EAI Endorsed Transactions
on Energy Web

Research Article **EAI.EU**

Adaptive Calibration Framework for Intelligent Meter Error Prediction Based on Machine Learning

Chunguang Wang*

Marketing Service Center of State Grid Fujian Electric Power Co., Ltd., Fuzhou, Fujian, 350013, China

Abstract

INTRODUCTION: Accurate measurement in intelligent metering systems is crucial for reliable energy monitoring, billing, and smart grid management. However, intelligent power meters often experience measurement inaccuracies due to environmental disturbances, hardware aging, and insufficient calibration mechanisms. These issues may lead to systematic errors and unreliable data in energy management systems. Therefore, improving the reliability and accuracy of intelligent metering devices through advanced data-driven techniques has become an important research focus.
OBJECTIVES: The main objective of this study is to develop an Adaptive Calibration Framework for Intelligent Meter Error Prediction using Machine Learning. The framework aims to detect abnormal data, predict measurement errors, and improve the accuracy and reliability of intelligent metering systems operating under diverse environmental conditions.
METHODS: The proposed framework integrates dynamic error detection, outlier filtering, and predictive modeling. Data collected from intelligent meters are first normalized using Min–Max scaling to ensure consistent training. Robust Principal Component Analysis (RPCA) with an improved distance function and adaptive thresholding based on the box plot method is used to detect and remove anomalous data. Subsequently, an Adaptive Cockroach Swarm Optimized Weighted Isolation Forest (ACSO-WIForest) model is employed to predict measurement errors by learning complex relationships between environmental stress factors and device performance.
RESULTS: The experimental evaluation of the proposed ACSO-WIForest model using datasets from intelligent meters in harsh environmental conditions shows its effectiveness. The model demonstrates strong predictive performance with a Mean Absolute Error (MAE) of 0.00156, Root Mean Square Error (RMSE) of 0.0040, and a low Mean Absolute Percentage Error (MAPE) of 0.221, indicating high prediction precision. With a coefficient of determination (R^2) of 0.998, the model achieves excellent fit and reliability, significantly outperforming existing methods in intelligent meter error prediction.
CONCLUSION: The results indicate that machine learning–based adaptive calibration can significantly improve the accuracy, reliability, and self-correcting capability of intelligent metering systems. This approach provides a practical solution for enhancing trustworthy energy monitoring and management infrastructures

Keywords: Adaptive Calibration, Machine Learning–Based Calibration, Intelligent Meter Error Prediction, Smart Metering Systems, ACSO-WIForest, Anomaly Detection, Robust Principal Component Analysis, Swarm Optimization

Received on 10 March 2026, accepted on 07 April 2026, published on 05 May 2026

doi: 10.4108/ew.12150

*Corresponding author. Email: chunguang_wang78@outlook.com, lunwen_zy@163.com

1. Introduction

An intelligent meter is a high-tech instrument that quantifies, monitors, and transmits energy consumption with great precision, allowing effective billing and management. Measurement accuracy in intelligent metering systems is essential to provide effective energy management, accurate billing, and integration of renewable energy sources into the modern power grid [1]. Smart meters are the central component of smart energy infrastructure, as they enable real-time monitoring of power consumption and furnish the utility organization and its consumers with real-time data. The reliability of these systems is compromised more often than not due to environmental interferences, hardware deterioration, aging of sensors, and inadequate calibration processes, which together lead to inaccuracy in measurements [2,3]. Conventional calibration techniques based mostly on static models, manual adjustment, or occasional checking are typically unsuitable in coping with the dynamic and heterogeneous conditions common in present-day energy networks. Such traditional techniques cannot typically accommodate quick environmental changes, leading to systemic errors and error propagation in energy management networks and loss of overall operational efficiency [4,5].

