Sustainable Computing: Informatics and Systems 49 (2026) 101284

Contents lists available at ScienceDirect

Sustainable Computing: Informatics and Systems

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

Explainable and counterfactual lasso regression for resilient micro gas turbine power prediction in smart grids

Sheikh Muhammad Saqib [a], Muhammad Amir khan [b], Tariq Shahzad [c], Muhammad Usman Tariq [d], Tehseen Mazhar [e,j,*], Habib Hamam [f,g,h,i]

[a] Department of Computing and Information Technology, Gomal University, Dera Ismail Khan 29050, Pakistan
[b] School of Computing Sciences, Faculty of Computer Science and Mathematics, Universiti Teknologi Mara, Shah Alam, 40450 Selangor, Malaysia
[c] Department of Computer Engineering, COMSATS University Islamabad, Sahiwal Campus, Sahiwal 57000, Pakistan
[d] Department of Marketing, Operations, and Information Systems, College of Business, Abu Dhabi University, Abu Dhabi, United Arab Emirates
[e] School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan
[f] Faculty of Engineering, Uni de Moncton, Moncton, NB E1A3E9, Canada
[g] School of Electrical Engineering, University of Johannesburg, Johannesburg 2006, South Africa
[h] International Institute of Technology and Management (IITG), Av. Grandes Ecoles, Libreville, BP 1989, Gabon
[i] Bridges for Academic Excellence - Spectrum, Centre-ville, Tunis, Tunisia
[j] Department of Computer Science and Information Technology, School Education Department, Government of Punjab, Layyah 31200, Pakistan

ARTICLE INFO

Keywords:
Explainable AI (XAI)
SHAP and LIME
Machine Learning Framework
Lasso Regression
Sustainable Smart Grids
Resilient Cyber-Physical Systems

ABSTRACT

Accurate prediction of electrical power output from micro gas turbines is essential for optimizing performance in microgrids and distributed power systems. This study introduces a novel and interpretable machine learning framework using Lasso regression applied to a newly published dataset titled Micro Gas Turbine Electrical Energy Prediction, available on Kaggle. The dataset captures time-series relationships between input control voltage and electrical power output, enabling effective modeling of micro turbine behavior. The proposed model relies on only two features, input voltage and time, ensuring computational efficiency while maintaining predictive performance. To support decision-making and model transparency, the framework incorporates Explainable AI (XAI) techniques such as SHAP and LIME, which reveal the influence of input features on predictions. Additionally, counterfactual analysis is integrated to explore how changes in inputs affect predicted outcomes. This allows users to define a minimum and maximum range for desired power outputs, providing actionable insight. The approach demonstrates high accuracy, with over 87 % of predictions falling into the low or category. By enabling interpretable and resource-efficient forecasting of local energy generation, the proposed framework contributes to the development of resilient and sustainable smart grid infrastructures. Most importantly, the proposed system is highly relevant for smart grid and microgrid operations, where transparent, accurate, and adaptive prediction of local generation units like micro gas turbines plays a critical role in maintaining system stability, load balancing, and energy efficiency.

1. Introduction

1.1. Context

Micro gas turbine engines have been widely considered based on their high thrust-to-weight ratio, low fuel rate, simple design, and low price from the last decade [1]. The efficiency of contemporary Combined Cycle Power Plants (CCPPs) normally surpasses 60 % with lower specific emissions while providing quick start-ups at reduced operation and maintenance costs. These characteristics place CCPPs among the most promising solutions for carbon emission reduction in power generation [2,3]. The cornerstone of CCPPs is the gas turbine, whose design optimization is crucial for fuel consumption and emission minimization [4]. One of the main performance indicators of these units is the heat rate (HR), which is the quantity of fuel energy used to produce a unit of electricity (kJ/kWh) [5].

* Corresponding author at: School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan
E-mail addresses: saqibsheikh4@gu.edu.pk (S.M. Saqib), amirkhan@uitm.edu.my (M.A. khan), tariq@cuisahiwal.edu.pk (T. Shahzad), Muhammad.kazi@adu.ac.ae (M.U. Tariq), tehseenmazhar719@gmail.com (T. Mazhar), Habib.Hamam@umoncton.ca (H. Hamam).

https://doi.org/10.1016/j.suscom.2025.101284
Received 15 July 2025; Received in revised form 30 October 2025; Accepted 13 December 2025
Available online 15 December 2025

Over the past few years, Artificial Intelligence (AI) application in energy system modeling has increased remarkably. Applications of machine learning [6,7] and deep learning [8,9] have shown significant potential in well-estimating key performance parameters in power systems. Many studies have been able to effectively apply the same to predict variables like power generation and heat rate, gaining a better understanding of system performance. Yet, in all these advancements, a very important modeling aspect is still not explored.

One of the current research gaps that were identified relates to the application of electrical power output (el_power) as an important input factor of the determinant of the heat rate in ANN models. Although previous researchers have used a model of power output and heat rate or used them independently, there has been no fully considered interdependence of both el_power and a determinant of HR. Moreover, the majority of the current research projects on micro gas turbine datasets are more aimed at predicting the performance of the turbine, neglecting the value of model transparency and the possibility of decision support. Namely, Explainable AI (XAI) techniques and counterfactual analysis are not well-integrated or not integrated at all, which is the key to understanding feature influence and providing recommendations on how to alter the system inputs to obtain the desired results.

To cover this gap, the proposed work makes use of a newly published dataset on kaggle, named Micro Gas Turbine Electrical Energy Prediction. This data set serves time-series regression and it measures the electrical power output of a commercial micro gas turbine in response to changes in the input controlling voltage with time. The dataset can be used in the modeling of a power grid especially in a small scale or distributed generation setting where it is important to predict power generation of a local plant such as micro gas turbines [10].

In contrast to the majority of previous efforts, the suggested method uses a simple Lasso regression model for el_power prediction. The model utilizes only two features—time and input voltage—thus ensuring simplicity and lower computational complexity, which is preferable in real-time applications over resource-limited environments.

This lightweight approach is particularly suited for integration into decentralized energy systems, where scalable, interpretable, and efficient AI solutions are needed to support resilient and sustainable grid operations.

What distinguishes the proposed work is its dual focus on prediction and decision support. In addition to predicting power output, the model integrates counterfactual analysis to answer "what-if" scenarios. Specifically, users are enabled to interact with an interface that allows adjustment of input values (time and voltage) within a defined minimum and maximum range, simulating new outcomes for el_power. This functionality is crucial for operators seeking to achieve specific power outputs by identifying the necessary adjustments in control settings. Moreover, XAI techniques such as SHAP or LIME are applied to interpret the model's predictions by highlighting the relative influence of each input feature. This increases the model's transparency and supports informed operational decisions.

Taken together, the proposed framework not only advances accurate power prediction in microgrids but also enhances operational adaptability, transparency, and trust—critical requirements for cyber-physical systems operating within intelligent, secure, and sustainable power grids.

This work's novelty lies in a unified, operator-facing pipeline that couples a lightweight Lasso predictor with formal explainability (SHAP/LIME) and operationally constrained counterfactual 'what-if' analysis. Unlike prior micro-turbine studies that emphasize accuracy with complex models but offer limited transparency or prescriptive guidance, our framework delivers (i) deployable simplicity for edge/embedded use, (ii) traceable, physics-consistent attributions for operator trust, and (iii) actionable counterfactuals screened by dataset bounds and empirical ramp-rate limits. We also make the evaluation scope explicit—best performance in steady/quasi-steady regimes—and pre-specify robust, group-aware/time-blocked cross-validation for future same-dataset benchmarking against higher-capacity baselines.

1.2. Key contributions

The proposed work offers the following key contributions:

- A lightweight Lasso regression model is developed to predict the electrical power output (el_power) of a micro gas turbine using only two input features—input voltage and time—ensuring simplicity and computational efficiency for resource-constrained and embedded deployments. This choice emphasizes stability and low latency over black-box complexity, aligning the model with on-device microgrid controllers.
- Integration of XAI techniques, specifically SHAP and LIME, provides transparency into the model's predictions by revealing the relative importance and influence of each input feature. In particular, SHAP confirms the physically consistent dominance of input_voltage (higher voltage → higher predicted power), while time exerts a smaller, context-dependent effect, enhancing operator trust in the model's behavior.
- Counterfactual analysis is incorporated to enable actionable insight through "what-if" scenarios, allowing users to explore how variations in input voltage and time affect predicted power output under bounded, operationally meaningful ranges—turning explanations into prescriptive adjustments for scenario planning and control.
- A unified, decision-oriented pipeline—prediction → explanation → counterfactual prescription—is introduced to address documented gaps where lightweight models typically lack formal XAI and prescriptive tooling. This integration targets operator-facing transparency and actionable guidance for resilience in microgrid operations.
- Real-world applicability of the framework is demonstrated through its relevance to smart grid and microgrid environments, where it supports performance optimization and operational planning for local generation units. The end-to-end design explicitly prioritizes interpretability and deployability to support routine decision-making by grid operators.

2. Literature review

2.1. Current research

The energy load forecast is essential for the economic development of a nation with proper energy management, energy consumption can be reduced [11]. In the first research works, statistical models were applied to discover underlying hidden statistical standards in renewable online time-series data [12], but recent studies used the latest deep learning models [13]. In particular, deep sequence models—artificial neural networks (ANN), convolutional neural networks (CNN), and long short-term memory (LSTM)—have been applied to micro-turbine power prediction and broader energy time series; in one study, LSTM performed comparably to CNNs but slightly worse than traditional ANNs, with reported MAPE around 20 % [14]. Subsequent works show that attention-enhanced LSTM variants can improve sequence modeling for industrial/energy data [15], and LSTM has also been used effectively for short-time load forecasting in power systems [16].

