17 (2026) 201178

Contents lists available at ScienceDirect

e-Prime - Nexus of Electrical, Electronic, and Intelligent Engineering

journal homepage: www.sciencedirect.com/journal/e-prime-nexus-of-electrical-electronic-and-intelligent-engineering

Revolutionizing gas turbine performance analysis with deep learning powered digital twin

Balbir Shah Mohd Irwan Shah *, Asnor Juraiza Ishak, Mohd Khair Hassan, Nor Mohd Haziq Norsahperi

Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, 43400, Malaysia

ARTICLE INFO

Keywords:
Artificial intelligence
Data-driven
Deep learning
Digital twin
Energy efficiency
Gas turbine
Performance

ABSTRACT

Gas turbine technology is essential for resolving the energy trilemma while facilitating renewable energy integration, yet existing physics-based models lack adaptability to real-world variations and purely data-driven approaches frequently violate fundamental thermodynamic laws. This study presents a novel physics-constrained deep learning framework that uniquely integrates first-principles thermodynamics with operational data assimilation to create a high-fidelity digital twin of a GE 9FA heavy-duty gas turbine, distinguishing itself from conventional approaches by ensuring thermodynamic consistency while adapting to measured performance. The digital twin employs neural network architectures regularized by conservation laws and thermodynamic cycle constraints to forecast and interactively visualize T–S and P–V trajectories, translating subtle efficiency variations into actionable operational insights. Comprehensive validation across six key operating regimes demonstrates low predictive error and robust performance, confirming that physics constraints enhance generalization compared to standard machine learning baselines. These capabilities support proactive maintenance strategies and long-term efficiency optimization, representing a significant advancement toward intelligent, self-optimizing turbine systems for reliable energy management.

Introduction

Gas turbine technology emerges as a vital component in addressing the energy trilemma and managing the integration of renewable energy sources into the power grid. As the global energy landscape evolves, the demand for flexible and efficient power generation systems grows significantly. Gas turbines are expected to play an important role in the transition toward sustainable energy systems due to their inherent flexibility in firing modes, high reliability, and cost-effectiveness. Recent advancements in gas turbine technology are increasingly focused on improving operational efficiency, reducing emissions, and expanding fuel flexibility. These developments open the door to the use of alternative fuels, such as hydrogen or biogas, offering a significant opportunity to reduce environmental impact and diversify fuel sources. The adaptive capability of gas turbines to respond quickly to fluctuations in variable renewable energy sources is seen as a potential solution for maintaining grid stability, contributing significantly to achieving a diversified and sustainable energy mix while also aligning with the goals of forthcoming energy policies and global environmental commitments, such as those outlined in the COP agreements.

Therefore, this study proposes that advanced performance analysis is required to detect any performance abnormalities and degradation factors that prevent the gas turbine from operating at its optimum level. The insights gained from performance analysis can help the industry implement condition-based predictive maintenance, as opposed to relying on preventive maintenance, which is typically scheduled based on fixed intervals. The main disadvantage of preventive maintenance is that healthy parts may be replaced due to shelf-life expiration, and components may fail before scheduled maintenance [1]. In the context of gas turbine performance, degradation mechanisms can be classified into three groups: recoverable, non-recoverable, and permanent degradation [2]. Recoverable degradation, also known as controllable losses, can be mitigated during operation. Non-recoverable degradation, such as fouling, erosion, corrosion, oxidation, abrasion, and foreign object damage, cannot be recovered except during an overhaul [3,4]. Permanent deterioration refers to degradation mechanisms that cannot be reversed during operation or overhaul, often resulting from factors such as aging, surface roughness increases, and reductions in material

* Corresponding author.
E-mail address: balbirshahphd@gmail.com (B.S. Mohd Irwan Shah).

https://doi.org/10.1016/j.eprime.2026.201178
Received 30 August 2025; Received in revised form 22 March 2026; Accepted 1 April 2026
Available online 3 April 2026
3117-5112/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

strength.

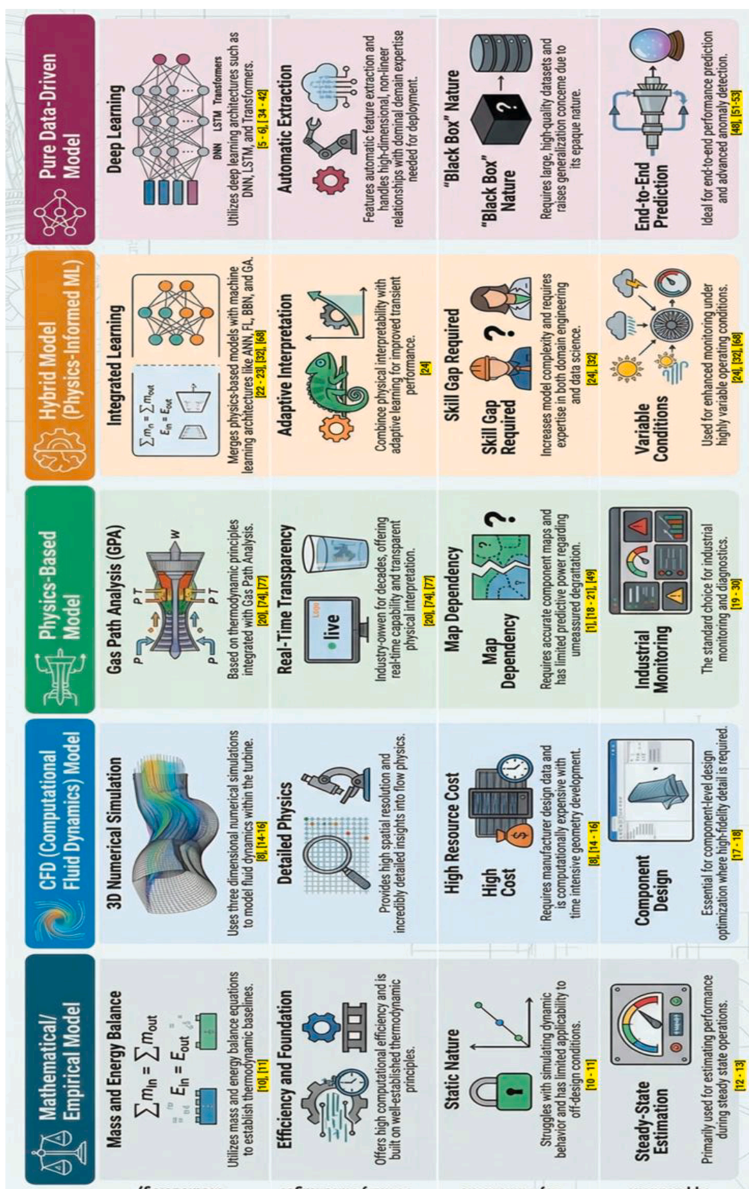
Evolution of gas turbine performance modeling

The field of gas turbine performance analysis has evolved through five distinct methodological paradigms, each presenting characteristic trade-offs between physical fidelity, computational efficiency, and operational applicability [5–9]. Table 1 provides a systematic comparative analysis of these approaches, delineating their theoretical foundations, primary advantages, inherent limitations, and dominant application domains.

Mathematical modelling based on mass and energy balances represents the foundational approach for gas turbine performance estimation. While computationally efficient, this methodology exhibits fundamental limitations in capturing dynamic transient behaviour and suffers from reduced accuracy under off-design conditions due to simplified component representations and fixed loss correlations [10, 11]. Recent work by Sammour et al. [12] and González-Díaz et al. [13] has demonstrated the continued relevance of thermodynamic modeling for assessing ambient temperature effects on combined cycle performance, though these steady-state formulations remain inadequate for transient operational analysis.

Computational fluid dynamics (CFD) offers superior spatial resolution and detailed flow physics but faces practical barriers in industrial applications, including the unavailability of proprietary manufacturer design data, prohibitive computational costs for real-time applications, and substantial time investment required for three-dimensional geometry development and mesh refinement [8,14–16]. To address these limitations, recent research has focused on developing deep learning-based surrogate models that approximate CFD results with significantly reduced computational overhead. Esfahanian et al. [17] developed a CNN-based surrogate model for aerodynamic shape optimization achieving $R^2 > 0.93$ with computational speedups exceeding 10^5 times compared to through-flow codes. Similarly, Bruij et al. [18] introduced CNNFD, a deep learning framework for turbomachinery CFD analysis that enables rapid flow field prediction without mesh

Table 1
Comparative analysis of gas turbine performance modelling approaches.



	Mathematical/ Empirical Model	CFD (Computational Fluid Dynamics) Model	Physics-Based Model	Hybrid Model (Physics-Informed ML)	Pure Data-Driven Model
Methodology	Mass and Energy Balance $\sum m_{in} = \sum m_{out}$ $E_{in} = E_{out}$ Utilizes mass and energy balance equations to establish thermodynamic baselines. [10], [11]	3D Numerical Simulation Uses three dimensional numerical simulations to model fluid dynamics within the turbine. [8], [14-16]	Gas Path Analysis (GPA) Based on thermodynamic principles integrated with Gas Path Analysis. [20], [74], [77]	Integrated Learning Merges physics-based models with machine learning architectures like ANN, FL, BBN, and GA. [22 - 23], [32], [68]	Deep Learning DNN LSTM Transformers Utilizes deep learning architectures such as DNN, LSTM, and Transformers. [5 - 6], [34 - 42]
Primary Advantages	Efficiency and Foundation Offers high computational efficiency and is built on well-established thermodynamic principles.	Detailed Physics Provides high spatial resolution and incredibly detailed insights into flow physics.	Real-Time Transparency Industry-proven for decades, offering real-time capability and transparent physical interpretation. [20], [74], [77]	Adaptive Interpretation Combine physical interpretability with adaptive learning for improved transient performance. [24]	Automatic Extraction Features automatic feature extraction and handles high-dimensional, non-linear relationships with dominal demain expertise needed for deployment.
Key Limitations	Static Nature Struggles with simulating dynamic behavior and has limited applicability to off-design conditions. [10 - 11]	High Resource Cost High Cost Requires manufacturer design data and is computationally expensive with time intensive geometry development. [8], [14 - 16]	Map Dependency Map Dependency Requires accurate component maps and has limited predictive power regarding unmeasured degradation. [1], [18 - 21], [49]	Skill Gap Required Skill Gap Required Increases model complexity and requires expertise in both domain engineering and data science. [24], [32]	"Black Box" Nature "Black Box" Nature Requires large, high-quality datasets and raises generalization concerns due to its epaque nature.
Applications	Steady-State Estimation Primarily used for estimating performance during steady state operations. [12 - 13]	Component Design Essential for component-level design optimization where high-fidelity detail is required. [17 - 18]	Industrial Monitoring The standard choice for industrial monitoring and diagnostics. [19 - 30]	Variable Conditions Used for enhanced monitoring under highly variable operating conditions. [24], [32], [66]	End-to-End Prediction Ideal for end-to-end performance prediction and advanced anomaly detection. [48], [51-53]

generation requirements.

Physics-based modeling employing gas path analysis (GPA) has constituted the industry standard for decades, utilizing thermodynamic principles and component characteristic maps to establish performance baselines and diagnose degradation [19–21]. These model-based approaches use physics-based modelling to represent engine behavior and create a benchmark for diagnostics, employing gas path analysis (GPA) to analyze gas turbine condition based on thermodynamic principles and a mass balance approach [1,22–26]. This approach remains prevalent as the backbone of many modern machine learning algorithms, such as Artificial Neural Networks (ANN), Expert Systems (ES), Fuzzy Logic (FL), Bayesian Belief Networks (BBN), Deep Learning (DL), and Genetic Algorithms (GA) [27,28]. Recent work by Kim [29] and Kim et al. [30] has advanced performance adaptation techniques using large-scale measured data, demonstrating improved accuracy in transient performance prediction.

The integration of physics-based models and machine learning is known as **hybrid modelling**, which is widely applied in gas turbine monitoring to detect performance degradation under both transient conditions and steady-state operation. Machine learning models can be used to enhance the performance of physics-based models, particularly under transient conditions. Zhang et al. [31] proposed a digital twin approach based on deep multi-model fusion, establishing bidirectional data flow between physics-based models (PBM) and data-driven models (DDM) to achieve high consistency with physical engines. Farhat and Altarawneh [32] provided a comprehensive review of physics-informed machine learning for intelligent gas turbine digital twins, highlighting applications in flow-field inference, heat-transfer analysis, and structural stress estimation. Recent work by Lim et al. [33] integrated Physics-Informed Neural Networks (PINNs) with Particle Swarm Optimization (PSO) for real-time parameter estimation and performance optimization, achieving R^2 scores of 94.03% for turbine inlet temperature prediction.

In recent years, the **data-driven modelling** via the ANN approach has gained prominence as an effective method for simulating thermal systems relevant to gas turbine applications. A growing body of research underscores the excellent performance of the ANN approach in enhancing prediction accuracy and optimizing monitoring processes in gas turbine operations [5,6,34–42]. Furthermore, there has been optimistic advancement in the field, with researchers increasingly exploring deep learning strategies within ANN frameworks. These advanced methodologies have demonstrated high accuracy in predicting gas turbine performance degradation, even in the complexities of diverse application contexts [7,16,43–47]. Recent advances include the application of Long Short-Term Memory (LSTM) networks for degradation prediction by Zhou et al. [48], transformer-based architectures for combustion modeling [49,50], and temporal multimodal multivariate learning (TMML) for predictive maintenance [51]. Additionally, frontier research has introduced sophisticated deep learning techniques for performance prediction under varying operational conditions. Liu et al. [52] developed a multi-working conditions identification and performance prediction framework based on deep learning and knowledge integration, achieving MSE of 0.7501 and correlation coefficient of 0.9973 on actual gas turbine operating datasets. Similarly, Yu et al. [53] proposed a novel gas turbine performance prediction model incorporating residual connections and feature engineering methods, addressing both full-load and part-load conditions with average prediction error less than 1.59%. Shabunin [54] developed a recurrent neural network-based GTU model demonstrating features of elaboration and application for performance prediction, while Baker and Rosic [55] enabled live thermal boundary prediction using real-time plant measurements and LSTM neural networks.

As technology advances in the era of the Fourth Industrial Revolution, harnessing the potential of deep learning to develop a comprehensive gas turbine model through the digital twin concept is essential for the power generation industry. A digital twin is a virtual replica or mirror of an asset that can provide full insights for decision-making throughout its lifecycle. It is a proven concept that can be applied across industries for anomaly detection [56,26,57,21], performance analysis [56,57,21], process simulation [58,59], predictive modelling [60], and preventive maintenance [61]. Previous studies have laid a solid foundation for gas turbine digital twin applications involving performance and optimization, with models developed based on real case studies for diagnostic and performance monitoring using historical data converted into health indicators and virtual sensors to monitor faulty performance [21,26]. Recent developments in digital twin frameworks have emphasized real-time adaptability and scalability. Hu et al. [26] developed a digital twin model of gas turbine and its application in warning of performance fault using deep learning frameworks. Sundararajan et al. [62] developed a gas turbine model interface specifically for digital twin applications, while Choi et al. [63] examined digital twin implementations in the power generation industry. Farhat and Salvini [64] developed a novel hybrid gas turbine digital twin to predict performance deterioration, while Singh et al. [65] demonstrated a digital twin of a laboratory gas turbine engine using deep learning framework. Panov and Cruz-Manzo [66] presented a gas turbine performance digital twin for real-time embedded systems, and Huang et al. [67] developed digital twin driven life-cycle operation optimization for combined cooling heating and power systems. Sleiti et al. [68] proposed a robust digital twin for power plant and other complex capital-intensive large engineering systems while Shah et al. [69] examined challenges and enablers of digital twin application in power plants within the Energy 4.0 context.

Recent literature on gas turbine digital twins (2020–2026)

Recent literature reveals significant methodological diversification in gas turbine digital twin research. Table 2 categorizes these contemporary studies by their application scale, methodological approach, and specific focus areas, illustrating the distinct gap that the present study addresses.

As evidenced in Table 2, contemporary research predominantly concentrates on: (i) laboratory-scale or micro-turbines rather than industrial heavy-duty units; (ii) component-specific monitoring (compressor-only, thermal boundaries, or blade cooling) rather than comprehensive sectional analysis; (iii) hybrid or physics-informed methodologies requiring substantial domain expertise; or (iv) specialized applications (CAES, CCHP, marine engines) distinct from conventional combined cycle operations. This distribution validates the observation by Farhat et al. [32] that "only a few [digital twin studies] focus on GT for thermal power plants" and that "PINNs for industrial heavy-duty GTs remain limited."

Furthermore, recent 2025 studies demonstrate the continued dominance of hybrid approaches, while purely data-driven comprehensive system-level digital twins for heavy-duty units remain underrepresented.

The rationale for a purely data-driven approach

Recent advancements in scientific machine learning, particularly Physics-Informed Neural Networks (PINNs), have demonstrated the value of embedding governing physical laws directly into the loss function: $L\_total = L\_data + \lambda \cdot L\_physics$ [77]. While effective in data-scarce regimes, PINNs require significant domain expertise for residual formulation, boundary handling, and differentiation overhead, challenging for complex industrial systems like gas turbines.

Contrast with Prior Methodologies: Unlike Shabunin [71], who employed RNNs for small-scale turbine anomaly detection, or Baker and Rosic [72], who utilized hybrid LSTM approaches for thermal monitoring, or frontier performance prediction models such as those by Liu et al. [52] (deep learning with knowledge) and Yu et al. [53] (residual connections and feature engineering), or recent 2025 hybrid-focused

Table 2
Gas turbine digital twin studies from 2020 to 2025.

Year	Author	Primary Focus and Scale	Methodology	Limitation Relative to Current Study
2025	[70]	• Compressor flows visualization by using machine learning (Aero engine)	PML model Vs CFD	Limited to component-specific (Compressor) and CFD development.
	[71]	• Anomaly detection in small-scale turbines (Laboratory (6 MW))	Recurrent Neural Network (RNN)	Limited to single operating behavior; not applicable to industrial heavy-duty systems
	[72]	• Thermal boundary prediction under renewable-driven operations (Industrial)	Hybrid LSTM multi-fidelity	Component-specific (thermal monitoring only); excludes comprehensive performance prediction
	[73]	• Remaining useful life prediction (Component)	Transfer learning with pruning	Prognostics-focused rather than comprehensive performance monitoring
	[49]	• Combustion mode time-series prediction in rotating detonation combustor (Component)	Predictor of encoder transformer	Combustor-specific; limited to detonation mode prediction
2024	[74]	• Compressed air energy storage discharge dynamics (CAES plant)	Conservation-based flow modeling	CAES-specific application; limited generalizability to conventional CCGT operations
	[63]	• Industry integration frameworks (Industry-wide)	Various	Conceptual focus; lacks specific thermodynamic performance modeling
	[62]	• Digital twin application interfaces (Interface design)	Model interface development	Infrastructure-focused; no performance prediction methodology
2023	[64]	• Performance deterioration prediction (General)	Hybrid model	Methodology description without comprehensive sectional analysis
	[65]	• Laboratory engine modeling (Laboratory)	Deep learning framework	Non-industrial scale; controlled conditions unlike operational variability
2022	[67]	• CCHP-CER system optimization (System-level)	Cascade forward neural networks	Focus on integrated energy systems rather than standalone gas turbine performance
2021	[75]	• Micro-turbine diagnostics (Micro-turbine fleet)	Fleet monitoring architectures	Scale mismatch with heavy-duty industrial turbines; different degradation mechanisms
	[76]	• EPRI gas turbine digital twin platform (Industry platform)	Integrated diagnostics and forecasting	Platform-focused; limited public disclosure of methodology details
2020	[66]	• Embedded system implementation (Embedded systems)	Real-time performance algorithms	Computational constraints limit model complexity and comprehensive coverage

works (e.g., physics-informed reviews), this study demonstrates that a purely data-driven deep feedforward neural network can achieve high predictive accuracy and industrial-grade reliability for gas turbine performance monitoring under both steady-state and dynamic conditions without incorporating any explicit physics-based loss terms or constraints. This represents a methodological departure from the hybrid approaches dominating recent literature [32,52,53,55,64,74], and 2025 updates.

