13040900

77. Shin D, Hamed A. Experimental investigation and parametric DOE appraisal of thermal barrier coating high temperature erosion. In: *54th AIAA Aerospace Sciences Meeting*; 2016.

doi: 10.2514/6.2016-1886

78. Kedir N, Gong C, Sanchez L, *et al*. Erosion in gas-turbine grade ceramic matrix composites (CMCs). *Am Soc Mech Eng*, 2018;14.

doi: 10.1115/GT2018-75827

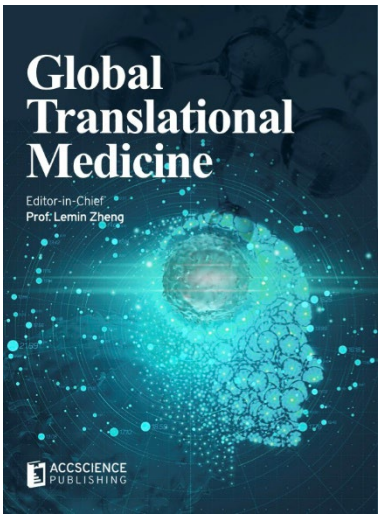
Our Journals

Tumor Discovery is a peer-reviewed and open-access journal that aims to present new cancer research with strong emphasis on fundamental and translational studies. *Tumor Discovery* covers topics, including but not limited to the following:

- Etiology and pathogenesis of cancer
- Mechanisms and molecular pathways underlying cancer initiation and progression
- Tumor metastasis
- Tumor evolution and heterogeneity
- Tumor microenvironment and tumor-host interactions
- Cancer genetics and genomics
- Cancer characterization using omics approaches
- Discovery and validation of cancer biomarker
- Discovery of new therapeutic targets
- New approaches of diagnostic and treatment modalities
- Statistical methods in cancer research

Global Translational Medicine is a quarterly journal that focuses on medicine, biological sciences, and biomaterials engineering. The goal of *Global Translational Medicine* is to provide a platform to researchers for showcasing their latest research works in translational medicine so as to advance the field towards the betterment of human health. Despite the advancement of omics and new technologies, the process of transforming these technologies and scientific research results into effective therapies and putting them into clinical use still has a long way to go. *Global Translational Medicine* provides a platform to fill the gaps in preclinical and inter-disciplinary research, to promote clinical translation of scientific research results, and to contribute to the conception of new and improved preventive measures as well as diagnostic and therapeutic techniques of diseases.

Global Translational Medicine covers the following themes: cardiovascular disease, metabolism/diabetes/obesity, neuroscience/neurology, cancer, biomaterials and their applications in medicine, proteomics/metabolomics, pharmacogenomics, biomarkers, bioinformatics and data mining, animal and clinical research, and medical methods arising from interdisciplinary crossover.



Start a new journal

Write to us via email if you are interested to start a new journal with AccScience Publishing. Please attach your CV, professional profile page and a brief pitch proposal in your email. We shall inform you of our decision whether we are interested to collaborate in starting a new journal.

Contact: info@accscience.com

https://accscience.com/journal/IJAMD

Contact
www.accscience.com
8 Burn Road, #15-03 Trivex, Singapore 369977
E-mail: editorial@accscience.com
Phone: +65 8182 1586