The study applies advanced deep learning architectures such as artificial neural networks (ANN), CNNs and short-term memory networks (LSTM) to predict the electrical power of micro gas turbines. The Kaggle/UCI-Gas-Turbine data set is used, which includes data from temporal and environmental parameter series. The approach focuses on complex model architectures capable of capturing nonlinear temporal dependencies. However, no XAI technique or counterfactual analysis is used to interpret the model outputs [17]. Extreme Learning Machines (ELM) and modified linear activation models (MLAM) are applied to simulate the performance of gas micro turbines using a simulated data set that reflects the characteristics of the turbine. The models are complex and provide reasonable generalization capacity, but the work does not involve XAI or counterfactual explanations for interpretability [18]. A light model has been designed using lasso regression in combination with temporal modeling techniques. The Kaggle/UCI-Gas-Turbine data set is used for training and testing. The emphasis is on simplicity and modeling with resource efficiency. Although the model is suitable for interpretability, no formal method of XAI or counterfactual analysis is included in the study [19].

Ensemble learning techniques and statistical models are used in an experimental turbine data set to increase forecasting accuracy. The modeling approach is complex and aims to extract collective strength from several learners. Despite performance evaluation, the study does not offer information through XAI or counterfactual methods [20]. The representation of the High Dimension Model, along with Artificial Neural Networks, is used to predict the performance of large-scale gas turbine systems. A real-world data set is used that captures extensive operational data. Although modeling is powerful and high in complexity, there is no evidence of XAI or counterfactual explanation in the analysis [21]. Machine learning techniques, including XGBoost models and transformer-based approaches such as SAINT, are used to predict turbine emissions using the Siemens testbed data set. The modeling structure is complex and deals well with the tabular temporal series data. However, the study lacks any XAI component or counterfactual analysis for model interpretability [22].

Beyond turbines, recent energy-forecasting studies report that gradient-boosted ensembles (e.g., XGBoost/LightGBM) often achieve strong accuracy and stability; many of these works integrate SHAP to provide global/local feature attributions [23–26]. Interpretable frameworks such as iXGB further extend this line by generating decision rules and counterfactuals to improve transparency and actionability [27]. Physics-based methods combined with machine learning are used to detect micro gas turbines, leveraging operational data from turbine fleets. The hybrid system is designed for real-time failure detection and monitoring. Although the model is robust and complex, there is no support for explainable results using XAI or counterfactual analysis [28]. An AI hybrid monitoring system is implemented to track the operation of micro gas turbines under smart grid environments using real-world fleet data. The approach is moderately to highly complex and supports dynamic monitoring. However, it does not integrate XAI frameworks or counterfactual interpretation tools [29]. Robust optimization methods, particularly $\ell\infty$ and $\ell1$ norms, are applied to solve the economic dispatch problem in a micro gas turbine setup within a Combined Heat and Power (CHP) system. The data is simulated for practical settings. The optimization models are mathematically complex, but the work does not include any XAI techniques or counterfactual exploration [30]. Standard regression techniques are applied to model the performance of experimental micro-turbojets, with a dataset derived from physical test environments. Though the models are simpler compared to deep learning, there is no use of XAI or counterfactual reasoning to support prediction analysis [31].

Notably, recent research in environmental modeling has successfully applied XAI tools, such as SHAP and agnostic models like Explainable Boosting Classifier, for transparency in predictive frameworks []. However, unlike the currently available literatures, the proposed model is dedicated to the employment of the lightweight mechanism resting on the implementation of the Lasso regression algorithm in terms of electrical power output (el_power) forecasting of micro gas turbines. This model is more concerned with ease and workability of deployment in less-resource settings. XAI methods like SHAP and LIME are included in order to provide transparency and assist in decision-making, and are also used in order to determine powerful features. Also, counterfactual analysis is used to create actionable information by presenting how input may change minorly to change predictions. This outlook closes the field in terms of the lack of prior research, as it integrates lightweight modeling and explicit interpretability. In summary, prior work emphasizes accuracy via deep or ensemble models [23–26] while interpretable rule-/counterfactual-based approaches are only beginning to appear [27]; our contribution is to pair a lightweight predictor with formal SHAP/LIME attributions and counterfactual "what-if" analysis tailored to micro-turbine deployment.

2.2. Research gap

Despite recent advances in machine learning and deep learning approaches for predicting the electrical performance of micro gas turbines, several critical gaps remain unaddressed in the current literature:

- Lack of Model Interpretability: Many existing models—including deep architectures such as ANN, CNN, and LSTM—prioritize predictive accuracy but neglect model interpretability. These black-box models do not provide transparent insights into how input features influence predictions, which limits their adoption in operational energy systems where explainability is crucial for trust and decision-making.
- Absence of Counterfactual Reasoning: While most prior studies focus solely on forecasting output values, they fail to incorporate counterfactual analysis. This omission restricts the ability to simulate "what-if" scenarios, thereby preventing users from understanding how changes in input parameters (e.g., voltage or time) could be adjusted to achieve desired power outputs.
- High Computational Complexity: Existing methods often involve complex model architectures or ensemble techniques, which are computationally intensive and unsuitable for real-time applications or deployment in resource-constrained environments such as embedded control systems in microgrids.
- Limited Focus on Lightweight and Interpretable Models: Although some studies explore simplified models such as Lasso regression, they typically do not integrate formal XAI tools (e.g., SHAP or LIME) or counterfactual methods. As a result, these models fall short of supporting operational decision-making with transparency and actionable feedback.

To bridge these gaps, the present study proposes a lightweight yet powerful Lasso regression framework that combines predictive modeling with XAI and counterfactual analysis. This integrated approach is specifically designed to enhance model transparency, support dynamic scenario planning, and ensure applicability in real-world microgrid and smart grid environments.

3. Proposed model

3.1. Data and splits

The proposed work introduces a regression-based machine learning approach to predict the electrical energy output of a micro gas turbine using historical sensor data. This framework combines data preprocessing, predictive modeling with Lasso Regression, interpretability through LIME (Local Interpretable Model-Agnostic Explanations), and counterfactual analysis to provide actionable insights as shown in Fig. 1.

The given dataset [32] is split into two major parts of the training set and testing set. The training data set is constructed using six separate files which are: ex_1.csv, ex_9.csv, ex_20.csv, ex_21.csv, ex_23.csv, and ex_24.csv. These individual files are read and they are grouped together to make one dataset that can be used to train models. The testing dataset consists of only two files namely; ex_4.csv and ex_22.csv. Just like the training files, these two test files would be read and joined into one data set to be used in testing the trained model. The system starts with the Micro Gas Turbine Electrical Energy Dataset that comprises of readings of three important variables that are time, input_voltage and el_power. Time and input voltage in this arrangement are independent variables and electrical power (el_power) is the target (dependent variable) that

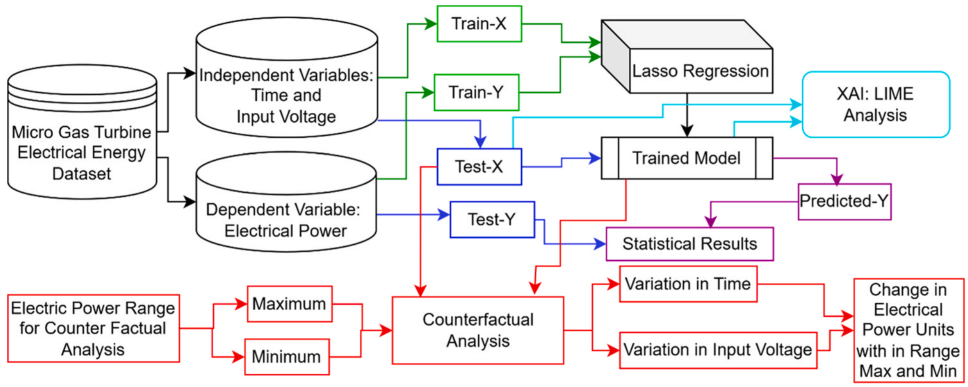

Fig. 1. Framework Integrating Lasso Regression, LIME Explainability, and Counterfactual Analysis for Electrical Power Prediction in Micro Gas Turbine Systems.

needs to be predicted. The data is divided into two, one to train and another one to test the model. We adopt a file-level hold-out to mirror deployment on unseen operating runs and to avoid temporal leakage between contiguous samples of the same run. While this choice aligns with operational usage, a single split can over- or under-estimate generalization; therefore, we treat it as a baseline estimate and complement it with a prespecified cross-validation protocol described below.

The dataset comprises 52,940 training records (from files ex_1.csv, ex_9.csv, ex_20.csv, ex_21.csv, ex_23.csv, ex_24.csv) and 18,285 test records (from ex_4.csv, ex_22.csv), for a total of 71,225 time-stamped instances with three columns: time, input_voltage, and el_power. These files represent distinct operating runs, providing variation across steady segments and ramp/step transitions; no additional environmental or plant-context features are included.

The schema of the dataset should be as follows (Table 1):

To build the predictive model, we use Lasso regression, which performs variable selection and L1 regularization to reduce overfitting. The estimator solves the optimization problem in (Eq-1):

$$\textit{Lasso Regression} = \underset{\beta}{Min}\left(\frac{1}{2n}\sum\nolimits_{i=1}^{n}(y_i - x_i\beta)^2 + \lambda\sum_{j=1}^{P}|\beta_j|\right) \tag{1}$$

where y_i is the observed electrical power, $\mathbf{x}_i$ is the feature vector (time, input_voltage), β is the coefficient vector, and $\lambda > 0$ controls the strength of regularization.

Once the model is trained, it predicts power output for unseen data, producing a vector of predictions y^. To make the model more interpretable, XAI techniques are applied using LIME, which generates human-understandable explanations for the predictions. LIME locally approximates the complex model f using a simpler model g, minimizing the following objective calculated in Eq-2:

$$\text{LIME} = \mathscr{L}(\text{f}, \text{g}, \pi_x) + \Omega(\text{g}) = \arg \underset{g\in G}{\min}\left[\sum\nolimits_{z\in Z}\pi_x(z)(f(z) - g(z))^2 + \Omega(g)\right] \tag{2}$$

In this equation, L(f,g,πx) measures how well g approximates f near instance xxx, πx defines the locality around x, and Ω(g) penalizes model complexity to ensure interpretability.