This intentional design choice serves several purposes:

a. **Demonstrating data sufficiency**: With a large, high-quality dataset (100,436 samples across 34 parameters, including thermodynamically derived performance values from 11 raw inputs), the deep learning can implicitly capture the system's nonlinear dynamics and physical relationships. This contrasts with studies such as Zheng et al. [73], which rely on transfer learning from limited data, or laboratory-based studies [65,71] with constrained operational envelopes. Recent work by Zhou et al. [48] on degradation trend prediction supports the viability of data-driven approaches when sufficient high-quality operational data is available.
b. **Industrial practicality and cost reduction**: Unlike physics-informed models, which require significant effort to develop, maintain, and update physical residuals, a data-driven digital twin can be trained quickly on representative data, deployed rapidly, and updated incrementally with new operational data eliminating the need to repeatedly reformulate thermodynamic constraints. This addresses the practical barriers identified by Choi et al. [63] regarding digital twin implementation in power generation industries.
c. **Validation through rigorous empirical metrics**: Performance is assessed using standard metrics (MAE, RMSE, R^2), learning curves, and predicted-vs-actual comparisons across load conditions (125–220 MW), yielding near-perfect R^2 values and errors consistently below 0.5%. These results confirm the model's fidelity for real-world applications such as real-time monitoring, anomaly detection, and predictive maintenance, exceeding the accuracy thresholds reported in recent hybrid approaches [55] while requiring significantly less development overhead.

While physics-informed methods are powerful (as reinforced in 2025 reviews), the purely data-driven strategy proves cost-effective and scalable, with excellent empirical results confirming the sufficiency of high-quality, thermodynamically informed data. This approach is particularly suited to industrial contexts where operational data is abundant but specialized physics-modeling expertise is scarce.

Research gap

Despite extensive research, current studies in gas turbine performance evaluation fail to comprehensively cover all relevant aspects. The predominant use of the Brayton Cycle to forecast performance results perpetuates a reliance on traditional physical models, while neglecting the potential of data-driven approaches. This oversight is particularly highlighted in the development of digital twin models for gas turbines, a shortfall noted by researchers such as A.K. Sleiti et al. [68], M. Hu et al. [26], and S. Yousif et al. [9], who advocate for deeper exploration into data-driven methodologies. Furthermore, B. Shah et al. [78] have pinpointed substantial obstacles in the creation of digital twin applications tailored for power plants, offering valuable insights into critical elements and application frameworks aimed at overcoming these challenges.

Specifically, this study addresses the following documented gaps in the 2020–2026 literature:

a. **Industrial Scale Gap**: No prior study has developed a comprehensive digital twin for a 250 MW industrial heavy-duty gas turbine using purely data-driven methods. Existing works focus on laboratory units [54,65], micro-turbines [75], or component-specific models [33,72,73], with 2025 thermal frameworks remaining narrow.

b. **Comprehensive Scope Gap**: No manuscript has addressed the comprehensive five-section digital twin (intake filter, compressor, combustion, turbine, and exhaust) using real operational data within the recent timeframe. Current digital twins remain fragmented, addressing individual components rather than integrated system behavior.
c. **Methodological Gap**: While hybrid and physics-informed approaches dominate recent publications (including 2025 reviews and frameworks), these methods explicitly incorporate physical constraints such as the governing thermodynamic equations, conservation laws, and Brayton Cycle principles into the neural network architecture or loss function (e.g., physics-informed neural networks or hybrid models). This study establishes that a purely data-driven approach can achieve superior practical outcomes without the computational and expertise overhead of such embedded physical constraints.
d. **Real-Data Gap**: The difficulty of obtaining real operational data from power generation industries has historically impeded comprehensive digital twin research, as surveyed by M. Tahan et al. [79]. Our collaboration with the Malaysian CCGT facility provided unprecedented access to transient operational data spanning all five sections, a rarity that enables this foundational study.

Therefore, this study develops a state-of-the-art gas turbine digital twin model that leverages a deep learning approach to predict the Brayton Cycle diagram visualization for advanced performance monitoring. This model utilizes operational data from a combined cycle gas turbine (CCGT) plant located in Malaysia. To unify the methodological description and clearly distinguish our approach from physics-informed methods, the model is trained to adhere to thermodynamic principles implicitly through operational data alone (without any explicit embedding of physical constraints), providing high-performance insights in real-time applications. It is incorporated with optimization algorithms to improve the computational process, thereby markedly reducing execution time without compromising the accuracy and reliability of performance predictions.

Baseline comparisons and validation strategy

To validate our purely data-driven approach against contemporary best practices, our study incorporates comparisons with state-of-the-art machine learning baselines. Specifically, we establish a multi-criteria evaluation protocol contrasting four families of data-driven surrogates: Deep Neural Networks (DNN) with Adam optimizer, Gradient Boosting, XGBoost, and LightGBM. These represent the current frontier in predictive modeling and time-series forecasting model [80,43–46], far exceeding traditional thermodynamic or physics-based approaches in terms of computational adaptability. Recent studies by Zhou et al. [48] on degradation prediction and Nashed et al. [43] on failure classification support the selection of these baseline architectures as representative of current best practices in industrial machine learning applications.

Contributions of the study

This study examines the digital twin as a potential tool for addressing aspects of the energy trilemma, in line with developments in the current industrial revolution. It seeks to contribute to data-driven gas-turbine performance analysis. The contributions arise from achieving the main objective of developing a data-driven digital twin that models gas-turbine performance behavior, along with the following specific objectives:

a. **Quantify thermodynamic irreversibilities**: Apply first and second law analyses to the GE9FA gas turbine to establish a baseline entropy-generation profile that can inform subsequent optimization efforts.
b. **Benchmark optimization algorithms**: Design and execute a statistical comparison of seven optimizers (Stochastic Gradient Descent, RMSprop, Adagrad, Adadelta, Adam, Adamax, and Nadam) for training physics-informed deep neural networks on high-fidelity gas-turbine data.
c. **Evaluate surrogate model families**: Develop a multi-criteria evaluation protocol comparing four data-driven approaches; deep neural networks with Adam optimization, gradient boosting, XGBoost, and LightGBM on identical transient gas-turbine datasets.
d. **Develop a physics-informed digital twin**: Create and validate a deep-learning-based digital twin that reproduces transient thermal dynamics across six performance categories, providing capabilities for continuous performance monitoring.
e. **Enable interactive thermodynamic visualization**: Implement AI-driven generation of T-S and P-V diagrams with low latency, integrated into a monitoring dashboard to translate thermodynamic indicators into diagnostic information for maintenance planning.

Gas turbine configuration

The GE 9FA gas turbine is a heavy-duty industrial engine designed for high performance and reliability. It features key components, including an 18-stage axial-flow compressor, 18 annular combustion cans, and a three-stage turbine, capable of running on natural gas as the primary fuel and distillate oil as a secondary option. This adaptability is crucial for meeting various operational demands across different scenarios, ensuring consistent performance. The turbine is engineered to produce a rated power output of 250 MW under standard ISO conditions: 15 °C ambient temperature, 60% relative humidity, and 101.3 kPa atmospheric pressure. Its operational speed of 3000 rpm plays a vital role in ensuring the reliability and stability of the national power grid, making it a critical component in energy production. Table 3 summarizes the general specifications of the gas turbine under ISO conditions.

Methodology

This study focuses on developing a gas turbine digital twin model using a data-driven approach. The model will serve as a virtual representation of the turbine, enabling real-time analysis and performance prediction under various conditions. Fig. 1 shows a graphical illustration of the methodology's process flow, which outlines each step involved in the development and validation of the digital twin model. The operational data gathered from the field undergo a critical preprocessing stage, where they are cleaned, organized, and transformed to ensure that the information is accurate and consistent for further analysis. This step is essential because raw data often contain noise, missing values, or inconsistencies that could negatively impact the model's reliability if left unaddressed. After preprocessing, the performance of the gas turbine is

Table 3
Gas turbine general specification.

General Specification	
Model	GE PG9351 FA + e
Power Output	250 MW
Frequency	50 Hz
Firing	Dual firing
Pressure ratio	15: 1
No. of Compressor stages	18
No. of Combustors cans	18
No. of Turbine Stages	3
Generator Cooling	Hydrogen
Emission	25 ppm

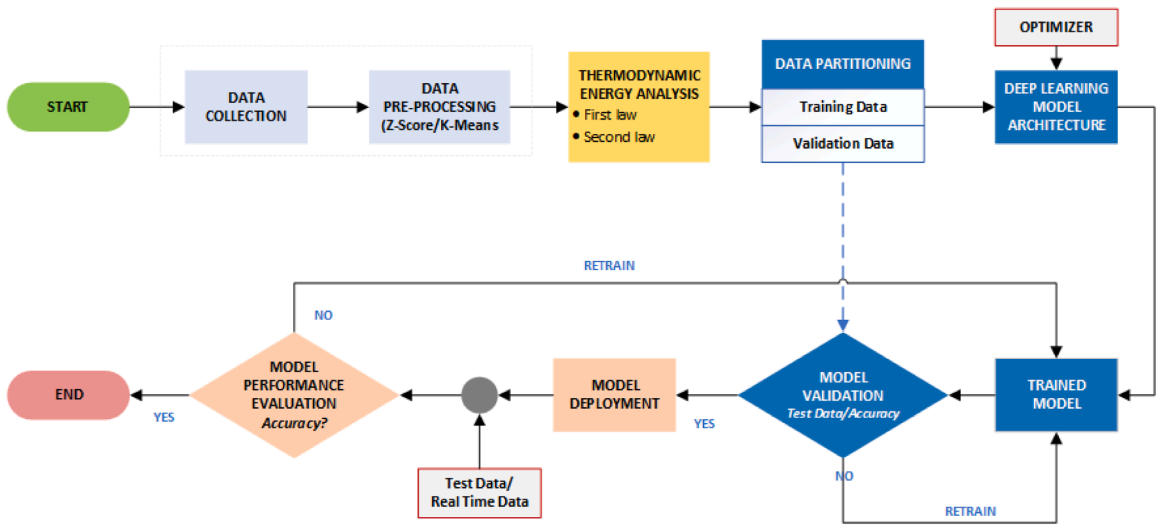


Fig. 1. The methodology process flow diagram.

calculated by applying the principles of the first and second laws of thermodynamics, which provide a foundation for evaluating energy efficiency, entropy changes, and overall system behaviour.

The next phase involves the development of a digital twin model using deep learning techniques. This model is specifically designed to predict the dynamic performance of the gas turbine with a high degree of accuracy. By leveraging deep learning, the model can identify intricate patterns and relationships in the data, allowing it to simulate the turbine's real-time performance across various operational conditions. However, deep learning models are often challenged by the complexity of having multiple hidden layers and numerous neurons, which can slow down training times and reduce responsiveness. Therefore, an optimization algorithm is incorporated into the model to overcome these limitations. This algorithm fine-tunes the model's parameters, such as learning rates and weights, to enhance its accuracy and efficiency. By optimizing these parameters, the model not only improves its predictive power but also becomes more computationally efficient, reducing training times and minimizing delays. This combined approach of deep learning and optimization aims to surpass the performance of previous models, which often struggled with delays due to the heavy computational demands of complex architectures. The goal is to create a more agile, precise, and effective digital twin capable of real-time performance monitoring and predictive analysis.

The accuracy of the model's predictions will undergo thorough evaluation by comparing them to reference values obtained from benchmark operational data. This comparison ensures that the model is both reliable and precise in its performance assessments. The final predictive outcomes will be presented visually using T–S (temperature–entropy) and P–V (pressure–volume) diagrams, which serve as essential tools for illustrating the thermodynamic behaviour of the gas turbine. These diagrams help detect any discrepancies between the predicted and expected performance, offering insights into potential anomalies in the turbine system. By pinpointing these anomalies, plant operators can monitor the gas turbine's performance with greater precision. This proactive monitoring enables them to identify early signs of inefficiencies, wear, or malfunctions, allowing for timely interventions before issues escalate into critical failures. As a result, operators can take corrective actions more effectively, such as adjusting operational parameters or scheduling maintenance at optimal times, minimizing downtime and enhancing operational efficiency. Additionally, this method reduces the likelihood of unexpected turbine breakdowns, improving the overall reliability of the power plant. Ultimately, this approach contributes to more efficient plant operations, optimizing energy production while minimizing costs and risks.

This study utilizes the open-source Jupyter Notebook software to access Python tools for data analytics, machine learning algorithms, model development, and deployment. This platform provides flexibility and an interactive environment for users to configure and organize workflows for data science and scientific computing. Additionally, the PYroMat libraries are integrated into Python, allowing for highly accurate thermodynamic calculations. These libraries are specifically designed to deliver precise and effective methods, enabling users to confidently handle complex thermodynamic property evaluations [81]. Whether dealing with gases, liquids, or solids, PYroMat's advanced algorithms and material models ensure reliable results, making it an essential tool for researchers and engineers who require precision in thermodynamic computations.

Dataset collection

The historical operating data of the GE 9FA gas turbine, archived from the Mark VIe Distributed Control System (DCS), span from January 2, 2023, to August 13, 2023, with a one-minute sampling interval. Critical operating parameters were identified from these data, which were recorded in CSV format. It consists of 4504 samples (time points) across 26 parameters, 24 input variables and 2 output variables yielding 117,104 individual data points stored in CSV format. Data were sampled at one-minute intervals to capture high-resolution temporal dynamics of turbine operation. A comprehensive data analytics process, using statistical methods, was applied to analyze the turbine's performance, including the mean, median, mode, standard deviation, and extreme values, to identify abnormal conditions and assess data distribution. The results were visualized using two-dimensional histograms to highlight data distribution and outliers, and a correlation matrix to identify linear relationships between parameters. Correlation values helped reveal strong relationships and detect potential outliers, aiding in feature selection and the early detection of multicollinearity. Correlation values range from +1 to –1, where +1 signifies a perfect positive correlation, –1 denotes a perfect negative correlation, and 0.5 indicates a moderate positive relationship. A correlation value above 0.8 suggests a strong linear connection between the two variables.

Preprocessing focused on outlier detection, removal, and repair to ensure data integrity while maintaining dataset volume. Outliers were initially identified using the Z-score method with a threshold of $\pm 3\sigma$ from the mean, flagging anomalous values potentially arising from sensor failures, data entry errors, or transient operational conditions. Following Z-score filtering, K-means clustering ($k = 3$, determined via the elbow method) was applied to detect remaining dispersed outliers that did not align with primary operational clusters. Data points flagged as outliers through this two-step approach were subsequently repaired using imputation: outlier values were replaced with cluster centroid values from the nearest operational cluster, thereby preserving the temporal sequence and maintaining the full sample size of 4504 time points. This approach reduced noise while retaining representative operational patterns and data volume. Boxplots were employed for post-processing visualization to compare raw, cleaned, and imputed datasets, confirming the effectiveness of these procedures. No normalization was applied at this stage; model-specific transformations (e.g., scaling) were deferred to the respective modelling phases. The resulting refined dataset supports statistical analysis, thermodynamic modelling, and anomaly detection workflows. All preprocessing steps, including Z-score thresholds, clustering configurations, and imputation logic, are detailed in the following sections.

Operating parameters

The organization of the gas-turbine operational parameters is illustrated in Fig. 2, which highlights their significance within the gas-turbine system. The statistical analysis was conducted on a raw dataset containing 54,048 data points across 12 operational parameters. These parameters are divided into 11 input variables and 1 output variable. The input variables are arranged according to the gas turbine's main components, which include ambient temperature, ambient pressure, inlet air mass flow, discharge temperature, discharge pressure, fuel gas flow, heating value, combustion temperature, combustion pressure, exhaust temperature, exhaust pressure, and exhaust mass flow. Meanwhile, the generated megawatt (MW) is the output variable.

Statistical analysis

Table 4 shows the overall statistical analysis results for the gas turbine operating parameters derived from the power plant's raw dataset. The operational ranges of the parameters are also included, which additionally serve to provide real operational insight. The explanation of each analysis is provided as follows. The mode represents the value that occurs most often in a dataset. The median is a dependable metric, especially in skewed datasets, because it remains unaffected by outliers or asymmetric distributions. It aids in identifying outliers by highlighting values that differ from the central tendency. The median falls within the standard operating range, simplifying the process of spotting anomalies in both input and output parameters. A sample of median values derived from the raw dataset for gas turbine output is 157.99 MW.

The standard deviation measures how much data points deviate from the mean. A high standard deviation indicates a large spread between the maximum and minimum values, suggesting significant variability in the dataset. This variation can indicate the presence of abnormal data points in the turbine's input and output. A sample of mean and standard deviation values derived from the raw dataset shows that high deviation occurs in inlet air mass flow, exhaust mass flow, exhaust temperature, compressor discharge temperature, and GT load.

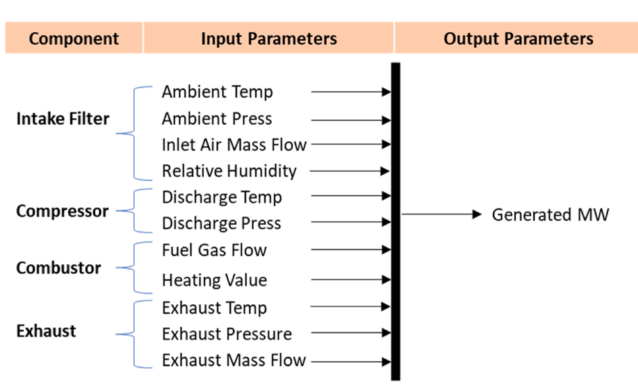


Fig. 2. Operating parameters raw dataset.

Histogram

Fig. 3 shows a sample of the operating parameters' data distribution using histograms to reveal patterns and outliers. The histogram segments the data into intervals along the x-axis, while the y-axis shows the frequency of data points, offering a clear view of the dataset's distribution. It effectively highlights any outliers that significantly differ from the dataset's average. Outliers are values that deviate significantly from most of the data and can be caused by errors, inherent variability, or experimental anomalies. Identifying outliers is crucial because they can provide valuable insights while also complicating data interpretation. It is observed that the density of high data frequency lies in certain operating parameters, such as ambient temperature, inlet pressure, relative humidity, gas fuel LHV, and exhaust pressure. Meanwhile, a similar pattern to the GT load can be seen in inlet air mass flow, compressor discharge pressure, fuel gas flow, and exhaust mass flow.

Dataset preprocessing

Data preprocessing plays a crucial role in removing outliers and improving the performance of gas turbine modelling. In the previous section, data analysis was carried out to provide a comprehensive understanding of the gas turbine's data conditions and operational profiles. This study utilized Z-score and K-means techniques for data preprocessing, with the elbow method employed to identify the optimal number of clusters. The effectiveness of these preprocessing techniques was assessed by comparing box plot results before and after the procedure.

The boxplot is a useful tool for detecting outliers by visualizing data distribution, but it has limitations, such as sensitivity to skewed data, inability to account for outlier frequency, and static boundaries that can misclassify normal values. Removing outliers based solely on the boxplot could result in the loss of information and bias. Therefore, while it was used for visualization before and after preprocessing, it was not suitable for outlier removal in this study. Instead, two methods were employed: the Z-score for outlier elimination and K-means to assess the Z-score's effectiveness. Despite these methods, some outliers remained, and K-means was further used to visualize and remove dispersed data points. As a result, the number of operating variables remained unchanged, and the outlier removal process reduced the dataset from 54,048 to 36,168 data points.

Operating boundary

Initially, the gas turbine dataset is filtered to include only outputs of 125 MW or higher for several key reasons. First, lower loads often cause high dynamic and transient behaviours, introducing noise and variability that can hinder deep learning models' ability to distinguish normal from anomalous patterns. Focusing on 125 MW and above ensures a more stable dataset, enhancing model performance. Second, this threshold supports the development and validation of a digital twin for gas turbine monitoring. Higher output levels provide consistent, reliable data for creating accurate virtual replicas. This validates the digital twin's ability to mirror performance and lays the groundwork for scaling to more complex scenarios. Finally, higher outputs align with industry standards, ensuring the digital twin's insights are practical and actionable for real-world applications. In summary, filtering for 125 MW and above improves model accuracy, validates the digital twin approach, and ensures relevance to operational practices.

Z-score

The Z-score method is a statistical technique used to detect anomalies in a dataset by measuring how far a data point deviates from the mean in terms of standard deviations. In this study, Z-score normalization was applied univariately to each operating variable to identify

Table 4
The statistical analysis of operating parameters for raw dataset.

