CAAI Transactions on Intelligence Technology

WILEY

ORIGINAL RESEARCH OPEN ACCESS

RES-AD: An AutoML-Driven Robust Ensemble Framework for Real-Time Anomaly Detection and Energy Forecasting in Smart Grids

Murad Ali Khan[1] | Naeem Iqbal[2,3] | Syed Shehryar Ali Naqvi[4] | Muhammad Faseeh[4] | Do-Hyeun Kim[1,5]

[1]Department of Computer Engineering, Jeju National University, Jeju, Republic of Korea | [2]Centre for Secure Information Technologies, Queen's University Belfast, Belfast, UK | [3]Momentum 1.0 Research, Queen's University Belfast, Belfast, UK | [4]Department of Electronics Engineering, Jeju National University, Jeju, Republic of Korea | [5]Advanced Technology Research Institute, Jeju National University, Jeju, Republic of Korea

Correspondence: Do-Hyeun Kim (kimdh@jejunu.ac.kr)

Received: 15 September 2024 | **Revised:** 30 November 2025 | **Accepted:** 21 May 2026

Keywords: anomaly detection | AutoML | cyber-attacks | energy forecasting | ensemble learning | smart grid | smart metres | sustainable energy management

ABSTRACT

Accurate load forecasting and reliable anomaly detection are critical for the stable operation of modern smart grids (SGs), which increasingly rely on cyber-connected infrastructures. However, the integration of smart metres and two-way communication exposes SGs to data integrity attacks that can manipulate consumption measurements, degrade forecast accuracy and compromise grid stability. This study proposes RES-AD, a robust AutoML-driven ensemble framework designed for real-time detection of abnormal consumption patterns and accurate short-term energy forecasting. RES-AD consists of an Optimal Ensemble Classifier (OEC) for anomaly detection and an Optimal Ensemble Regressor (OER) for forecasting, both selected and optimised automatically using H2O AutoML. Five types of synthetic data integrity attacks are injected to simulate realistic adversarial conditions, whereas SMOTE and correlation-based feature selection ensure balanced and informative training data. The revised framework further incorporates stealthy adversarial attack simulation evaluated using FGSM and demonstrates strong robustness with only a 1.7% degradation under perturbation, further mitigated through adversarial augmentation. Additionally, real-time readiness is validated through latency and memory profiling, where the OEC ensemble requires only 23.36 ms per inference with a footprint below 0.05 MB, making it suitable for edge or SCADA-integrated deployment. Experimental results using public residential smart metre data show that OEC achieves 99.7% (hourly) and 99.9% (daily) detection accuracy, whereas OER improves energy forecasting with a MAPE reduction of 11.11% and 11.72%, respectively. These outcomes demonstrate that RES-AD provides a scalable, attack-aware and deployment-ready solution that enhances both the reliability and resilience of modern smart grid operations.

1 | Introduction

The SG is innovative because it uses advanced technologies, such as computers, control technologies and automation, to meet electricity demands and improve grid efficiency. It integrates energy generation, substation integration, transmission, distribution and consumption using bidirectional, cyber-secure communication technologies and computational intelligence [1]. ICT enhances control and communication services, turning traditional power grids into intelligent ones.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *CAAI Transactions on Intelligence Technology* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology and Chongqing University of Technology.

Although SG improves energy reliability and efficiency, they are also more vulnerable to cyber-attacks due to the large number of smart devices and access points outside conventional administrative zones [2]. A cyber-attack could cause appliance damage or disrupt energy demand by generating false data [3–5]. It is crucial to detect cyber-attacks early to prevent significant damage and enhance resilience with continuous monitoring.

Anomalies, or abnormal data, are typical indicators of potential cyber-attacks. Anomaly detection is widely used in intrusion detection and in preventing industrial damage. In SG, anomalies or outliers can affect model performance, requiring distinct techniques for detection [6]. The transition from traditional grids to SG, although beneficial, introduces risks from cyber-attacks. Hackers can manipulate energy consumption data from smart metres, resulting in inaccurate predictions of future electricity demand and potential supply disruptions. This issue is discussed in the literature, highlighting SG's growing vulnerabilities. For instance, a study by [7] demonstrates how SG's increasing interconnectivity and complexity introduce vulnerabilities to cyber-attacks targeting data integrity. Syrmakesis et al. [8] discuss resilient frequency control methods in interconnected power systems to counter cyber-attacks, especially those manipulating control signals. Similarly, Yang et al. [9] focus on the risks posed by deception cyber-attacks in distributed control systems, a growing concern as more SG technologies are deployed. Furthermore, [10] reviews cyber-attack's dynamic impacts on state estimation in SG, underscoring the need for robust security measures to ensure grid stability.

In this paper, we used a publicly available unsupervised energy load dataset for the experiment. The first step is preprocessing, in which we format the data to be readable for classification and regression algorithms. Then, to handle outliers in the data, we used DBSCAN to detect outliers across the entire dataset and refilled them using mean imputation to maintain the temporal pattern. The purpose of eliminating outliers from the dataset in the initial stage is to prevent classification algorithms from being affected by them: masking outliers with attack records and sometimes showing outliers as attacks, and sometimes showing attacks as outliers. After detecting and removing outliers from the dataset using DBSCAN, the next step is to add five types of attacks to the dataset. These simulated attacks (Scaling, Ramp, Random, Pulse and Scaling FIX) are considered data-integrity attacks, a subset of cyberattacks that aim to compromise the accuracy of SG data. These attacks, often associated with FDIA [11, 12], DoS attacks [13, 14] and other cyber-attack techniques [15–17], manipulate energy consumption measurements to disrupt forecasting models and grid stability. Using these attacks, we labelled our dataset as 1 for attack data and 0 for normal data.

The dataset used for the experiments consists of 50,000 records, of which 47,909 are labelled as regular data and 2091 are labelled as attack data. As we know, classification algorithms perform better on balanced datasets, but in real life, anomalies are significantly less common than expected. Similarly, in our case, the data is imbalanced, so we need to use techniques to address the imbalance. Therefore, we used the SMOTE technique to increase the minority class data (attacks) by generating synthetic samples. Finally, we developed a RES-AD framework based on AutoML and Soft Voting, and despite our limited sample size, accuracy, recall, precision and $F1$ score showed encouraging results. Moreover, the main contributions of this study are summarised as follows:

1. A novel RES-AD framework is proposed, integrating anomaly detection and energy forecasting within a unified ensemble learning pipeline specifically designed for smart grid environments.
2. H2O-based AutoML is employed to automate model selection, hyperparameter optimisation and performance ranking, thereby reducing manual tuning effort and enhancing model generalisation.
3. Five realistic data integrity attacks are simulated and SMOTE-based class balancing is applied to convert the originally unsupervised dataset into a robust supervised learning setting.
4. A comprehensive preprocessing pipeline is developed, incorporating DBSCAN-based outlier identification, temporal feature engineering and correlation-driven feature selection to improve data quality and model performance.
5. The proposed OEC and OER are rigorously validated on hourly and daily consumption datasets, achieving state-of-the-art anomaly detection performance ($F1$-score up to 99.9%) and significant forecasting improvements (MAPE reduction exceeding 11%).
6. Scalability, deployment feasibility and integration pathways of RES-AD are analysed, highlighting its suitability for real-time operation in large-scale smart grid infrastructures.

The rest of the paper is organised as follows: Section 2 reviews related work. Section 3 presents the proposed RES-AD framework. Section 4 describes data preparation and attack simulation. Section 5 explains AutoML-based model selection, voting ensemble and implementation environment. Section 6 defines evaluation metrics. Section 7 presents experimental results and analysis. Section 8 discusses key insights and the scalability of the system. Finally, Section 9 concludes the study, and Section 9.1 outlines future directions.

2 | Related Work

Anomalies, also known as 'abnormal data', are data points that deviate from the norm. Detecting these anomalies is crucial, as they can negatively impact model performance. Anomaly detection is commonly applied in intrusion detection, industrial damage prevention, and medical diagnosis. Anomalies can be point-wise, contextual, or collective. Point-wise anomalies refer to individual data points that deviate from the rest, whereas contextual anomalies occur when data for a specific day differs from that of other days. Collective anomalies are groups of data points that deviate together [18]. This study proposes an anomaly prediction model based on an encoder–decoder architecture with an RNN. The model predicts anomalies by identifying high reconstruction errors when reconstructing time series data. The difference between the input and reconstructed time series is measured using the Manhattan distance and the edit distance. Using IEEE 39 bus simulation data and power demand data from the University of California, Riverside, the model achieved 95% accuracy in predicting anomalies based on the encoder–decoder framework [19].

This research introduces a novel OE-IDS system [20] that leverages AutoML and soft voting to efficiently detect intrusions in network environments, addressing challenges such as limited accuracy, imbalanced data and large-scale network profiling. OE-IDS achieves higher detection rates and outperforms traditional ML-based systems. Additionally, the study focuses on an anomaly detection model capable of handling concept drift, such as changes in household behaviour [21]. Utilising an LSTM network, it predicts energy consumption patterns and distinguishes between anomalies and regular behavioural changes. The framework, based on an RNN, detects anomalies in real-time. Experiments show improved trend recognition and anomaly detection, even with shifting energy consumption patterns. In ref. [22], the author proposed a model with two methods: a data-driven approach for anomaly detection and a moving split-window mechanism for real-time analysis. A statistical MSR was also introduced to measure data correlations and filter parameter exchanges between utilities. In another study by ref. [23], a distributed multiagent maximum likelihood technique was used for anomaly detection in SG, aiming to reduce computational time while preserving data privacy. The model showed high accuracy in detecting anomalies in power transmission systems. Similarly ref. [24], developed a real-time framework using data from smart metres to predict abnormal events at both customer and lateral levels.

Anomaly detection in SG is a challenging task. In ref. [25], an anomaly prediction technique based on a LOF is proposed, using Lagrangian interpolation for missing values and testing on four data groups. The model performed well in predicting abnormal events. In ref. [19], collective anomaly detection was achieved using frequent item-set mining and clustering techniques to predict and visually represent 10 different types of anomalies in smart metres. Additionally [26], proposed a k-Means algorithm for predicting attacks like flooding, which can cause DoS. The model outperformed existing clustering algorithms in identifying outliers in traffic data between SG and homes.

In ref. [27], a novel technique for SG security is proposed, focussing on protecting devices like circuit breakers and relays. The model's reliability was tested using k-fold cross-validation, with practical rules making the system controller's tasks more efficient. In ref. [28], an anomaly prediction algorithm was used to safeguard customer data privacy in intelligent networks. The OPTICS density-based technique was implemented, combined with the LOF index, to successfully identify malicious data. Similarly, in ref. [29], various ML techniques were applied to detect anomalies in power stations, distinguishing between natural disturbances, anomalies and cyberattacks. The study evaluated the performance of these algorithms using the IEEE 3-bus system dataset, demonstrating their efficiency in predicting cyber-attacks and assessing accuracy and processing power.

Anomaly detection in SG is challenging due to an imbalanced class distribution, with fewer attacks than regular traffic. In ref. [30], the author examines resampling methods for anomaly detection using a hierarchical three-layer communication system and the ADFA-LD dataset, which includes advanced attacks. The study shows that applying resampling techniques can improve intrusion detection model performance. Similarly, in ref. [31], an anomaly-detection framework for high-dimensional SG is developed to distinguish between cyberattacks and natural disturbances. Experiments using the IEEE 2848, 39 and 118 bus systems demonstrate high efficiency, with 99% accuracy, 98% TPR and FPR below 2%. In ref. [32], the study focuses on SG within the Industry 4.0 domain and evaluates the efficiency and effectiveness of ML and DL models in SG applications. The paper also highlights emerging challenges, including scalability, cybersecurity and large-scale data management. In ref. [33], an anomaly detection technique, MLAD, is proposed for predicting load under cyber-attacks. The method uses projected load data with k-means clustering to reconstruct reference data, followed by naive Bayes classification to identify attack patterns. Dynamic programming is then applied to assess the frequency and characteristics of the cyberattacks.

In ref. [34], the author formulates the online intrusion prediction problem as a POMDP and proposes a robust live-detection technique using RL for POMDPs. This RL-based algorithm effectively detects cyber-attacks in SG. In ref. [35], a model is proposed to enhance end-layer security in smart energy systems using advanced ML and security mechanisms. In ref. [36], the authors explore online threats in SG and provide literature-based solutions for anomaly detection, aiming to identify security trigger points for future researchers. A comparison of existing solutions by [37] discusses cybersecurity challenges in SG networks and proposes a revised IoT-based security model. In another study [38], high-frequency signals and power line modems are used as network sensors to detect anomalies in distribution grids. The study presents algorithms and simulation results that confirm the effectiveness of this approach.