Next, counterfactual analysis is performed to explore "what-if" scenarios by modifying input features and observing the effect on predicted power output. The first step is to compute the range of the target variable, electrical power, defined by Eq-3:

$$P_{\text{range}} = [P_{\min}, P_{\max}] \tag{3}$$

Where P_{min} and P_{max} calculated in Eq-4 and Eq-5.

$$P_{\min} = \underset{i}{Min}(y)_i \tag{4}$$

$$P_{\max} = \underset{i}{Max}(y)_i \tag{5}$$

Synthetic variations in the independent variables are introduced to understand how changes in time and input_voltage would alter the predicted outcome. The effect of such changes is captured by Eq-6:

$$\Delta P = f(\text{time} + \delta_t, \text{voltage} + \delta_v) - f(\text{time}, \text{voltage}) \tag{6}$$

In this study, synthetic variations are screened by empirical feasibility rules (dataset-observed bounds and ramp-rate thresholds) before acceptance as candidate counterfactuals.

This expression quantifies the sensitivity of power output to changes in input features.

For evaluating model performance, statistical error metrics are calculated. The prediction error for a single record is given by Eq-7:

$$\text{Error}_i = y_i - \hat{y}_i \tag{7}$$

and the Mean Absolute Error (MAE) for each instance is calculated in Eq-8:

$$\text{MAE}_i = |y_i - \hat{y}_i| \tag{8}$$

These errors are aggregated over all test instances to compute average performance metrics such as MAE, MSE, and R^2.

The derived ‘energy per time’ metric provides an intuitive operational indicator equivalent to instantaneous power output normalized over time intervals, allowing grid operators to compare performance across different micro-turbines or usage cycles. Similarly, when adapted as ‘distance traveled per hour’ in analogous transportation or mobile-generation contexts, this measure aligns with common user-facing performance dashboards. Expressing outputs in such normalized, rate-based metrics facilitates direct interpretation by technicians and operators for scheduling, maintenance, and efficiency benchmarking.

3.2. Robustness protocol

To provide split-invariant estimates while respecting data structure, we specify two complementary procedures: (1) Group-aware K-fold (files as groups) so that folds do not mix samples from the same file/run across train and validation; (2) Blocked time-series CV within each file (non-overlapping chronological blocks) to preserve temporal order. Hyperparameter selection (λ for Lasso) is performed within each training fold only. We report fold-wise metrics and their dispersion

Table 1
Structure of Training and Testing Sets for Lasso-Based Power Prediction Model.

Records	Rows	Features
Total Records	71,225	3
Training Records	52,940	3
Test Records	18,285	3

(median and IQR) alongside the primary file-hold-out results in future releases to ensure comparability. Accordingly, the file-hold-out results should be interpreted as a deployment-oriented baseline, with robustness to data partitioning assessed via the group-aware and time-blocked CV procedures described in Methods.

4. Results and discussions

4.1. Prediction of electrical power

The proposed model utilizes Lasso regression to predict electrical power (el_power) generated by a micro gas turbine system, and its performance is evaluated using a combination of time-series and scatter plot analyses. The model was trained using a dataset consisting of 52,940 records and tested on an unseen set of 18,285 records. Both datasets comprise three features: time, input_voltage, and el_power, as detailed in Table 2 (Training Dataset) and Table 3 (Testing Dataset).

To assess model effectiveness, two visualizations were developed: a time-series comparison and a scatter plot of actual vs. predicted values, offering insights into both trend-following behavior and predictive accuracy.

Fig. 2 is a time-series plot that directly compares the actual electrical power output (blue solid line) with the predicted electrical power output (red dashed line) over the Sample Index of the test dataset. This visualization is crucial for understanding the model's ability to track the dynamic changes in electrical power.

The plot reveals that the model effectively captures the general trends and magnitudes of the el_power variations. For instance, during periods where the actual el_power is relatively stable (e.g., around Sample Index 0–2500, 2500–5000, 7500–10000), the predicted values closely align with the actual values. This indicates good performance in steady-state conditions.

More importantly, the plot highlights the model's response to significant changes in el_power. We observe instances of step changes (e.g., around Sample Index 5000 and 7500) and ramp-like changes (e.g., from Sample Index 10000 onwards, exhibiting a triangular or sawtooth pattern). The predicted values generally follow these transitions, demonstrating the model's capability to adapt to varying operational states of the micro gas turbine. However, it is also evident that the predicted values, particularly during the sharp transitions (steps and ramps), appear to be slightly smoother or to have some lag compared to the actual values. This suggests that while the model captures the overall pattern, it might not perfectly replicate the high-frequency fluctuations or instantaneous responses present in the actual data. Nonetheless, the close visual alignment indicates a strong predictive capability. Taken together, these visuals indicate close alignment primarily in steady segments, with noticeable lag/smoothing during steps and ramps; thus, aggregate alignment should not be interpreted as uniformly high accuracy under rapidly changing conditions.

Fig. 3 is a scatter plot illustrating the relationship between the actual el_power values and the predicted el_power values from the test set. Each point on the scatter plot represents a single data point from the test set, with its x-coordinate being the actual el_power and its y-coordinate being the predicted el_power. The black diagonal line represents the ideal scenario where Predicted El_Power = Actual El_Power.

Table 2
Sample from Train Dataset.

time	input_voltage	el_power
810.0703	10	1228.792
811.0694	10	1223.042
818.84	3.445	1253.79
819.84	3.445	1349.837
2794.7	3.368421	1180.645
2795.74	3.368421	1270.054

Table 3
Sample from Test Dataset.

time	input_voltage	el_power
758.4258	3	1206.993
759.4251	3	1255.072
7912.18	5.394141	1453.099
7913.2	5.394141	1531.88
7914.22	5.394141	1439.963
4316.919	7.5	1495.634
4317.919	7.5	1443.061
4318.918	7.5	1513.495

The concentration of data points tightly clustered around the black diagonal line in Fig. 2 signifies a high degree of accuracy in the model's predictions. If the model were perfectly accurate, all points would lie exactly on this line. The observed spread indicates the presence of some prediction error, which is inherent in real-world modeling.

The distinct horizontal and vertical bands of points in the scatter plot correspond to the different power levels and transitions observed in Fig. 1. For instance, the clusters of points at lower el_power values (e.g., around 1000–1500) and higher el_power values (e.g., around 2500–3000) reflect the discrete operating points of the micro gas turbine. The points forming diagonal trends between these clusters represent the periods of ramping up or down in el_power. The close proximity of these clusters and trends to the ideal line further reinforces the notion that the Lasso model effectively learned the underlying relationship between the input features and the electrical power output across various operating conditions.

4.2. Error categorization and interpretation

Fig. 4 is the heatmap visualizes prediction errors of the Lasso model over time, with the x-axis representing time samples. The "Label: 0" on the vertical axis simply indicates that the heatmap is displaying a single series of error values. White/gray areas indicate accurate predictions (errors near zero), while prominent red and blue vertical bands highlight periods of significant positive (over-prediction) and negative (under-prediction) errors, respectively. These high-error zones align with rapid changes in the actual electrical power, suggesting the model's primary limitation lies in accurately capturing abrupt transitions or dynamic shifts in the micro gas turbine's output. Consistent with this visualization, we treat predictions during rapid ramps/steps as less reliable than those in steady segments; this behavior is expected for a linear, two-feature baseline and is transparently surfaced by the residual/heatmap analyses. For operational use, we therefore flag these intervals as caution zones and position the model for supervisory monitoring/planning rather than fast closed-loop control during rapid ramps/steps.

Table 4 presents a detailed comparison between the actual and predicted values of electrical power (el_power) for a sample of test records selected from Table 3. The error is measured using absolute and percentage error, providing a clearer view of the model's predictive reliability across different operational conditions. To systematically interpret the model's performance, the percentage error is categorized into three groups: Low Error (<20 %), Moderate Error (20–40 %), and High Error (>40 %).

Based on the complete test dataset comprising 18,285 records, 87.05 % of the predictions fall into the Low Error (<20 %) category, indicating that the proposed Lasso regression model achieves high predictive accuracy across the majority of test scenarios. Meanwhile, 7.83 % of the records are classified as having Moderate Error (20–40 %), and only 5.12 % exhibit High Error (>40 %) as shown in Fig. 4. This distribution confirms that the model performs robustly under typical operating conditions, while a small portion of predictions deviate more significantly, likely due to rapid dynamic changes or unseen voltage patterns in the test data.

We employ the above XAI-driven diagnostic to flag ramp-zone errors

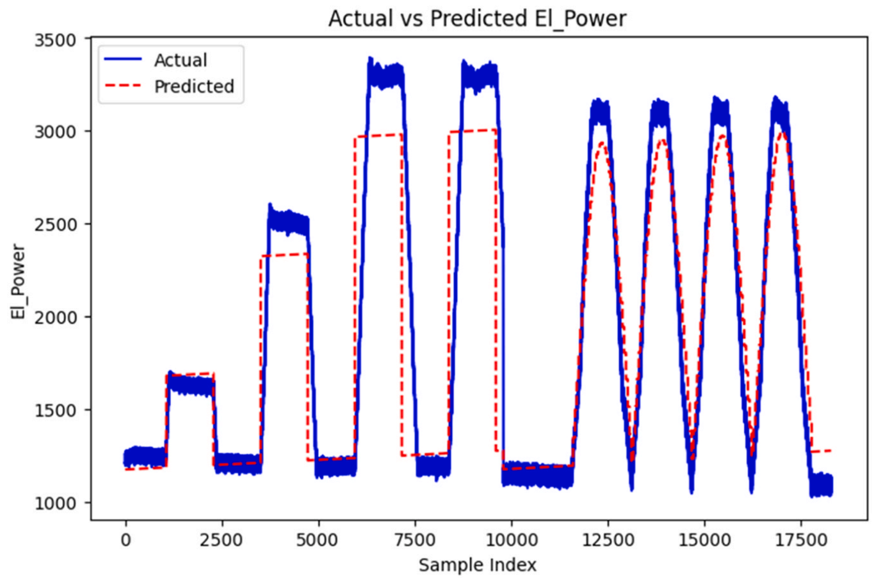

Fig. 2. Actual vs predicted El_Power.