Intelligent metering systems have new opportunities to become more accurate and resistant with the emergence of machine learning (ML). The ML algorithms are capable of handling large volumes of data produced by meters, identifying patterns, discovering anomalies, and predicting measurement errors, so corrections to errors can be made proactively[6]. Advanced approaches, including deep learning and ensemble models, have been applied for real-time error mitigation, offering significant advantages over traditional methods. These algorithms capture nonlinear relationships between environmental variables such as temperature, humidity, and voltage fluctuations and meter readings, which is crucial in complex power systems where multiple factors simultaneously affect measurement accuracy. Using predictive ML features, smart meters are capable of autonomously correcting readings, enhancing their accuracy and stability in challenging operating environments [7,8].

Accurate calibration of intelligent meters remains a significant challenge in modern smart grid infrastructures due to the combined effects of sensor drift, environmental variability, and inherent measurement noise. Over time, meter components degrade, leading to systematic errors, while dynamic environmental conditions such as temperature, humidity, and voltage fluctuations introduce additional uncertainty in measurements. These challenges are further amplified in large-scale smart grid deployments, where massive volumes of heterogeneous data increase the complexity of maintaining consistent accuracy. Inaccurate meter readings can result in billing discrepancies, inefficient energy management, and reduced reliability of grid operations.

Notwithstanding these accomplishments, machine learning application in meter error prediction has several challenges. Power system data is usually large, high-dimensional, diverse, and noisy, making it difficult to build robust models. Additionally, most previous researches do not properly account for environmental factors affecting the accuracy of measurements, reducing the potential for real-world applications of these models. Severe environmental fluctuations or sensor malfunctions may lead to the deterioration of predictive performance of conventional ML models, which is why mechanisms that can be autonomously recalibrated in real-time are necessary [9,10]. Recent studies have highlighted the growing role of machine learning in intelligent meter calibration and error prediction. Advanced models such as attention-based LSTM and hybrid optimization techniques have demonstrated improved accuracy by capturing nonlinear relationships between environmental factors and meter performance. These approaches emphasize the importance of adaptive, data-driven calibration frameworks for enhancing reliability and robustness in modern smart metering systems.

Machine learning-based models, although effective, often struggle with noisy and high-dimensional data, limited generalization across diverse operating conditions, and insufficient integration of adaptive optimization mechanisms. Moreover, many existing methods do not adequately prioritize critical features influencing calibration accuracy, leading to reduced prediction performance and increased false detections. These limitations highlight the need for a robust, adaptive, and feature-aware calibration framework. To address this gap, the proposed ACSO-WIForest model integrates adaptive swarm optimization with weighted anomaly detection, enabling efficient parameter tuning, improved feature selection, and accurate intelligent meter error prediction under varying environmental conditions.

To overcome these issues, this research introduces an Adaptive Calibration Framework that combines dynamic error detection, outlier rejection, and predictive modeling to improve the accuracy of meters. The framework utilizes Robust Principal Component Analysis (RPCA) with an optimized distance function and adaptive thresholding to identify and eliminate irregular data due to extreme weather conditions. In addition, an Adaptive Cockroach Swarm Optimized Weighted Isolation Forest (ACSO-WIForest) model is employed to forecast meter measurement errors based on detecting sophisticated interactions between environmental stressors and metering performance. Through this combination of techniques, a self-correcting, smart calibration system is achieved that guarantees accurate and reliable energy monitoring under adverse and variable operating conditions.

1.1 Significant contribution

- This research aims to develop an adaptive machine learning framework that predicts and corrects intelligent meter errors for reliable energy management.
- Gathered intelligent meter datasets from the Kaggle open-source platform, covering diverse environmental and device measurement scenarios.
- Handled missing or incomplete data entries and applied Min-Max scaling, normalizing all features for consistent model training.
- Developed a dynamic framework integrating error detection, outlier filtering, and predictive modeling for meters.
- RPCA with optimized distance function and adaptive thresholding effectively removes extreme measurement anomalies.
- The model integrates ACSO with WIForest, optimizing anomaly detection and error prediction for intelligent meter calibration.
- Introduced the ACSO-WIForest model, learning environmental-meters relationships, enhancing error prediction accuracy and robustness.
- Validated on harsh-environment meter data, demonstrating superior MAPE, RMSE, R^2 and MAE compared to existing methods.
- Ensures reliable metering under sparse datasets while enhancing accuracy, self-correction, and infrastructure robustness.