Recent studies have made significant contributions to the domains of reliability assessment and cyberattack detection in distributed and DC microgrid systems. In ref. [39], the authors present an analytical methodology for reliability assessment and failure analysis in distributed power systems. The approach utilises outage data and system design metrics to evaluate system reliability and minimise failure rates, providing a structured framework for reliability modelling. In ref. [40], a novel cyberattack detection method is proposed for smart DC-microgrids that combines the Hilbert–Huang Transform for time-frequency signal analysis with blockchain technology to enhance data integrity and detect FDIAs. Furthermore, the work in ref. [41] introduces a deep ML approach that integrates wavelet singular value decomposition to detect cyber intrusions in DC microgrids. The model effectively distinguishes between normal and attack scenarios by analysing the transformed feature space, thereby improving detection accuracy and system resilience. These methodologies provide valuable insights for enhancing the reliability and security of modern power grids, offering complementary perspectives to our proposed RES-AD framework. To better understand the current landscape of anomaly detection and load forecasting in SG, we present a comparative overview of recent studies in Table 1.

2.1 | Identified Gaps in Prior Work and Motivation for RES-AD

Although existing studies have explored anomaly detection and forecasting in smart grids, several technical gaps remain

TABLE 1 | Overview of existing work using the electricity load diagrams 2011–2014 dataset or similar SG datasets.

Study	Year	Task	Methodology	Dataset used	Key limitation
Fengming et al. [18]	2017	Anomaly detection	RNN encoder-decoder	IEEE 39-bus simulation data	No attack simulation, limited generalisation
Cui et al. [33]	2019	Forecasting under attacks	MLAD + Naive Bayes + K-means	Electricity load diagrams 2011–2014	No ensemble learning, lacks detection generality
Karimipour et al. [31]	2019	Anomaly detection	Feature selection + SVM	Electricity load diagrams 2011–2014	Static thresholding, lacks attack adaptation
Altamimi et al. [42]	2025	Anomaly detection	Residual analysis + forecast baseline	Electricity load diagrams 2011–2014	No labelled attacks, unsupervised only
Zheng et al. [43]	2025	Anomaly detection	SVM + RF fusion	Electricity load diagrams 2011–2014	Limited to detection, not forecasting
Shabad et al. [44]	2023	Anomaly detection	Isolation forest + PCA	Electricity load diagrams 2011–2014	No class balancing or temporal features
Xu et al. [11]	2017	Intrusion detection	Sparse matrix + topology inference	IEEE 14-bus simulation data	Not validated on real-world consumption data

unresolved. First, most prior work relies on manually selected models and hyperparameters, with limited use of AutoML-based optimisation, resulting in suboptimal performance. Second, existing anomaly detection solutions rarely incorporate temporal dependency encoding, such as hour-of-day, weekday–weekend, or seasonal patterns, despite their strong influence on load behaviour. Third, none of the reviewed studies integrate robust preprocessing strategies, such as DBSCAN-driven outlier detection combined with SMOTE-based class balancing, which are essential for handling noisy, imbalanced SG datasets. Fourth, the literature lacks attack-aware data transformation pipelines, and prior works do not simulate realistic FDIA-style consumption attacks to generate labelled datasets for supervised learning. Fifth, adversarial robustness is unexplored mainly; existing frameworks do not evaluate model stability under adversarial perturbations or noise-induced degradation. Finally, most prior research treats anomaly detection and load forecasting as isolated tasks, rather than combining them into a unified, automated and attack-resilient pipeline.

These gaps highlight the need for a comprehensive framework capable of automated model optimisation, robust preprocessing, temporal pattern learning and adversarial resilience, all of which form the foundation for the proposed RES-AD system.

2.2 | Novelty of the Proposed RES-AD Framework

Although ensemble methods and AutoML have individually been explored in prior work, the novelty of the proposed RES-AD framework lies in its integrated, attack-aware and temporally informed architecture, explicitly designed for smart grid environments. Unlike existing studies, RES-AD combines (i) AutoML-optimised model selection, (ii) a hybrid ensemble architecture using Soft Voting for classification and Stacking for regression, (iii) a temporal-feature–enhanced preprocessing pipeline and (iv) a DBSCAN + SMOTE–driven noise and imbalance correction strategy tailored to SG consumption data. Furthermore, RES-AD introduces a realistic attack simulation module that generates labelled FDIA-style attacks, enabling supervised learning on previously unsupervised datasets. The framework also integrates adversarial robustness evaluation, demonstrating resilience under FGSM-style perturbations. Collectively, these components address significant limitations in prior work, including a lack of automation, limited temporal modelling, the absence of adversarial evaluation and fragmented anomaly/forecasting approaches, establishing RES-AD as a unified and robust solution for modern SG anomaly detection.

3 | Proposed Framework

This section outlines the methodology for the proposed RES-AD model to accurately predict various energy theft attacks and detect anomalies in SG data. The task of SG intrusion detection has posed significant challenges for researchers due to the intricate structure of SG and the utilisation of advanced technologies.

3.1 | Overview of the Proposed System

The suggested model uses hourly/daily SG data and employs preprocessing techniques to transform it into a format suitable for the classifier. The Block model of the proposed work is given in Figure 1 and consists of preprocessing, adding attacks, labelling the data, feature selection by statistical techniques, AutoML, Soft Voting and performance analysis. In the first step, null records are removed from the data. Since some features in the dataset are alphanumeric, they are encoded as numeric values using a label encoder, and missing values are imputed using the mean.

In the given Figure 1, after labelling data by using attacks, feature selection is done with the target column, and with the help of AutoML, optimal models are selected. The experiments were conducted using an ensemble model with Soft Voting. The model's performance was analysed using Accuracy, Precision, Recall and $F1$ score metrics.

3.2 | Suggested Flow Model

This section presents a flow diagram of the proposed ensemble anomaly detection mechanism based on automated learning and soft voting in an SG. Figure 2 illustrates the flow structure of the proposed research study.

In Figure 2, the process involves preprocessing, feature engineering, feature selection, adding attacks and labelling the data. Two parallel experiments are conducted: one using AutoML and a stacking ensemble on a dataset without attacks, and another on a dataset with attacks but no target label. Model performance is evaluated using metrics such as R^2 score, MAE, MSE, RMSE and MAPE. After labelling the data with attacks, AutoML is used to select the optimal model. In the final experiment, the Soft Voting ensemble is evaluated using metrics such as accuracy, precision, recall and $F1$ score.

3.3 | Algorithm for Proposed Ensemble Model

The proposed RES-AD Algorithm 1 follows a structured five-stage workflow designed to ensure robust anomaly detection in smart grid environments. First, all missing values in the raw energy dataset are imputed using attribute-wise mean replacement to maintain data consistency. Second, the preprocessed data is augmented with five realistic data integrity attacks—Scaling, Ramp, Random, Pulse and Fixed Scaling—to

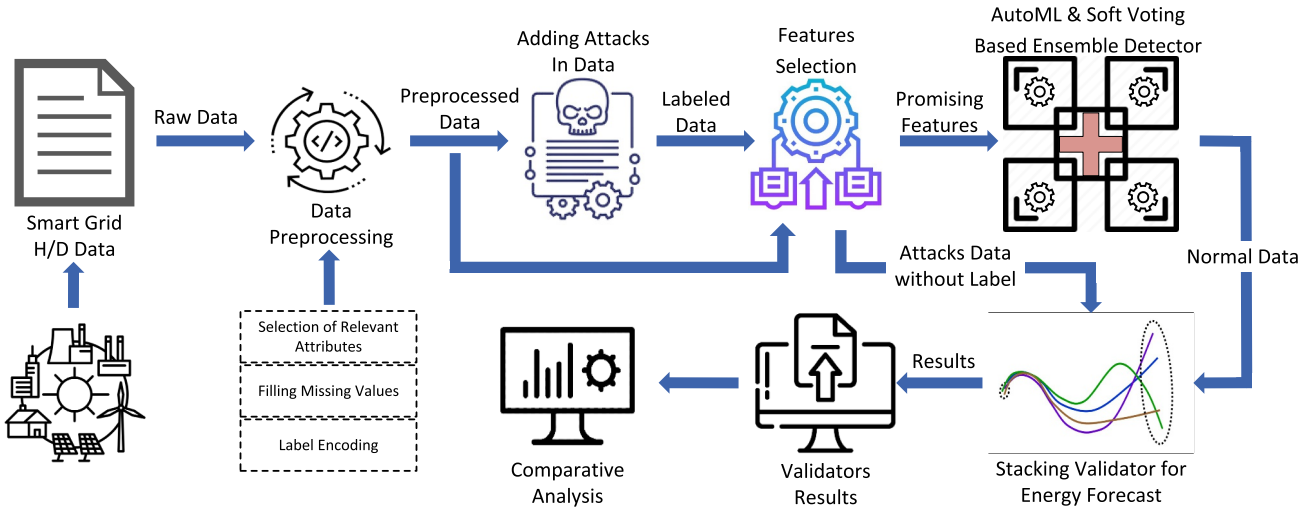


FIGURE 1 | Block diagram of the proposed RES-AD model.

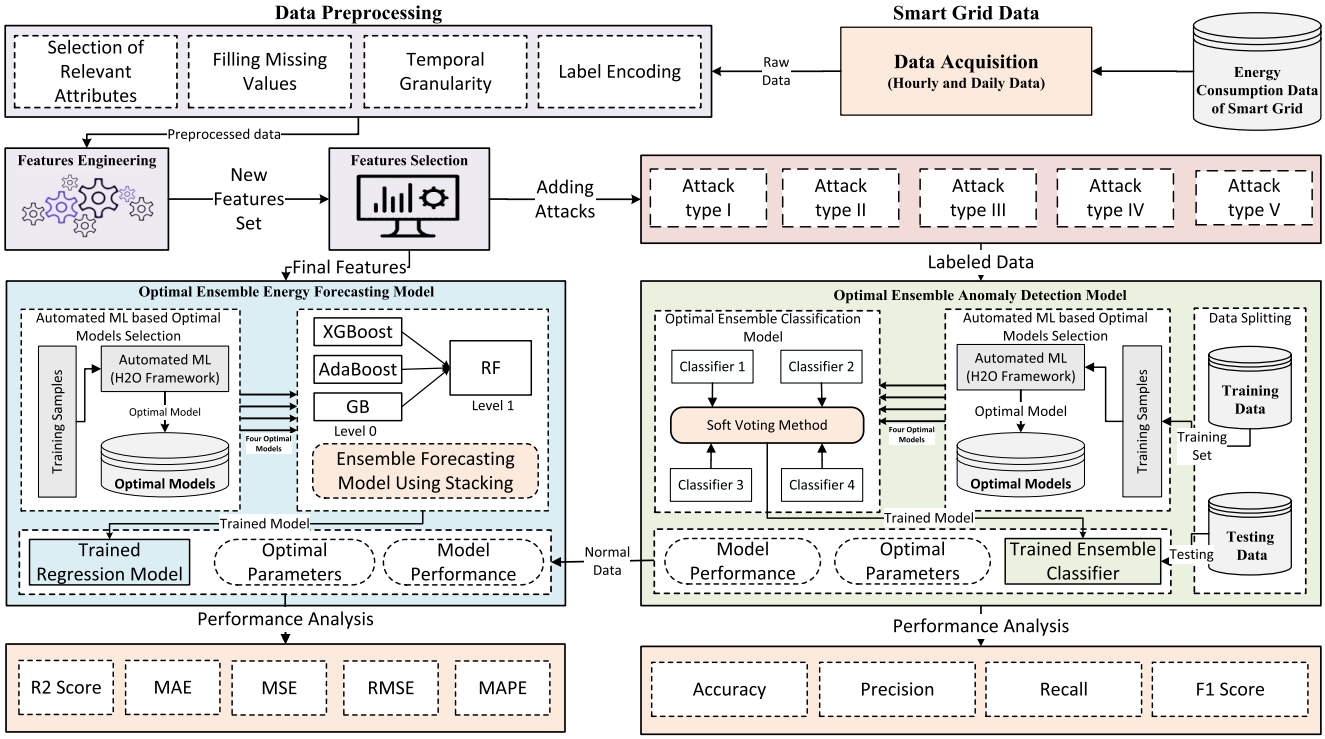


FIGURE 2 | Flowchart depicting the suggested ensemble model.

ALGORITHM 1 | OE-IDS algorithm (modular and classification-oriented).