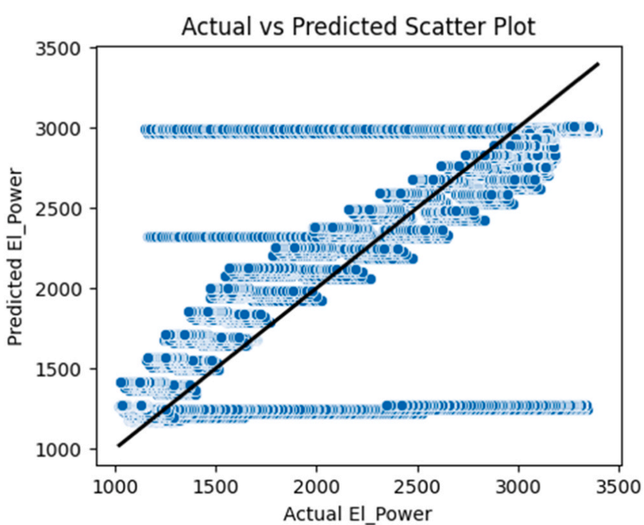

Fig. 3. Actual vs predicted scatter plot.

as expected failure modes for the linear baseline and to separate them from potential data or operational issues in steady segments.

4.3. Performance of the Lasso regression model for electrical power prediction

The performance of the proposed Lasso regression model for electrical power prediction is summarized in Table 5. The R^2 score of approximately 0.78 indicates that nearly 78 % of the variance in the actual power output (el_power) is explained by the model, which we interpret as substantial but not complete variance explanation under the evaluated split. The Explained Variance Score is 0.7814, consistent with R^2; however, errors are not uniform across operating conditions, particularly during rapid transitions.

The Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), recorded as 143,915.62 and 379.36 respectively, reflect the magnitude of the average prediction errors in squared and original units. These magnitudes should be understood alongside the residual/heatmap evidence that highlights larger deviations during sharp ramps/steps.

The Mean Absolute Error (MAE) of 217.81 quantifies the average absolute deviation in the original units; this error remains non-trivial during dynamic intervals even when steady segments are well tracked.

Importantly, the Mean Absolute Percentage Error (MAPE) is 10.97 %, i.e., an average ≈ 11 % relative deviation at the dataset level; in power-system operations such margins can be material—especially at change points—so these values should be interpreted with caution.

The Median Absolute Error of 124.00 is notably lower than the MAE, indicating that the majority of predictions are even closer to the actual values, and that a few outliers may be inflating the mean. Lastly, the Max Error of 2096.65 suggests that in the worst-case scenario, the model made a large deviation, which is consistent with the higher errors observed during abrupt transitions in the residual/heatmap analyses.

Overall metrics (e.g., MAPE ≈ 10.97 %; $R^2 \approx 0.78$) should thus be interpreted as a blend of strong performance in steady segments and concentrated errors during abrupt transitions—a trade-off made explicit by our XAI-supported error analyses.

Overall, these metrics support using the model as a lightweight, interpretable baseline for supervisory monitoring and planning in steady or quasi-steady regimes, rather than as a sole controller during fast transients.

Our choice of Lasso is intentional: it prioritizes low latency, stability, and interpretability for embedded or resource-constrained microgrid controllers, where transparent behavior is essential for operator trust and routine decision support. As documented in the residual and heatmap analyses, transient spikes are the principal error mode, whereas steady regimes exhibit residuals near zero and a tight actual–predicted alignment.

In line with the error visualizations, the principal limitation is reduced fidelity at rapid ramps/steps; consequently, results are most reliable in steady segments, and we present the method as a transparent baseline to support monitoring and planning rather than fast closed-loop control.

All reported metrics reflect this single public dataset and file-level split; generalization beyond these operating runs and feature constraints should be interpreted with caution (Fig. 5).

4.4. Residuals curve of lasso regression model for electrical power prediction

Fig. 6 and Fig. 7, collectively termed "Residuals Curve for Electrical Power Prediction," analyze the errors (residuals) of the Lasso model. Fig. 6, plots these errors over time, revealing that while predictions are

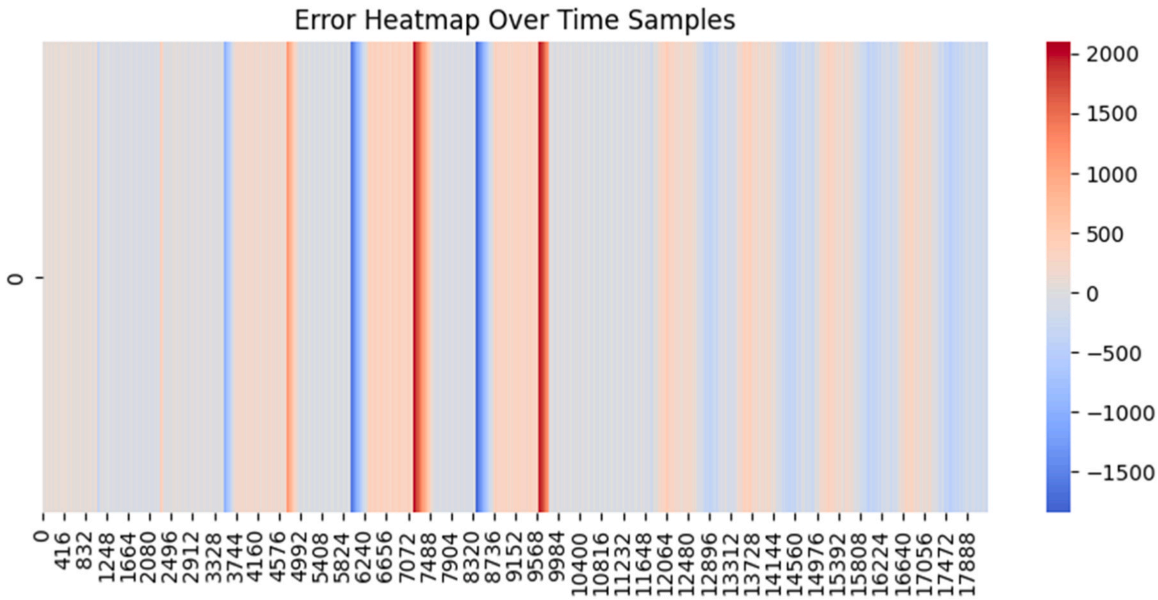

Fig. 4. Error heatmap over time samples.

Table 4
Error percentage for each records from sample test data.

time	Input voltage	el_power	Predicted	Error	Absolute Error	Percentage Error	Error_Category
758.4258	3	1206.993	1172.3662752478897	34.62715831924038	34.62715831924038	2.868877108710012	Low Error (<20 %)
759.4251	3	1255.072	1172.3767361067405	82.69538544179954	82.69538544179954	6.5888950939144335	Low Error (<20 %)
7912.18	5.394141	1453.099	1839.8796914637417	-386.78	386.7801918744117	26.617598587276518	Moderate Error (20 < and > 40 %)
7913.2	5.394141	1531.88	1839.8903690550912	-308.01	308.01049343665113	20.106700162263262	Moderate Error (20 < and > 40 %)
7914.22	5.394141	1439.963	1839.9010466464406	-399.938	399.93827877094054	27.77420970133857	Moderate Error (20 < and > 40 %)
4316.919	7.5	1495.634	2323.5109876221286	-827.877	827.8767800373687	55.35289149171534	High Error (>40 %)
4317.919	7.5	1443.061	2323.5214484809794	-880.46	880.4604261242894	61.013388379542874	High Error (>40 %)

Table 5
Lasso Regression Evaluation Metrics for Electrical Power Prediction.

Metric	Value
R^2 Score (Accuracy)	0.7799
Mean Squared Error (MSE)	143,915.62
Root Mean Squared Error (RMSE)	379.36
Mean Absolute Error (MAE)	217.81
Mean Absolute Percentage Error (MAPE)	10.97 %
Median Absolute Error	124.00
Max Error	2096.65
Explained Variance Score	0.7814

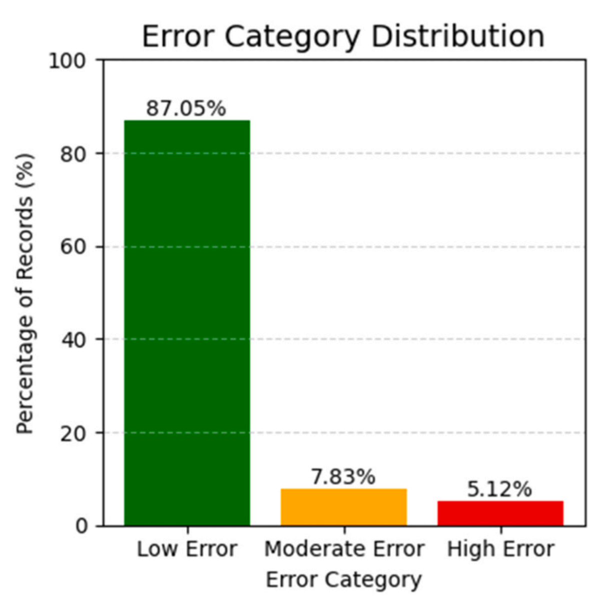

Fig. 5. Error Category Distribution.

accurate during stable periods (residuals near zero), significant positive and negative spikes occur during rapid increases and decreases in electrical power, respectively, indicating the model struggles with dynamic transitions. Fig. 7, "Residuals Distribution," a histogram, shows that most errors are clustered around zero, confirming overall good performance, but the presence of wider tails and secondary peaks highlights the larger, systematic errors that occur during those aforementioned rapid changes in power output.

4.5. Conceptual contrast with common baselines

- Ordinary Linear Regression (OLS): Shares the linear form of Lasso but lacks L1 regularization, making coefficients less stable under multicollinearity and potentially more sensitive to noise. Lasso's

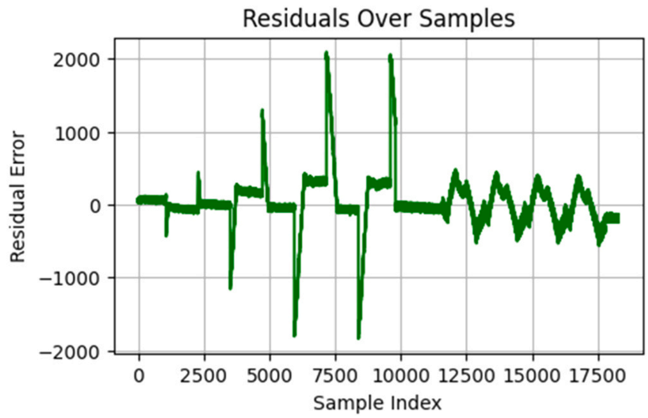

Fig. 6. Residuals over samples for electrical power prediction.