The research is structured as follows: **Section 1** discusses the background of the research and identifies the main challenges. **Section 2** reviews related work and limitations. **Section 3** presents the proposed Adaptive Calibration Framework with ACSO-WIForest. **Section 4** explains evaluation metrics and experimental validation. **Section 5** concludes with findings and future directions.

2. Previous Work

The measurement error (ME) of the smart meter (SM), which was used to improve prediction accuracy across varying regions, was analyzed [11]. Environmental stresses influencing ME were identified using Pearson correlation and least squares methods, and outliers were removed. A Harris Hawks Optimization–Bidirectional Long Short-Term Memory (HBO-BiLSTM) model predicted ME effectively, outperforming conventional machine learning methods with high-altitude field data. Results demonstrated strong predictive ability, though generalization to all regions may be limited. Accurate prediction of measurement errors for smart electricity meters (SEMs) in extreme cold environments was addressed [12] to enhance energy efficiency by using Improved Kernel Support Vector Regression (IKSVR) model. Parameters were optimized. Results on high-altitude datasets showed superior performance, though applicability to other extreme environments remained limited.

Existing machine learning approaches for intelligent meter error prediction, such as support vector regression, neural networks, and recurrent models, primarily focus on capturing temporal patterns or nonlinear relationships between environmental factors and meter readings. However, these methods often lack robustness in handling noisy, high-dimensional data and may require extensive parameter tuning. Optimization-based and hybrid models improve prediction accuracy but typically do not incorporate adaptive mechanisms for real-time calibration or effective feature prioritization. In contrast, the proposed ACSO-WIForest model introduces a hybrid framework that integrates adaptive swarm optimization with a feature-weighted anomaly detection approach.

The impact of outliers and small sample sizes on SM failure rate prediction was addressed [13] to enhance reliability analysis. SM errors were predicted, optimized by an adaptive Gauss genetic algorithm (AGG), and used in a proportional FR (PFR) model to assess batch reliability. Results on datasets from four companies showed improved performance, though generalization to larger populations may be limited. Reliability assessment of electric metering devices under harsh environments was suggested for predictive maintenance support [14]. Key environmental factors were extracted and fused, while outliers were detected using a bi-directional method. Multi-stress fusion nonlinear degradation (MFND) predicted measurement errors considering cyclic variations and environmental impacts. High-dry-hot area datasets showed strong performance, though applicability to other regions remained uncertain. To improve error prediction in SM energy management under extreme environments, an accurate prediction of SM error was investigated [15]. A multisource fusion framework using extracted temporal, temperature, and humidity features was enhanced. Experiments on high dry-heat datasets achieved a low Root Mean Squared value (RMSE) (0.0311), outperforming advanced models, though scalability and validation across diverse environments remained limited.

Degradation of SM accuracy from aging and faults was detected [16] through a measurement error estimation method using a modified neural network. A relationship model linking correction coefficient, network loss, and energy consumption was established, and improved activation and iteration strategies enhanced parameter estimation. Case studies verified effectiveness, though broader validation and real-time adaptability remained limited. Measurement error evaluation of power metering equipment (PME) under extreme stresses was examined [17] to enhance design and accuracy. Outliers were identified using an improved local outlier factor (ILOF) with optimized distance and adaptive threshold methods. Kernel support vector regression (KSVR) then fused errors with stress features. Experiments showed superior performance under small samples, though broader validation across diverse conditions was limited.

The influence of environmental conditions on electric metering device accuracy was analyzed [18] to develop an effective correction model. Significant environmental factors were selected using mutual information, and a Gated Recurrent Unit (GRU)-attention model with particle swarm optimization tuned hyperparameters for error adjustment. Experimental validation across seasons achieved strong performance with RMSE as low as 1.24, though applicability to other extreme regions required further research. In large-scale IoT networks, a flexible smart-metering architecture for unified device monitoring across diverse technologies was analyzed [19]. The system enabled reliable telemetry, remote management, and multilayer security. Experiments demonstrated effective interoperability over Zigbee, Wi-Fi and Sigfox. However, the approach was limited by potential network congestion and scalability constraints in extremely dense deployments.