Input: ER_{data}: raw energy consumption data, θ: optimal hyperparameters
Output: Class labels indicating normal or abnormal consumption
Step 1: Data Preprocessing
foreach $f \in ER_{data}$ **do**
  $f \leftarrow FillMissingValues(mean(f))$
end
$FI_{data} \leftarrow ER_{data}$
Step 2: Attack Injection (Five Data Integrity Attacks)
foreach $r_{sample} \in Rand(FI_{data})$ **do**
  $A_1 \leftarrow ScalingAttack(r_{sample})$ // Eq. (1)
  $A_2 \leftarrow RampAttack(r_{sample})$ // Eq. (2)
  $A_3 \leftarrow RandomAttack(r_{sample})$ // Eq. (3)
  $A_4 \leftarrow PulseAttack(r_{sample})$ // Eq. (4)
  $A_5 \leftarrow FixedScaling(r_{sample})$ // Eq. (5)
  $FI_{data} \leftarrow FI_{data} \cup \{A_1, A_2, A_3, A_4, A_5\}$
end
Step 3: Labeling, Balancing, and Feature Reduction
$LE_{data} \leftarrow LabelEncoding(FI_{data})$
$B_{data} \leftarrow SMOTE(LE_{data})$
$FR_{data} \leftarrow CorrelationSelection(B_{data})$
$SC_{label} \leftarrow FR_{data}.class$
$FP_{data} \leftarrow FR_{data} - class$
Step 4: AutoML Model Selection (No K-Fold)
$O_{models} \leftarrow AutoMLClassifier(FP_{data}, SC_{label})$
// AutoML returns top models and tuned hyperparameters
Step 5: Ensemble Classification (OEC)
foreach *model* $\in O_{models}$ **do**
  $Pred_{model} \leftarrow model(FP_{data})$
end
$Label_0 \leftarrow \frac{1}{4}\sum Pred_{normal}$
$Label_1 \leftarrow \frac{1}{4}\sum Pred_{attack}$
$\hat{y} \leftarrow \arg\max(Label_0, Label_1)$
return $\hat{y}$

generate labelled abnormal samples. Third, the class imbalance introduced by attack injection is addressed using SMOTE, which synthesises minority-class instances. In the fourth step, correlation-based feature selection is applied to remove low-contributing attributes and retain only the most informative features for classification. Finally, an AutoML module automatically identifies the top-performing models and their optimised hyperparameters, which are then combined into an OEC for final anomaly detection.

4 | Data Preparation and Analysis

This research used a publicly available dataset from the UCI ML Repository, **Electricity Load Diagrams 2011-2014** [45]. The dataset contains electricity consumption data measured in kW, recorded every 15 min from 370 smart metres between 2011 and 2014. Each column in the dataset corresponds to a unique smart metre, with some smart metres being introduced after 2011. For these smart metres, consumption values were previously considered zero.

The first column of the dataset contains the date and time in the string format 'YYYY-MM-DD HH:MM:SS', corresponding to Portuguese local time. Each day consists of 96 measurements (24 h × 4 per hour). However, there are specific time changes due to DST:

1. In March, during the time change from Standard Time to Daylight Saving Time (with 23 h in the day), the values between 1:00 AM and 2:00 AM are set to zero for all metres.
2. In October, when the time changes back from Daylight Saving Time to Standard Time (with 25 h in the day), the values between 1:00 AM and 2:00 AM aggregate the consumption from 2 hours into a single data point.

To convert kW to kWh, divide by 4, since each value represents energy consumption over 15 min. Subsequently, the consumption is converted to MWh by dividing the kWh values by 1000.

For our experiments, we used data from a randomly selected user, with consumption values in kW presented as floating-point values across all columns, except the time column.

4.1 | Data Preprocessing

Preprocessing is essential for converting raw data into a format suitable for machine learning algorithms, ensuring more accurate decision-making. The first step is to transform the 15-min consumption data into hourly and daily consumption to align with typical smart metre reporting intervals. Feature engineering is applied to create new features, including temporal features such as hour of the day, day of the week and month of the year, which capture time-of-day and seasonal variations. These features help the model recognise usage patterns such as peak consumption times (e.g., evenings) and seasonal changes (e.g., higher energy use in winter).

Null records are removed, and alphanumeric data (e.g., categorical labels) is encoded into numeric values using a label encoder. Missing values are handled by filling them with the data's mean. To handle outliers, we used DBSCAN to detect them in the dataset. Instead of removing outlier records entirely, outlier consumption values are treated as missing values and imputed using the data's mean. This approach ensures that extreme data points do not distort the model's performance while retaining the data for analysis. The imputation process helps preserve the integrity of the dataset without the risk of removing valuable information.

The DBSCAN parameters, epsilon (ϵ) and min_samples, were selected based on the density of the data points in the feature space. Specifically, ϵ was chosen by examining the distribution of distances between data points to determine the appropriate neighbourhood radius for detecting clusters. The min_samples parameter was set to 5, reflecting the requirement for a minimum number of points to form a valid cluster. These parameters were validated using domain knowledge and cross-validation on a smaller subset of the data to ensure that DBSCAN could effectively identify outliers without overly broad clustering.

4.1.1 | Handling Temporal Trends, DST Anomalies and Missing Values

To account for temporal trends such as time of day and seasonality, we included features that represent different parts of the day (e.g., morning, evening) and seasons (e.g., summer, winter). These features were generated using the hour of the day, day of

the week and month of the year, allowing the model to learn from seasonal consumption patterns and daily usage cycles.

The dataset also contains DST anomalies, which can introduce inconsistencies in time series data. These anomalies were handled by adjusting timestamps to prevent duplicated or missing hours during the spring forward and fallback transitions, thereby maintaining consistency in the time series.

For handling missing values, we used mean imputation for numerical data and SMOTE for class balancing, generating synthetic samples to address class imbalances. The DBSCAN algorithm was used to detect outliers, ensuring the model trained on clean, meaningful data.

4.1.2 | *Attack Simulation and Labelling*

In this study, the unsupervised energy consumption data contains only normal records without attack data. After preprocessing, the data was labelled by adding different types of data integrity attacks (e.g., Scaling Attack, Ramp Attack, Random Attack, Pulse Attack and Fixed Scaling Attack) to simulate realistic attack scenarios. The attacks were introduced to analyse their influence on the stability of the smart grid and market operations, providing insights into potential vulnerabilities in real-world SG environments.

1. Attack Type I: Scaling Attack: The first type of data integrity attack considered in this study is a *Scaling Attack* [13], which manipulates the original energy consumption data by applying a scaling factor, denoted by λ_p. This factor can either increase or decrease energy consumption. The scaling attack is mathematically represented in Equation (1):

$$AEC_t = \left(1 + \lambda_p\right) \cdot ECM_t, \quad \forall\, t = t_\alpha, \tag{1}$$

   where AEC_t represents the anomalous energy consumption at time t, ECM_t is the original energy consumption measurement at time t and λ_p is the scaling parameter. The term t_α denotes the time the attack occurs.
2. Attack Type II: Ramp Attack: The second type of data integrity attack, referred to as a *Ramp Attack* [46], gradually alters the original energy consumption measurements over time. This alteration follows a ramp function, where the energy consumption increases or decreases progressively as time t advances. The ramp attack is represented by the Equation (2):

$$AEC_t = ECM_t + (\lambda_r \cdot t), \quad \forall\, t = t_\alpha. \tag{2}$$

3. Attack Type III: Random Attack: The third type of data integrity attack, known as a *Random Attack* [47], manipulates energy consumption data by adding a randomly generated value within a specified range. This attack is defined in Equation (3):

$$AEC_t = ECM_t + \lambda_r \cdot \text{rand}(lower_{bound}, upper_{bound}), \quad \forall\, t_s < t < t_e. \tag{3}$$

4. Attack Type IV: Pulse Attack: The fourth type of data integrity attack, called a *Pulse Attack* [35], involves introducing sudden, short-duration spikes or drops in energy consumption at specific timestamps. This attack is designed to simulate extreme but brief changes in energy consumption. The pulse attack is represented in Equation (4):

$$AEC_t = ECM_t + \lambda_p \cdot \text{pulse}(t), \quad \forall\, t = t_\alpha. \tag{4}$$

5. Attack Type V: Fixed Scaling Attack: The fifth type of data integrity attack is the *Fixed Scaling Attack* [48], where a constant scaling factor is applied to the energy consumption data, resulting in a persistent bias. The Equation (5) represents the fixed scaling attack:

$$AEC_t = \left(1 + \lambda_f\right) \cdot ECM_t, \quad \forall\, t = t_\alpha. \tag{5}$$

4.1.3 | *SMOTE and Class Imbalance*

Data integrity attacks can occur as part of cyberattacks, specifically FDIA, DoS or man-in-the-middle attacks, in the context of SG. They don't fully represent the broad range of cyberattacks that could affect an SG, such as hacking into communication networks or taking control of devices. However, they are certainly cyberattack-related because they manipulate data to disrupt grid operations. Since the Electricity Load Diagrams 2011-2014 dataset lacks attack data and is unsupervised, we added attack types. However, these attacks are fewer than the normal records, leading to data imbalance. To address this, we used SMOTE to balance the class distribution by generating synthetic samples for the minority class. SMOTE helps prevent the model from becoming biased toward the majority class. The SMOTE process is shown in Equation (6). The optimised SMOTE parameters used for data balancing are presented in Table 2.

$$S_1 = A_1 + GAP \cdot Diff. \tag{6}$$

4.2 | Features Analysis

Understanding the relevance of input features is essential for developing an efficient and robust ML-based anomaly detection framework. In this study, a detailed feature analysis was conducted to quantify the contributions of both statistical and temporal features for predicting energy consumption behaviour. Pearson's correlation analysis was used to evaluate the strength of association between each feature and the target variable, as illustrated in Figure 3. The results show that not all temporal features contribute equally to model performance. In particular, the month and date exhibit a very weak correlation with the consumption and provide minimal discriminatory power. Therefore, these two features were excluded from the final feature set to reduce noise and prevent unnecessary complexity. Conversely, other temporal features, such as day of the week and quarter, demonstrate meaningful correlation patterns and

TABLE 2 | SMOTE parameters for data balancing.

Parameter	Optimised value
Sampling strategy	Auto
K neighbours (*k_neighbors*)	5
Random state	42
Number of jobs (*n_jobs*)	−1

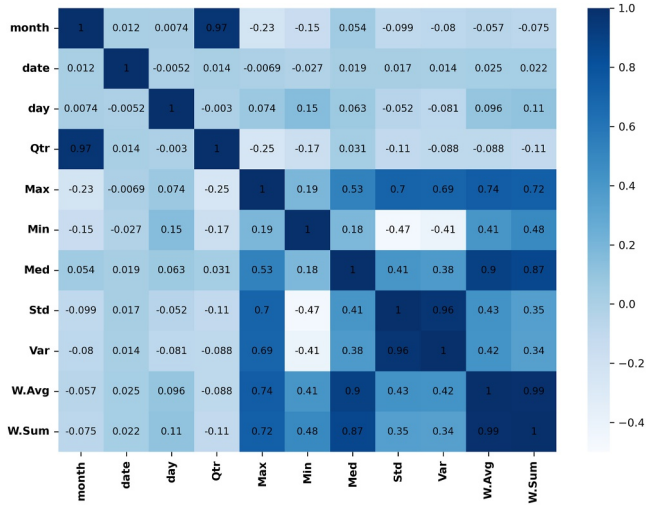


FIGURE 3 | Pearson correlation analysis.

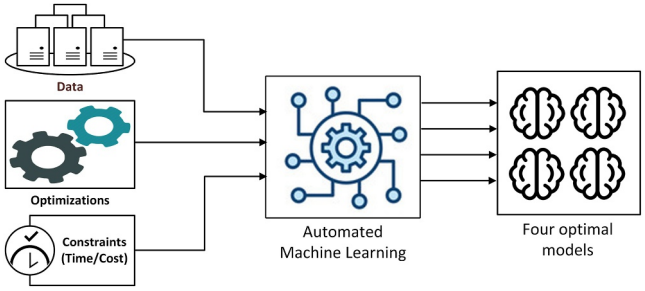


FIGURE 4 | Selection of optimal classification models using AutoML.

were retained, as they capture behavioural cycles (weekday–weekend differences) and broader seasonal variations that influence consumption patterns.

Additionally, statistical features such as maximum, minimum, median, standard deviation and variance show stronger correlations and therefore play a more dominant role in model decision-making. These features effectively capture the distribution and variability of consumption across time windows, thereby contributing substantially to anomaly identification. By evaluating feature contributions through correlation and temporal relevance, we ensured that the model captures essential temporal dependencies while avoiding redundant or weakly informative features. This aligns with the reviewer's concern by clarifying which temporal attributes meaningfully influence the model and how feature selection was guided by quantitative analysis.

5 | AutoML and Voting Classifier

AutoML is employed to extract four optimal models with their corresponding hyperparameters. Using AutoML, the following four optimal models are identified: AdaBoost, GB, XGBoost and RF Classifier. These four models are then applied individually to two distinct datasets: the hourly dataset and the daily dataset, using the extracted hyperparameters.

5.1 | AutoML-Based Model Selection and Tuning

Instead of manually selecting optimal models by testing each one individually, AutoML was used to automate the process. AutoML prepares the data and returns the best models along with their hyperparameters. As shown in Figure 4, AutoML identified four optimal models for the dataset: AdaBoost, GB, XGBoost and RF, each with its respective set of hyperparameters.