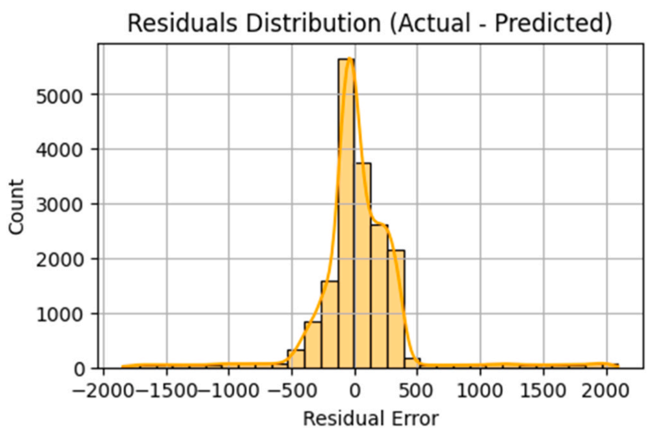

Fig. 7. Residuals distribution for electrical power prediction.

shrinkage yields simpler, more stable coefficient profiles with similar inference cost.

- DecisionTreeRegressor: Captures nonlinearity with axis-aligned splits and is intuitively interpretable at the single-tree level, but it can be high-variance and piecewise-constant, leading to unstable rules and jagged responses around ramps/steps. Local explanations may change substantially with small data perturbations.
- RandomForestRegressor: Reduces variance via ensembling and often improves accuracy over a single tree, but global interpretability declines and inference/footprint increase with the number and depth of trees. Post-hoc explanations (e.g., SHAP) are feasible yet heavier computationally than for sparse linear models.
- Rationale for Lasso in this study: We prioritize transparent coefficients, low latency, and small memory footprint on constrained devices. L1 regularization offers robustness to correlated inputs and supports straightforward SHAP/LIME attributions and counterfactual 'what-if' probes without surrogate modeling or heavy explanation overhead.

5. Explainable AI for of lasso regression model

To gain deeper insights into the behavior of the Lasso regression model for electrical power prediction, XAI techniques were employed. Specifically, SHAP (SHapley Additive exPlanations) analysis was utilized for a global understanding of feature importance across the entire dataset, while LIME (Local Interpretable Model-agnostic Explanations) provided localized explanations for individual predictions.

5.1. SHAP analysis for global feature importance

The SHAP summary plot in Fig. 8 provides critical insights into the global importance of features for the Lasso model's electrical power predictions. This plot visualizes the impact of each feature on the model's output, with high feature values (red points) and low feature values (blue points) pushing the prediction higher or lower, respectively. It clearly demonstrates that input_voltage is the most influential feature, as higher input voltages strongly correlate with higher predicted electrical power, a physically intuitive relationship for the micro gas turbine. In contrast, the time feature exhibits a significantly smaller and less consistent impact, suggesting its primary role might be contextual rather than directly predictive of power output in a global sense.

5.2. LIME analysis for global feature importance

While SHAP provides global insights into feature importance, LIME (Local Interpretable Model-agnostic Explanations) was utilized to provide local, instance-specific explanations for individual predictions of the Lasso model. To illustrate this, four distinct records from the test dataset—'Record: 0′, 'Record: 1067', 'Record: 5951', and 'Record: 14763'—were selected for detailed LIME analysis. Each LIME plot visualizes how specific feature values within that record contribute positively (orange) or negatively (blue) to the model's prediction relative to a baseline, with the length of the bar indicating the magnitude of impact.

XAI-driven diagnostic (Record 5951). The LIME panel for Record 5951 indicates a strong positive contribution from high input_voltage and a much smaller positive contribution from time, yet the actual output is far lower than predicted.We interpret this mismatch by (i) localizing the error on the residual heatmap, which shows high-error bands aligned with rapid ramps/steps; (ii) checking attribution consistency against the global SHAP summary, which confirms input_voltage as the dominant driver; and (iii) examining the counterfactual band ($\pm$100 units around the current prediction) to see whether feasible, small input changes would reconcile the discrepancy in this regime. Taken together, these steps suggest a transient-regime miss: the linear, two-feature baseline appropriately links higher voltage to higher predicted power, but during rapid dynamics other (unmodeled) factors likely govern the realized output, yielding a large residual.

For the contrast with deep learning black-box models, deep architectures (e.g., CNNs, LSTMs, Transformers) can achieve higher raw

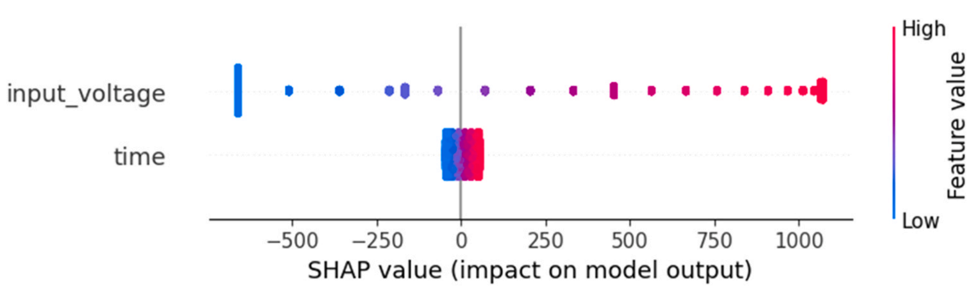

Fig. 8. SHAP analysis for global feature importance.

accuracy in some energy time-series tasks—especially under strongly nonlinear dynamics—but their internal decision logic is opaque and harder to audit, certify, or deploy on embedded controllers; post-hoc explanation is heavier, and inference latency/memory are typically higher. In contrast, the sparse Lasso baseline offers transparent coefficients, consistent global/local attributions (via SHAP/LIME), and actionable counterfactuals that operators can directly inspect and use for "what-if" planning. We therefore position the present approach as a complementary, interpretable option for supervisory monitoring and planning, with hybrid/nonlinear extensions reserved for future transient-focused upgrades.

Fig. 9 presents a LIME explanation for a specific data point, 'Record: 0′. It illustrates that for this record (where input_voltage = 3.00 and time = 758.43), the condition input_voltage < = 3.00 is the primary factor driving the predicted electrical power downwards (negative contribution, blue bar). The time feature (time <= 3086.93) also contributes negatively, though to a lesser extent, further reducing the prediction for this particular instance. This highlights how low input voltage locally influences the model's output.

Fig. 10 illustrates the local explanation for a specific data point, 'Record: 1067'. For this instance, an input_voltage value falling within the range of 4.84 < input_voltage … (which 5.00 satisfies) makes a significant positive contribution (orange bar) to the predicted electrical power (1678.59). Conversely, the time feature (specifically, time <= 3086.93, which 1824.67 satisfies) contributes negatively, albeit to a much lesser extent (blue bar). This local analysis reinforces the strong positive influence of higher input voltage on the electrical power prediction for this specific record.

Fig. 11 displays the LIME explanation for a specific data point, 'Record: 5951'. For this instance (where input_voltage = 10.00 and time = 6705.24), the condition input_voltage > 7.42 makes a very strong positive contribution (orange bar) to the predicted electrical power, pushing it significantly higher towards the predicted value of 2967.34. The time feature (specifically, 5431.03 < time <= 7709.84) also contributes positively, but to a much smaller extent. This instance notably highlights a large discrepancy between the high predicted value and a much lower actual value (1160.04), suggesting that while the model correctly identifies high voltage as a strong predictor, other unmodeled factors or specific operational conditions might be influencing the actual output at this particular time.

Fig. 12 presents the LIME explanation for a specific data point, 'Record: 14763'. For this instance (where input_voltage = 4.22 and time = 6399.34), the input_voltage in the range 3.00 < input_voltage … contributes negatively (blue bar) to the predicted electrical power. Conversely, the time feature (specifically, 5431.03 < time <= 7709.84) makes a small positive contribution (orange bar). This local explanation illustrates how, for this particular record, the input_voltage value is the primary driver influencing the prediction downwards, while time has a minor upward influence.

Diagnostic workflow for XAI use. For any record exhibiting unusually large error: (1) localize the error on the residual heatmap/time series to identify whether it occurs in a ramp/step zone; (2) cross-check attributions for consistency—compare the record's LIME bars with the global SHAP ranking (e.g., voltage dominance); and (3) probe locality with the predefined counterfactual band (±100 units) to assess whether small, feasible input shifts would close the gap or whether the miss is regime-driven. We use Record 5951 to illustrate this procedure.

5.3. Understanding Changes in Output Electrical Power via Counterfactual Inputs

As the problem at hand is a regression task, the predicted output—electrical power (el_power)—is a continuous variable. Therefore, the counterfactual analysis can be guided by specifying a desired range of predicted values, allowing us to explore how the output can be increased or decreased by modifying the input features, namely time and input voltage.

This interpretability-enhancing approach aligns with prior applications of counterfactual reasoning in transparent energy-related modeling frameworks, such as those introduced in [33].

In the proposed interface (highlighted with a red guideline in Fig. 1), users are allowed to define the minimum and maximum target values for el_power predictions. For example, a range such as (current_el_power + 100, current_el_power + 120) instructs the system to generate counterfactual instances where the model predicts a power output within that upper-shifted interval. This enables the discovery of feature configurations that lead to a desired increase in power output.

The ±100 band is used not only to illustrate sensitivity but also as a local probe in high-error records (e.g., 5951) to check whether small, feasible input adjustments can recover the target prediction in that neighborhood or whether the discrepancy reflects regime limits of the linear model.

Importantly, the desired_range is flexible and adjustable. For a more conservative analysis, one may set the range to (current_el_power − 100, current_el_power + 100), effectively exploring model behavior and feature influence within a ± 100 unit margin around the current prediction.