A scalable Big Data framework was developed [20] to enhance short-term load forecasting (STLF) for residential buildings. Data from SM and weather sensors was analyzed using multiple ML algorithms. Real-world validation showed improved forecasting accuracy with automated best-algorithm selection, though scalability and performance in larger heterogeneous grids required further investigation. The impact of environmental conditions on power metering device accuracy was evaluated [21] to improve calibration. Influential factors were identified using mutual information (MI), and a symmetry attention LSTM model was applied to correct measurement errors. Validation with error prediction showed over 10% error reduction in spring and consistent improvements across seasons, though generalization to broader environments remained limited.

Measurement error assessment for SEMs under extreme natural stresses was conducted [22] to improve equipment quality and reduce grid costs. Outliers were identified using optimized kernel density estimation (OKDE), while a modified double-kernel support vector regression (MKSVR) fused measurement error with multiple stress features. High dry-heat region datasets showed improved evaluation performance, though results may be limited under larger or more diverse samples. SM error detection in low-voltage energy systems was investigated [23] to improve accuracy and efficiency. Abnormal data were filtered using orthogonal matching pursuit and Bollinger Bands, while a recursive model estimated meter errors through linear equations. Experiments showed high accuracy and low loss. Despite strong results, performance under larger, more complex systems required further validation. Improving dynamic error prediction by analyzing the quantitative relationship between grid parameters and gateway energy meter errors was examined [24]. Thirusubramaniyan presented a machine learning-based approach for anomaly detection in IoT-enabled financial systems to identify irregular patterns effectively. The proposed ACSO-WIForest framework adopts this strategy for smart meter error detection, demonstrating improved accuracy, robustness, and reduced false positives [25]. A Nonlinear AutoRegressive model with eXogenous inputs (NARX)-based nonlinear modeling approach was established, and the Hammerstein–Wiener estimator achieved the highest fitting accuracy. Results showed about 81% prediction accuracy under varying loads, though further validation across wider operating conditions remained necessary. The work applies an improved particle swarm optimization approach for efficient scheduling in microgrid systems, enhancing resource optimization. The proposed ACSO-WIForest framework adopts a similar optimization strategy, demonstrating improved parameter tuning, convergence efficiency, and prediction performance [26].

3. Proposed Method

To design an adaptive ML framework that accurately predicts and self-corrects intelligent meter errors, ensuring reliable energy management, is the objective. The methodology focuses on achieving the objective by integrating data collection from intelligent meters and Kaggle sources, preprocessing using Min-Max normalization for consistency, anomaly detection via RPCA with adaptive thresholding, and the proposed ACSO-WIForest model for accurate error prediction through adaptive calibration and multi-source data fusion. The methodology framework is illustrated in Figure 1.

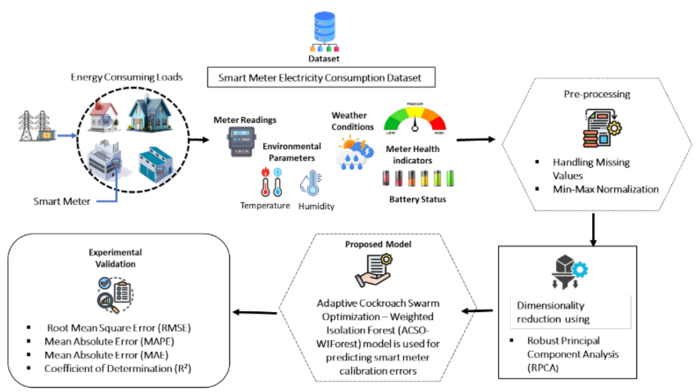

Figure 1. Proposed Adaptive Calibration Framework for Intelligent Meter Error Prediction using ACSO-WIForest

The overall system workflow of the proposed adaptive calibration framework follows a structured sequence of data processing, optimization, and prediction stages. Initially, raw intelligent meter data are collected from real-time sources and benchmark datasets. In the preprocessing stage, missing values are handled, and Min–Max normalization is applied to ensure consistent feature scaling. Subsequently, Robust Principal Component Analysis is employed to decompose the data into low-rank and sparse components, effectively removing noise and anomalous measurements.