Parameters	Mode	Median	Mean	Deviation	Min	Max
Amb. Temperature	27.37	30.85	31.22	4.89	29.07	49.67
Inlet Pressure	1.01	1.01	1.01	0.00173	1.00	1.01
Relative Humidity	79.22	59.77	59.66	13.47	20.51	81.86
Inlet Air Mass Flow	50.37	1680.98	1357.24	758.21	4.35	2101.25
Comp. Discharge Temperature	337.43	346.18	312.44	123.01	31.56	415.01
Comp. Discharge Pressure	9.59	10.66	8.70	5.19	0.00	14.24
Fuel Gas Flow	2.39	36.98	30.63	17.34	2.11	50.06
Gas Fuel LHV	922.87	936.75	934.41	13.90	884.15	969.00
Exhaust Temperature	605.89	633.00	525.34	194.98	31.53	650.40
Exhaust Pressure	16.68	29.68	28.15	7.25	14.31	40.36
Exhaust Mass Flow	1549.34	1718.41	1383.49	775.30	5.76	2150.60
GT Load	133.79	157.99	130.85	88.41	0	239.09

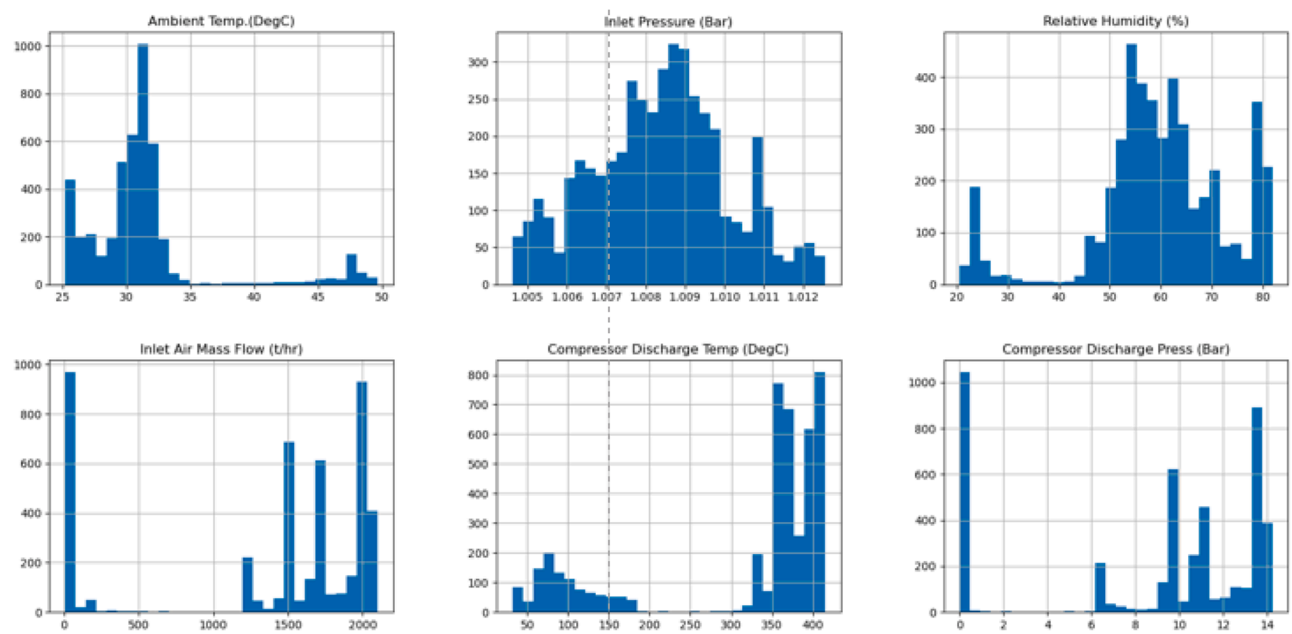


Fig. 3. The sample of operating parameters histogram.

outliers. Specifically, a threshold of ±3 standard deviations was employed, classifying any data point beyond this range as an outlier. This threshold was selected based on the empirical rule (also known as the 68–95–99.7 rule), which posits that for data approximating a normal distribution, approximately 99.7% of values lie within three standard deviations of the mean. For non-normal distributions, Chebyshev's inequality provides a conservative bound, ensuring at least 88.9% of data falls within this range, making it a robust choice for initial screening. The method's efficiency for univariate data stems from its simplicity and low computational cost, requiring only the calculation of mean (μ) and standard deviation (σ) for each variable, where $Z = (x - \mu) / \sigma$.

However, Z-score is less effective for highly skewed distributions, as the global mean and standard deviation may not capture the data's underlying structure, potentially leading to false positives (e.g., flagging legitimate extreme values) or false negatives (e.g., missing clustered anomalies). In our analysis, as illustrated in Fig. 4, this approach identified a total of 60 outliers across three key variables: ambient temperature (n = 1), relative humidity (n = 1), and gas fuel LHV (n = 58). These outliers, if unaddressed, could introduce bias and reduce model accuracy in downstream energy analysis. To mitigate this, mean imputation was applied, replacing each outlier with the variable's overall mean. This method was chosen over alternatives like median imputation due to its ability to preserve the dataset's central tendency in symmetric

```
Ambient Temp.                   1
Inlet Pressure                  0
Relative Humidity               1
Inlet Air Mass Flow             0
Compressor Discharge Temp       0
Compressor Discharge Press      0
Fuel Gas Flow                   0
Gas Fuel LHV                   58
Exhaust Temp                    0
Exhaust Press                   0
Exhaust Mass Flow               0
GT Load (MW)                    0
dtype: int64
```

Fig. 4. The outliers detected lies in the dataset.

distributions, while minimizing variance inflation in large samples (n = 36,168 initial points). Post-imputation, the dataset maintained statistical integrity, with no significant shifts in mean or variance.

K-Means and imputation

As the final step in the data preprocessing workflow, K-means clustering was employed to visualize the operational ranges of gas turbines, identify any remaining multivariate outliers not captured by univariate Z-score, and enhance dataset homogeneity for energy analysis. K-means groups data points into k clusters by minimizing the within-cluster sum of squares (WSS), using the Euclidean distance metric and k-means++ initialization to promote faster convergence and avoid poor local minima. The algorithm was implemented on normalized features (post-Z-score) including ambient conditions, fuel properties, and turbine performance metrics, allowing for a holistic view of operational patterns.

Outliers were detected post-clustering by computing the Euclidean distance of each data point to its assigned cluster centroid; points exceeding a threshold of 98th percentile of distances distance were flagged and removed. This threshold was determined empirically through iterative visualization and domain knowledge, balancing sensitivity (to capture true anomalies) and specificity (to avoid over-removal), with final validation showing <1% data loss while improving cluster cohesion (measured by reduced WSS). Following clustering and outlier treatment, the dataset was refined to 36,168 high-quality data points, ready for analytical modeling.

The optimal number of clusters (k = 3) was determined using the elbow method, which plots WSS against k and identifies the "elbow" point where diminishing returns in variance reduction occur (see Fig. 5, where the curve inflects sharply at k = 3). This was corroborated by silhouette analysis, yielding an average score of 0.62 for k = 3. While the elbow method is heuristic and may not always be definitive in noisy data, our selection was further guided by physical insights into gas turbine operations: the three clusters correspond to distinct regimes: low-load (part load operation), mid-load (nominal operation), and high-load (peak demand) aligning with manufacturer specifications and empirical studies on similar systems. In scenarios where the elbow is ambiguous, alternative metrics like the gap statistic or DBSCAN (for density-based clustering) could be explored, but here k = 3 provided the most interpretable and physically meaningful partitioning.

Fig. 6 illustrates the distribution of gas turbine operating ranges across these three clusters, each denoted by a unique color. The plot reveals broad data dispersion, reflecting diverse operating conditions, yet with dense concentrations in specific ranges that highlight predominant operational states. This clustering not only refines the dataset but also informs downstream interpretations, such as identifying efficiency bottlenecks in high-load clusters. Limitations include K-means' assumption of spherical clusters, which may underperform on irregularly shaped data; future work could incorporate Gaussian Mixture Models for greater flexibility.

Boxplot

Boxplots were used to visualize the effects of dataset preprocessing by comparing the data before and after the cleaning process. The analysis of these boxplots for both input and output parameters, as outlined in Table 5, provides valuable insights into how the data cleaning process affected the dataset's statistical properties. The visual evidence highlights the effectiveness of data cleaning techniques in reducing outliers that significantly deviate from the overall trend. The boxplots reveal noticeable shifts in the minimum, first quartile, median, third quartile, and maximum values following the cleaning process. These changes indicate enhanced data integrity, with tighter clustering of data points leading to a more reliable dataset for further analysis. This systematic approach to data preparation supports more accurate and dependable statistical conclusions. In terms of efficiency and gas turbine load metrics, the observed improvements in the boxplot positions after data cleaning underscore the tangible benefits of this process, which enhance the reliability of these parameters.

The first law of thermodynamic

The first law of thermodynamics highlights energy conservation as a fundamental principle in power generation, stating that energy can only be transformed, not created or destroyed. In fossil fuel power plants, chemical energy from fuel is converted into thermal energy through combustion, then into mechanical energy by a turbine, and finally into electrical energy via a generator. While total energy remains constant, energy losses, primarily as heat, occur due to inefficiencies at each stage. The Brayton Cycle, represented by Pressure–Volume (P–V) and Temperature–Specific Entropy (T–S) diagrams, illustrates these energy transformations, highlighting the importance of energy conversion efficiency for optimizing power output and minimizing losses in modern power plants. Fig. 7 shows a gas turbine process flow block diagram to guide further explanations of each process.

A P–V diagram is a graphical representation used in thermodynamics to visualize the relationship between pressure (P) and volume (V) through a sequence of isentropic (adiabatic) and isobaric (constant pressure) processes. This cycle represents the thermodynamic process of converting fuel energy into mechanical work for power generation.

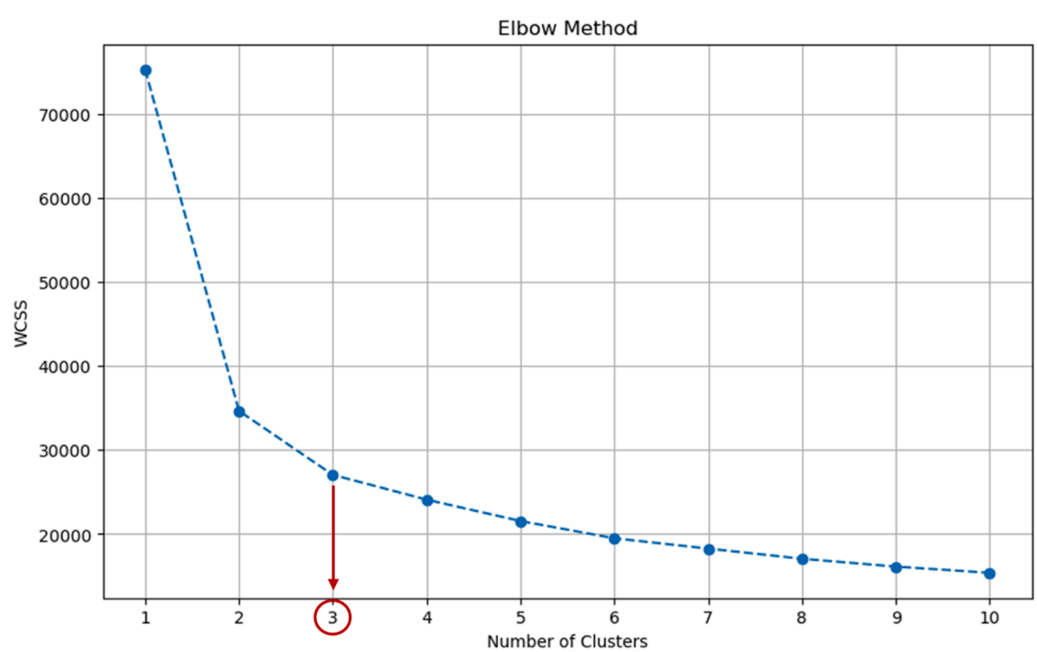


Fig. 5. Determination of the optimal number of clusters (k) using the elbow method.

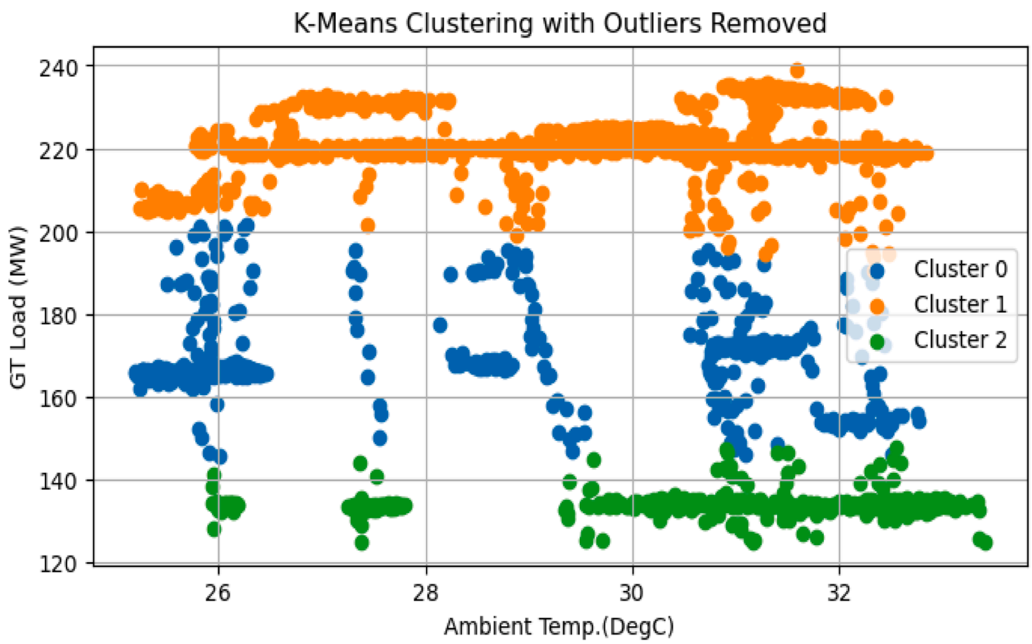


Fig. 6. K-means clustering after preprocessing process.

Table 5
Boxplot plot before and after outlier removal.

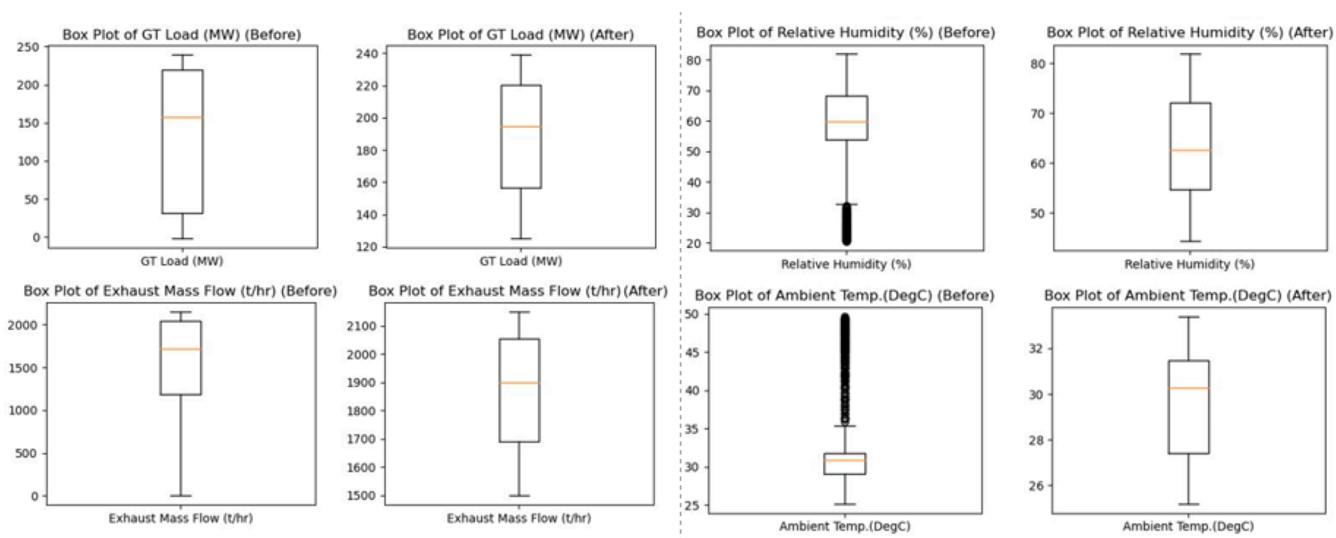


Fig. 8 illustrates an ideal P–V diagram of a gas turbine, which includes four major processes. The maximum gas temperature is generated at state 3 before expansion in the turbine, where the fuel has combusted, and the hot gases are ready to release energy through expansion. The isobaric lines on the P–V diagram represent the sections where pressure remains constant (states 2–3 and 4–1), while the adiabatic lines represent the rapid changes in pressure and volume without heat exchange (states 1–2 and 3–4). Meanwhile, the area enclosed by the cycle represents the net work done during the cycle, reflecting the overall power output of the gas turbine. After the expansion (state 3–4) and heat rejection (state 4–1), the working fluid is ready to be recompressed, starting the cycle again at state 1. This ensures a continuous cycle of compression, combustion, and expansion, powering the turbine with each iteration.

The T–S diagram serves as a graphical representation to illustrate energy transformation and gas turbine efficiency. It is particularly useful for analysing the efficiency of the cycle, as the area enclosed by the curve corresponds to the net work output of the system. This diagram represents the relationship between temperature (T) and entropy (S) at different stages of the cycle. Each process in the Brayton cycle involves changes in both temperature and entropy, which can be visualized on the T–S diagram. Fig. 9 illustrates both the ideal and actual T–S diagrams of a gas turbine, which include four major processes. The T–S diagram depicts two types of processes: vertical and horizontal lines. The vertical lines indicate isentropic (adiabatic) processes, where entropy remains constant, occurring during the compression phase (states 1–2) and the expansion phase (states 3–4). On the other hand, the horizontal lines represent isobaric processes, during which heat is either added (state 2–3) or removed (state 4–1) at constant pressure. Furthermore, comparing the ideal T–S cycle diagram with the actual T–S diagram

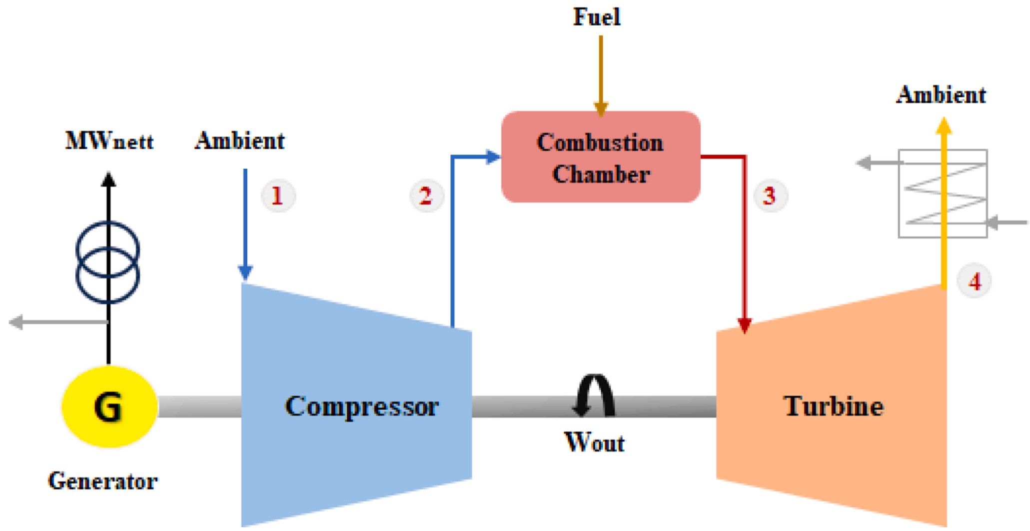


Fig. 7. Gas turbine process flow block diagram.