To run AutoML on a dataset, several key parameters must be specified. The first is *y*, which represents the index of the response column. The second is the training frame, which is used to select the training data. We use a simple train-validation-test split with 80%, 10% and 10% of the data for training, validation and testing, respectively. Additionally, two stopping parameters are crucial: `max_runtime_secs`, which sets the time limit for AutoML (with the default being 0 for no limit) and `max_models`, which defines the maximum number of models to train for the leaderboard. The `score_validation_set parameter`, when set to `TRUE`, ensures that the validation set is scored during training. Other optional parameters like *x* (input feature columns), `validation_frame` (used for validation data), `leader_board_frame` (for leaderboard generation) and `blending_frame` (for blending models), provide further customisation and flexibility in the AutoML process.

Additionally, for optimising the performance of AutoML, parameters such as `parallelism` (*n_threads*) and `stopping criteria`, including stopping_metric (e.g., Decrease in AUC) and `stopping_tolerance` (e.g., 0.001), help control training behaviour and ensure efficient model selection.

The best hyperparameters for the identified models and the AutoML H_2O input parameters are shown in the Table 3.

5.2 | Overfitting Risk in AutoML Models

One of the risks when using AutoML frameworks for training models, particularly when synthetic data is involved, is the potential for overfitting. AutoML aims to identify the best-performing models by exploring a wide range of algorithms and hyperparameters. However, this approach can sometimes lead to models that are too tailored to the training data, including synthetically generated attack data, which may not generalise well to real-world unseen data.

To mitigate the risk of overfitting, we employed the following strategies:

1. Train-Validation-Test Split: We used a standard 80-10-10 train-validation-test split, ensuring that the model is evaluated on data it has not seen during training. The validation set is used to tune the model, and the final model performance is assessed on a held-out test set.
2. Regularisation Techniques: We applied regularisation methods, such as early stopping and hyperparameter tuning (e.g., max depth, learning rate), which help prevent models from fitting excessively to noise or outliers in the synthetic data.
3. Synthetic Data Augmentation: By augmenting the synthetic attack data with a variety of adversarial scenarios

TABLE 3 | AutoML input parameters and best hyperparameters for regression and classification models returned by AutoML.

AutoML H_2O parameters	
Parameters	**Optimised values**
Max runtime (seconds)	0 (no limit)
Max models (`max_models`)	50
Stopping criteria (`stopping_metric`)	Decrease in AUC
Stopping tolerance (`stopping_tolerance`)	0.001
Score validation set (`score_validation_set`)	TRUE
Parallelism (`n_threads`)	4
Best hyperparameters for models	
XGBoost	
Learning rate (`eta`)	0.01 (regression)/0.05 (classification)
Maximum depth (`max_depth`)	6 (regression)/5 (classification)
Number of estimators (`n_estimators`)	100 (regression)/150 (classification)
Subsample rate (`subsample`)	0.8 (regression)/0.85 (classification)
Minimum child weight (`min_child_weight`)	1 (both)
Gamma (`gamma`)	0.1 (both)
AdaBoost	
Learning rate (`eta`)	0.05 (regression)/0.1 (classification)
Number of estimators (`n_estimators`)	50 (regression)/100 (classification)
Base estimator (`base_estimator`)	DecisionTreeClassifier (max depth 3) (both)
Learning rate schedule	Constant (regression)/adaptive (classification)
Gradient boosting	
Learning rate (`eta`)	0.1 (regression)/0.05 (classification)
Number of estimators (`n_estimators`)	100 (regression)/150 (classification)
Maximum depth (`max_depth`)	5 (regression)/6 (classification)
Random forest	
Number of estimators (`n_estimators`)	200 (both)
Maximum depth (max_depth)	10 (regression)/15 (classification)
Minimum samples split (`min_samples_split`)	2 (regression)/5 (classification)
Minimum samples leaf (`min_samples_leaf`)	1 (regression)/2 (classification)

(e.g., different attack types, attack intensities and variations), we exposed the AutoML models to a broader range of attack behaviours, improving the ability of the models to generalise to unseen adversarial patterns.

Despite these measures, we acknowledge that overfitting remains a potential risk, particularly when synthetic data does not adequately capture the full complexity of real-world attacks. Future work will focus on incorporating real-world datasets alongside synthetic data to validate the framework's generalisability further and reduce overfitting risks.

5.3 | Safeguards Against Adversarial Attacks on AutoML Models

The vulnerability of AutoML-generated models to adversarial attacks is an important consideration, particularly when training on synthetically generated data. Since adversarial inputs are specifically crafted to mislead models, there is a potential risk that AutoML might overfit to the synthetic attack data, making it more susceptible to real-world adversarial manipulations.

To address this, we implemented several safeguards:

1. Adversarial Training: We included adversarially perturbed samples in the training data, allowing the AutoML model to learn from both clean and adversarial examples. This ensures that the model becomes more robust to adversarial inputs, improving its generalisation to real-world attacks.
2. Ensemble Learning: By using an ensemble approach, the model aggregates predictions from multiple algorithms, reducing the likelihood that any single model will become overly sensitive to adversarial perturbations.
3. Evaluation with Adversarial Examples: After training, we tested the model on adversarial examples to assess its performance in the presence of adversarial manipulation. If significant degradation in performance is observed, we plan

to incorporate additional adversarial defence mechanisms in future versions of the framework, such as adversarial detection or robust optimisation techniques.

These steps aim to improve the robustness of AutoML models against adversarial attacks and reduce the risk of overfitting. However, we recognise that as adversarial techniques become more sophisticated, continuous adaptation of the framework will be necessary to ensure its resilience against emerging threats.

6 | Performance Analysis

In this study, two types of experiments were conducted. The first experiment uses regression to predict energy consumption without adding attacks using hourly and daily consumption datasets. The second experiment uses classification to predict anomalies after adding attacks to hourly and daily consumption data.

6.1 | Evaluation of Regression Models

Regression is a type of ML used to find the relationship between input and target features. In other words, Regression can be used for an ML problem in which we need to predict continuous values, such as energy consumption, price, Rating, temperature, etc.

1. R^2 Score: The R^2 score is important to evaluate the Regression-based ML model's performance. It is the coefficient of determination, called the goodness-of-fit statistic and is known as *R*-squared. A dataset's variance is measured by the number of predictions it explains. In simple terms, it is the difference between the actual values in the test data and the model's predicted values. Equation (7) shows how to calculate the R^2 score.

$$R2Squared = 1 - \frac{\mathrm{SSr}}{\mathrm{SSt}}. \tag{7}$$

SSR is the sum of the residuals squared and *SSt* represents the sum of the squared distances between the data and the mean. The value can range from 0 to 1, since it is a percentage.

2. MAE: MAE is a straightforward metric that measures the absolute difference between actual and predicted values. For example, in a Linear Regression model, the error (or MAE) is the difference between the actual and predicted values. To calculate the MAE for the entire dataset, the absolute differences across all records are summed and divided by the total number of records. The goal is to minimise the MAE, as it represents the model's loss. Equation (8) shows how to calculate the MAE value.

$$MAE = \frac{1}{s}\sum_{j=1}^{s}|x_i - y_i|, \tag{8}$$

here y_i is the value predicted by the model, x_i is the actual value in the data and s is the total number of samples.

3. MSE: MSE is a commonly used performance metric that differs from MAE by calculating the squared difference between actual and predicted values. Although MAE calculates the absolute difference, MSE squares the differences, eliminating negative values and preventing term cancelation. This makes MSE particularly useful for highlighting larger errors. Equation (9) shows how to calculate the MSE value.

$$MSE = \frac{1}{s}\sum_{j=1}^{s}(x_i - y_i)^2. \tag{9}$$

In the above Equation (9), y_i is the value predicted by the model, x_i is the actual value in the data and s is the entire sample.

4. RMSE: RMSE measures the error between actual and predicted values by calculating the difference between the true and predicted values. A lower RMSE indicates that the predicted values are closer to the actual values, indicating better model performance. Equation (10) shows how to calculate the RMSE value.

$$RMSE = \sqrt{\frac{\Sigma_{j=1}^{s}(x_i - y_i)^2}{s}}. \tag{10}$$

In the above Equation (10), y_i is the value predicted by the model, x_i is the actual value in the data and s is the total number of samples.

5. MAPE: After calculating the absolute error for each data sample, the MAPE is found by summing all the absolute errors and dividing the total by the number of data entries. For example, if the dataset has 369 entries, the sum is divided by 369 to obtain the MAPE. Equation (11) shows how to calculate the MAPE value.

$$MAPE = \frac{1}{s}\sum_{t=1}^{s}\left|\frac{A_t - F_t}{A_t}\right|, \tag{11}$$

where F_t is the forecast value by model, A_t is the actual value in the data and s is the total number of fitted samples.

6.2 | Evaluation of Classification Models

Confusion matrices provide a valuable framework for evaluating model performance by comparing predicted and actual classifications. They help identify critical values such as True Positives, True Negatives, False Positives and False Negatives, which serve as the foundation for calculating key performance metrics. Below are some of the essential metrics used to assess a model's effectiveness in classification tasks:

1. True Positive Rate (T_{pr}): T_{pr} is calculated by $T_{pr} = T_p/(T_p + F_n)$. T_{pr} is the probability that a real anomaly will be predicted as an anomaly by the model. The detection rate and sensitivity of the model are alternative names used to determine the T_{pr}.
2. True Negative Rate (T_{nr}): T_{nr} is calculated by $T_{nr} = T_n/(T_n + F_p)$. This represents the probability that a benign (nonmalicious) instance in the SG data is correctly classified as normal by the proposed RES-AD model. This metric is also known as the anomaly detection system's specificity.
3. False Positive Rate (F_{pr}): This is calculated by $F_{pr} = F_p/(F_p + T_n)$. When the RES-AD model classifies a

normal packet in the SG network as anomalous, it is referred to as a false positive rate F_{pr}, also known as a false alarm.

4. False Negative Rate (F_{nr}): This is calculated by $F_{nr} = F_n/(F_n + T_p)$. This scenario occurs when the model incorrectly classifies a legitimate event as an anomaly within the SG network, resulting in a false positive, also known as a false alarm.

Now, performance metrics such as correctness rate, exactness, sensitivity and harmonic mean can be computed using T_p, T_n, F_p and F_n as follows:

Accuracy is calculated as the ratio of correct predictions to total predictions.

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}. \tag{12}$$

Precision is defined by dividing $True_p$ by the sum of $True_p$ and $False_p$.

$$Precision = \frac{T_p}{T_p + F_p}. \tag{13}$$

Recall is defined as the ratio of $True_p$ and the sum of $True_p$ and $False_n$.

$$Recall = \frac{T_p}{T_p + F_n}. \tag{14}$$

Furthermore, Precision and recall are transformed into a harmonic mean to determine the $F1$ score. The basic formula for the $F1$ score is given in Equation (15).

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}. \tag{15}$$

7 | Experimental Results

This study conducts various experiments using the proposed ensemble and traditional ML models on the Electricity Load Diagrams 2011–2014 dataset. The dataset is divided into two sub-datasets: one with hourly data and the other with daily consumption data.

7.1 | Anomaly Detection Model Results

In this section, four types of results are discussed: hourly and daily validation data results, and hourly and daily testing results.

7.1.1 | *Hourly/Daily Consumption Validation Results*

Table 4 shows the validation results of hourly and daily data. The proposed OEC achieved significantly better results on the validation data. Based on the provided hourly consumption data, our proposed OEC achieves the following results: 99.0% accuracy, 98.8% precision, 99.3% recall and 99.1% $F1$ score. Similarly, for the daily consumption data, the results of our proposed OEC are as follows: 99.6% accuracy, 99.3% precision, 99.9% recall and 99.6% $F1$ score.

7.1.2 | *Hourly/Daily Consumption Results Using Testing Samples*

Table 5 shows the classification results of Hourly/Daily consumption data. All the selected conventional ML models and the proposed OEC achieved significantly better results. On the given hourly consumption data set, the proposed OEC results in accuracies of 99.5%, 99.7%, 99.4% and 99.5% for accuracy, precision, recall and $F1$ score, respectively. The results of the proposed OEC for daily consumption data in terms of accuracy, precision, recall and $F1$ score are 100%, 99.9%, 99.8% and 99.9%, respectively.

7.2 | Energy Consumption Forecasting

In this section, regression results for hourly and daily data, with and without attacks, are presented in a results table and a line chart.

7.2.1 | *Energy Consumption With Anomaly Data*

Two types of data, hourly and daily energy consumption, are used for prediction experiments. The first experiment uses four conventional ML models and the proposed OER model to predict hourly consumption. At level 0, AdaBoost, GB and XGB are applied, whereas RF is used at level 1. Table 6 presents the results of hourly predictions with attack data, including individual model performances and the proposed OER model. The second experiment uses the same models to predict daily energy consumption with anomalies, and the results are analysed using R^2 score, MAE, RMSE and MAPE.