To demonstrate the effectiveness of counterfactual generation in this study, a desired output range of −100 to + 100 units relative to the current prediction was specified. This range enables the model to explore counterfactuals that result in predictions close to the original output, within a ± 100 unit tolerance. The purpose of setting such a bounded range is to investigate minimal yet meaningful changes in input features, specifically time and input_voltage, that lead to slightly different but realistic outcomes. This facilitates a more interpretable understanding of how the model responds to subtle variations in input conditions.

To ensure practical relevance, we constrain counterfactual proposals to (i) observed input domains for time and input_voltage (no extrapolation), and (ii) empirically plausible change rates. In particular, we estimate the distribution of ramp rates $|\Delta(\text{input\_voltage})/\Delta(\text{time})|$ from the training set and reject candidates exceeding a high-percentile threshold (e.g., 95th), thereby enforcing actuator-consistent 'slew-rate' limits. Time adjustments are further limited to forward-feasible windows consistent with operational scheduling granularity.

To validate this approach, five counterfactual examples were extracted for each of the following four selected records: Record-0, Record-1067, Record-5951, and Record-14763. These records were also analyzed using LIME explanations to identify influential features. The corresponding counterfactual scenarios illustrate the local sensitivity of the model and highlight which input feature changes most significantly influence the output prediction, all while keeping the predicted

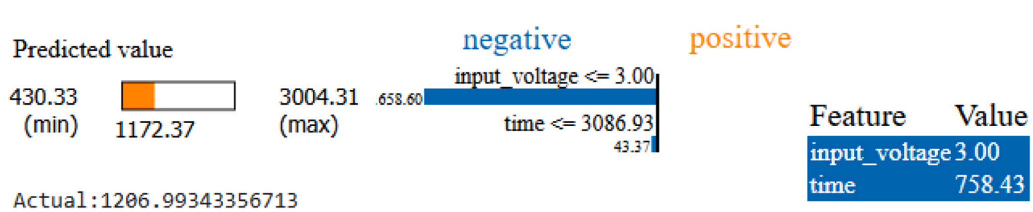

Fig. 9. Lime analysis for 'Record: 0'.

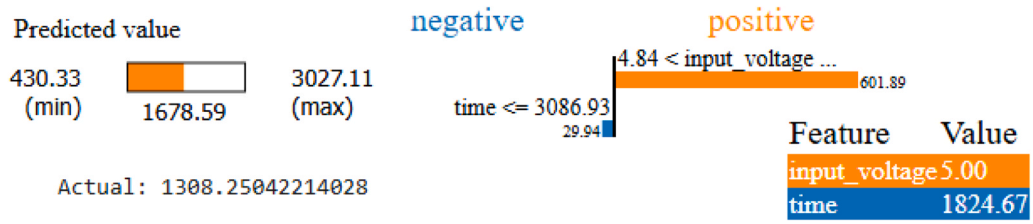

Fig. 10. Lime analysis for 'Record: 1067'.

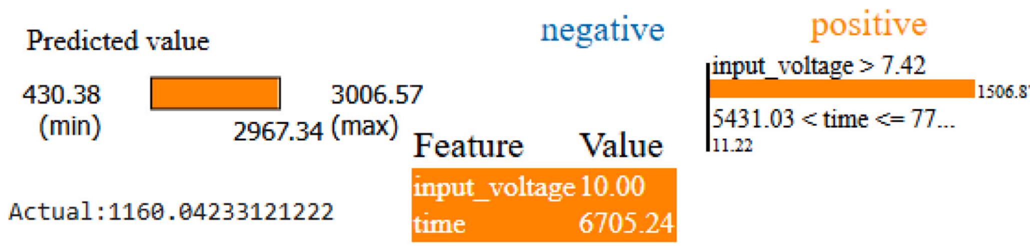

Fig. 11. Lime analysis for 'Record: 5951'.

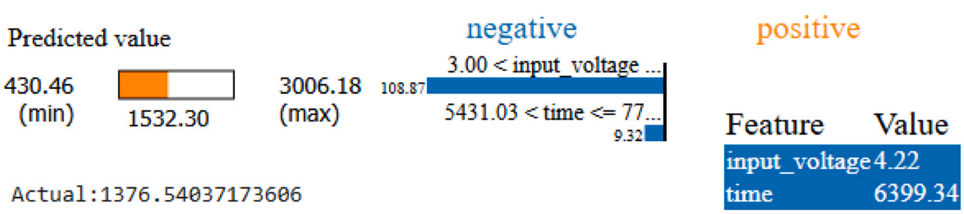

Fig. 12. Lime Analysis for 'Record: 14763'.

electrical power (el_power) within a controlled and interpretable range. This contributes to both model transparency and practical utility in operational settings where slight adjustments to system inputs are preferable to drastic interventions [33].

Table 6 presents the counterfactual analysis results for Record-0. The actual input values consist of a time of 758.43, an input voltage of 3.0, and a predicted electrical power (el_power) of 1172.0. The five counterfactual instances demonstrate alternative input configurations that result in slightly higher power predictions, all within the desired range of ±100 units from the actual value. These counterfactuals suggest that moderate increases in time or slight changes in input_voltage (e.g., to 3.3) can lead to improved predicted power outputs, indicating the model's sensitivity to these input features in the local region around Record-0.

Table 7 presents the counterfactual analysis for Record-1067, where the actual input (time = 1824.67, input_voltage = 5.0) results in a predicted electrical power of 1679.0. The five counterfactuals show how adjusting the input_voltage downward to values between 3.1 and 3.9, and varying time significantly (e.g., up to 10039.88 or down to 1404.40), leads to reduced power predictions ranging from 1294.28 to 1406.31. These variations highlight that lowering the input_voltage and altering time can considerably decrease the predicted el_power, demonstrating the model's responsiveness to these features in the vicinity of Record-1067.

Table 7
Counterfactual Analysis for Record-1067.

Analysis	time	input_voltage	el_power
Actual	1824.674805	5.0	1679.0
Counterfactual-1	10039.880000	3.1	1294.279785
Counterfactual-2	1824.674815	3.9	1406.306763
Counterfactual-3	5530.290000	3.4	1321.332031
Counterfactual-4	1404.400000	3.6	1327.647583
Counterfactual-5	7263.030000	3.4	1339.470703

Table 8 presents the counterfactual analysis for Record-5951, where the actual input (time = 6705.24, input_voltage = 10.0) yields a predicted electrical power of 2967.0. The five counterfactuals demonstrate that by reducing the input_voltage to between 3.0 and 3.1 and varying the time (e.g., to 1644.29 or 6533.43), the predicted el_power significantly decreases, with values ranging from 1181.64 to 1257.76. This analysis confirms that the model associates lower input_voltage and different time conditions with a substantial drop in power output, illustrating strong feature sensitivity in this region.

Table 6
Counterfactual analysis for Record-0.

Analysis	time	input_voltage	el_power
Actual	758.425842	3.0	1172.0
Counterfactual-1	4863.28	3.3	1289.596436
Counterfactual-2	8121.05	3.0	1249.439941
Counterfactual-3	9401.76	3.0	1262.846680
Counterfactual-4	7159.91	3.0	1239.378418
Counterfactual-5	8912.87	3.0	1257.728882

Table 8
Counterfactual analysis for Record-5951.

Analysis	time	input_voltage	el_power
Actual	6705.237305	10.0	2967.0
Counterfactual-1	1644.290000	3.0	1181.639648
Counterfactual-2	4136.950000	3.1	1232.486572
Counterfactual-3	6533.430000	3.1	1257.573486
Counterfactual-4	6705.237258	3.0	1234.618896
Counterfactual-5	6551.550000	3.1	1257.763184

Table 9
Counterfactual analysis for Record-14763.

Analysis	time	input_voltage	el_power
Actual	6399.339844	4.215537	1532.0
Counterfactual-1	6399.34	3.2	1280.922974
Counterfactual-2	7335.92	3.6	1389.740112
Counterfactual-3	5109.20	4.0	1465.443115
Counterfactual-4	1982.79	3.6	1333.702393
Counterfactual-5	4238.11	3.6	1357.311523

Table 9 presents the counterfactual analysis for Record-14763, where the actual input (time = 6399.34, input_voltage = 4.215537) results in a predicted electrical power of 1532.0. The five counterfactuals involve reducing the input_voltage to values between 3.2 and 4.0 and modifying the time to a range between 1982.79 and 7335.92. These adjustments lead to predicted el_power values between 1280.92 and 1465.44. The analysis indicates that slight reductions in input_voltage and changes in time can moderately decrease the predicted power, reflecting the model's localized sensitivity to both input features in the context of Record-14763.

5.4. Operationally-constrained counterfactuals

We employ a feasibility mask that screens generated candidates against input bounds and ramp-rate limits; infeasible points are discarded and the nearest feasible alternative within the desired output band is retained. These constraints produce counterfactuals that reflect plausible operator actions over realistic timelines. Consistent with our scope (see §4.3), the counterfactuals are intended for supervisory insight and scenario planning rather than direct fast closed-loop control during rapid ramps/steps.

5.5. Comparison with similar studies

Most of the existing models in Table 10, including deep learning techniques such as ANN, CNN, and LSTM, as well as ensemble and statistical approaches like ELM, MLAM, XGBoost, HDMR, and SAINT, focus primarily on improving the prediction accuracy of micro gas turbine power output using various datasets. These models are typically complex and lack features that support Explainable AI (XAI) or counterfactual analysis. As a result, they provide accurate results but offer little insight into how predictions are made or how inputs could be adjusted to achieve different outcomes.

In contrast, the proposed model uses a lightweight Lasso regression approach that includes both XAI and counterfactual analysis. This allows for better interpretability and actionable feedback. It addresses a major gap in the literature by enabling users to not only make predictions but also understand the reasoning behind them and explore how changes in input values can influence the predicted power output.