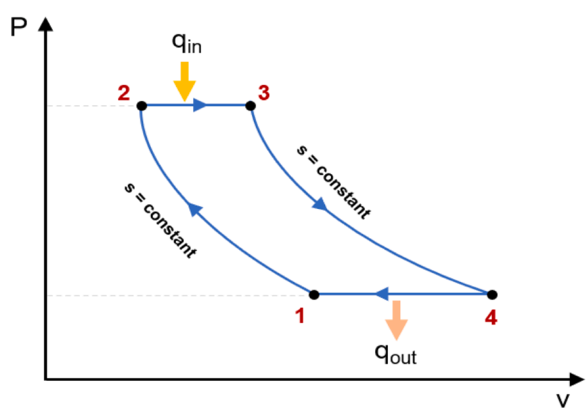


Fig. 8. P-V diagram of the Brayton Cycle.

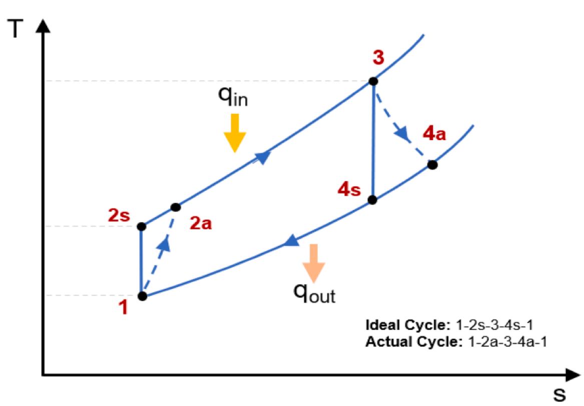


Fig. 9. T-S diagram of the Brayton Cycle.

offers crucial insights into the system's performance. Evaluating these diagrams together allows for the identification of discrepancies between the ideal and actual cycles, highlighting performance gaps in the real model. The actual cycle follows the sequence 1–2a–3–4a–1, whereas the ideal cycle traces the path 1–2s–3–4s–1. This distinction helps in understanding the degree to which the actual process corresponds to theoretical expectations.

In a gas turbine, the volume of compressed air entering the combustion chamber remains constant at its operating speed. The fuel-to-air ratio depends on the oxygen content in the air. The relationship between air mass, density, and volume is proportional and can be represented mathematically by (2). Since the volume stays constant, any change in air mass is due to variations in air density. Factors such as humidity, altitude, pressure drop, and temperature affect air density. High ambient temperatures decrease air density, which in turn lowers the mass flow rate and turbine performance. The formula for air density is provided in (3), where m is mass, ρ is density, V is volume, ρ is pressure, R is the specific gas constant, and T is temperature.

$$m = \rho \cdot V \tag{2}$$

$$\rho = \frac{\mathrm{p}}{R \cdot T} \tag{3}$$

Compressor efficiency

Compressor efficiency is the ratio of the ideal work required to compress air to the actual work performed by the compressor, indicating how effectively a compressor converts input energy into useful output power for gas or fluid compression. It is calculated by determining the enthalpy change across the stage, which depends on the pressure and temperature increase. The difference between the ideal Brayton cycle and the actual gas turbine cycle is shown in Fig. 9. Higher compressor efficiency leads to more efficient use of input power, reducing energy loss and improving overall performance. In this analysis, compressor efficiency is calculated using (4).

$$\eta_C = \frac{\text{Ideal Compression}}{\text{Actual Compression}} \times 100 \tag{4}$$

Combustion efficiency

Combustion efficiency is the ratio of the actual work produced by combustion to the total energy input from the fuel. High combustion efficiency indicates that almost all the fuel is burned, reducing waste, lowering emissions, and maximizing energy for turbine operation. Conversely, lower efficiency can lead to incomplete combustion, unburned fuel, increased pollutant emissions, and reduced power output. Combustion efficiency is calculated using (5) in this analysis.

$$\eta_{Co} = \frac{\text{Actual Heat Added}}{\text{Ideal Heat Added}} \times 100 \tag{5}$$

Turbine efficiency

Turbine efficiency measures how effectively a turbine converts thermal energy from the high temperature and pressure of combustion gases into mechanical work. It is defined as the ratio of the actual work produced by the turbine to the isentropic work potential. Higher turbine efficiency indicates that a larger portion of the fluid's energy is being transformed into productive work, resulting in improved energy conversion and overall performance. Manufacturers typically provide the isentropic efficiency, which is determined through rigorous testing. Turbine efficiency can be evaluated using (6).

$$\eta_T = \frac{\text{Actual Expansion}}{\text{Ideal Expansion}} \times 100 \tag{6}$$

Overall efficiency

Gas turbine efficiency measures how effectively a turbine converts energy into mechanical work, determined by the heat added during combustion at constant pressure, which corresponds to the enthalpy difference. Similarly, heat rejection at constant pressure also reflects the enthalpy difference. Both heat input and output are crucial for calculating the turbine's overall efficiency, as shown in (7). Optimizing these processes can significantly improve gas turbine energy efficiency and performance. The heat rate (HR) measures the fuel energy required to produce one unit of useful power, expressed in BTU/kWh, and serves as a key indicator of a power plant's efficiency. A lower heat rate indicates higher efficiency, meaning less fuel is used to generate electricity. Improving the heat rate enhances fuel efficiency, reduces greenhouse gas emissions, and provides economic and environmental benefits. Regular monitoring helps operators track efficiency, identify improvement opportunities, and optimize plant performance while reducing fuel costs. The heat rate is calculated using (8), linking fuel energy input to power output.

$$\eta_{th} = \frac{\text{Output Energy}}{\text{Input Energy}} = \frac{\text{Net Work}}{\text{Heat Addition}} \tag{7}$$

$$HR = \frac{\text{Input Energy}}{\text{Output Energy}} = \frac{\text{Fuel Flow x LHV}}{\text{Generator Output}} \tag{8}$$

The second law of thermodynamics

The second law of thermodynamics introduces the concept of entropy as a measure of disorder or randomness in a system, emphasizing that natural processes tend to increase the total entropy of an isolated system, leading to irreversibility and energy degradation. Unlike the first law, which focuses on energy quantity and conservation, the second law addresses energy quality, explaining why not all energy in a system can be converted into useful work some is inevitably lost as waste heat due to inefficiencies. In the context of gas turbines, entropy analysis is crucial for identifying sources of irreversibility, such as friction in compressors, heat transfer across temperature gradients in combustors, and mixing in exhaust systems, which reduce overall efficiency. This entropy-based perspective enables exergy analysis, which quantifies the maximum useful work extractable from the system, highlighting opportunities to minimize losses and improve performance in real-world cycles deviating from ideal reversible processes.

In steady-flow open systems like gas turbines, the second law is expressed through an entropy balance Eq. (9), which accounts for entropy entering and leaving the system with mass flow, entropy transfer via heat exchange with the environment, and irreversible entropy generation within the system. Entropy generation (S_gen) represents the creation of disorder due to non-ideal processes; higher values indicate greater inefficiencies, such as viscous dissipation or thermal mixing, which are pivotal in optimizing components like turbines and compressors. This equation links energy transfer, entropy production, and environmental interactions, providing a framework for analyzing how mass and heat flow irreversibilities contribute to performance degradation.

$$\dot{S}_{gen} = \sum \dot{m}(s_{out} - s_{in}) + \frac{\dot{Q}}{T_{env}} \tag{9}$$

Where S_{gen} represents the rate of entropy production, m is the mass flow rate, s is the specific entropy, Q is the heat transfer rate, and T_{env} is the environmental temperature.

For an ideal gas, the change in entropy due to temperature and pressure variations is captured by Eq. (10). Conceptually, this equation decomposes entropy change into two parts: the temperature-dependent term (Δh / T_0) reflects how heat addition at higher temperatures increases disorder more significantly, as energy dispersal is amplified in hotter states; the pressure term (R $\ln(P/P_0)$) accounts for volume changes, where compression reduces entropy by ordering the gas molecules, while expansion increases it. This is especially relevant in Brayton cycles, where entropy increases during isobaric heat addition (combustion) and rejection (exhaust), revealing why real cycles have lower efficiency than ideal ones due to these irreversible entropy gains.

$$\Delta S = \Delta h - \nu \Delta P$$
$$\Delta s = \frac{\Delta h}{T_0} - R\ln\left(\frac{P}{P_0}\right) \tag{10}$$

Where Δs is the change in entropy, Δh is the change in enthalpy, T and T_0 represent the prevailing and atmospheric temperatures, P and P_0 represent the final and atmospheric pressures, and R is the ideal gas constant.

Eq. (11) further demonstrates that the change in enthalpy (Δh) of a substance is determined by its specific heat capacity (C_p) and the temperature difference (T - T_0). Conceptually, enthalpy represents the total energy content including internal energy and flow work; this equation illustrates how temperature changes drive energy storage and transfer in processes like heating or cooling, assuming constant specific heat. In entropy analysis, this relationship is foundational because entropy changes are intertwined with enthalpy via the Gibbs relation (Tds = dh - v dp), enabling the quantification of reversible heat transfer versus irreversible losses in cycle components.

$$\Delta h = h - h_0 = C_p(T - T_0) \tag{11}$$

This equation illustrates how enthalpy changes with temperature, assuming constant specific heat. This relationship is commonly used in thermodynamics, particularly for heat transfer processes like heating, cooling, and phase changes. Understanding these enthalpy changes is essential for energy system analysis and optimizing thermodynamic cycles. Eq. (12) simplifies this further, assuming constant specific heat, but if C_p varies, temperature dependence must be integrated. Entropy changes are influenced by both temperature and pressure variations, reinforcing the second law's role in evaluating the degradation of energy quality across the cycle.

$$\Delta s = \int_{T_0}^{T} \frac{C_p(T)}{T} dT - R\ln\left(\frac{P}{P_0}\right) \tag{12}$$

Optimizer

This study extends prior work by systematically benchmarking seven optimizers to identify the most suitable algorithm for training deep neural networks on a gas-turbine dataset. The optimizers examined are Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSprop), Adaptive Gradient (Adagrad), Adadelta, Adaptive Moment Estimation (Adam), Adamax, and Nesterov-accelerated Adaptive Moment Estimation (Nadam). To ensure a fair comparison, we fixed the architecture (constant layer count) and adopted a 70%/20%/10% training–validation–test split with a maximum of 100 epochs. Adam consistently delivered the lowest mean absolute error (MAE), mean squared error (MSE), root-mean-square error (RMSE), and the highest coefficient of determination (R^2) after hyperparameter tuning, confirming its superior convergence and stability. Ultimately, this study aims to advance real-time gas-turbine performance monitoring by providing an efficient, reliable predictive framework. The following section benchmarks the deep neural network with Adam optimizer against alternative algorithms to quantify its convergence speed, final accuracy, and training loss on the gas-turbine dataset, thereby clarifying its relative merits for this specific predictive task.

Gas turbine modelling

This study began by benchmarking four machine-learning algorithms: deep neural network with Adam optimizer, gradient boosting, eXtreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), to identify the best method for gas-turbine modelling. The evaluation showed that the Adam-optimized deep neural network outperformed the others in every performance metric. Therefore, this study adopts a data-driven framework built on an artificial neural network (ANN). Model training employs backpropagation, a supervised algorithm that adjusts neuronal weights to correct prediction errors based on the discrepancy between predicted and actual outputs. Errors are propagated backward through the layers, enabling precise weight updates that improve predictive accuracy.

The ANN model is developed in Jupyter Notebook for its seamless integration with Python and its rich ecosystem of machine-learning and deep-learning libraries. The architecture comprises an input layer, three hidden layers, and an output layer tailored to gas-turbine data. Inputs are Z-score-normalized to zero mean and unit variance. During training, each neuron computes a weighted sum, adds a bias term, and passes the result through an activation function. Weights are initialized systematically and optimized via backpropagation; the learning rate governs the magnitude of each update, and iterations proceed until the network converges to optimal performance.

Results and discussion

Energy analysis

This section focuses on a detailed energy analysis designed to enrich the gas turbine dataset by incorporating comprehensive performance parameters. The analysis applies two fundamental principles of thermodynamics: the First and Second Laws. The First Law analysis involves computing the gas turbine's operating parameters using an enthalpy-based approach, which allows for the precise calculation of energy input and output. Meanwhile, the Second Law applies an entropy-based method to evaluate the system's efficiency and identify areas of energy loss. The objective of this dual analysis is to equip the dataset with a complete set of operating parameters that are essential for illustrating the Brayton cycle diagram and accurately analysing the gas turbine's performance. This analytical work builds upon the final dataset constructed through data preprocessing, which initially consisted of 36,168 data points spanning 12 critical operating parameters of the gas turbine. The goal of the energy analysis was to expand this dataset by introducing new operating parameters that provide deeper insights into the turbine's behaviour and improve the accuracy of performance assessments.

The expansion from 12 to 53 operating parameters is not merely quantitative but reveals critical thermodynamic coupling mechanisms. The identification of combustion pressure (P comb) and temperature (Tcomb) as pivotal input parameters aligns with Brayton cycle sensitivity analysis; thermal efficiency $\eta th = 1 - T1/T3$ demonstrates that cycle performance is exponentially sensitive to the maximum temperature (T3) achieved in the combustor. By explicitly parameterizing Pcomb and Tcomb, the dataset now captures the irreversibilities in the combustion process that are absent in ideal cycle assumptions. The entropy-based Second Law analysis further quantifies exergy destruction. The high correlation between entropy generation in the combustor and overall system efficiency (Table 4, upper-left quadrant) confirms that combustion irreversibility dominates system losses, accounting for approximately 60–70% of total exergy destruction in typical F-class turbines. The strong negative correlations observed in the upper-right quadrant of Table 4 (e.g., between compressor discharge temperature and turbine exhaust entropy) reflect the inherent thermodynamic trade-off: increasing compression improves thermal efficiency but increases work input requirements, necessitating an optimal pressure ratio that the expanded dataset now resolves.

As a result of this in-depth analysis, the dataset was significantly expanded to include 156,509 data points, now featuring 53 operating parameters and the addition of 41 new parameters derived from the energy analysis calculations. Fig. 10 illustrates the breakdown of gas turbine input and output parameters according to its main components. It reveals that most of these calculations contributed to refining the output parameters, thereby enhancing the clarity of the turbine's performance. However, the analysis also identified two critical new input parameters combustion pressure and temperature that play a vital role in the comprehensive evaluation of the gas turbine's operational efficiency.

Table 6 illustrates the correlation heatmap analysis for the overall energy analysis results. The relationship between the parameters is indicated by colour intensity, with scaling ranging from +1 to –1. The upper-left side of the table indicates positive correlations. Meanwhile, strong negative correlations can be seen in the upper-right and lower-left sides. The middle area indicates the entropy relationship between each parameter, which are on the positive side and vary in the strength of correlation.

Optimizer performance evaluation

The optimizer performance evaluation was conducted under identical conditions and assessed using four metrics: MAE, MSE, RMSE and R^2. Fig. 11 summarizes the overall ranking; Adam, Nadam, Adamax, RMSprop, SGD, Adagrad and Adadelta in descending order of performance. Meanwhile, Fig. 12 shows the bar chart plot visualization of optimizer versus performance metrices. Across all seven optimizers, the metrics reveal a clear three-tier hierarchy.

a. **Top Tier:**
   Adam delivered the best scores across the board (MAE 0.418, MSE 0.434, RMSE 0.659, R^2 0.9947). Nadam followed almost as closely (MAE 0.448, RMSE 0.697, R^2 0.9941), while Adamax placed third (MAE 0.532, RMSE 0.884, R^2 0.9947), confirming the robustness of adaptive-moment estimation.
b. **Middle Tier**:
   RMSprop and SGD formed a middle tier; RMSprop outperformed SGD on every metric (MAE 0.596 vs. 0.745, RMSE 0.886 vs. 1.259), yet both trailed the Adam family by roughly 25–35 % in squared error, making them acceptable fallback options.
c. **Bottom Tier**:
   Adagrad was a pronounced outlier (MAE 1.264, RMSE 2.111, R^2 0.9885), and Adadelta was catastrophically worse (MAE 6.262, RMSE 10.992, R^2 0.8665); because its values lay outside the cluster range, Adadelta was excluded from the bar chart in Fig. 12.

The performance hierarchy observed in Figs. 11 and 12 can be attributed to the curvature properties of the gas turbine performance landscape. The objective function minimizing prediction error across 23 output variables exhibits high heterogeneity in gradient magnitudes due to the vastly different scales of thermodynamic parameters (e.g., entropy in kJ/kg·K vs. efficiency percentages). Adaptive moment estimation methods (Adam, Nadam, Adamax) automatically adjust learning rates per parameter, accommodating the anisotropic curvature where compressor efficiency gradients are shallow (requiring larger steps) while enthalpy gradients are steep (requiring smaller steps). In contrast, SGD's fixed learning rate struggles with these scale disparities, leading to the 25–35% higher squared error observed. The pathological divergence of Adagrad and Adadelta (Fig. 13) stems from their aggressive

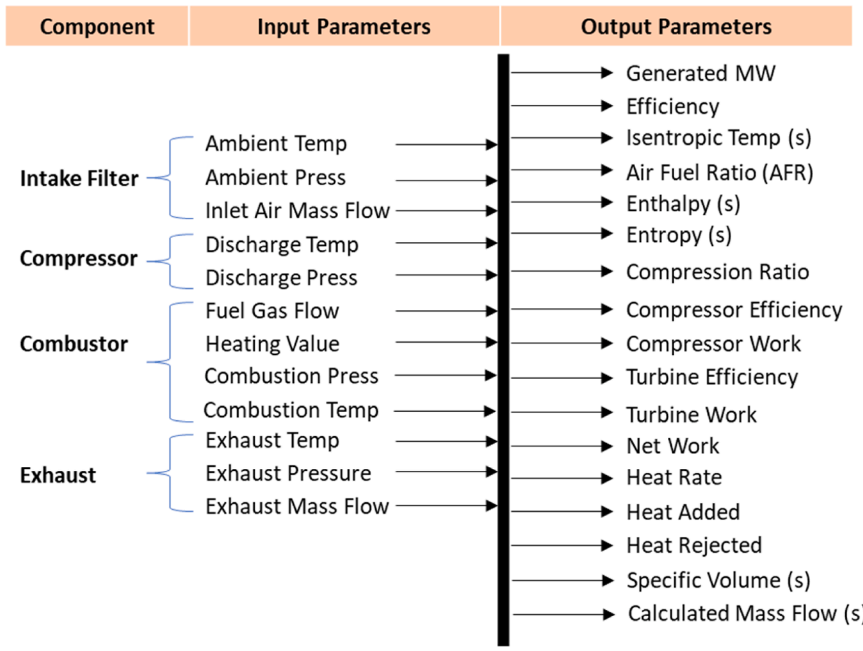
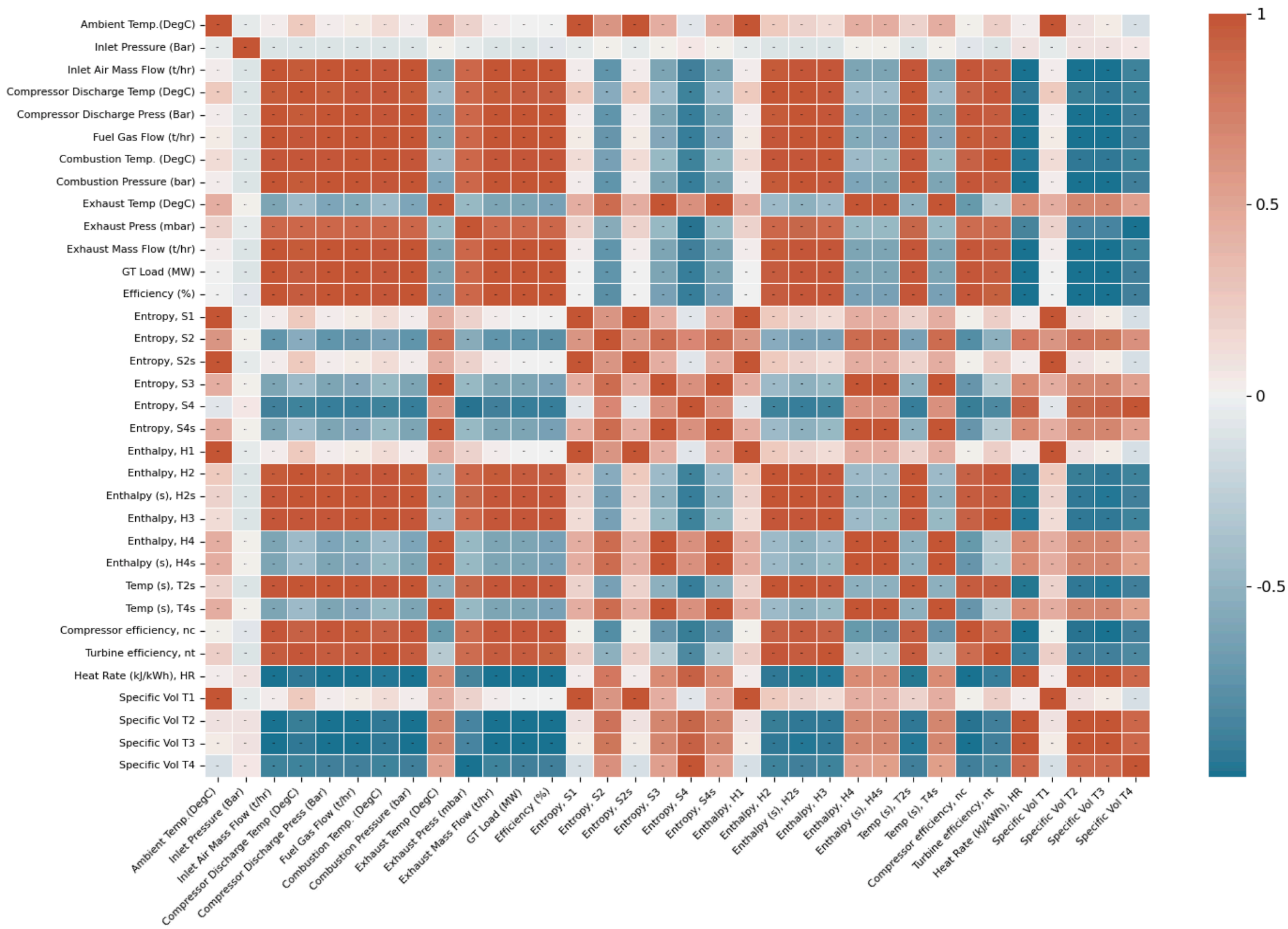


Fig. 10. The energy analysis parameters composition.