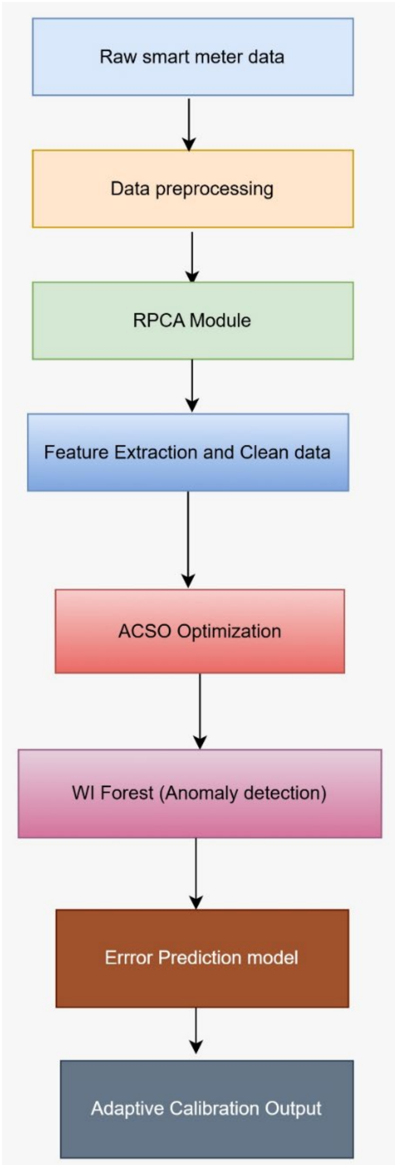

Figure 2. Flowchart of the Proposed Adaptive Calibration Framework Showing Data Preprocessing, RPCA-Based Anomaly Detection, and ACSO-WIForest Prediction Modules

A flowchart of the proposed adaptive calibration framework is illustrated in Figure 2. The diagram provides a clear representation of the interaction between preprocessing, RPCA-based anomaly detection, and the ACSO-WIForest prediction modules, enabling better understanding of the overall system architecture.

3.1 Smart Meter Consumption and Environmental Data Collection

Smart Meter (SM) Electricity Consumption Dataset is gathered from Kaggle https://www.kaggle.com/datasets/ziya07/smart-meter-electricity-consumption-dataset. It provides 30-minute interval SM electricity consumption data enriched with weather conditions and anomaly labels. It includes features such as electricity consumed (kWh), temperature, humidity, wind speed, rolling average consumption, and anomaly detection results from Isolation Forest. Designed for anomaly detection, predictive modeling, energy efficiency analysis, and smart grid management, it supports real-time forecasting and optimization, making it valuable for machine learning researchers, data scientists, and energy analysts. The dataset used in this study is obtained from the Kaggle open-source platform and consists of smart meter electricity consumption records integrated with environmental parameters. The dataset contains time-series measurements collected at 30-minute intervals, capturing variations in electricity consumption along with influencing factors such as temperature, humidity, and wind speed. It comprises a large number of samples representing diverse operating conditions, including normal and anomalous scenarios. The data reflect metering environments where fluctuations arise due to environmental stress and load variability. For model development, the dataset is divided into training and testing subsets to ensure reliable evaluation of predictive performance. This structured dataset enables effective learning of nonlinear relationships between environmental factors and meter errors, supporting robust calibration and anomaly detection in intelligent metering systems.

3.2 Preprocessing

Preprocessing is the step of preparing raw data for analysis. This research utilizes Min-Max normalization for scaling values uniformly and handling missing values to ensure consistent, reliable, and unbiased model training.