In Figure 5a,b, the hourly and daily actual energy consumption and predicted energy consumption are highlighted with the help of a line chart. The blue line shows the actual energy consumption in the original data, and the remaining lines show the predicted energy consumption from four selected ML algorithms and the proposed OER model.

7.2.2 | *Anomaly Detection Based Energy Consumption Results*

This section uses the same data as before, but with anomalies removed, to predict hourly and daily energy consumption. The first experiment uses four conventional ML models and the proposed OER model to predict hourly consumption. AdaBoost, GB and XGB are applied at level 0, with RF at level 1. Table 7 presents the results for hourly predictions without attack data, including individual models and the proposed OER model. The second experiment, using daily consumption data without anomalies, applies the same models and analyses results using R^2, MAE, MSE, RMSE and MAPE.

Figure 6a illustrates the energy forecasting results obtained using OER and other regression models, following anomaly detection and removal performed by the proposed OEC model. The results indicate a low difference between actual and forecasted energy consumption, demonstrating the proposed OER's accuracy and individual learners' accuracy. However, AdaBoost shows slightly higher forecasting errors than the proposed model and other learners. This leads to higher MAPE for AdaBoost, whereas GB, XGB and RF learners

TABLE 4 | Hourly/daily consumption validation results.

Data	Model	Accuracy (%)	Precision (%)	Recall (%)	*F*1 score (%)
Hourly consumption	AdaBoost	98.3 ± 0.0061	98.0 ± 0.0109	98.6 ± 0.0076	98.3 ± 0.0061
	GB	98.8 ± 0.0037	98.3 ± 0.0086	99.2 ± 0.0037	98.8 ± 0.0035
	XGB	99.2 ± 0.0017	99.1 ± 0.0041	99.3 ± 0.0037	99.2 ± 0.0017
	RF	99.0 ± 0.0033	98.9 ± 0.0053	99.2 ± 0.0040	99.1 ± 0.0028
	Proposed OEC	99.0 ± 0.0032	98.8 ± 0.0063	99.3 ± 0.0045	99.1 ± 0.0031
Daily consumption	AdaBoost	99.4 ± 0.0031	99.0 ± 0.0067	99.8 ± 0.0034	99.4 ± 0.0031
	GB	96.5 ± 0.0043	94.8 ± 0.0077	98.5 ± 0.0021	96.6 ± 0.0041
	XGB	99.6 ± 0.0021	99.2 ± 0.0033	99.9 ± 0.0023	99.6 ± 0.0021
	RF	99.6 ± 0.0015	99.4 ± 0.0027	99.7 ± 0.0023	99.6 ± 0.0014
	Proposed OEC	99.6 ± 0.0015	99.3 ± 0.0028	99.9 ± 0.0021	99.6 ± 0.0016

TABLE 5 | Hourly/daily consumption testing results.

Data	Model	Accuracy (%)	Precision (%)	Recall (%)	*F*1 score (%)
Hourly consumption	AdaBoost	0.991 ± 0.0039	0.9901 ± 0.0063	0.990 ± 0.0037	0.990 ± 0.0040
	GB	0.990 ± 0.0039	0.990 ± 0.0063	0.9904 ± 0.0037	0.990 ± 0.0040
	XGB	0.994 ± 0.0029	0.996 ± 0.0047	0.993 ± 0.0031	0.994 ± 0.0029
	RF	0.992 ± 0.0029	0.993 ± 0.0040	0.991 ± 0.0039	0.992 ± 0.0032
	Proposed OEC	0.995 ± 0.0036	0.997 ± 0.0046	0.994 ± 0.0033	0.995 ± 0.0038
Daily consumption	AdaBoost	0.997 ± 0.0032	0.996 ± 0.0044	0.988 ± 0.0033	0.997 ± 0.0031
	GB	0.977 ± 0.0062	0.961 ± 0.0097	0.995 ± 0.0064	0.978 ± 0.0068
	XGB	0.999 ± 0.0010	0.996 ± 0.0014	0.993 ± 0.0021	0.994 ± 0.0010
	RF	0.998 ± 0.0018	0.997 ± 0.0018	0.997 ± 0.0027	0.998 ± 0.0017
	Proposed OEC	1.000 ± 0.0013	0.999 ± 0.0014	0.998 ± 0.0027	0.999 ± 0.0015

TABLE 6 | Performance analysis of energy consumption forecasting with anomaly data.

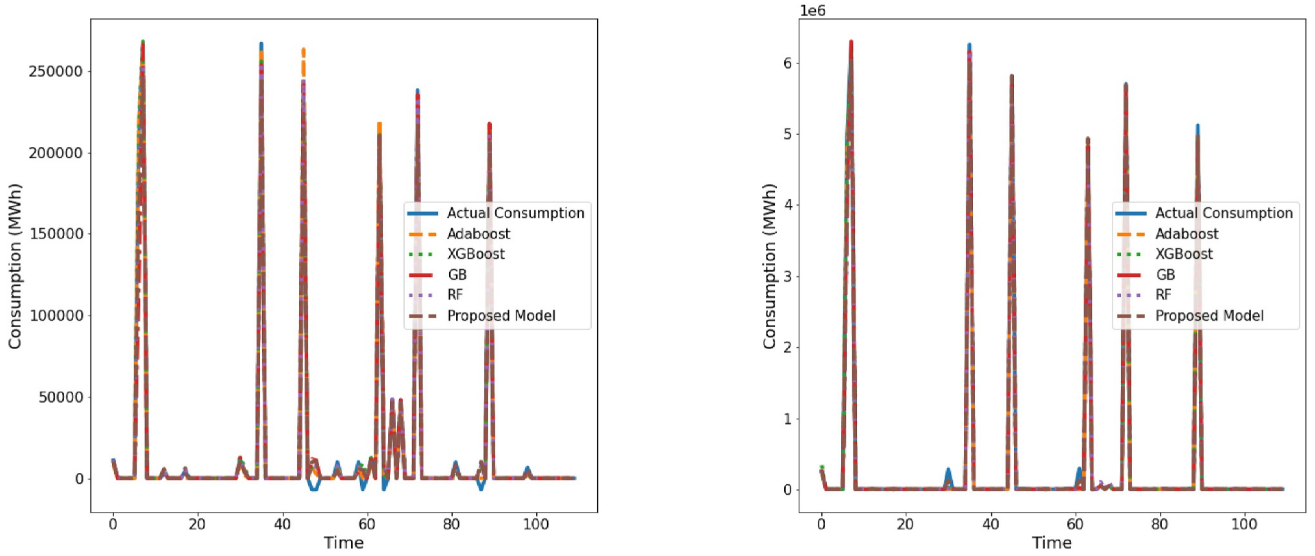
Data	Model	R^2 score	MAE	RMSE	MAPE
Hourly (with attacks)	AdaBoost	0.994	1736.612	4278.008	24.04
	GB	0.989	1645.012	6176.904	13.23
	XGB	0.992	1776.331	5299.780	24.04
	RF	0.993	1570.757	4795.751	15.13
	Proposed OER	0.995	1426.789	4170.641	13.21
Daily (with attacks)	AdaBoost	0.998	16,873.262	58,135.778	37.25
	GB	0.995	19,043.985	97,879.501	13.09
	XGB	0.998	15,691.920	60,435.899	18.07
	RF	0.996	22,998.909	83,663.979	14.90
	Proposed OER	0.999	12,738.155	50,513.053	12.58

produce lower errors and perform better for hourly energy forecasting.

Similarly, after anomaly removal by the proposed OEC model, Figure 6b illustrates the energy forecasting performance of OER alongside other regression models on daily data. The results show that the OEC detector improves the forecasting performance of both the proposed OER and individual models, with forecasting errors below 5%. This highlights the OEC detector's effectiveness in supporting efficient energy forecasting and aiding policymakers in making informed energy supply and demand decisions. The proposed OER model demonstrates superior MAPE performance compared to individual models, making it a more efficient and reliable forecasting solution for SG management.

7.3 | Stealthy Adversarial Attack Simulation

To further evaluate the robustness of the proposed RES-AD framework against modern stealth attacks, we simulated adversarial perturbations on the hourly consumption data using the FGSM. Adversarial attacks of this nature are specifically



(a) Hourly energy consumption with attacks.

(b) Daily energy consumption with attacks.

FIGURE 5 | Comparison between actual and predicted energy consumption values under simulated attack scenarios: (a) Hourly and (b) daily.

TABLE 7 | Performance analysis of energy consumption forecasting without anomaly data.

Data	Model	R^2 score	MAE	RMSE	MAPE
Hourly (without attacks)	AdaBoost	0.888	3.422	4.758	3.02
	GB	0.929	2.655	3.790	2.33
	XGB	0.903	2.729	4.428	2.40
	RF	0.826	3.522	5.937	3.23
	Proposed OER	0.931	2.491	3.730	2.10
Daily (without attacks)	AdaBoost	0.950	47.564	75.362	1.78
	GB	0.959	28.499	68.357	1.10
	XGB	0.956	29.097	70.225	1.14
	RF	0.971	28.662	57.287	1.04
	Proposed OER	0.977	22.896	50.519	0.86

designed to subtly modify data, making it difficult for conventional machine learning detectors to identify the manipulation.

In this simulation, a small, imperceptible perturbation was applied to the input features x by computing the gradient of the loss function $\mathcal{L}(x, y)$ with respect to x, as shown in Equation (16):

$$\delta = \epsilon \cdot \text{sign}(\nabla_x \mathcal{L}(x, y)), \tag{16}$$

where δ is the adversarial perturbation and ϵ is the perturbation strength (in this case, $\epsilon = 0.01$). $\nabla_x \mathcal{L}(x, y)$ represents the gradient of the loss function with respect to the input features x and x and y are the input feature vector and the corresponding true label, respectively.

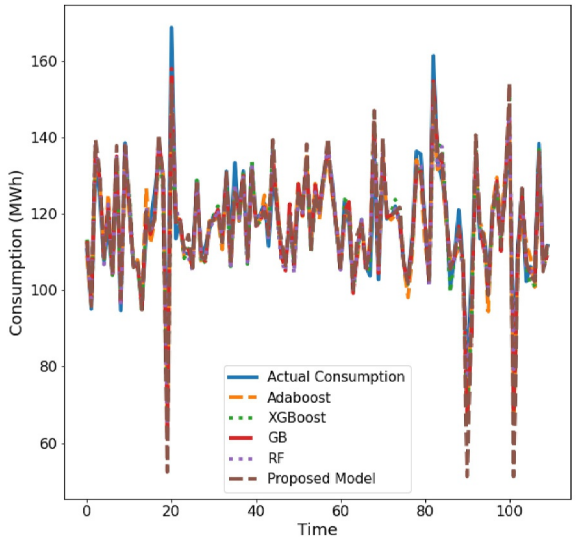
The perturbation δ is then added to the original data, resulting in adversarial examples $\tilde{x}$, shown in Equation (17):

$$\tilde{x} = x + \delta = x + \epsilon \cdot \text{sign}(\nabla_x \mathcal{L}(x, y)). \tag{17}$$

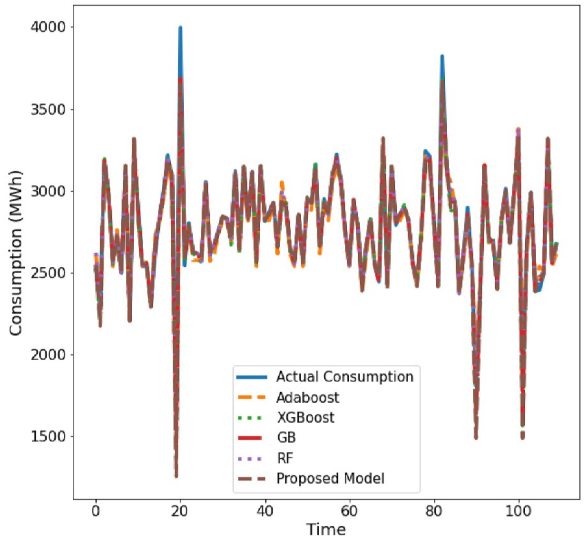
These perturbed samples were designed to preserve the overall load profile while subtly shifting feature correlations in such a way that they evade conventional detection models. The goal was to simulate how a small adversarial perturbation could go unnoticed, yet still disrupt model performance.

When tested under this adversarial stress scenario, the proposed OEC maintained a strong $F1$-score of 97.8% and recall of 98.2%, indicating that the model can handle small-scale stealth perturbations. However, there was a 1.7% performance drop compared with the baseline dataset, highlighting the system's resilience to such attacks. These results suggest that although RES-AD remains reliable at detecting anomalies under moderate adversarial manipulation, further model refinement is required to withstand larger-scale or more sophisticated adversarial attacks.

To evaluate how adversarial training can improve the model's robustness, we introduced adversarial augmentation into the



(a) Hourly forecasting without attacks.



(b) Daily forecasting without attacks.

FIGURE 6 | Energy consumption forecasting using the proposed OEC detector and regression models under normal (nonattack) conditions: (a) Hourly and (b) daily.