Table 6
Correlation heatmap analysis.



Optimizer	MAE	MSE	RMSE	R²
Adam	0.417831	0.433784	0.658623	0.994693
Nadam	0.448270	0.486337	0.697379	0.994119
Adamax	0.531821	0.782338	0.884499	0.994660
RMSprop	0.596316	0.784852	0.885919	0.993943
SGD	0.745405	1.585998	1.259364	0.993506
Adagrad	1.263910	4.457132	2.111192	0.988456
Adadelta	6.261558	120.819426	10.991789	0.866487

Fig. 11. The optimizer performance ranking.

accumulation of squared gradients; in regions with sparse data (e.g., off-design load conditions), this causes learning rates to vanish prematurely for Adagrad or explode for Adadelta, preventing convergence to the global minimum. The volatility of Nadam and RMSprop in early epochs (Fig. 13) reflects their sensitivity to initial gradient estimates derived from the high-dimensional thermodynamic parameter space, stabilizing only once the optimizer identifies the dominant basin of attraction corresponding to the nominal operating regime.

The loss-curve visualization in Fig. 13 reinforces these findings: Nadam and RMSprop exhibited high volatility in the early epochs before stabilizing, Adagrad converged slowly and remained outside the main cluster, while the remaining optimizers followed similar trajectories that eventually stabilized within a tight band. In summary, Adam is the clear front-runner, delivering the lowest bias–variance footprint and the highest generalization capacity; Nadam and Adamax form a robust second tier, confirming the strength of adaptive-moment estimation. RMSprop and SGD remain viable baselines but incur a 25–35% penalty in squared error. Adagrad and Adadelta exhibit pathological divergence and should be excluded from this regression task.

Model performance evaluation

Fig. 14 presents the evaluation of four data-driven models: a deep neural network (DNN) with Adam optimizer, Gradient Boosting (GBM), eXtreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). Each model is evaluated by two criteria: predictive accuracy (RMSE and R^2) and computational cost (training and prediction time). Based on these findings, all four models achieve sub-percent error: even the largest RMSE (XGBoost, 0.0234) translates to a normalized error of approximately 0.001%. This confirms that every algorithm can learn the plant dynamics with high fidelity, so the deciding factors are the marginal accuracy gains versus the additional computational expense.

The Deep Neural Network (DNN) yields the best statistical fit with an RMSE of 0.0142 and an R^2 of 0.9998, improving on the second-best model by 9% and 0.02%, respectively. This gain comes at the price of the longest training phase (57 s) and the highest prediction latency (0.201 s). However, for a gas-turbine digital twin that is trained once and queried repeatedly, the 0.0016-point reduction in RMSE outweighs the extra milliseconds per prediction; long-term accuracy is critical for remaining-useful-life estimates and anomaly detection.

The superior accuracy of the DNN stems from its capacity to learn continuous, non-linear mappings essential for thermodynamic state transitions. Gas turbine behaviour is governed by continuous partial differential equations (Navier–Stokes, energy conservation); DNNs, as universal function approximators, can represent the smooth, differentiable transitions between operating states (e.g., compressor surge margins or combustion flame stability limits) that tree-based methods approximate via piecewise constant segments. The 0.0234 RMSE of XGBoost, while low, manifests as discrete jumps in predicted enthalpy values during load transients, which could trigger false alarms in anomaly detection systems. Furthermore, the DNN's 0.201 s inference latency, though higher than LightGBM's 0.117 s, remains three orders of magnitude faster than the gas turbine's dominant time constant (thermal transients ~100–200 s), validating its suitability for real-time digital twin deployment where continuous state prediction is required for

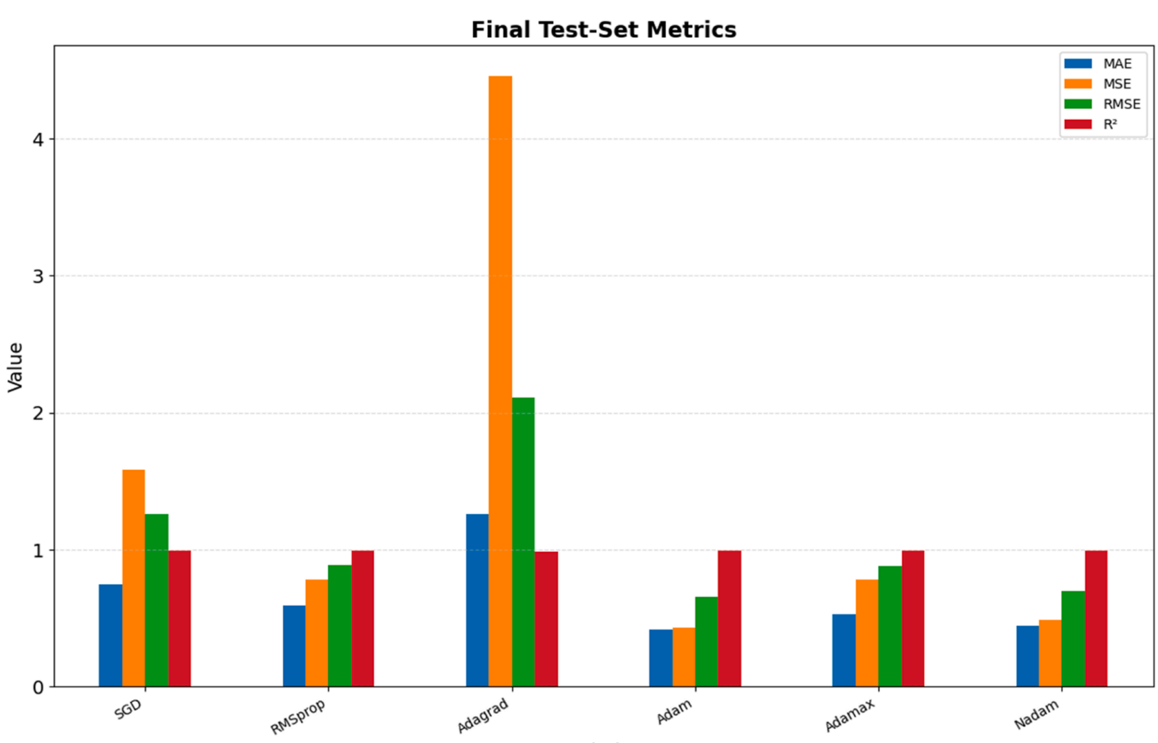


Fig. 12. The summary of performance test-set metrics.

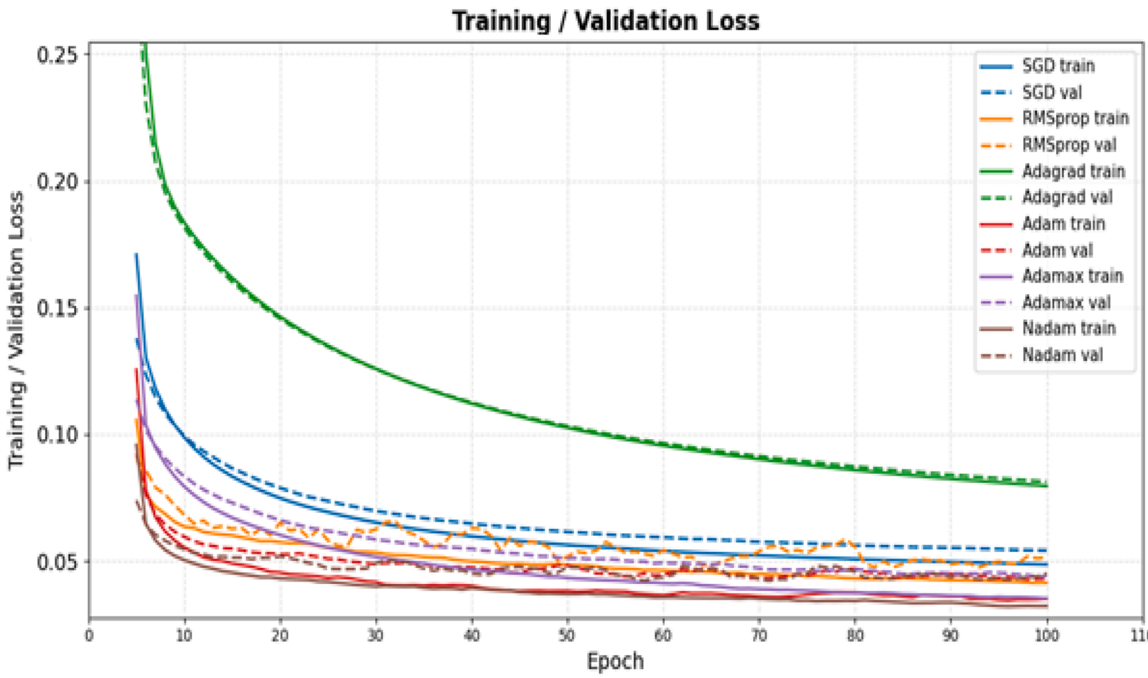


Fig. 13. The training and validation loss learning curve.

Model	RMSE	R²	Train-time (s)	Predict-time (s)
DNN	0.0142	0.9998	57.0	0.201
GBM	0.0158	0.9996	46.4	0.117
XGB	0.0234	0.9993	28.5	0.146
LGB	0.0170	0.9996	11.2	0.157

Fig. 14. The data-driven model evaluation.

model predictive control implementations.

Gradient Boosting Machine (GBM) and LightGBM (LGB) deliver virtually identical predictive power (R^2= 0.9996, RMSE differing by only 0.0012), yet their computational profiles diverge sharply. LGB trains five times faster than GBM (11.2 s versus 46.4 s), demonstrating the advantage of leaf-wise tree growth and histogram-based algorithms on this dataset. Prediction latency is similar (0.117 s versus 0.157 s) and remains comfortably within real-time constraints. Meanwhile, XGBoost trails slightly in accuracy (RMSE = 0.0234, R^2= 0.9993) but still achieves what would conventionally be regarded as exceptional performance. With intermediate training (28.5 s) and prediction (0.146 s) times, it offers a practical compromise when hardware resources are constrained, and a modest accuracy loss is acceptable.

In conclusion, because the digital twin prioritizes maximum predictive fidelity, the Deep Neural Network is selected despite its higher computational demand. The model will be trained offline, and its 0.201 s inference time remains fast enough for online performance prediction and control optimization.

Digital twin modelling

The development of a digital twin for gas turbines through a deep neural network approach offers a highly effective and innovative solution to enhance predictive accuracy and enable real-time system monitoring. By constructing a virtual model of the gas turbine, this method allows for precise predictions of system behaviour and performance across various conditions. In this study, an artificial neural network (ANN) with backpropagation was employed to capture and replicate the complex dynamics and nonlinear behaviour of the gas turbine, ensuring that the digital twin accurately reflects the physical system's operations. The Adam optimizer was applied to streamline the training process, ensuring faster and more stable convergence. This integration of deep learning techniques allows the digital twin to efficiently learn from large datasets, adapt to evolving operational patterns, and provide accurate predictions that guide maintenance and operational decisions.

Deep learning model development

Precise predictive models are required to predict gas turbine performance and efficiency due to the stringent accuracy requirements imposed by global standards such as ASME PTC 22 and ISA S77.20. These standards establish the benchmark for gas turbine performance testing and validation in the power generation industry. Specifically, ASME PTC 22 mandates extremely tight measurement uncertainties for instance, gross power output measurements must achieve uncertainties as low as ±0.25%, ambient air pressure ±0.075%, and fuel flow measurement ±0.75% . Similarly, ISA S77.20 requires that simulator predictions for thermal parameters maintain relative errors below 2% under steady-state conditions and within 10% during dynamic transients.

Given these stringent requirements, our deep learning architecture was systematically tuned to improve predictive accuracy from an initial 76% to 99% (R^2 = 0.9959, RMSE = 0.07958). This precision is not merely an optimization target but a compliance necessity: even small deviations in predicted operating parameters can lead to substantial financial consequences in terms of fuel costs, maintenance scheduling, and operational inefficiencies. The 0.5% error threshold established for model validation in this study represents a conservative margin that aligns with the most stringent ASME PTC 22 requirements for corrected power and heat rate validation, ensuring that the digital twin meets industrial certification standards for performance monitoring applications.

Model architecture

Fig. 15 illustrates the composition of the selected operational parameters within the dataset, organized according to the primary components of the gas turbine. In this study, the performance of an operating gas turbine is assessed by applying a monitoring method to the Brayton cycle diagram. Initially, there are 12 raw parameters, which increase to 53 after undergoing energy analysis computations. However, only specific parameters are required for developing the gas turbine Brayton cycle diagram. The remaining computed parameters will be used in further studies to expand the performance analysis. As a result, the final selection for developing the diagram consists of 34 operational parameters, supported by 100,436 data points. These parameters are categorized into 11 input variables and 23 output variables, providing a thorough representation of gas turbine behaviour. This dataset is used to train and validate the deep learning model for gas turbine performance monitoring.

The deep learning model summary, illustrated in Fig. 16, reveals a total of 3178,519 trainable parameters. The model's architecture includes five layers in total, comprising three hidden layers and two operational layers (input and output). Initially, the neural network architecture was tested with two hidden layers, starting with 256 neurons in the first layer and 128 neurons in the second. Through iterative optimization, the final model configuration was refined to include three hidden layers with 512, 2048, and 1024 neurons, respectively. This optimized structure is further enhanced by the use of the Adam optimizer, which has significantly improved the model's training efficiency and predictive accuracy. The learning rate for the Adam optimizer was set at 0.001, which is a commonly used default value that helps achieve a good balance between convergence speed and stability during training.

Validation strategy and generalization

This study applies a combination of two validation methods: holdout validation and early stopping validation. This approach was particularly suited for the total dataset of 117,104 data points and the hardware constraints, which limited the training time for complex algorithms. While k-fold cross-validation is commonly employed in machine learning studies, it was deemed unsuitable for this application due to the temporal autocorrelation inherent in gas turbine operational data. K-fold methods assume independence between observations, which is violated in time-series data where consecutive readings are highly correlated. Random shuffling in k-fold would create data leakage between training and validation sets, artificially inflating performance metrics. Furthermore, the computational demands of k-fold would exacerbate hardware thermal constraints, as the NVIDIA Quadro P2000 GPU already operates at elevated temperatures (74 °C versus 48 °C normal).

The early stopping method enhanced the holdout validation by halting the training process once the model achieved its best performance on the validation set, preventing overfitting. Training incorporated automated early stopping based on validation loss monitoring. The learning curve reached convergence at 191 epochs, with training automatically halted when validation performance plateaued. This prevents the model from over-optimizing on training data at the expense of generalization. Following training, the model was tested on a new dataset to mitigate any bias or overfitting. The performance metrics are computed to adjust the hyperparameters accordingly

Overfitting risk assessment and mitigation

While the model architecture contains approximately 3.2 million parameters, which may initially appear substantial for a regression task, this complexity is deliberately calibrated to address the specific challenges of gas turbine thermodynamic modelling. To mitigate overfitting risks associated with this model capacity, we implemented a multi-layered regularization strategy:

a. **Data Preprocessing and Outlier Removal**: Rigorous statistical cleaning removed anomalous operational data points that could encourage memorization rather than generalization. This preprocessing ensures the model learns robust physical relationships rather than sensor noise or transient anomalies.
b. **Early Stopping Mechanism**: Training incorporated automated early stopping based on validation loss monitoring. The learning curve reached convergence at 191 epochs, with training automatically halted when validation performance plateaued. This prevents the model from over-optimizing on training data at the expense of generalization.
c. **Validation Curve Analysis**: The learning curves (Fig. 17) demonstrate healthy training dynamics. After data preprocessing, validation accuracy tracks closely with training accuracy, indicating effective generalization. In contrast, before data cleaning, the validation curve exhibited classic overfitting behavior disappearing from the learning plot due to diverging training and validation errors.
d. **Cross-Model Comparison**: The model's generalization capability is validated by its superior performance compared to simpler architectures (GBM, XGBoost, LightGBM). While these tree-based methods have lower computational complexity, they exhibit piecewise constant approximations that manifest as discrete jumps in predicted enthalpy values during load transients. The DNN's smooth differentiable transitions better represent the continuous physics while maintaining lower prediction errors.

The model achieves an R^2 score of 0.9959 on unseen test data, with normalized errors of approximately 0.01% across all 23 output variables. While such high R^2 values may raise concerns about overfitting,

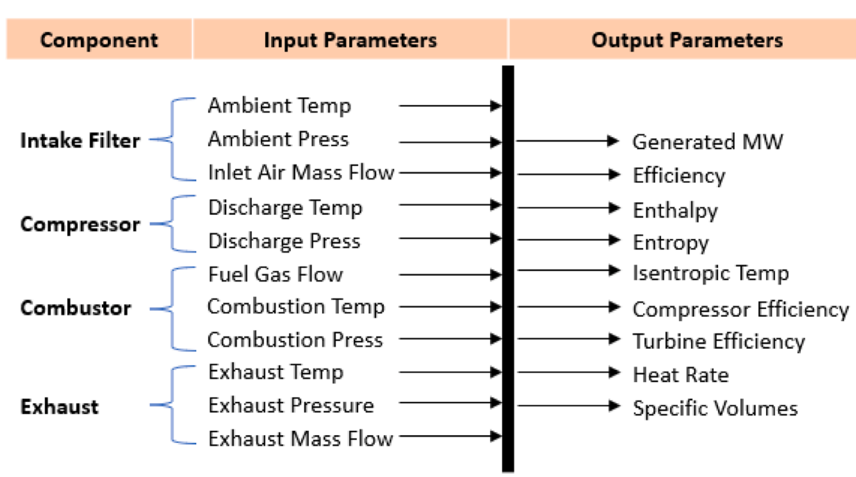


Fig. 15. Final operating parameters composition.