Noise and abnormal consumption values are handled using RPCA, which separates normal patterns from sparse anomalies. Missing readings are treated based on their type: MCAR values are removed, while MAR and MNAR values are imputed using statistical mean and feature-based estimation. Abnormal values are further filtered using adaptive thresholding based on box plot analysis to eliminate extreme deviations. Finally, all features are normalized using Min–Max scaling to ensure uniform data distribution and improve model convergence. These preprocessing steps collectively enhance data reliability and support accurate intelligent meter error prediction.

The data preprocessing pipeline is designed to ensure high-quality input for reliable model training. Initially, raw smart meter data is normalized using Min–Max scaling to transform all features into a uniform range, eliminating scale imbalance and improving convergence. Missing values are handled using statistical imputation methods based on data characteristics to maintain dataset completeness.

3.2.1 Min–max scaling

In this research, normalization scales input features such as voltage, current, power factor, and load variations into a uniform 0,1 range using the Min–Max transformation. This step eliminates scale disparities among attributes, ensuring balanced contribution of each feature to the learning process. It enhances the convergence of the proposed ACSO–WIForest model and improves meter error prediction accuracy. The mathematical representation is given in equation (1).

$$D_{norm} = \frac{D-D}{D_{max}-D_{max}} \tag{1}$$

Where (D) is the unique data value, (D_{min}) is the minimum value, (D_{max}) is the maximum value, (D_{norm}) is the normalized value (between 0 and 1).

The result of using this normalization ensures balanced, noise-free data, enabling faster training, robust predictions, and reliable meter calibration.

3.2.2 Handling missing values

Handling missing values involves detecting and imputing absent data for dataset completeness and accuracy. In intelligent metering, missingness arises from sensor failures, communication errors, or disturbances. Missing Completely at Random (MCAR) assumes randomness, Missing at Random (MAR) imputes from observed correlations, and Missing Not at Random (MNAR) depends on hidden factors, as described in equations (2-4). Proper handling ensures robust preprocessing, reliable calibration, and accurate meter error prediction.

MCAR

$$S(W|Z,\emptyset) = S(W,\emptyset) \tag{2}$$

Where (S) is the conditional probability of the target variable (W) (e.g., a meter reading) (Z) is the set of observed features in the dataset (e.g., environmental factors like temperature, humidity, voltage). and (W) is the dependent variable or target value (e.g., meter reading or label), $(\emptyset)$ (empty set) represents missing values in the dataset, and $S\ (W,\emptyset)$ is the overall probability of (w) and missing values, independent of (z).

MAR

$$S(W|Z,\emptyset) = S(W_c \mid \emptyset) \tag{3}$$

(W_c) is the subset of observed dependent variable values related to missing data. The probability of missingness, $S\ (W_c,\emptyset)$ is the probability of missingness depends only on the observed part (W_c) of the data.

MNAR

$$S(W|Z,\emptyset) = S(W \mid c,d,\emptyset) \tag{4}$$

$S\ (W \mid c,d,\emptyset)$ is the probability of (W) depending on both observed (c) and unobserved (d) factors.

Correctly handling MCAR, MAR, and MNAR ensures robust preprocessing, improves model training, and enables accurate intelligent meter error prediction.

The outcome of handling missing values with imputation, class-specific means, and predictive modeling ensures complete datasets, enhancing preprocessing and accurate meter error prediction.

3.2.3 Dimensionality Reduction using Robust Principal Component Analysis (RPCA)

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms correlated variables into uncorrelated components, preserving maximum data variance. RPCA is employed as a key preprocessing step in the proposed adaptive calibration framework to effectively handle noisy, corrupted, and anomaly-prone data from intelligent meters. In contrast to the traditional PCA, which is particularly sensitive to outliers, RPCA is specifically adapted to handle data in the presence of sudden disturbances, including voltage variations, temperature swings, electromagnetic distortion, or defective sensor measurements. Not addressed, these radical anomalies may also severely skew model training and diminish predictions.