TABLE 8 | Performance of RES-AD under different adversarial perturbations and adversarial augmentation.

Scenario	ϵ	Perturbed	Accuracy	Precision	$F1$
Clean	0.000	0	99.50	99.70	99.50
$attack_{only}$	0.002	100	99.65	99.85	98.92
$attack_{only}$	0.005	100	99.30	99.80	97.81
$attack_{only}$	0.010	100	98.85	99.50	96.35
$attack_{only}$ $with_{train_aug}$	0.005	100	99.23	99.30	99.15

training process. Specifically, we perturbed 10% of the training samples (using $\epsilon = 0.005$) and added them to the training set. We then tested the model using 5% perturbed attack samples in the test set.

As shown in Table 8, after applying adversarial augmentation, the $F1$-score improved significantly, recovering to 99.15%, compared to 99.50% on the clean test set. This indicates that adversarial augmentation effectively restores performance after perturbation, demonstrating that the model can handle moderate adversarial manipulation when trained with augmented data.

To further assess the robustness of the RES-AD framework, we evaluated its performance under varying levels of noise and adversarial perturbations. As the intensity of noise and perturbation increased, there was a gradual decline in $F1$-score, highlighting the sensitivity of the model to adversarial manipulations. The model's performance degradation demonstrates the importance of adversarial training and ensemble diversity to ensure robustness in the face of real-world data irregularities.

Although RES-AD demonstrates strong resilience against **small-scale adversarial perturbations, the inclusion of larger-scale adversarial samples or more sophisticated multivector attacks (e.g., replay attacks or protocol-level exploits) could expose further vulnerabilities in the system. Currently, the dataset used in this work does not contain network-level data or protocol exploits like IEC 61850 or MQTT traffic, which would require future work focused on protocol-aware intrusion detection.

7.4 | Comparative Evaluation With Deep Learning Models

To further strengthen the evaluation, we implemented three representative DL architectures frequently used in recent SG anomaly detection literature: TabNet, a CNN-based classifier and a GNN constructed using a kNN graph. All DL models were trained under the same experimental environment used for RES-AD, with identical preprocessing, temporal feature engineering, DBSCAN outlier handling and SMOTE class balancing. This ensures a fair and consistent comparison across all models.

Despite using the same features, balancing strategy and preprocessing pipeline, the DL models exhibit noticeably lower performance as shown in Table 9. TabNet struggles with the dataset's statistical structure, CNN performs moderately but fails to capture the broader distributional patterns introduced by synthetic attacks and the GNN is limited by the sparse relational structure of consumption data. In contrast, the AutoML-optimised OEC within RES-AD consistently delivers superior accuracy, precision, recall and $F1$ score. These results indicate that for this dataset and attack configuration, the RES-AD architecture is more effective and stable than modern DL-based alternatives.

7.5 | Real-Time Inference Latency and Resource Utilisation

To support the claim of real-time readiness, we evaluated the inference latency, memory usage and model size of each AutoML-

TABLE 9 | Performance comparison between RES-AD and recent deep learning models under the same experimental setup.

Model	Accuracy	Precision	Recall	*F*1 score
RES-AD (OEC)	0.995	0.997	0.994	0.995
TabNet	0.187	0.167	0.982	0.285
CNN-based model	0.972	1.000	0.833	0.909
GNN (kNN graph)	0.745	0.342	0.586	0.432

TABLE 10 | Latency, memory usage and model size of RES-AD base models and ensemble.

Model	Latency (ms)	Memory (MB)	Model size (MB)
AdaBoost	9.567	0.01	0.08
GB	0.160	0.01	0.25
XGBoost	0.178	0.01	0.54
RF	6.395	0.02	3.88
OEC	23.367	0.03	9.46

selected base learner and the final OEC ensemble. All measurements were obtained under the same runtime environment used for model training and testing. As shown in Table 10, the individual models exhibit sub-10 ms inference times, with the ensemble achieving real-time performance at 23.36 ms per sample. Memory usage remains below 0.05 MB for all models, demonstrating the lightweight nature of the RES-AD components and their suitability for edge or near-edge deployment.

These results confirm that the RES-AD ensemble maintains low computational overhead and supports real-time anomaly detection in smart grid environments. The lightweight model sizes enable deployment on edge devices such as Raspberry Pi, industrial gateways and microservice containers within SCADA/Kafka processing pipelines.

7.6 | Comparison With SOTA Methods

To benchmark the effectiveness of the proposed RES-AD framework, we compare it with SOTA methods recently introduced in the domain of SG anomaly detection and forecasting. Table 11 summarises key methodologies, quantitative performance metrics, and highlights for each approach, focussing on papers published after 2023. For example, Zheng et al. achieved an AUC of 0.9653 using SVM and RF fusion. At the same time, Turowski et al. reported an RMSE of 0.15 in forecasting tasks by comparing anomaly compensation versus raw data processing. Korba et al. developed a two-level detection model reaching a 98% detection rate with only 1.5% false alarms. Although these methods demonstrate promising performance in specific aspects (e.g., detection accuracy or compensation efficiency), they often address anomaly detection or forecasting in isolation, lack synthetic attack simulation or omit automated model tuning. In contrast, the proposed RES-AD framework integrates AutoML-optimised ensemble learning, synthetic attack injection and SMOTE-based balancing into a unified, end-to-end pipeline, achieving superior detection (*F*1-score up to 100%) and forecasting performance (MAPE of 11.11% hourly and 11.72% daily).

8 | Discussion

Figure 7a presents the testing results of the proposed ensemble learner compared with the individual base models. The proposed ensemble learner achieves 99.5% accuracy, which represents a 0.1%–0.4% improvement over the best-performing individual models (GB at 99.4% and XGBoost at 99.2%). Similarly, the ensemble shows higher precision (99.7%), recall (99.4%) and *F*1 score (99.5%), outperforming the strongest baseline models by 0.1%–0.6% across all metrics in hourly energy consumption data.

Figure 7b presents the testing results of the proposed OEC compared with individual learners for classifying normal and abnormal daily energy consumption records. The proposed OEC achieves 100% precision, 99.9% recall, 99.9% *F*1 score and 99.8% accuracy, outperforming AdaBoost, XGB and RF by 1.0%–2.2% across the reported metrics. Compared to the Gradient Boosting learner, which records an *F*1 score of 96.1%, the OEC delivers a 3.8% absolute improvement, highlighting the stronger robustness and reliability of the proposed ensemble model when applied to daily consumption patterns.

Figure 8 shows a significant reduction in error rates for hourly and daily energy consumption forecasting after anomalous records were identified and removed. The blue bar represents hourly consumption, and the orange bar represents daily consumption, highlighting the differences in MAPE before and after removing abnormal data. This visualisation underscores the importance of identifying anomalous records to improve predictive model accuracy.

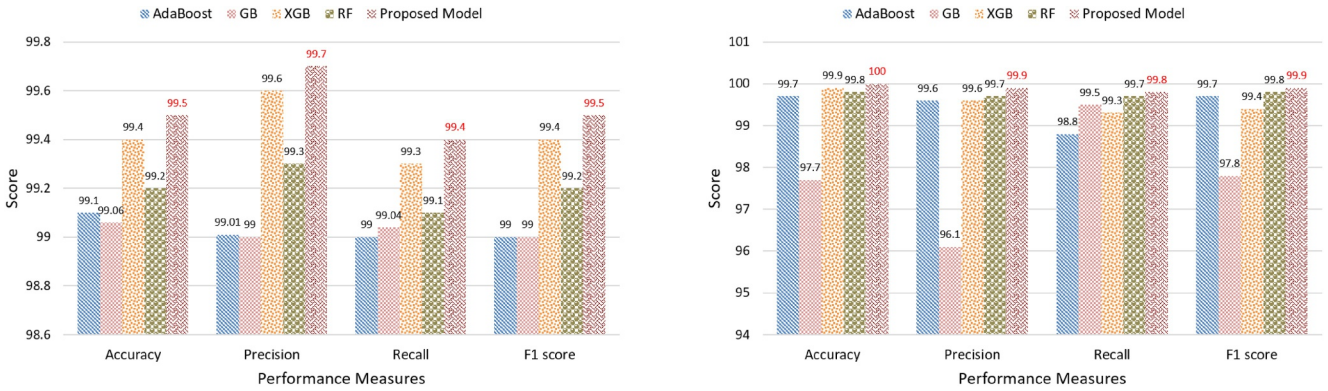
For hourly energy consumption, AdaBoost, GB, XGB, RF and the proposed OER showed significant reductions in MAPE: 21.02% for AdaBoost, 10.9% for GB, 21.64% for XGB, 11.9% for RF and 11.11% for the OER model. For daily energy consumption data, the reductions were even more pronounced, with 35.47% for AdaBoost, 11.98% for GB, 16.93% for XGB, 13.86% for RF and 11.72% for OER. These results highlight the importance of data cleaning, as removing abnormal data significantly improves model accuracy and reliability, aiding in precise energy management and planning.

Figure 9 shows that the proposed OEC model consistently outperforms baseline models for hourly and daily energy consumption predictions. For hourly energy prediction, the OEC model demonstrated a notable improvement, outperforming AdaBoost by 0.043 in the R^2 score and even surpassing the strong-performing GB model by 0.002. The most significant improvement was observed compared to the RF model, with a margin of 0.105, highlighting the OEC model's superior ability to generalise and handle the complexities of hourly energy data. Similarly, the OEC model maintained its advantage for daily energy prediction, with R^2 improvements of 0.027 over AdaBoost, 0.018 over GB and 0.021 over XGB. Even against the relatively strong RF model, the OEC model achieved a higher R^2 score by 0.006, underscoring its overall precision and reliability.

These results show that the proposed OEC model consistently outperforms existing ML techniques across hourly and daily energy consumption data. Its superior predictive capabilities make it a reliable and powerful tool for energy forecasting,

TABLE 11 | Comparison with SOTA methods for SG anomaly detection and forecasting.

Study	Methodology	Performance metrics	Description
Zheng et al. [43]	SVM + RF fusion	AUC: 0.9653; TPR: 64.2%	Multimodel fusion for anomaly detection and correction in power metring data using structured SG datasets.
Turowski et al. [49]	Detection versus compensation strategies for anomaly-affected forecasting	RMSE: 0.15; MAE: 0.12	Found compensation strategy more effective than detection-only or raw-data forecasting in anomaly-prone SG time series.
Rogers [50]	Physics-aware data augmentation + fault detection via prediction models	Accuracy: +8.6% (bus); +4.19% (line)	Data augmentation improved infrastructure fault detection under limited fault sample availability.
Korba et al. [51]	Two-level anomaly detection with regression decision trees	Detection rate: 98%; false alarm rate: 1.5%	Designed cyberattack detection framework targeting overload scenarios in SG AMI using two-stage regression logic.
Sen et al. [52]	Benchmarking of IDS models for SG intrusion detection	Detection rate: 95%; false positive rate: 2%	Provided a reproducible environment to benchmark various anomaly-based IDS methods across common cyberattack patterns.
Proposed RES-AD	AutoML + OEC + OER + SMOTE + attack injection	*F*1-score: 99.5%–100%; MAPE: 11.11% (hourly), 11.72% (daily)	Unified and fully automated pipeline for anomaly detection and forecasting; includes synthetic attack injection, AutoML optimisation and class balancing.



(a) Hourly energy consumption data.

(b) Daily energy consumption data.

FIGURE 7 | Anomaly detection results of the proposed OEC and individual learners using (a) hourly and (b) daily energy consumption data.

suitable for applications where accuracy is critical. The OEC model's consistent performance enhances the efficiency of energy management and planning.

8.1 | Scalability and Real-Time Deployment Considerations

The RES-AD framework was designed with scalability and deployment practicality in mind. Using H2O AutoML enables distributed model training across multiple nodes, supporting large-scale SG infrastructures with thousands or millions of smart metres. Although model training is performed offline due to computational cost, the resulting models are lightweight and suitable for real-time inference. These models can be deployed in edge servers or cloud microservices, enabling continuous monitoring of high-frequency data. Furthermore, the modular design of the RES-AD pipeline makes it compatible with existing SCADA systems [53]. It facilitates integration with real-time data streaming platforms such as Apache Kafka or Apache Flink, which would be explored in future implementations.

8.2 | Integration Into Real-World SG Infrastructure

Integrating the RES-AD framework into existing SG infrastructures can be achieved by aligning with modular communication protocols and SCADA systems currently in operation

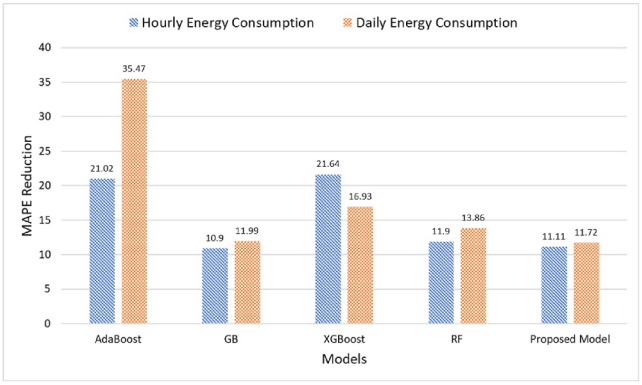


FIGURE 8 | Reduction of error in hourly and daily energy consumption prediction after removing abnormal data.