5.6. Mitigation strategies for fast transitions

To address the model's lag in rapid ramps/steps without abandoning transparency, we consider: (i) a regime-aware hybrid in which a light change-point/ramp detector routes steady segments to Lasso and transient segments to a small transient specialist (e.g., compact GRU/LSTM or gradient-boosted regressor), with SHAP/LIME applied to both; (ii) transition-sensitive feature design for linear baselines—lagged input_voltage terms, moving averages, first differences Δ(input_voltage), estimated ramp-rate $|\Delta V/\Delta t|$, and simple interaction terms; (iii) a ramp-aware loss that up-weights errors when $|\Delta V/\Delta t|$ exceeds a threshold, emphasizing fidelity during changes; and (iv) prediction filtering and residual gating (e.g., exponential/Kalman smoothing and control-chart thresholds) to flag suspect outputs during detected transitions. These options preserve operator-facing interpretability while improving responsiveness where the present baseline is weakest.

5.7. Relevance of electrical power prediction from micro gas turbine to power grid and microgrid modeling

Predicting electrical power (el_power) from micro gas turbines is exceptionally applicable and treasured in energy grid, microgrid, and dispensed era (DG) contexts. In those structures—in particular in microgrids—accurate forecasting of neighborhood era output is essential for maintaining supply-call for stability, ensuring grid stability, and optimizing operational efficiency. The dataset used in this study captures the time-series dating between enter manipulate voltage and the resulting electric electricity output of a micro fuel turbine. This correctly displays the conduct of real-global microturbines used as controllable, dispatchable devices inside modern grid infrastructures. By predicting el_power, energy management structures (EMS) and controllers could make knowledgeable selections for load dispatch, frequency regulation, and integration with other allotted sources like solar panels and battery garage. Moreover, in dispensed grid configurations, understanding how electricity output varies with manipulate inputs allows operators modify settings in actual time to meet demand, optimize fuel usage, and decrease emissions.

Such intelligent forecasting capabilities are especially important given the increasing cybersecurity and resilience challenges faced by modern smart grid infrastructures, as highlighted in recent AI and

Table 10
Comparison with Similar Studies.

Used Technique	Used Dataset	Model Complex	XAI	Counterfactual Analysis
ANN, CNN, LSTM [17]	Kaggle/UCI micro-gas-turbine	Yes	No	No
ELM, MLAM predictors [18]	Simulated micro-gas-turbine data	Yes	No	No
Lasso + temporal models [19]	Kaggle/UCI dataset	No	No	No
Ensemble & statistical models [20]	Experimental turbine dataset	Yes	No	No
HDMR + ANN [21]	Large-scale GT data	Yes	No	No
XGBoost, Transformer (SAINT) [22]	Siemens testbed turbine	Yes	No	No
ML anomaly detection [28]	Fleet turbine operational data	Yes	No	No
AI hybrid model monitoring [29]	Micro GT fleet data	Yes	No	No
Robust optimization [30]	Simulated data	Yes	No	No
Regression models [31]	Experimental turbojet data	Yes	No	No
ANN vs. LSTM vs. CNN [14]	Micro-turbine / energy time series	Yes	No	No
LSTM (attention) [15]	Industrial / energy time series	Yes	No	No
Short-time load forecasting LSTM [16]	Power-system load	Yes	No	No
XGBoost (maintenance / forecasting) [23]	Wind-turbine SCADA / energy informatics	Yes	Yes (SHAP)	No
ML vs. DL for wind-turbine power [24]	Wind-turbine environmental factors	Yes	Yes	No
XGBoost + feature engineering [25]	Short-term wind-power forecasting	Yes	No	No
XGBoost + SHAP (CCPP) [26]	Combined-cycle power plant	Yes	Yes (SHAP)	No
EBM / iXGB (rules + counterfactuals) [27]	Interpretable tabular forecasting	Yes (interpretable)	Yes	Yes
Proposed Model	Lasso Regression with SHAP/LIME + Counterfactuals	No	Yes	Yes

blockchain-based security frameworks [34].

A important enhancement delivered in this work is the software of counterfactual evaluation. This approach permits grid operators and selection-makers to discover hypothetical eventualities through enhancing input variables—which includes input voltage or operation time—and gazing the corresponding impact at the expected electric power output. This is in particular treasured in real-time grid packages, wherein brief changes are regularly necessary to cope with renewable intermittency, surprising load changes, or call for reaction strategies. Counterfactual motives not simplest guide operational flexibility but additionally provide obvious, actionable guidance for attaining preferred electricity stages within predefined limits. Therefore, the prediction of el_power from micro fuel mills—augmented with counterfactual evaluation—gives a powerful device for advancing explainable, adaptive, and intelligent electricity systems.

Recent research has also demonstrated that advanced hybrid approaches, such as Quantum-Inspired Particle Swarm Optimization (QPSO) combined with RNNs, can significantly enhance forecasting accuracy in dynamic settings like educational buildings [35]. While these complex models are well-suited for large-scale demand forecasting, they often require considerable computational resources and lack interpretability. In contrast, the lightweight and explainable nature of the proposed Lasso-based framework offers a practical alternative for embedded and distributed environments, particularly within microgrid CPS infrastructures. This balance between efficiency, interpretability, and responsiveness positions our model as a complementary tool in the broader landscape of smart grid optimization.

5.8. Interpretable forecasting for cyber-resilient distributed energy systems

In decentralized and digitally integrated energy systems, interpretability supports operational trust and manual validation of AI-driven forecasts. While this work does not experimentally evaluate cybersecurity or anomaly-detection performance, the same transparency that aids operator understanding can, in future studies, serve as a foundation for detecting data inconsistencies or sensor faults. In contrast to black-box models, explainable frameworks like Lasso regression augmented with SHAP and LIME allow stakeholders to distinguish between legitimate deviations and malicious or anomalous behavior. This interpretability is crucial in cyber-physical settings where data manipulation, spoofing, or misconfigured sensors could compromise system integrity. By providing both global and localized feature attributions, the proposed model supports resilience-oriented applications such as fault detection, secure control loop validation, and trust calibration for autonomous energy agents operating in distributed smart grid environments. Future work will formalize this concept by coupling SHAP/LIME outputs with threshold-based anomaly indicators and controlled spoofing tests to quantify detection sensitivity and false-alarm rates within microgrid CPS contexts.

6. Conclusion

6.1. Recapitulation

This study presents a lightweight, interpretable, and decision-oriented framework for predicting the electrical power output of a micro gas turbine using a newly released Kaggle dataset titled Micro Gas Turbine Electrical Energy Prediction. The dataset, which captures high-resolution temporal relationships between input voltage and power output, is particularly well-suited for real-time modeling in distributed power systems, microgrids, and broader smart grid infrastructures.

The proposed model leverages Lasso regression due to its simplicity, regularization capability, and computational efficiency—qualities that make it appropriate for deployment in resource-constrained or embedded environments. Using only two input features (time and input voltage), the model achieved dataset-level predictive performance, with 87 % of test instances falling within the low error category (<20 %). Evaluation metrics such as MAE, MSE, and R^2 validate both the model's accuracy and its generalization capacity. However, the residual and heatmap analyses indicate reduced fidelity during rapid ramps/steps; therefore, performance should be interpreted as strongest in steady or quasi-steady regimes.

A distinguishing contribution of this work is its integration of Explainable AI (XAI) techniques—SHAP and LIME—for feature attribution and model transparency. These tools help identify which inputs most influence the predictions, enabling stakeholders to interpret and trust the model's output. Moreover, counterfactual analysis is employed to simulate "what-if" scenarios by adjusting inputs within defined limits, thereby supporting actionable insights and scenario-based operational planning. In line with the above, we position the framework as a transparent, lightweight baseline intended for supervisory monitoring and planning rather than fast closed-loop control during highly dynamic transitions.

In the context of cyber-physical power systems, such transparency is critical for operational decision-making, security assurance, and resilience enhancement. Explainability empowers grid operators to diagnose prediction outcomes, validate control strategies, and detect anomalies with greater confidence. Meanwhile, counterfactual reasoning offers a tangible method for stress-testing system responses to input variations—facilitating adaptive control, failure prevention, and proactive tuning of decentralized units such as micro gas turbines. While these CPS applications are conceptually motivated here, they are not experimentally evaluated in this study. Together, these capabilities contribute to the development of intelligent, robust, and operator-friendly CPS infrastructures in modern power grids.

Collectively, the proposed framework advances the state of the art by coupling predictive modeling with interpretability and user-oriented counterfactual reasoning, while explicitly acknowledging its limited reliability under rapid transients and offering a deployable baseline for power prediction in smart energy systems.

6.2. Limitations and future work

While the current framework demonstrates useful predictive accuracy and transparency, several limitations warrant future exploration. First, the model relies exclusively on two input features, which simplifies interpretation but may constrain the capture of complex or nonlinear dependencies in more dynamic environments. Incorporating additional contextual features—such as ambient temperature, turbine load conditions, or fuel properties—could further enhance performance and realism.

Second, although SHAP and LIME provide valuable explanations, their computational cost can increase significantly for large-scale deployments or real-time operations. Future work could explore lightweight explanation methods or approximation techniques to maintain scalability. Although we reference cyber-physical resilience conceptually, the current study does not perform security validation; subsequent work will empirically assess how explainability aids anomaly detection and secure control validation in real CPS environments.

Third, the counterfactual analysis is currently limited to bounded linear adjustments of the input space. We now constrain counterfactuals by dataset-observed bounds and empirical ramp-rate (slew-rate) limits to improve operational plausibility; expanding this to include domain-specific constraints or multi-objective optimization could provide more realistic and actionable planning recommendations.

Fourth, fidelity is reduced during rapid transients due to the linear form and minimal feature set; future extensions will incorporate lagged variables, moving averages, interaction terms, and nonlinear or hybrid components to better capture dynamic responses while preserving explainability.

A fifth limitation of the present evaluation is reliance on a single file-level split; although operationally motivated, it can bias generalization estimates, which is why we now specify complementary group-aware and time-blocked cross-validation procedures for robustness.

Finally, the proposed model should be validated in field scenarios or integrated into operational microgrid control systems to assess its real-world applicability and reliability under diverse operating conditions. This would support further refinement of the framework and promote adoption in industrial applications.