An intelligent metering dataset (M) is decomposed by RPCA into two distinct components: a low-rank matrix (H), representing the true underlying measurement patterns, and a sparse matrix (Q), capturing the irregular anomalies. Mathematically, this is expressed as equations (5 & 6).

$$M = H + Q \tag{5}$$

Equation (2) minimizes noise by decomposing data into low-rank structure and sparse anomalies simultaneously.

$$min\ (||H||_* + \lambda||Q||_1) \qquad subject\ to\ H + Q = V \tag{6}$$

Where (H) captures essential metering patterns, while (Q) isolates anomalies and noise. This decomposition filters corrupted data, ensuring refined inputs for ACSO-WIForest, enhancing feature extraction, training efficiency, and prediction accuracy under challenging environmental conditions in intelligent metering systems.

3.3 Adaptive Cockroach Swarm Optimization (ACSO) Enhanced Weighted Isolation Forest (WIForest) for Adaptive Calibration and Accurate Error Prediction in Smart Meters

The novelty of the proposed approach lies in the integration of Adaptive Cockroach Swarm Optimization (ACSO) with Weighted Isolation Forest (WIForest) to address the limitations of conventional anomaly detection methods in intelligent metering systems. Traditional Isolation Forest models treat all features equally and rely on static parameters, which reduces their effectiveness in high-dimensional and noise-prone smart meter data. The proposed WIForest enhances anomaly detection by assigning adaptive weights to critical features, improving sensitivity to calibration-relevant deviations.

The hybrid method of ACSO-WIForest incorporates the feature-aware WIForest with ACSO to enhance intelligent meter error prediction. WIForest prioritizes critical calibration features, while ACSO dynamically tunes tree parameters and feature weights. This hybrid framework accurately detects anomalies, minimizes calibration bias, and enables scalable, self-correcting calibration, improving reliability and prediction accuracy in diverse metering environments.

To improve clarity, the proposed Adaptive Cockroach Swarm Optimized Weighted Isolation Forest (ACSO-WIForest) model is summarized into two main stages: optimization and prediction.

Stage 1: Optimization using ACSO

1. Initialize a population of candidate solutions representing WIForest parameters such as number of trees, subsample size, and feature weights.
2. Evaluate each candidate solution using an objective function based on prediction error.
3. Update candidate solutions through adaptive swarm behaviors including swarming, chasing, dispersion, and ruthless replacement.
4. Dynamically adjust step size and movement strategies to balance exploration and exploitation.
5. Identify the optimal parameter set that minimizes calibration and prediction error.

Stage 2: Prediction using WIForest

6. Construct the Weighted Isolation Forest using optimized parameters obtained from ACSO.
7. Perform weighted sampling and feature-based splitting to build multiple isolation trees.
8. Compute weighted path lengths for each data instance across all trees.
9. Calculate anomaly scores to identify abnormal readings and potential meter errors.
10. Classify readings as normal or erroneous based on the defined threshold and generate final predictions.

The cooperation between ACSO and WIForest is established through an adaptive parameter optimization process, where ACSO dynamically tunes the key parameters of the WIForest model to enhance prediction accuracy. Initially, WIForest is constructed with parameters such as the number of trees, subsample size, tree depth, and feature weights. These parameters directly influence anomaly detection performance and calibration accuracy. ACSO treats each candidate solution as a parameter set for WIForest and evaluates its fitness based on prediction error.

The parameter initialization and optimization process of the ACSO-WIForest model are defined to ensure reproducibility and transparency. The number of trees t is set to 100 and subsample size ψ to 256 based on empirical evaluation. Feature weights are initialized in the range [0.5, 1.0] to emphasize important metering attributes. For ACSO, the population size is set to 30, maximum iterations to 50, and step size to 0.2. The ACSO employs an adaptive search strategy that balances exploration and exploitation. Initially, candidate solutions are randomly generated within defined bounds. Early iterations promote global exploration, while later iterations focus on refining solutions using the global best and individual best positions. Adaptive dispersion and replacement mechanisms help maintain diversity and avoid premature convergence. The optimization objective is to minimize prediction error using RMSE. At each iteration, candidate parameters are evaluated by training the WIForest model and computing RMSE on validation data. The process continues until convergence or maximum iterations are reached. All experiments use fixed random seeds and consistent data splits to ensure reproducibility.