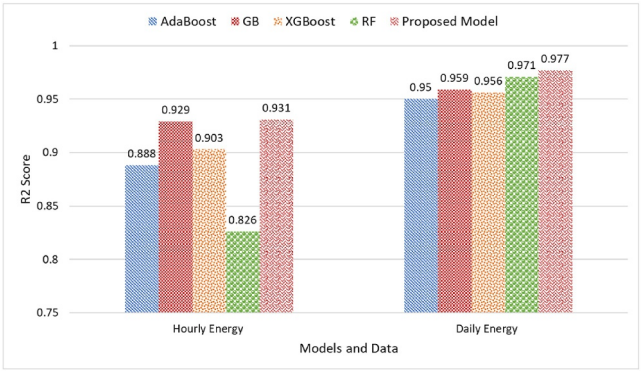


FIGURE 9 | Comparison of the proposed OEC model with the base learner in terms of R^2 score.

[51, 54]. The trained OEC and OER models are computationally lightweight and can be deployed as containerised services (e.g., using Docker) in edge servers or cloud environments, which aligns with the microservice-based architectures increasingly adopted in SG deployments [55, 56]. This enables scalable and distributed real-time anomaly detection and prediction. Nevertheless, challenges such as interoperability with legacy infrastructure, managing latency in high-frequency data streams and ensuring secure and resilient communication pipelines remain critical risks [57, 58]. Security-by-design principles should be enforced to address these issues, including data encryption, access control and tamper-resistant logging mechanisms [59]. Furthermore, RES-AD's modularity enables it to support standardised industrial interfaces (e.g., IEC 61850, MQTT) for seamless integration [60, 61]. Periodic model retraining using updated grid data is recommended to retain robustness against evolving consumption patterns and cyberattack strategies.

To illustrate how these components operate in practice, Figure 10 presents the deployment workflow of the RES-AD framework within a real-world SG environment. The diagram shows how consumption data flows from smart metres through the edge gateway and Kafka/MQTT stream layer into the EMS/SCADA server, where the trained OEC model performs real-time anomaly detection. It also highlights the alerting mechanism and the retraining loop triggered through drift detection to ensure sustained model robustness during live operation.

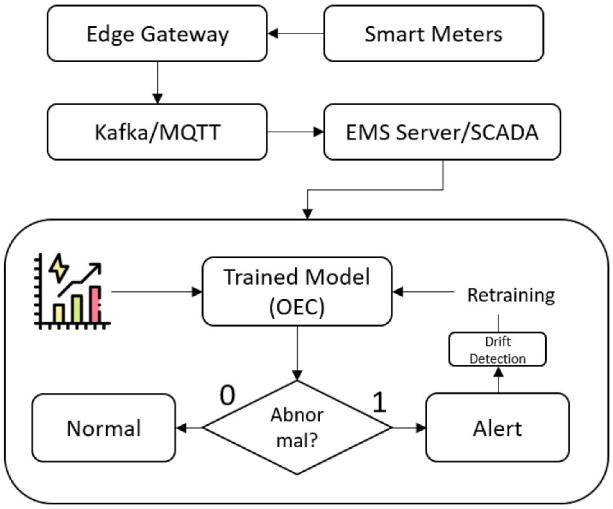


FIGURE 10 | Deployment architecture of the RES-AD framework.

8.3 | Limitations

Although the RES-AD framework demonstrates strong performance, several limitations remain. First, the evaluation relies solely on the UCI ELDi residential smart metre dataset, which captures consumption patterns from 2011 to 2014 within a single climatic region. As a result, the generalisability of the framework to commercial, industrial or climate-diverse environments has not yet been fully validated. These sectors often exhibit substantially different load dynamics commercial and industrial profiles show higher stability and predictable operational cycles, whereas residential consumption displays stronger seasonal and behavioural variability. Future work will incorporate heterogeneous datasets across multiple sectors and regions to assess cross-domain robustness.

Second, the dataset provides only hourly and daily granularities. Although the model performs effectively at these resolutions, its behaviour under coarser (monthly) or finer (sub-hourly) time scales remains unexplored. Evaluating RES-AD across multiple temporal granularities will help determine its sensitivity to aggregation effects and its suitability for real-time or long-term planning scenarios.

Finally, although this study introduces a preliminary assessment of stealthy adversarial perturbations, more sophisticated multivector cyberattacks, such as coordinated replay attacks, protocol-level exploits (e.g., IEC 61850, MQTT) or network-layer adversaries, were not included due to the lack of publicly available datasets. Extending RES-AD to incorporate protocol-aware features and more advanced adversarial defence techniques remains an important direction for future research.

9 | Conclusion

This study presented RES-AD, a robust AutoML-driven ensemble framework designed to enhance anomaly detection and energy forecasting in smart grid environments. The framework integrates an OEC for attack-aware anomaly detection and an OER for accurate forecasting, both selected and optimised using H_2O AutoML. Five realistic data integrity attacks were injected into the

originally unsupervised dataset to transform it into a supervised learning setting and rigorously evaluate resilience under malicious conditions. Experimental results on the UCI ELDi dataset demonstrate that the OEC achieves high detection accuracy for both hourly (99.7%) and daily (99.9%) consumption patterns. After anomaly removal, the OER considerably improves the forecasting performance, reducing MAPE by 11.11% and 11.72% for hourly and daily predictions, respectively. Furthermore, RES-AD maintains strong robustness against stealthy adversarial perturbations, with only a minor 1.7% degradation under FGSM attacks, an effect substantially mitigated through adversarial augmentation. Comparisons with recent deep learning models (TabNet, CNN and GNN) under identical preprocessing conditions confirm that RES-AD consistently outperforms these architectures in stability and accuracy. To support claims of deployment readiness, we also quantified inference latency, memory footprint and model size, showing that the OEC ensemble maintains real-time performance with a 23.36 ms inference time and sub-0.05 MB memory usage. A deployment architecture highlights how RES-AD integrates with Kafka/SCADA pipelines for real-time monitoring, alerting and automated retraining. Overall, the RES-AD framework offers a scalable, adversarially robust and deployment-ready solution for modern smart grids. By combining AutoML-guided model selection, ensemble learning and comprehensive attack simulation, it significantly enhances the reliability, security and forecasting stability of smart grid operations.

9.1 | Future Directions

Future research will extend the RES-AD framework to address a broader spectrum of advanced and stealthy cyberattacks that are not covered in this study. These include adversarially crafted evasion samples, replay- and time-shift-based manipulations and network-layer exploits targeting industrial protocols such as IEC 61850 and MQTT. Such attack types are not included in the current work due to the absence of communication-layer traces and protocol metadata in the employed dataset; however, their potential to exploit temporal dependencies or protocol vulnerabilities poses significant risks for smart grid security. Upcoming extensions will incorporate adversarial training strategies, protocol and communication-aware intrusion features and blended multivector attack scenarios to further enhance the resilience of RES-AD. Additionally, future work will explore deployment on heterogeneous grid environments, integration with real-time drift detection and evaluation using commercial and industrial datasets to improve generalisability and operational robustness.

Nomenclature

AD	abnormal detection
AdaBoost	adaptive boosting
ADFA-LD	Australian Defence Force Academy Linux Dataset
AMI	advanced metring infrastructure
AUC	area under the curve
AutoML	automated machine learning
DBSCAN	density-based spatial clustering of applications with noise
DC	direct current
DL	deep learning
DoS	denial of service
DST	daylight saving time
DST	daylight saving time
FDIA	false data injection attacks
FGSM	fast gradient sign method
FNR	false negative rate
FPR	false positive rate
GB	gradient boosting
GNN	graph neural network
HPO	hyperparameter optimisation
ICT	information and communication technology
IDS	intrusion detection system
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IoT	internet of things
KNN	K-nearest neighbours
kW	kilowatts
kWh	kilowatt-hours
LOF	local outlier factor
LSTM	long short-term memory
MAE	mean absolute error
MAPE	mean absolute percentage error
ML	machine learning
MLAD	machine learning-based anomaly detection
MQTT	message queueing telemetry transport
MSE	mean squared error
MWh	megawatt-hours
OEC	optimal ensemble classifier
OER	optimal ensemble regressor
OPTICS	ordering points to identify the clustering structure
PCA	principal component analysis
POMDP	partially observable Markov decision process
R^2	coefficient of determination
RES	robust ensemble strategy
RF	random forest
RL	reinforcement learning
RMSE	root mean squared error
RNN	recurrent neural network
SMOTE	synthetic minority oversampling technique
SOTA	state-of-the-art

SVM	support vector machine
TPR	true positive rate
UCI	University of California, Irvine
XGB	extreme gradient boosting

Acknowledgements

This research was supported by the Regional Innovation System & Education (RISE) program through the Jeju RISE centre, funded by the Ministry of Education (MOE) and the Jeju Special Self-Governing Province, Republic of Korea, (2026-RISE-17-001), This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT), (No. RS-2024-00405278). Any correspondence related to this paper should be addressed to Do-Hyeun Kim.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Available on request.

References

1. X. Fang, S. Misra, G. Xue, and D. Yang, "Smart Grid—The New and Improved Power Grid: A Survey," *IEEE Communications Surveys and Tutorials* 14, no. 4 (2011): 944–980, https://doi.org/10.1109/surv.2011.101911.00087.

2. H. Karimipour and V. Dinavahi, "Robust Massively Parallel Dynamic State Estimation of Power Systems Against Cyber-Attack," *IEEE Access* 6 (2017): 2984–2995, https://doi.org/10.1109/access.2017.2786584.

3. Z. Huang, C. Wang, T. Zhu, and A. Nayak, "Cascading Failures in Smart Grid: Joint Effect of Load Propagation and Interdependence," *IEEE Access* 3 (2015): 2520–2530, https://doi.org/10.1109/access.2015.2506503.

4. M. X. Cheng, M. Crow, and Q. Ye, "A Game Theory Approach to Vulnerability Analysis: Integrating Power Flows With Topological Analysis," *International Journal of Electrical Power & Energy Systems* 82 (2016): 29–36, https://doi.org/10.1016/j.ijepes.2016.02.045.

5. H. Karimipour and V. Dinavahi, "On False Data Injection Attack Against Dynamic State Estimation on Smart Power Grids," in *2017 IEEE International Conference on Smart Energy Grid Engineering (SEGE)* (IEEE, 2017), 388–393.

6. M. Fahim and A. Sillitti, "Anomaly Detection, Analysis and Prediction Techniques in IoT Environment: A Systematic Literature Review," *IEEE Access* 7 (2019): 81664–81681, https://doi.org/10.1109/access.2019.2921912.

7. A. G. Miles, Y. Chakhchoukh, and S. S. Alam, "Robust Distribution State Estimation for Reliable Locational Marginal Pricing Under Cyber-Attacks," *IEEE Transactions on Power Systems* 39, no. 3 (2023): 4963–4974, https://doi.org/10.1109/tpwrs.2023.3325219.

8. A. D. Syrmakesis, H. H. Alhelou, and N. D. Hatziargyriou, "A Novel Cyberattack-Resilient Frequency Control Method for Interconnected Power Systems Using SMO-Based Attack Estimation," *IEEE Transactions on Power Systems* 39, no. 4 (2023): 5672–5686, https://doi.org/10.1109/tpwrs.2023.3340744.

9. S. Yang, K.-W. Lao, H. Hui, and Y. Chen, "Secure Distributed Control for Demand Response in Power Systems Against Deception Cyber-Attacks With Arbitrary Patterns," *IEEE Transactions on Power Systems* 39, no. 6 (2024): 7277–7290.

10. B. R. Amin, S. Taghizadeh, M. S. Rahman, M. J. Hossain, V. Varadharajan, and Z. Chen, "Cyber Attacks in Smart Grid–Dynamic Impacts, Analyses and Recommendations," *IET Cyber-Physical Systems: Theory & Applications* 5, no. 4 (2020): 321–329, https://doi.org/10.1049/iet-cps.2019.0103.

11. R. Xu, R. Wang, Z. Guan, L. Wu, J. Wu, and X. Du, "Achieving Efficient Detection Against False Data Injection Attacks in Smart Grid," *IEEE Access* 5 (2017): 13787–13798, https://doi.org/10.1109/access.2017.2728681.

12. Y. Li, X. Wei, Y. Li, Z. Dong, and M. Shahidehpour, "Detection of False Data Injection Attacks in Smart Grid: A Secure Federated Deep Learning Approach," *IEEE Transactions on Smart Grid* 13, no. 6 (2022): 4862–4872, https://doi.org/10.1109/tsg.2022.3204796.