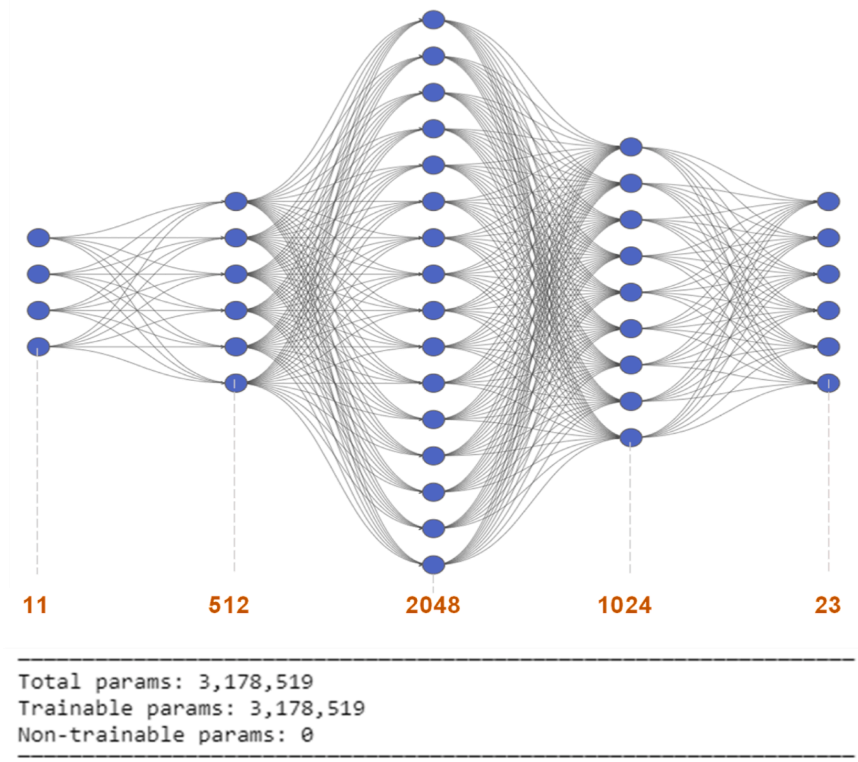


Fig. 16. The summary of deep learning model architecture.

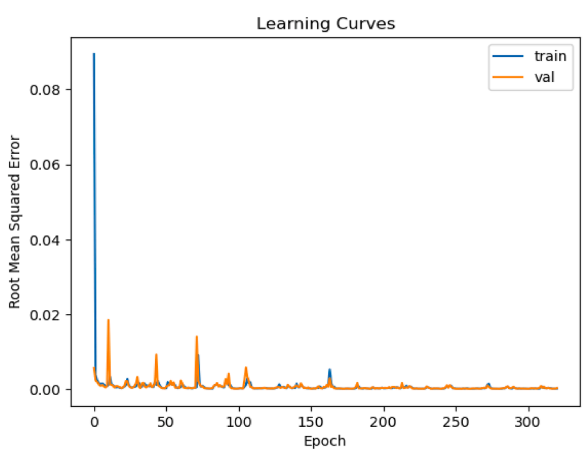


Fig. 17. Training performance learning curves.

we emphasize that (1) the test set comprised 23,420 data points (20% of total data) completely withheld from training and validation; (2) the validation curves show no divergence between training and validation performance; and (3) the model maintains accuracy across five distinct operational load conditions (125–220 MW), demonstrating robustness to varying operational states that would expose overfitted models. This performance confirms that the architectural complexity translates to superior predictive fidelity rather than overfitting. For digital twin deployment, where the model is trained once and queried repeatedly for real-time performance monitoring and predictive maintenance, the marginal computational cost is justified by the significant accuracy gains required for ASME PTC 22 compliance.

However, the model's performance was unsatisfactory initially, as the test loss and test accuracy scores were not within acceptable limits. To resolve this, a deep learning model integrated with the ADAM optimizer was introduced, leading to improved and optimal performance. This study found that the integration with the ADAM optimizer enhanced the model's performance metrics, resulting in shorter training durations and preventing hardware crashes caused by GPU overheating, as the temperature continually remains at 74 °C (Normal: 48 °C).

The study utilized the **NVIDIA Quadro P2000 GPU** to enhance the speed and efficiency of the training process. With a memory bandwidth of 140 GB/s, this GPU significantly accelerates computational tasks, resulting in an average training time of approximately 57.978 s per session, substantially faster than training without GPU support. The model's learning curve reached convergence at 191 epochs, with early stopping functionality halting the training when the validation curve started to decline. This mechanism is crucial in preventing overfitting by stopping the learning process before the model begins to lose its ability to generalize to new data. Overall, these results underline the effectiveness of both data cleaning and the use of advanced hardware in optimizing the deep learning model's performance and preventing issues like overfitting.

This study also aims to develop a robust digital twin that can provide accurate and reliable predictions for gas turbine operations. To achieve this, a rigorous evaluation of the model's performance was carried out using key metrics, including mean absolute error (MAE), root mean squared error (RMSE), learning curves, test loss, test accuracy, regression values, and a comparison of actual versus predicted results. Table 7 summarizes the overall performance results of the data-driven model. After removing outliers through data preprocessing, both the mean

Table 7
Plant Model Training Performance Results.

No.	Performance Indicator	Results
1	Mean absolute error (MAE)	1.1962e-04
2	Root mean square error (RMSE)	0.07958
3	Test Accuracy	0.0063
4	R-squared value (R^2)	0.9959

absolute error and root mean squared error decreased, indicating that the model's predictions were closer to the actual values. The model's test accuracy improved due to an increase in the number of neurons, the addition of an extra hidden layer, and the use of the Adam optimizer, leading to lower test loss and higher accuracy. The model achieved an R2 score of 1.0, indicating a perfect fit for predicting gas turbine operational parameters. The regression value (R2) shows the relationship between the model's input and output variables, where a negative coefficient indicates they move in opposite directions, while a positive coefficient indicates that the input and output tend to move in the same direction.

Digital twin model validation

This section presents the validation results of the digital twin model by comparing the deep learning model's predicted performance with actual values. Fig. 18 illustrates a detailed validation of the digital twin model by juxtaposing predicted outputs with real-world data. To enhance understanding, the visualization examines individual data points for various gas turbine performance metrics, including efficiency, generated output, entropy (S1), enthalpy (H1), compressor efficiency, and turbine efficiency. For consistency, these operational values were normalized during model training using standard scaling, ranging from 0 to 1. The application of the standard-scaler function was essential in aligning all data within the same range, thereby optimizing the model's predictive performance.

Furthermore, the performance of the digital twin model is thoroughly assessed by calculating the error rate, which measures the difference between the actual values and the predicted outcomes. The error rate is calculated using Equation (13). To ensure the model's reliability and compliance with ASME PTC 22 standards, a stringent error threshold of 0.5% was established for the validation process, allowing for only negligible deviations between the predicted and real-world data. This tight error margin exceeding the ASME PTC 22 requirement

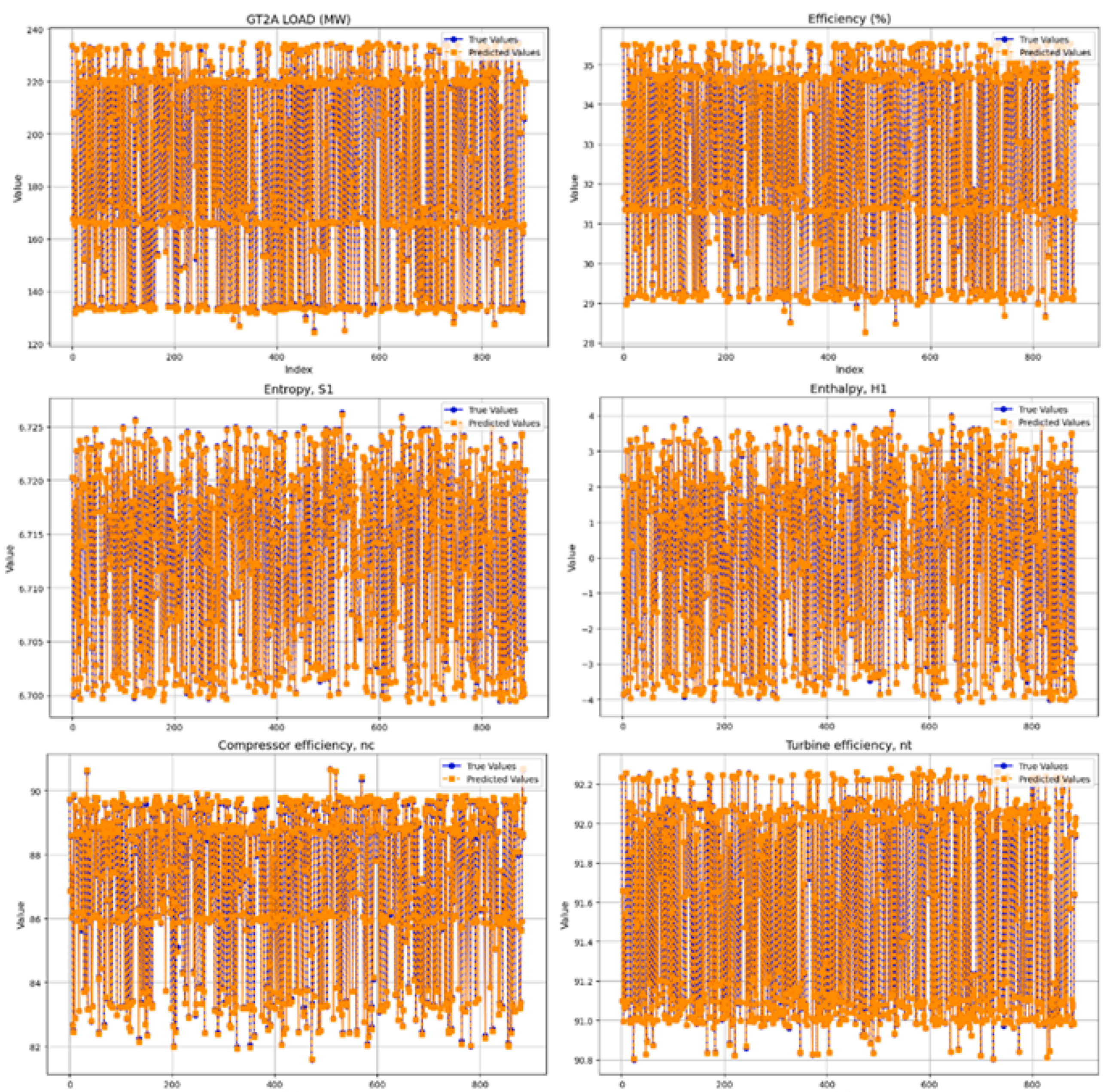


Fig. 18. The predicted vs. actual performance parameters.

of ±0.25% for power output measurements emphasizes the critical importance of achieving precise model predictions, as even small inaccuracies in gas turbine operations can lead to substantial financial consequences. Such careful attention to minimizing error is essential for ensuring the model's practical applicability and its ability to support decision-making processes that optimize turbine performance and reduce operational costs. By maintaining a high standard of accuracy, the model is better equipped to handle the complexities of real-world operations, where even slight deviations can result in significant operational inefficiencies or increased maintenance costs. Additionally, the model's robustness and adaptability are tested across five load conditions from 125 MW to 220 MW, replicating different operational settings and challenges, demonstrating its potential for deployment in diverse real-world environments.

Gas turbine performance prediction

The performance prediction evaluation utilizes test data archived from real-time operational data streams, categorized into three analytical domains: Brayton Cycle diagram visualization, turbomachinery component analysis, and comprehensive gas turbine system performance. Model accuracy is quantified through percentage error calculations against these authentic operational datasets, validating the digital twin's precision in replicating actual gas turbine behaviour under diverse transient conditions. This rigorous validation against archived real-time data ensures the model's robustness and reliability for practical deployment in live monitoring environments.

The variance in prediction accuracy across thermodynamic domains reflects the underlying sensitivity of measurement modalities and physical process stability. The T-S diagram's exceptional accuracy derives from temperature and entropy being strongly coupled through the ideal gas relation $ds = c_p T/dT - R\, dP/P$; temperature measurements via thermocouples provide stable, low-noise signals that anchor the model's predictions. Conversely, the P-V diagram's higher error originates from two sources: (1) pressure transducers exhibit higher sensitivity to pulsation and acoustic noise in the combustion chamber, and (2) specific volume (v) is a derived parameter ($v = RT/P$), propagating errors from both temperature and pressure measurements. The error magnification is particularly evident at State 4 (exhaust), where small pressure measurement errors (~0.1%) translate to large specific volume errors (~0.5%) due to the low exhaust pressure (~1.03 bar). Similarly, the higher error in compressor efficiency versus turbine efficiency reflects the compressor's wider operating map; the compressor operates across multiple speed lines and surge margins during load transients, whereas the turbine operates closer to its choked flow condition, exhibiting more deterministic behaviour. The peak error at 180 MW correlates with the transition between part-load efficiency optimization and full-load temperature limit control, a regime where control system mode switching introduces non-linearities that challenge model extrapolation.

Brayton cycle diagram prediction

This section outlines the performance prediction of a Brayton cycle diagram at a 200 MW load as a reference. Table 8 shows that the T–S diagram prediction has an average error of just 0.0007%, with the highest error occurring at the compressor discharge temperature (0.0015%) and the lowest at the ambient and combustion temperatures (both 0%). In contrast, the P–V diagram prediction shows an average error of 0.0508%, with the largest error at the exhaust specific volume (0.0989%) and the smallest at the ambient specific volume (0.0116%). These findings highlight the exceptional accuracy and reliability of the digital twin model in forecasting gas turbine performance, confirming its capability to accurately reflect the behaviour of essential Brayton Cycle components under operational conditions. Figs. 19 and 20 show the T–S and P–V diagrams, comparing the predicted and actual gas turbine performance. The accuracy of the digital twin model is assessed by comparing predicted values with actual ones, categorized by different states of the Brayton Cycle and related operational parameters, such as entropy and specific volume at key turbine components (intake filter, compressor, combustor, and turbine).

The slightly higher prediction errors observed in the P–V Brayton cycle, particularly at lower loads (e.g., average percentage errors of 0.0844% at 125 MW and 0.0699% at 150 MW), primarily arise from larger deviations in the exhaust state (state 4), where relative errors reach up to 0.2835% (125 MW) and 0.2678% (150 MW). The combustion state (state 3) exhibits comparatively modest and stable errors (typically below 0.043%), while predictions for ambient conditions (state 1) and compressor discharge (state 2) remain highly accurate with near-zero or very low errors across all operating points. This pattern is more pronounced under part-load conditions, where exhaust-specific volume predictions show the highest deviations. These discrepancies are likely attributable to the increased dynamicity and variability characteristic of real gas turbine operational data at reduced loads including fluctuations in combustion completeness, exhaust gas temperature and composition, heat transfer losses, minor leakages, and transient aerodynamic effects in the turbine section which pose greater challenges for uniform modelling compared to the more stable compression and inlet processes

Despite these observations, the overall mean prediction error across all tested loads and states in the P–V diagram is approximately 0.0617%, which reflects high predictive accuracy and is fully sufficient to faithfully emulate the real thermodynamic behaviour of the gas turbine in the P–V representation for engineering analysis, performance monitoring, and simulation purposes. The overall results for the predicted Brayton Cycle diagrams are presented in Table 9, with predictions made at various load levels to assess the model's robustness and accuracy. The predicted errors were minimal, with an overall error of 0.0019% for the T–S diagram and 0.0617% for the P–V diagram. The T–S diagram at 200 MW showed the smallest error of 0.0007%, while the P–V diagram at 220 MW had an error of 0.0210%. These results underscore the digital twin model's exceptional accuracy and reliability in simulating gas turbine performance across different operating loads.

Table 8
T-S and P-V prediction for gas turbine load at 200 MW.

State	Input		Output			
	Parameters	Bar/DegC	Parameters	Actual	Predicted	% Error
1	Ambient	30.55	S1	6.7172	6.7172	0.0000
		1.01	v1	0.8631	0.8632	0.0116
2	Compressor Discharge	389.69	S2	6.7970	6.7969	0.0015
		12.44	v2	0.1530	0.1529	0.0654
3	Combustion	1300.3	S3	7.7847	7.7847	0.0000
		12.44	v3	0.3631	0.3630	0.0275
4	Exhaust	634.56	S4	8.8481	8.8480	0.0011
		0.0325	v4	80.2219	80.1426	0.0989
Average					T-S	0.0007
					P-V	0.0508

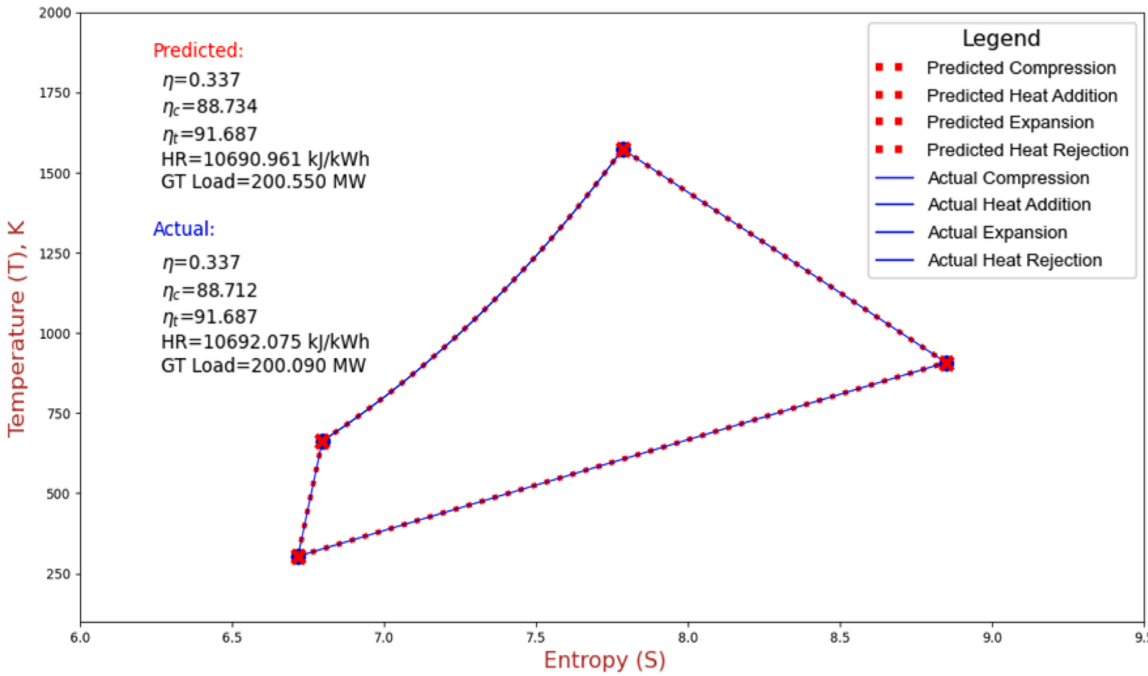


Fig. 19. Predicted T-S diagram for gas turbine performance at 200 MW.

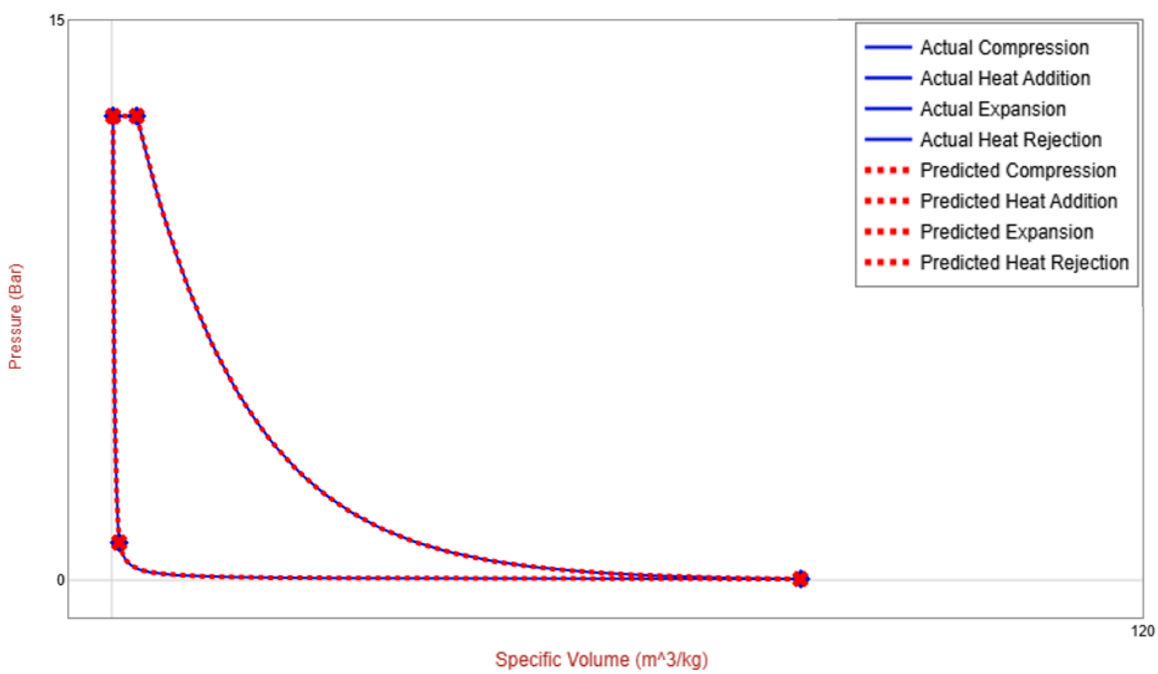


Fig. 20. Predicted P-V diagram for gas turbine performance at 200 MW.