Next steps (actionable roadmap):

- Data & external validation: Curate at least two additional micro-turbine datasets from distinct sites/operating policies plus one public proxy dataset; evaluate external generalization under the same protocol.
- Feature engineering (linear baseline): Introduce lagged input_voltage ($k = 1\ldots8$), moving averages, first differences ΔV, ramp-rate $|\Delta V/\Delta t|$, and simple interaction terms; quantify incremental effects via ablations (ΔMAPE/ΔRMSE).
- Benchmarking & CV: Run group-aware K-fold (files as groups) and blocked time-series CV; benchmark OLS, DecisionTree, RandomForest, compact GRU/LSTM, and XGBoost against Lasso with identical metrics, plus model size and inference latency.
- Regime-aware hybrid: Prototype a lightweight change-point/ramp detector that routes steady segments to Lasso and transients to a compact specialist; retain SHAP/LIME for both branches.
- Online adaptation: Evaluate sliding-window/forgetting-factor Lasso for gradual drift; report stability–responsiveness trade-offs.
- Uncertainty & calibration: Add conformal prediction intervals; report coverage and width, stratified by ramp intensity.
- Operational counterfactuals: Extend feasibility with plant datasheet limits (e.g., voltage slew, thermal constraints) and multi-objective trade-offs (power vs. actuator effort).
- CPS & anomaly studies: Design controlled spoofing/fault-injection tests; pair SHAP/LIME thresholds with residual alarms and report detection and false-alarm rates.
- Deployment & monitoring: Package an edge-oriented reference implementation meeting memory/latency targets, plus a simple dashboard for attributions, counterfactuals, drift monitoring, and alerting.

Author statement

We have revised our manuscript according to the reviewers comments

CRediT authorship contribution statement

Sheikh Muhammad Saqib: Writing – review & editing, Visualization, Validation, Methodology. **Amir Khan Muhammad:** Visualization, Validation, Resources. **Tariq Shahzad:** Visualization, Validation, Resources. **Muhammad Usman Tariq:** Investigation, Formal analysis, Data curation. **Tehseen Mazhar:** Writing – review & editing, Writing – original draft, Visualization. **Habib Hamam:** Visualization, Supervision, Project administration.

Funding

None.

Declaration of Competing Interest

The authors have no conflict of interest

Data availability

Data will be made available on request.

References

[1] W.E. Bill Forsthoffer, Gas turbine performance Forsthoffer's Rotating Equip. Handbooks, 2005, pp. 395–406, 10.1016/b978-185617472-5/50033-2.
[2] J. Kotowicz, M. Job, M. Brzeczek, The characteristics of ultramodern combined cycle power plants, Energy 92 (2015) 197–211, https://doi.org/10.1016/j.energy.2015.04.006.
[3] R. Wittenburg, M. Hübel, H. Prause, C. Gierow, M. Reißig, E. Hassel, Effects of rising dynamic requirements on the lifetime consumption of a combined cycle gas turbine power plant, Energy Procedia 158 (2019) 5717–5723, https://doi.org/10.1016/j.egypro.2019.01.562.
[4] Y. Hu, et al., Gas turbine-based system taking advantage of LNG regasification process for multigeneration purposes; techno-economic-environmental analysis and machine learning optimization, Process Saf. Environ. Prot. 179 (2023) 10–26, https://doi.org/10.1016/j.psep.2023.08.068.
[5] C.A. Saleel, Forecasting the energy output from a combined cycle thermal power plant using deep learning models, Case Stud. Therm. Eng. 28 (2021), https://doi.org/10.1016/j.csite.2021.101693.
[6] A.A. Galarza-Chavez, J.L. Martinez-Rodriguez, R.F. Domínguez-Cruz, E. López-Garza, A.B. Rios-Alvarado, Multi-step wind energy forecasting in the Mexican Isthmus using machine and deep learning, Energy Rep. 13 (2025) 1–15, https://doi.org/10.1016/j.egyr.2024.11.074.
[7] E. Losi, L. Manservigi, P.R. Spina, M. Venturini, Data-driven generative model aimed to create synthetic data for the long-term forecast of gas turbine operation, J. Eng. Gas. Turbines Power 147 (3) (2025), https://doi.org/10.1115/1.4066360.
[8] K. Manatura, et al., Gas turbine heat rate prediction in combined cycle power plant using artificial neural network, Therm. Sci. Eng. Prog. 59 (2025), https://doi.org/10.1016/j.tsep.2025.103301.
[9] A. A, S.A. Alharbi, A. Chinnathambi, R.R. P. J, B. Gavurová, Experimental investigation and machine learning prediction of thrust, fuel consumption, and emissions for micro gas turbine engine fueled with biofuel and hydrogen- A comparative study of linear regression and LSTM model, J. Taiwan Inst. Chem. Eng. (2025), https://doi.org/10.1016/j.jtice.2025.106104.
[10] R. De Robbio, Micro gas turbine role in distributed generation with renewable energy sources, Energies 16 (2) (2023), https://doi.org/10.3390/en16020704.
[11] R. Chandrasekaran, S.K. Paramasivan, Advances in deep learning techniques for short-term energy load forecasting applications: a review, Arch. Comput. Methods Eng. 32 (2) (2025) 663–692, https://doi.org/10.1007/s11831-024-10155-x.
[12] K. Sarkar, Load and Renewable Energy Forecasting Using Deep Learning for Grid Stability, *arXiv Prepr.*, 2025, doi: https://doi.org/10.48550/arXiv.2501.13412.
[13] A. Ahmed, et al., Novel deep neural network architecture fusion to simultaneously predict short-term and long-term energy consumption, PLoS One 20 (1) (2025), https://doi.org/10.1371/journal.pone.0315668.
[14] F. Liu, et al., An ensemble model based on adaptive noise reducer and over-fitting prevention LSTM for multivariate time series forecasting, IEEE Access 7 (2019) 26102–26115, https://doi.org/10.1109/access.2019.2900371.
[15] H. Abbasimehr, Reza Paki, Improving time series forecasting using LSTM and attention models, J. Ambient Intell. Humaniz. Comput. 13 (2021) 673–691, https://doi.org/10.1007/s12652-020-02761-x.
[16] T. Ciechulski, S. Osowski, High precision LSTM model for short-time load forecasting in power systems, Energies 14 (2021) 2983, https://doi.org/10.3390/en14112983.
[17] E. Bakiş, E. Acar, Deep learning-based time series prediction of micro gas turbine power output. IEEE Glob. Energy Conf. 2024, GEC 2024, 2024, pp. 270–275, https://doi.org/10.1109/GEC61857.2024.10880993.
[18] A.V. Thomas Jayachandran, H. Omar, A. Tkachenko, A. Krishnakumar, Machine learning predictor for micro gas turbine performance evaluation, Aeronaut. Aerosp. Open Access J. 4 (4) (2020) 172–180, https://doi.org/10.15406/aaoaj.2020.04.00120.
[19] P. Bielski, A. Eismont, J. Bach, F. Leiser, D. Kottonau, K. Böhm, Knowledge-guided learning of temporal dynamics and its application to gas turbines. Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems, 2024, pp. 279–290, 10.1145/3632775.3661967.
[20] Y. Liu, et al., Machine learning based predictive modelling of micro gas turbine engine fuelled with microalgae blends on using LSTM networks: an experimental approach, Fuel 322 (2022), https://doi.org/10.1016/j.fuel.2022.124183.
[21] Z. Liu, I.A. Karimi, Gas turbine performance prediction via machine learning, Energy 192 (2020), https://doi.org/10.1016/j.energy.2019.116627.
[22] R. Potts, R. Hackney, G. Leontidis, Tabular machine learning methods for predicting gas turbine emissions, Mach. Learn. Knowl. Extr. 5 (3) (2023) 1055–1075, https://doi.org/10.3390/make5030055.
[23] H. Dhungana, A machine learning approach for wind turbine power forecasting for maintenance planning, Energy Inf. 8 (2025) 2, https://doi.org/10.1186/s42162-024-00459-4.
[24] M. Abdelsattar, et al., Evaluating machine learning and deep learning models for predicting wind turbine power output from environmental factors, PLOS ONE 20 (2025), https://doi.org/10.1371/journal.pone.0317619.
[25] H. Zheng, Y. Wu, A XGBoost model with weather similarity analysis and feature engineering for short-term wind power forecasting, Appl. Sci. (2019), https://doi.org/10.3390/app9153019.
[26] Y. Shang, Xgboost-based prediction of net electrical energy output of combined cycle power plants and model interpretability study, Highlights Sci. Eng. Technol. (2025), https://doi.org/10.54097/khhwdr12.
[27] M.R. Islam, M.U. Ahmed, S. Begum. iXGB: Improving the Interpretability of XGBoost Using Decision Rules and Counterfactuals, 2024, pp. 1345–1353, 10.5220/0012474000003636.

[28] 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. 7 (2021), https://doi.org/10.3389/fmech.2021.676853.
[29] R. Banihabib, M. Assadi, The role of micro gas turbines in energy transition, Energies 15 (21) (2022), https://doi.org/10.3390/en15218084.
[30] J.F. Rist, M.F. Dias, M. Palman, D. Zelazo, B. Cukurel, Economic dispatch of a single micro-gas turbine under CHP operation, Appl. Energy 200 (2017) 1–18, https://doi.org/10.1016/j.apenergy.2017.05.064.
[31] H. Aygun, O.O. Dursun, K. Dönmez, O. Sahin, S. Toraman, Prediction of performance characteristics of an experimental micro turbojet engine using machine learning approaches, Energy 313 (2024), https://doi.org/10.1016/j.energy.2024.133997.
[32] Dataset, 2025, ⟨https://www.kaggle.com/datasets/shivamarora384/micro-gas-turbine-electrical-energy-prediction⟩.
[33] M.A. Khan, et al., Decoding 'Eligibility Unknown': transparent classification and feature-based reclassification in CAFV analysis, Intern. J. Electr. Power Energy Syst. 170 (2025) 110894, https://doi.org/10.1016/j.ijepes.2025.110894.
[34] Y.Y. Ghadi, et al., A hybrid AI-Blockchain security framework for smart grids, Sci. Rep. 15 (2025) 20882, https://doi.org/10.1038/s41598-025-05257-w.
[35] S. Khan, et al., Optimizing load demand forecasting in educational buildings using quantum-inspired particle swarm optimization (QPSO) with recurrent neural networks (RNNs): a seasonal approach, Sci. Rep. 15 (2025) 19349, https://doi.org/10.1038/s41598-025-04301-.