13. A. Huseinović, S. Mrdović, K. Bicakci, and S. Uludag, "A Survey of Denial-of-Service Attacks and Solutions in the Smart Grid," *IEEE Access* 8 (2020): 177447–177470, https://doi.org/10.1109/access.2020.3026923.

14. M. K. Hasan, A. A. Habib, S. Islam, N. Safie, S. N. H. S. Abdullah, and B. Pandey, "Ddos: Distributed Denial of Service Attack in Communication Standard Vulnerabilities in Smart Grid Applications and Cyber Security With Recent Developments," *Energy Reports* 9 (2023): 1318–1326, https://doi.org/10.1016/j.egyr.2023.05.184.

15. İ. Avcı, "Investigation of Cyber-Attack Methods and Measures in Smart Grids," *Sakarya University Journal of Science* 25, no. 4 (2021): 1049–1060, https://doi.org/10.16984/saufenbilder.955914.

16. K. N. Qureshi, R. Hussain, and G. Jeon, "A Distributed Software Defined Networking Model to Improve the Scalability and Quality of Services for Flexible Green Energy Internet for Smart Grid Systems," *Computers & Electrical Engineering* 84 (2020): 106634, https://doi.org/10.1016/j.compeleceng.2020.106634.

17. R. Younisse and M. AlKasassbeh, "Sgid: A Semi-Synthetic Dataset for Injection Attacks in Smart Grid Systems," in *2024 15th International Conference on Information and Communication Systems (ICICS)* (IEEE, 2024), 1–4.

18. Z. Fengming, L. Shufang, G. Zhimin, W. Bo, T. Shiming, and P. Mingming, "Anomaly Detection in Smart Grid Based on Encoder-Decoder Framework With Recurrent Neural Network," *Journal of China Universities of Posts and Telecommunications* 24, no. 6 (2017): 67–73, https://doi.org/10.1016/s1005-8885(17)60243-7.

19. B. Rossi, S. Chren, B. Buhnova, and T. Pitner, "Anomaly Detection in Smart Grid Data: An Experience Report," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (IEEE, 2016), 2313–2318.

20. M. A. Khan, N. Iqbal, H. Jamil, D.-H. Kim, and D. H. Kim, "An Optimized Ensemble Prediction Model Using Automl Based on Soft Voting Classifier for Network Intrusion Detection," *Journal of Network and Computer Applications* 212 (2023): 103560, https://doi.org/10.1016/j.jnca.2022.103560.

21. G. Fenza, M. Gallo, and V. Loia, "Drift-Aware Methodology for Anomaly Detection in Smart Grid," *IEEE Access* 7 (2019): 9645–9657, https://doi.org/10.1109/access.2019.2891315.

22. X. He, Q. Ai, R. C. Qiu, W. Huang, L. Piao, and H. Liu, "A Big Data Architecture Design for Smart Grids Based on Random Matrix Theory," *IEEE Transactions on Smart Grid* 8, no. 2 (2015): 674–686, https://doi.org/10.1109/TSG.2015.2445828.

23. R. Moslemi, M. Davoodi, and J. M. Velni, "A Distributed Approach for Estimation of Information Matrix in Smart Grids and Its Application for Anomaly Detection," in *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)* (IEEE, 2020), 1–7.

24. R. Moghaddass and J. Wang, "A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data," *IEEE*

Transactions on Smart Grid 9, no. 6 (2017): 5820–5830, https://doi.org/10.1109/tsg.2017.2697440.

25. Y. Liu, R. Yuan, S. Zheng, K. Yan, and H. Miao, "An Abnormal Detection of Positive Active Total Power Based on Local Outlier Factor," in *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)* (IEEE, 2021), 180–183.

26. D. M. Menon and N. Radhika, "Anomaly Detection in Smart Grid Traffic Data for Home Area Network," in *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)* (IEEE, 2016), 1–4.

27. S. S. Rawat, V. A. Polavarapu, V. Kumar, E. Aruna, and V. Sumathi, "Anomaly Detection in Smart Grid Using Rough Set Theory and K Cross Validation," in *2014 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2014]* (IEEE, 2014), 479–483.

28. F. Fathnia and M. H. J. D. Bayaz, "Anomaly Detection in Smart Grid With Help of an Improved Optics Using Coefficient of Variation," in *Electrical Engineering (ICEE), Iranian Conference on* (IEEE, 2018), 1044–1050.

29. M. Panthi, "Anomaly Detection in Smart Grids Using Machine Learning Techniques," in *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)* (IEEE, 2020), 220–222.

30. C. Promper, D. Engel, and R. C. Green, "Anomaly Detection in Smart Grids With Imbalanced Data Methods," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (IEEE, 2017), 1–8.

31. H. Karimipour, S. Geris, A. Dehghantanha, and H. Leung, "Intelligent Anomaly Detection for large-scale Smart Grids," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)* (IEEE, 2019), 1–4.

32. T. Kotsiopoulos, P. Sarigiannidis, D. Ioannidis, and D. Tzovaras, "Machine Learning and Deep Learning in Smart Manufacturing: The Smart Grid Paradigm," *Computer Science Review* 40 (2021): 100341, https://doi.org/10.1016/j.cosrev.2020.100341.

33. M. Cui, J. Wang, and M. Yue, "Machine Learning-Based Anomaly Detection for Load Forecasting Under Cyberattacks," *IEEE Transactions on Smart Grid* 10, no. 5 (2019): 5724–5734, https://doi.org/10.1109/tsg.2018.2890809.

34. M. N. Kurt, O. Ogundijo, C. Li, and X. Wang, "Online Cyber-Attack Detection in Smart Grid: A Reinforcement Learning Approach," *IEEE Transactions on Smart Grid* 10, no. 5 (2018): 5174–5185, https://doi.org/10.1109/tsg.2018.2878570.

35. S. N. Islam, Z. Baig, and S. Zeadally, "Physical Layer Security for the Smart Grid: Vulnerabilities, Threats, and Countermeasures," *IEEE Transactions on Industrial Informatics* 15, no. 12 (2019): 6522–6530, https://doi.org/10.1109/tii.2019.2931436.

36. C. Lopez, A. Sargolzaei, H. Santana, and C. Huerta, "Smart Grid Cyber Security: An Overview of Threats and Countermeasures," *Journal of Energy and Power Engineering* 9, no. 7 (2015): 632–647, https://doi.org/10.1109/ISAP63260.2024.10744349.

37. S. Shapsough, F. Qatan, R. Aburukba, F. Aloul, and A. Al Ali, "Smart Grid Cyber Security: Challenges and Solutions," in *2015 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)* (IEEE, 2015), 170–175.

38. F. Passerini and A. M. Tonello, "Smart Grid Monitoring Using Power Line Modems: Anomaly Detection and Localization," *IEEE Transactions on Smart Grid* 10, no. 6 (2019): 6178–6186, https://doi.org/10.1109/tsg.2019.2899264.

39. M. Ghiasi, N. Ghadimi, and E. Ahmadinia, "An Analytical Methodology for Reliability Assessment and Failure Analysis in Distributed Power System," *SN Applied Sciences* 1, no. 1 (2019): 44, https://doi.org/10.1007/s42452-018-0049-0.

40. M. Ghiasi, M. Dehghani, T. Niknam, A. Kavousi-Fard, P. Siano, and H. H. Alhelou, "Cyber-Attack Detection and Cyber-Security Enhancement in Smart DC-Microgrid Based on Blockchain Technology and Hilbert Huang Transform," *IEEE Access* 9 (2021): 29429–29440, https://doi.org/10.1109/access.2021.3059042.

41. M. Dehghani, T. Niknam, M. Ghiasi, N. Bayati, and M. Savaghebi, "Cyber-Attack Detection in DC Microgrids Based on Deep Machine Learning and Wavelet Singular Values Approach," *Electronics* 10, no. 16 (2021): 1914, https://doi.org/10.3390/electronics10161914.

42. E. Altamimi, A. K. Al-Ali, A. Al-Ali, and Q. Malluhi, "Distribution-Based Residual Analysis of Load Forecasting for Enhanced Anomaly Detection," *SSRN Electronic Journal* (2025), https://ssrn.com/abstract=5230288.

43. S. Zheng, M. Zhu, and Y. Liu, "Research on Anomaly Detection and Correction of Power Metering Data Based on Machine Learning Algorithm," *Science and Technology for Energy Transition* 2025, no. 1 (2025): 25–35, https://doi.org/10.2516/stet/2024106.

44. P. K. R. Shabad, A. Alrashide, and O. Mohammed, "Anomaly Detection in Smart Grids Using Machine Learning," *ResearchGate* (2023), https://www.researchgate.net/publication/356141365.

45. A. Trindade, "Electricity Load Diagrams 2011–2014," *UCI Machine Learning Repository* (2015), https://doi.org/10.24432/C58C86.

46. L. Che, X. Liu, Z. Shuai, and J. Zhao, "The Impact of Ramp-Induced Data Attacks on Power System Operational Security," *IEEE Transactions on Industrial Informatics* 15, no. 9 (2019): 5064–5075, https://doi.org/10.1109/tii.2019.2895058.

47. K. Sun, I. Esnaola, S. M. Perlaza, and H. V. Poor, "Stealth Attacks on the Smart Grid," *IEEE Transactions on Smart Grid* 11, no. 2 (2019): 1276–1285, https://doi.org/10.1109/tsg.2019.2935353.

48. P. Wang and M. Govindarasu, "Multi-Agent Based Attack-Resilient System Integrity Protection for Smart Grid," *IEEE Transactions on Smart Grid* 11, no. 4 (2020): 3447–3456, https://doi.org/10.1109/tsg.2020.2970755.

49. M. Turowski, O. Neumann, L. Mannsperger, et al., "Managing Anomalies in Energy Time Series for Automated Forecasting," in *Energy Informatics Academy Conference* (Springer, 2023), 3–29, https://doi.org/10.1007/978-3-031-48649-4_1.

50. J. Rogers, *Data Augmentation for Smart Grid Anomaly Detection* (University of Rhode Island, 2024), https://doi.org/10.23860/diss-1659.

51. A. A. Korba, N. Tamani, Y. Ghamri-Doudane, and N. E. I. Karabadji, "Anomaly-Based Framework for Detecting Power Overloading Cyberattacks in Smart Grid AMI," *Computers & Security* 96 (2020): 101896, https://doi.org/10.1016/j.cose.2020.101896.

52. Ö. Sen, S. Glomb, M. Henze, and A. Ulbig, "Benchmark Evaluation of Anomaly-Based Intrusion Detection Systems in the Context of Smart Grids," in *2023 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)* (IEEE, 2023), 1–6, https://doi.org/10.1109/ISGTEUROPE56780.2023.10407262.

53. C. A. Fonseca, "Scada System of Pipelines," in *Handbook of Pipeline Engineering* (Springer, 2024), 1–28.

54. M. A. Mohamed, T. Chen, W. Su, and T. Jin, "Proactive Resilience of Power Systems Against Natural Disasters: A Literature Review," in *IEEE Access*, vol. 7 (IEEE, 2019), 163778–163795, https://doi.org/10.1109/ACCESS.2019.2952362.

55. A. Wali and F. Alshehry, "A Survey of Security Challenges in Cloud-Based SCADA Systems," *Computers* 13, no. 4 (2024): 97, https://doi.org/10.3390/computers13040097.

56. N. S. Nafi, K. Ahmed, M. A. Gregory, and M. Datta, "A Survey of Smart Grid Architectures, Applications, Benefits and Standardization,"

Journal of Network and Computer Applications 76 (2016): 23–36, https://doi.org/10.1016/j.jnca.2016.10.003.

57. B. N. Jørgensen and Z. G. Ma, "Cybersecurity and Resilience of Smart Grids: A Review of Threat Landscape, Incidents, and Emerging Solutions," *Applied Sciences* 16, no. 2 (2026): 981, https://doi.org/10.3390/app16020981.

58. W. Kang, L. Zhang, Z. Hu, and Z. Xia, "A Secure and Efficient Data Aggregation Scheme for Cloud–Edge Collaborative Smart Meters," *International Journal of Electrical Power & Energy Systems* 162 (2024): 110270, https://doi.org/10.1016/j.ijepes.2024.110270.

59. G. Areo, Blockchain for Secure Call Logging in Asterisk-Based Voip Systems (2024).

60. H.-J. Jun and H.-S. Yang, "Performance of the Xmpp and the Mqtt Protocols on IEC 61850-Based Micro Grid Communication Architecture," *Energies* 14, no. 16 (2021): 5024, https://doi.org/10.3390/en14165024.

61. L. Tightiz and H. Yang, "A Comprehensive Review on IoT Protocols' Features in Smart Grid Communication," *Energies* 13, no. 11 (2020): 2762, https://doi.org/10.3390/en13112762.