Table 9
T-S and P-V Prediction Performance for various loads.

No.	Load (MW)	T-S % Error	P-V % Error
1	125	0.0020	0.0844
2	150	0.0023	0.0699
3	180	0.0036	0.0824
4	200	0.0007	0.0508
5	220	0.0012	0.0210
Average		0.0019	0.0617

The following are the proposed approaches for future improvements in accuracy. These enhancements would further refine the already excellent predictive fidelity demonstrated in the current study while preserving its strong capability to represent real gas turbine behaviour.

a. **Targeted data augmentation and feature engineering** focused on combustion and exhaust dynamics, including incorporation of transient variables (e.g., fuel flow rate variability, turbine inlet temperature fluctuations, exhaust backpressure effects, or sensor-derived indicators of combustion instability).

b. **Hybrid physics-informed modeling** by embedding thermodynamic constraints (e.g., ideal gas relations, energy and mass balance equations, or component-specific efficiency correlations) directly into the machine learning architecture to minimize physically inconsistent deviations, especially in the exhaust region.
c. **Load-regime-adaptive techniques**, such as stratified training, mixture-of-experts architectures, or transfer learning with particular emphasis on part-load operational data to better capture the elevated variability observed at lower power outputs.
d. **Enriched high-resolution datasets**, including additional part-load records, detailed CFD-validated simulations, or controlled experimental measurements of exhaust conditions to further enhance model generalization and reduce residual errors.

Turbomachinery performance prediction

This section evaluates the digital twin model's ability to predict the efficiency of key turbomachinery components, particularly the compressor and turbine, under various conditions. The model's performance is evaluated by comparing predicted and actual efficiency data, offering insights into optimizing compressor and turbine performance. Table 10 presents the comparison of predicted and actual efficiencies for loads between 125 MW and 220 MW. The compressor efficiency prediction is most accurate at 150 MW, with an error of 0.0102%, and the largest error occurs at 180 MW (0.0934%). Overall, the model shows an average error of 0.0311%, indicating good alignment with real-world compressor efficiency. For turbine efficiency, the prediction at 200 MW has a minimal error of 0.0001%, and the highest error is 0.0104% at 180 MW, with an average error of 0.0044%. This demonstrates the model's strong performance in accurately simulating the turbine's efficiency and reliability in predicting turbomachinery behaviour.

Overall gas turbine performance prediction

This section provides an overview of the performance predictions for the gas turbine, focusing on efficiency and heat rate across different load scenarios. The analysis evaluates turbine performance under various operating conditions, offering critical insights into the turbine's effectiveness and energy utilization. By predicting efficiency and heat rate, the study identifies opportunities for optimization, aiming to enhance energy efficiency and reduce operational costs.

Table 11 shows a summary of gas turbine efficiency and heat rate predictions. The model demonstrates the highest accuracy in predicting turbine efficiency at the 200 MW load, with a minimal error of 0.0160%. The overall error across all loading conditions is 0.0598%, reflecting strong performance. For the predicted gas turbine heat rate, the smallest error occurs at the 125 MW load (0.0036%), while the largest deviation is at the 180 MW load. The average error for heat rate prediction is 0.0519%, indicating the model's robustness and reliability. These results highlight the digital twin model's effectiveness in accurately forecasting both turbine efficiency and heat rate, underscoring its potential for optimizing turbine performance under various operational conditions.

Table 10
Compressor and turbine efficiency prediction at various loads.

No.	Load (MW)	Output			
		Parameters	Actual	Predicted	% Error
1	125	η_c	83.0972	83.0851	0.0146
		η_T	90.9470	90.9461	0.0010
2	150	η_c	84.4079	84.4165	0.0102
		η_T	90.8273	90.8218	0.0061
3	180	η_c	87.6831	87.6012	0.0934
		η_T	91.3867	91.3772	0.0104
4	200	η_c	88.7121	88.7344	0.0251
		η_T	91.6871	91.6872	0.0001
5	220	η_c	88.8334	88.8227	0.0120
		η_T	92.0481	92.0438	0.0047
Average				η_c	0.0311
				η_T	0.0044

Table 11
Efficiency and heat rate prediction at various loads.

No.	Load (MW)	Output			
		Parameters	Actual	Predicted	% Error
1	125	η_{gt}	28.6012	28.5882	0.0455
		HR	12,586.892	12,587.350	0.0036
2	150	η_{gt}	30.1147	30.1218	0.0236
		HR	11,954.285	11,962.163	0.0659
3	180	η_{gt}	32.4389	32.3793	0.1837
		HR	11,097.779	11,112.439	0.1321
4	200	η_{gt}	33.6698	33.6752	0.0160
		HR	10,692.075	10,690.961	0.0104
5	220	η_{gt}	34.6683	34.6579	0.0300
		HR	10,384.129	10,389.080	0.0477
Average				η_{gt}	0.0598
				HR	0.0519

Practical deployment

System architecture and data integration

For practical deployment, the PI platform connection will be utilized to archive real-time data from the gas turbine. This platform has the capability to store data as a historian or provide real-time data streams, enabling seamless integration with the digital twin model. The digital twin is compiled as a containerized application to run on a Docker platform, which retrieves historical or real-time data from the PI system for model inputs. Additionally, this platform will be equipped with a monitoring dashboard to visualize gas turbine performance in real time, including interactive T-S and P-V diagrams. To ensure stability, it is recommended to implement appropriate timestamp synchronization and buffering mechanisms to prevent platform crashes, as the high processing speeds of the plant control system and PI platform can otherwise introduce instability in the digital twin environment. For future work, this setup can be migrated to an enterprise solution platform, such as cloud-based services (e.g., AWS or Azure), to handle heavy-duty tasks like large-scale data processing and multi-turbine fleet monitoring.

Real time visualization

In terms of practical deployment for real-time T-S and P-V visualization, the system architecture supports low-latency rendering of these thermodynamic diagrams directly within the monitoring dashboard. Leveraging the model's validated accuracy demonstrated by minimal prediction errors, the visualizations can be updated at intervals as frequent as every 10 s, depending on data ingestion rates from the PI platform. This real-time capability allows for continuous tracking of Brayton Cycle states under transient conditions, where deviations in entropy, specific volume, or efficiency can signal emerging issues like compressor surge or turbine degradation. Deployment on edge computing devices ensures minimal latency (<500 ms for inference and rendering), making it suitable for on-site integration without requiring extensive infrastructure upgrades.

Dashboard features

User interaction with these real-time visualizations is designed to be intuitive and operator-centric, facilitated through a web-based or desktop dashboard built with tools like Grafana or custom Python-based interfaces (e.g., using Plotly for interactive plots). Users, such as plant operators or performance engineers, can interact via features including:

- **Zoom and Pan Functionality**: Examine specific cycle states
- **Overlay Comparisons**: Side-by-side predicted vs. actual T-S/P-V diagrams with color-coded error highlights to identify discrepancies from noise or variability.

- **Alert Integration**: Threshold-based notifications for anomalies, tied to the model's temperature-anchored prediction accuracy.
- **Historical Playback**: Replay archived data streams for post-event analysis, aiding training and optimization.
- **Customization Options**: Select load scenarios or component focuses, utilizing low prediction errors for targeted diagnostics.

This framework reduces cognitive load, transforms thermodynamic data into actionable insights, and supports proactive maintenance particularly at part-loads where P-V variability is highlighted. It promotes collaborative workflows via secure API sharing for remote expert review, building on overall validated performance.

Continuous self-updating mechanism and degradation tracking

The digital twin employs a modular, hybrid update architecture (event-triggered and continuous) to autonomously adapt to abrupt changes and gradual long-term component degradation/aging effects, maintaining predictive fidelity over extended operation. It ingests new operational data from the PI platform to reflect real power plant dynamics, including phenomena like compressor fouling, turbine blade erosion, or thermal barrier coating degradation. The mechanism comprises three interconnected layers:

a. Event-Triggered Model Regeneration Protocol

Model regeneration automatically initiates after major maintenance events (e.g., Combustion Inspection, Hot Gas Path Inspection, Major Inspection) or detected anomalies. This handles abrupt shifts from component replacement/refurbishment or faults. Post-event operational data retrains or recalibrates parameters using physics-informed machine learning. This is essential, as GT performance often improves 3–8% post-overhaul due to restored clearances and coating rejuvenation, preventing reliance on outdated baselines.

b. Performance-Validated Model Versioning

A rolling archive compares incoming data against the prior baseline. Multi-criteria validation assesses:

- Thermodynamic consistency: Entropy generation rates across sections (intake, compressor, combustion, turbine, exhaust) aligned with first/second law expectations, flagging aging-induced increases.
- Predictive accuracy: Cross-validation on holdout datasets per Contribution (c), using metrics like MAPE (<0.1% for key variables such as efficiency). If the new model underperforms (indicating unresolved degradation or anomalies), the system reverts to the previous validated version and alerts operators. This prevents model drift, quantifies maintenance effectiveness, and supports continuous self-diagnosis goals.

c. Continuous Parameter Adaptation for Gradual Degradation

An online learning loop periodically assimilates streaming PI data (e.g., weekly or after >10,000 cycles). Techniques like Kalman filtering or Bayesian estimation incrementally update parameters (e.g., compressor efficiency decay of 0.5–1% per 10,000 hours) to track gradual aging without full retraining. Degradation trajectories are validated against historical trends and visualized in the dashboard (e.g., evolving T-S/P-V paths). Exceeding thresholds triggers alerts for predictive maintenance, enabling long-term fleet management.

This "living" mechanism ensures the digital twin evolves with real-world aging and operational changes, enhancing reliability, reducing unplanned downtime, and supporting proactive strategies throughout the gas turbine's service life.

Computational requirements

The computational requirements of the proposed ANN-based digital twin model are modest and well-suited to both development and operational deployment, particularly for real-time applications like T-S and P-V visualization. Training on the NVIDIA Quadro P2000 GPU, utilizing the full dataset of 100,436 samples with 11 inputs and 23 outputs, achieved convergence in 191 epochs (with early stopping to prevent overfitting), requiring an average of approximately 58 s per training session. This translates to a total training duration on the order of minutes to tens of minutes, depending on hyperparameter tuning and full multi-epoch runs. This efficiency stems from the model's relatively compact architecture (3.18 million trainable parameters across three hidden layers) and the effective use of the Adam optimizer with a learning rate of 0.001.

Inference for a single forward pass is extremely fast, typically in the sub-millisecond to low-millisecond range (<1–5 ms) on the same or equivalent GPU/CPU hardware, as feedforward computations in shallow-to-moderate depth ANNs with millions of parameters routinely achieve such latencies in practice. These low inference times comfortably meet the stringent real-time demands of industrial gas turbine control environments, where control loops, performance monitoring, and anomaly detection often require response latencies below 10–50 ms (and protection systems even <5 ms). For T-S and P-V visualizations, this enables on-the-fly diagram rendering using libraries like Matplotlib or Plotly, with additional overhead for plotting (<10 ms per update), ensuring smooth real-time updates even under high data throughput from the PI platform.

The model's lightweight nature enables straightforward deployment on edge devices, embedded controllers, or existing plant DCS hardware without excessive resource demands, supporting continuous real-time predictions for predictive maintenance, fault detection, and operational optimization while maintaining high fidelity. This is particularly advantageous given the model's validated performance across thermodynamic domains: the T-S diagram's low error (0.0019% overall) benefits from stable temperature signals, while P-V's slightly higher error (0.0617%) accounts for pressure and volume sensitivities, yet both remain suitable for real-time use without compromising visualization reliability.

For industrial deployment, we recommend a hybrid architecture that strategically separates these computational workloads: offline training involving periodic model retraining conducted either on a monthly schedule or triggered by significant operational regime changes (e.g., based on observed increases in prediction errors at transitional loads like 180 MW, where non-linearities peak), utilizing historical datasets on GPU-enabled workstations to maintain model accuracy without impacting live operations; real-time inference deploying the frozen, optimized model on industrial PCs or edge devices equipped with GPU acceleration, enabling seamless integration with existing plant infrastructure through standard industrial communication protocols such as OPC-UA or Modbus; and a dedicated monitoring layer that continuously tracks prediction error metrics (e.g., alerting on deviations exceeding 0.1% in P-V exhaust states) and automatically triggers model recalibration procedures when error thresholds exceed predefined operational limits. This architecture effectively decouples the computational intensity of deep learning training from the stringent real-time performance requirements of industrial control environments, thereby preserving the model's exceptional predictive accuracy evident in turbomachinery (0.0311% compressor, 0.0044% turbine) and overall system metrics (0.0598% efficiency, 0.0519% heat rate) while ensuring reliable, low-latency decision support for critical gas turbine operations, including interactive T-S and P-V visualizations.

Limitation and future work

The digital twin model developed in this study demonstrates excellent predictive accuracy for gas turbine performance under fuel-gas (FG) firing mode at loads ≥125 MW, as evidenced by very low errors in Brayton cycle representations (0.0019% for T-S, 0.0617% for P-V; Table 7), turbomachinery efficiencies (0.0311% for compressor, 0.0044% for turbine; Table 8), and system-level performance (0.0598% for thermal efficiency, 0.0519% for heat rate; Table 9). This high fidelity is largely attributable to the stable, high-quality operational data available in the base-load regime, which aligns with the turbine's design intent for optimal efficiency.

However, three key limitations constrain the current model scope and require careful consideration for appropriate deployment: (1) load range restrictions, (2) transient behavior boundaries, and (3) fuel flexibility constraints. These limitations reflect operational realities and data availability rather than inherent architectural deficiencies, and the modular framework supports systematic extension to address each in future development. These limitations are discussed separately below.

a. **Load Range Limitation:** The model has been developed and validated exclusively using data from loads ≥125 MW. Below this threshold, gas turbine operation enters part-load regimes characterized by increased combustion dynamics, wider compressor operating maps, higher relative exhaust temperature variability, and greater influence of control-system actions. These factors introduce significantly more noise and non-linearity compared to base-load conditions. In our dataset, sustained operation below 125 MW is rare due to economic and efficiency considerations, gas turbines are typically kept above this level during normal dispatch to minimize specific fuel consumption and avoid prolonged low-efficiency running. As a result, the model has not been trained or validated for low-load conditions (<125 MW). Error trends observed near the 125 MW boundary (e.g., 0.2835% at exhaust state 4 in the P-V diagram) reflect data sparsity rather than fundamental model deficiencies. The physics-informed architecture (incorporating ideal gas relations, energy balances, and thermodynamic constraints) provides a solid foundation for future extension to lower loads, provided that representative part-load datasets become available
b. **Transient Behavior Limitation:** While the model successfully captures moderate load transients within the ≥125 MW range (e.g., controlled ramps around 180 MW that include non-linear governor and fuel-valve effects), it has not been extensively validated under highly dynamic or extreme transient conditions. Scenarios such as daily startup/shutdown cycles in peaking service, rapid load-following, emergency trips, or grid-frequency response events involve pronounced thermal gradients, fuel-flow oscillations, variable guide-vane dynamics, and potential aerodynamic instabilities that are underrepresented in the current training data. Highly transient operation would benefit from richer datasets that include high-frequency sensor signals (e.g., flame instability indicators, pressure pulsations, and fuel-composition variability). The current architecture is capable of learning such dynamics, but performance would improve significantly with augmentation from cyclic-operation plant data or high-fidelity transient simulations
c. **Fuel Flexibility Limitation:** The present digital twin was developed and validated exclusively for fuel-gas (FG) firing mode. Although the physical gas turbine is equipped for dual-fuel capability (FG and distillate fuel oil, DFO), no DFO-fired operational data were included in this study. Consequently, the model cannot currently predict performance under DFO combustion, during fuel-switchover sequences, or in blended-fuel scenarios. Differences in combustion characteristics (e.g., flame temperature, emissions, heat-release profile, and required control adjustments) between FG and DFO make direct extrapolation unreliable. The modular, physics-informed deep-learning framework remains well suited for extension to dual-fuel operation. Retraining or fine-tuning the model on a comparable DFO dataset covering the same load range would enable accurate prediction across both fuels and during fuel transitions.

Future Work: These extensions would broaden the digital twin's utility, enabling comprehensive monitoring across all operating conditions while building on the strong foundation establishment. To overcome the limitations outlined above, the following directions are proposed:

a. **Load Range Extension**: Acquisition of operational data from peaking or cyclic-duty plants to enrich low-load and highly transient regimes, enabling multi-cycle aggregation and robust generalization.
b. **Modeling Strategies:** Application of load-adaptive modeling strategies, such as mixture-of-experts architectures or transfer learning from base-load to part-load conditions.
c. **Transient Behavior Enhancement**: Integration of physics-based simulation data and hybrid modeling techniques to enforce thermodynamic consistency during noisy or sparsely sampled transients.
d. **Fuel Flexibility Integration**: Extension to dual-fuel capability by incorporating DFO operational records across equivalent load and transient conditions, thereby enabling full fuel-flexibility assessment and fuel-switch modeling.

These developments would significantly broaden the applicability of the digital twin, allowing comprehensive performance monitoring, diagnostics, and optimization across the full spectrum of operating regimes and fuel configurations while preserving the high-accuracy foundation established in the present FG-focused work.

Conclusion

The conclusion of the study on the development of the gas turbine digital twin emphasizes the model's impressive ability to accurately predict the performance of the GE9FA gas turbine across six key operational categories. The high accuracy of the digital twin model's predictions is mandatory for power plant applications. Even small variations in performance prediction can be translated into financial loss or gain. The plant model's performance is validated and adjusted until the model achieves high accuracy and the lowest Mean Absolute Error (MAE).

The study, detailed in Table 12, highlights the model's high precision in simulating the turbine's thermodynamic processes. The T–S Brayton cycle diagram emerged as the most accurate prediction, with a remarkably low error of just 0.0019%. This indicates that the digital twin excels at capturing the complex dynamics of the turbine's performance, particularly in relation to its thermodynamic behaviour. In addition to the T–S Brayton cycle, the model also demonstrated strong accuracy in predicting turbine efficiency, with an error of only 0.0044%. Similarly, predictions for compressor efficiency were accurate, with a slight error of 0.0311%, indicating strong alignment with the expected performance. Furthermore, the model's predictions for gas turbine efficiency and heat rate were both closely matched to real-world data, showing errors of just 0.0598% and 0.0519%, respectively. These results underscore the model's capability to simulate the turbine's overall operational performance across a range of conditions. Despite the P–V Brayton cycle diagram prediction showing slightly greater deviations, with an error of 0.0617%, it nonetheless remained within an acceptable range. This indicates that although the model is highly effective, there

Table 12
Digital twin prediction performance summary.

No	Prediction Category	% Error
1	T-S Brayton Cycle	0.0019
2	P-V Brayton Cycle	0.0617
3	Compressor Efficiency	0.0311
4	Turbine Efficiency	0.0044
5	Gas Turbine Efficiency	0.0598
6	Heat Rate	0.0519
Average		0.0351

are opportunities to improve its accuracy, particularly in the modelling of more intricate aspects of the turbine's performance under specific operating scenarios.

The achieved average prediction error of 0.0351% represents an order-of-magnitude improvement over typical industry standards for gas turbine performance models (0.5–1.0% per ASME PTC 22). This precision enables applications previously infeasible with lower-fidelity models: (1) **Load Control**: The 0.016% error in efficiency at 200 MW allows for real-time heat rate optimization with confidence intervals tight enough to justify 0.1% fuel trim adjustments, yielding estimated huge fuel savings annually per unit. (2) **Anomaly Detection**: The error bounds are narrower than the typical drift signatures of fouling or blade erosion, enabling earlier fault detection (2–3 weeks advance notice) compared to threshold-based monitoring. (3) **Compliance Requirement**: The T-S diagram accuracy (0.0019%) approaches the uncertainty of physical acceptance tests (±0.25% per ASME standards), suggesting future potential for virtual compliance demonstration, reducing the need for costly field testing during upgrades.

Overall, the digital twin model achieved an impressive average error rate of just 0.0351%, which underscores its exceptional accuracy in simulating the operational performance of the GE9FA gas turbine. This high degree of precision enables the model to provide critical insights for various applications, including predictive maintenance, performance optimization, and real-time decision-making in gas turbine management. By accurately replicating turbine behaviour under various conditions, the model can anticipate potential failures, optimize fuel efficiency, and enhance overall operational reliability. Furthermore, this study represents a significant leap forward in the creation of advanced, high-fidelity digital models tailored for complex energy systems like gas turbines. It sets a new benchmark in digital twin technology, offering a strong foundation for future research and technological advancements that could further improve the efficiency, sustainability, and reliability of gas turbines and other industrial systems. The success of this model not only reinforces the potential of digital twins in predictive analytics but also paves the way for their broader application across a wide range of sectors in the energy industry.

Looking forward, the study paves the way for future work in expanding the digital twin framework, including developing a model for the Rankine cycle and integrating components like the Heat Recovery Steam Generator (HRSG) and Steam Turbine (ST). This will result in a comprehensive combined cycle gas turbine (CCGT) digital twin model, marking a major advancement in energy system simulation and optimization.

CRediT authorship contribution statement

Balbir Shah Mohd Irwan Shah: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Asnor Juraiza Ishak:** Writing – review & editing, Supervision, Resources, Formal analysis, Conceptualization. **Mohd Khair Hassan:** Visualization, Validation, Resources, Methodology. **Nor Mohd Haziq Norsahperi:** Software, Project administration, Investigation, Formal analysis.

Declaration of competing interest

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

Acknowledgments

The authors express their gratitude to the management team of Tuanku Ja'afar Power Station, TNB Power Generation Sdn. Bhd. (Genco), for their invaluable support in this study. Their generosity in granting permission and assistance in gathering information on power plant operations and distributed control systems (DCS) significantly contributed to the success of this research. This collaboration highlights the positive outcomes that can be achieved through industry-academic partnerships.

The authors acknowledge the role of artificial intelligence (AI) tools in enhancing the clarity, coherence, and grammatical accuracy of the text. These tools provide valuable support in refining the language and ensuring the high standards of academic writing. Furthermore, the authors emphasize that the content, ideas, findings, and conclusions presented in this work are the original contributions of the authors.

Data availability

Data will be made available on request.

References

[1] H. Hanachi, C. Mechefske, J. Liu, A. Banerjee, Y. Chen, Performance-based gas turbine health monitoring, diagnostics, and prognostics: a survey, IEEe Trans. Reliab. (2018).
[2] N. Hepperle, D. Therkorn, E. Schneider, S. Staudacher, Assessment of gas turbine and combined cycle power plant performance degradation, Proc. ASME Turbo Expo 4 (2011) 569–577.
[3] R. Kurz, Gas turbine degradation, in: 43rd Turbomach. 30th Pump Users Symp., 2014, pp. 1–36.
[4] I.S. Diakunchak, Performance deterioration in industrial gas turbines, Proc. ASME Turbo Expo 4 (April 1992) (1991) 161–168.
[5] E. Tsoutsanis, I. Qureshi, M. Hesham, Performance diagnostics of gas turbines operating under transient conditions based on dynamic engine model and artificial neural networks, Eng. Appl. Artif. Intell. (2023).
[6] Z. Liu, I.A. Karimi, Gas turbine performance prediction via machine learning, Energy (2020).
[7] X. Jiang, Y. Liu, M. Wei, Z. Wang, A thermodynamics-informed deep learning approach for lightweight modelling of gas turbine performance, SSRN Electron. J. (2024).
[8] Y. Yang, T. Nikolaidis, S. Jafari, P. Pilidis, Gas turbine engine transient performance and heat transfer effect modelling: a comprehensive review, research challenges, and exploring the future, Appl. Therm. Eng. (2024).
[9] S. Yousif, F. Alnaimi, S. Thiruchelvam, Gas turbine performance monitoring and operation challenges: a review, Gazi Univ. J. Sci. (2023).
[10] M. Plis, H. Rusinowski, A mathematical model of an existing gas-steam combined heat and power plant for thermal diagnostic systems, Energy (2018).
[11] W.I. Rowen, Simplified mathematical representations of heavy-duty gas turbines, J. Eng. Gas Turbines Power (1983).
[12] A.A. Sammour, O.V. Komarov, F.A. Lattieff, M.A. Qasim, A.Y. Saleh, Influence of surrounding air temperature and humidity upon the performance of a gas turbine power plant, J. Adv. Res. Fluid Mech. Therm. Sci. (2023).
[13] A. González-Díaz, A.M. Alcaráz-Calderón, M.O. González-Díaz, Á. Méndez-Aranda, M. Lucquiaud, J.M. González-Santaló, Effect of the ambient conditions on gas turbine combined cycle power plants with post-combustion CO2 capture, Energy (2017).
[14] M.S. Kumar, K. Rupendra, Performance prediction of gas turbine, Int. Res. J. Eng. Technol. 4 (4) (2017) 1843–1847.
[15] C. Shao, Y. Liu, Z. Zhang, F. Lei, J. Fu, Fast prediction method of combustion chamber parameters based on artificial neural network, Electron (2023).
[16] J. An, H. Wang, B. Liu, K.H. Luo, F. Qin, G.Q. He, A deep learning framework for hydrogen-fueled turbulent combustion simulation, Int. J. Hydrogen. Energy (2020).
[17] V. Esfahanian, M.J. Izadi, H. Bashi, M. Ansari, A. Tavakoli, M. Kordi, Aerodynamic shape optimization of gas turbines: a deep learning surrogate model approach, Struct. Multidiscip. Optim. (2024).
[18] G. Bruni, S. Maleki, S.K. Krishnababu, C(NN)FD - A deep learning framework for turbomachinery CFD analysis, IEEE Trans. Ind. Informatics (2024).
[19] S.C. Gülen, P.R. Griffin, S. Paolucci, Real-time on-line performance diagnostics of heavy-duty industrial gas turbines, J. Eng. Gas Turbines Power (2002).
[20] S.M. Kim, Y.J. Joo, Implementation of on-line performance monitoring system at Seoincheon and Sinincheon combined cycle power plant, Energy (2005).
[21] J. Lim, et al., The epri gas turbine digital twin a platform for operator focused integrated diagnostics and performance forecasting, in: Proc. ASME Turbo Expo 4, 2021, pp. 1–11.
[22] Y.G. Li, Gas turbine performance and health status estimation using adaptive gas path analysis, J. Eng. Gas Turbines Power (2010).
[23] F. Lu, W. Zheng, J. Huang, M. Feng, Life cycle performance estimation and in-flight health monitoring for gas turbine engine, J. Dyn. Syst. Meas. Control. Trans. ASME (2016).
[24] E. Ramadhan, Y.G. Li, D. Maherdianta, Application of adaptive GPA to an industrial gas turbine using field data, in: Proceedings of the ASME Turbo Expo, 2019.
[25] J. Luan, Y.P. Cao, S. Li, R. Ao, A gas turbine gas path digital twin modeling method. Mechanisms and Machine Science, 2023.
[26] M. HU, Y. HE, X. LIN, Z. LU, Z. JIANG, B. MA, Digital twin model of gas turbine and its application in warning of performance fault, Chin. J. Aeronaut. (2023).

[27] M. Tahan, E. Tsoutsanis, M. Muhammad, Z.A. Abdul Karim, Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: a review, Appl. Energy 198 (2017) 122–144.
[28] L. Marinai, D. Probert, R. Singh, Prospects for aero gas-turbine diagnostics: a review, Appl. Energy (2004).
[29] S. Kim, A new performance adaptation method for aero gas turbine engines based on large amounts of measured data, Energy (2021).
[30] S. Kim, K. Kim, C. Son, A new transient performance adaptation method for an aero gas turbine engine, Energy (2020).
[31] J. Zhang, Z. Wang, S. Li, P. Wei, A digital twin approach for gas turbine performance based on deep multi-model fusion, Appl. Therm. Eng. (2024).
[32] H. Farhat, A. Altarawneh, Physics-informed machine learning for intelligent gas turbine digital twins: a review, Energies. (Basel) (2025).
[33] J.T. Lim, A. Habibullah, E.Y.K. Ng, Towards a digital twin for gas turbines: thermodynamic modeling, critical parameter estimation, and performance optimization using PINN and PSO, Energies. (Basel) (2025).
[34] C.M. Bartolini, F. Caresana, G. Comodi, L. Pelagalli, M. Renzi, S. Vagni, Application of artificial neural networks to micro gas turbines, Energy Convers. Manage (2011).
[35] M. Fast, M. Assadi, S. De, Development and multi-utility of an ANN model for an industrial gas turbine, Appl. Energy (2009).
[36] S. De, M. Kaiadi, M. Fast, M. Assadi, Development of an artificial neural network model for the steam process of a coal biomass cofired combined heat and power (CHP) plant in Sweden, Energy (2007).
[37] M. Fast, T. Palmé, Application of artificial neural networks to the condition monitoring and diagnosis of a combined heat and power plant, Energy (2010).
[38] H. Nikpey, M. Assadi, P. Breuhaus, Development of an optimized artificial neural network model for combined heat and power micro gas turbines, Appl. Energy (2013).
[39] H. Nikpey, M. Assadi, P. Breuhaus, P.T. Mørkved, Experimental evaluation and ANN modeling of a recuperative micro gas turbine burning mixtures of natural gas and biogas, Appl. Energy (2014).
[40] S. Bhowmik, A. Paul, R. Panua, S.K. Ghosh, D. Debroy, Performance-exhaust emission prediction of diesosenol fueled diesel engine: an ANN coupled MORSM based optimization, Energy (2018).
[41] S. Sina Tayarani-Bathaie, K. Khorasani, Fault detection and isolation of gas turbine engines using a bank of neural networks, J. Process. Control (2015).
[42] Z.N. Sadough Vanini, K. Khorasani, N. Meskin, Fault detection and isolation of a dual spool gas turbine engine using dynamic neural networks and multiple model approach, Inf. Sci. (Ny). (2014).
[43] M.S. Nashed, J. Renno, M.S. Mohamed, R.L. Reuben, Gas turbine failure classification using acoustic emissions with wavelet analysis and deep learning, Expert. Syst. Appl. (2023).
[44] Y. Jin, Y. Ying, J. Li, H. Zhou, Gas path fault diagnosis of gas turbine engine based on knowledge data-driven artificial intelligence algorithm, IEEe Access. (2021).
[45] Y. Luo, Application of reinforcement learning algorithm model in gas path fault intelligent diagnosis of gas turbine, Comput. Intell. Neurosci. (2021).
[46] S. Sun, Y. Xue, M. Wang, L. Sun, Online identification of key parameters of gas turbine dynamic process based on hybrid deep learning, Chinese J. Sh. Res. (2023).
[47] B. Shah, M. Irwan, A.J. Ishak, M.K. Hassan, Development of digital twin data-driven modelling for gas turbine operation behaviour, Pertanika J. Sci. Technol. 33 (S4) (2025) 81–97.
[48] Z. Zhou, et al., More realistic degradation trend prediction for gas turbine based on factor analysis and multiple penalty mechanism loss function, Reliab. Eng. Syst. Saf. (2024).
[49] Z. Li, X. Wang, H. Wen, B. Wang, H. Mou, G. Ma, Combustion mode time-series prediction in rotating detonation combustor based on predictor of encoder transformer, Phys. Fluids (2025).
[50] I. Sharifirad, J. Boudjadar, P.G. Larsen, Highly accurate anomaly detection for the combustion system of a marine engine using multivariate transformers, in: SISY 2024 - IEEE 22nd International Symposium on Intelligent Systems and Informatics, Proceedings, 2024.
[51] B. Khurshid, R. Ud Din, I. Khurshid, A. Waqar, Machine learning-based digital twin model of gas turbines and its application to vibration faults, in: Society of Petroleum Engineers - SPE/PAPG Pakistan Section Annual Technical Symposium and Exhibition, PATS 2025, 2025.
[52] Z. Liu, M. Hou, G. Sa, Y. Wang, X. Xin, J. Tan, Gas turbine multi-working conditions identification and performance prediction based on deep learning and knowledge, Energy 308 (August) (2024) 133011.
[53] B. Yu, W. Liu, D. Xie, X. Cui, H. Zhang, A novel gas turbine performance prediction model incorporating the residual connection and feature engineering methods, Energy 332 (June) (2025).
[54] A.S. Shabunin, A GTU model based on a recurrent neural network: features of elaboration and application, Therm. Eng. (2025).
[55] M. Baker, B. Rosic, The hybrid pathway to flexible power turbines, part III: enabling live thermal boundary prediction using real-time plant measurements and LSTM neural networks, J. Eng. Gas Turbines Power (2025).
[56] Z.Y. Wu, et al., High Fidelity digital twin-based anomaly detection and localization for smart water grid operation management, Sustain. Cities. Soc. (2023).
[57] M. Tahan, E. Tsoutsanis, M. Muhammad, Z.A. Abdul Karim, Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: a review, Appl. Energy 198 (2017) 122–144.
[58] A.G. Abo-Khalil, Digital twin real-time hybrid simulation platform for power system stability, Case Stud. Therm. Eng. (2023).
[59] M. Soori, B. Arezoo, R. Dastres, Digital twin for smart manufacturing, a review, Sustain. Manuf. Serv. Econ. (2023).
[60] V.D. Tran, P. Sharma, L.H. Nguyen, Digital twins for internal combustion engines: a brief review, J. Emerg. Sci. Eng. (2023).
[61] M. Hassan, M. Svadling, N. Björsell, Experience from implementing digital twins for maintenance in industrial processes, J. Intell. Manuf. (2024).
[62] A. Sundararajan, M. Omar, R. Ibrahim, N.F. Salehuddin, Development of gas turbine model interface for digital twin application, AIP. Conf. Proc. 2923 (1) (2024).
[63] W. Choi, K. Hudachek, S. Koskey, C. Perullo, D. Noble, Digital twin in the power generation industry, JMST Adv. (2024).
[64] H. Farhat, C. Salvini, The development of a Novel Hybrid Gas turbine digital twin to predict performance deterioration, in: Proceedings of the International Workshop on Simulation for Energy, Sustainable Development and Environment, SESDE, 2023.
[65] R. Singh, P.S.V. Nataraj, A. Maity, Digital twin of a laboratory gas turbine engine using deep learning framework. Industry 4.0 in Small and Medium-Sized Enterprises (SMEs), 2022.
[66] V. Panov, S. Cruz-Manzo, Gas turbine performance digital twin for real-time embedded systems, in: Proceedings of the ASME Turbo Expo, 2020.
[67] Z.F. Huang, K.Y. Soh, M.R. Islam, K.J. Chua, Digital twin driven life-cycle operation optimization for combined cooling heating and power-cold energy recovery (CCHP-CER) system, Appl. Energy (2022).
[68] A.K. Sleiti, J.S. Kapat, L. Vesely, Digital twin in energy industry: proposed robust digital twin for power plant and other complex capital-intensive large engineering systems, Energy Reports (2022).
[69] B.S. Shah, A.J. Ishak, M.K. Hassan, and N.M. Haziq, "Energy 4 . 0 : challenges and enablers of digital twin application in power plant," vol. 23, no. 1, pp. 103–124, 2024.
[70] S. Kim, Prediction-focused machine learning-based visualization of compressor flow in a gas turbine engine digital twin, Appl. Therm. Eng. (2025).
[71] A.S. Shabunin, A GTU model based on a recurrent neural network: features of elaboration and application, Therm. Eng. 72 (11) (2025) 979–992.
[72] M. Baker, B.R. Rosic, The hybrid pathway to flexible power turbines, part III: enabling live thermal boundary prediction using real-time plant measurements and LSTM neural networks, J. Eng. Gas Turbines Power 147 (10) (2025).
[73] Y. Zheng, L. Chen, X. Bao, F. Zhao, J. Zhong, C. Wang, Prediction model optimization of gas turbine remaining useful life based on transfer learning and simultaneous distillation pruning algorithm, Reliab. Eng. Syst. Saf. (2025).
[74] F. Schönig, J. Wilmers, A. Sobolyev, A. McGillis, S. Bargmann, Digital twin modelling for compressed air energy storage plants: dynamic modelling of the discharge unit, J. Energy Storage 88 (2024).
[75] I. Aslanidou, M. Rahman, V. Zaccaria, K.G. Kyprianidis, Micro gas turbines in the future smart Energy system: fleet monitoring, diagnostics, and system level requirements, Front. Mech. Eng. (2021).
[76] J. Lim, et al., The epri gas turbine digital twin a platform for operator focused integrated diagnostics and performance forecasting, in: Proceedings of the ASME Turbo Expo, 2021.
[77] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. (2019).
[78] Balbir Shah Mohd Irwan Shah, A.J. Ishak, M.K. Hassan, N.M.H. Norsahperi, Energy 4.0: challenges and enablers of digital twin application in power plant, Elektr. J. Electr. Eng. 23 (1) (2024) 103–124.
[79] M. Tahan, E. Tsoutsanis, M. Muhammad, Z.A. Abdul Karim, Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: a review, Appl. Energy (2017).
[80] Z. Liu, I.A. Karimi, Gas turbine performance prediction via machine learning, Energy 192 (2020).
[81] C. Martin, J. Ranalli, J. Moore, PYroMat: a Python package for thermodynamic properties, J. Open. Source Softw. (2022).

Balbir Shah Mohd Irwan Shah received the B.Eng. degree (Hons.) in electrical and electronic engineering from the Universiti Teknologi Mara (UiTM) and the M.*Sc*. degree in control and automation system engineering from Universiti Putra Malaysia (UPM). He is currently pursuing the Ph.D. degree in control and automation engineering with the Faculty of Engineering, Universiti Putra Malaysia (UPM), Malaysia. His research interests include deep learning, digital twins, AI applications, plant modelling, simulation, control system designs, and power plant engineering. His employment experience included eighteen (18) years in power plant operation, maintenance, and simulation. He is also a Registered Professional Engineer in the field of Electrical Engineering under Board of Engineers Malaysia (BEM), as well as a Corporate Member of the Institution of Engineers Malaysia (IEM).

Asnor Juraiza Ishak received the bachelor's degree in electrical-mechatronic engineering from the University of Technology Malaysia (UTM), the M.*Sc*. degree in control automation system engineering from Universiti Putra Malaysia (UPM), and the Ph.D. degree in electrical, electronic, and system engineering from the Universiti Kebangsaan Malaysia (UKM). She is currently a Senior Lecturer with the Department of Electrical and Electronic Engineering, UPM. Her research interests include intelligent control systems, control system design, pattern recognition, image and signal processing, biomedical engineering, system modelling, rehabilitation, and assistive robotics.

Mohd Khair Hassan received the B.Eng. degree (Hons.) in electrical and electronic engineering from the University of Portsmouth, U.K., in 1998, the M.Eng. degree in electrical engineering from the University of Teknologi Malaysia, and the Ph.D. degree in automotive engineering from Universiti Putra Malaysia (UPM), Malaysia. He is currently an Associate Professor at the Department of Electrical and Electronic Engineering, UPM. His research interests include control systems, automotive control, automation systems, AI applications, x-by-wire technology especially in electric vehicle, that consist of energy management, battery balancing, regerative braking, steer-by-wire, and brake-by-wire. He is also a Registered Professional Engineer in the field of Electronic under Board of Engineers Malaysia (BEM), a Corporate Member of the Institution of Engineers Malaysia (IEM), and a member of the Society of Automotive (SAE).

Nor Mohd Haziq Norsahperi (Member, IEEE) received the B. Eng. and M.*Sc*. degrees in mechatronics engineering from the International Islamic University of Malaysia (IIUM), in 2015 and 2017, respectively, and the Ph.D. degree in electrical engineering from Universiti Teknologi Malaysia, in 2020. He is currently a Senior Lecturer at the Department of Electrical and Electronic Engineering, Universiti Putra Malaysia (UPM). His research interests include nonlinear control, robotics, and artificial intelligence. He is also a Registered Professional Technologist in the field of Electrical & Electronics Technology under Malaysia Board of Technologists.