The effectiveness of the Adaptive Cockroach Swarm Optimization (ACSO) algorithm is driven by key swarm intelligence principles that ensure efficient convergence. The algorithm maintains a balance between exploration and exploitation, where initial iterations emphasize global search through random dispersion, allowing the swarm to explore diverse regions of the solution space. As iterations progress, adaptive parameter tuning reduces step size and enhances local search, enabling exploitation around promising solutions.

3.3.1 Weighted Isolation Forest (WIForest)

Isolation Forest detects anomalies by isolating data points through random partitioning. IForest improved in this research by giving characteristics priority and increasing the accuracy of intelligent meter error prediction. WIForest is an enhanced anomaly detection method that assigns weights to features, prioritizing critical attributes to improve detection accuracy. It reduces bias toward irrelevant dimensions, handles high-dimensional, imbalanced, or noisy datasets effectively, and focuses on calibration-relevant features in SM data, such as voltage/current mismatches and consumption deviations, minimizing false alarms.

In WIForest, (t) represents the number of isolation trees (iTree) built independently using random subsamples (ψ). More iTrees improve anomaly detection diversity, stabilize scores by reducing variance, and enhance reliability. (iTree = single isolation tree).

Advantages: Feature-aware detection, reduced false positives, scalable for large datasets, and integrates with ACSO for optimized parameter tuning.

Weighted Sampling & Splitting

Step 1 – Weighted Sampling

Each tree is built on a subsample (ψ). Feature selection probability is proportional to its weight, emphasizing critical attributes.

Step 2 – Weighted Splitting:

Splits are biased toward high-weight features (e.g., load imbalance, meter error signals), isolating anomalies more effectively.

Scoring Phase

Step 3 – Weighted Path Length:

For each reading, path length $v_q(r)$ considers weighted contributions of splits, reflecting feature importance in calibration. It is represented in equation (7).

Step 4 – Anomaly Score

$$K_{q\,(r,s)} = 2^{-\frac{G(v_q(r))}{b\,(s)}} \tag{7}$$

Where $v_q(r)$ is the weighted path length, $G(v_q(r))$ is the average over all trees, and $b(s)$ is the normalization factor (same as IForest).

$K_{q\,(r)} \approx 1$: strong anomaly → likely meter calibration error.

$K_{q\,(r,s)} \approx 0.5$: borderline anomaly → requires further validation.

$K_{q\,(r,s)} \ll 0.5$: normal → valid meter reading.

Feature-aware detection in WIForest emphasizes critical attributes to identify subtle calibration errors, reduces false positives by ignoring normal variations, and ensures scalability, making it effective for large-scale SM datasets. ACSO dynamically tunes parameters (t, ψ) and feature weights, ensuring scalable, efficient, and optimized anomaly detection in SM datasets.

The configuration of the Weighted Isolation Forest (WIForest) is designed to enhance anomaly detection accuracy and reduce false positives in complex smart meter datasets. The anomaly scoring mechanism is based on the weighted path length, where data points that are isolated with shorter paths are assigned higher anomaly scores, indicating potential calibration errors.

3.3.2 Adaptive Cockroach Swarm Algorithm (ACSO)

The Cockroach Swarm Algorithm (CSO) is a population-based optimization method inspired by cockroach behaviors like swarming, chasing, dispersing, and seeking food. It iteratively updates candidate solutions to locate optimal values. The ACSO enhances traditional CSOA by adaptively tuning parameters like step size, perception distance, and migration speed during iterations. This adaptive mechanism improves convergence, prevents premature stagnation, and ensures robustness. In the research, ACSO optimizes the WIForest predictive model, effectively capturing meter–environment relationships, minimizing errors, and strengthening anomaly detection in noisy, harsh metering conditions.

Adaptive Swarming–Chase Behavior

The adaptive swarm ensures balance between exploration (large steps early) and exploitation (smaller steps later), where cockroaches move toward individual and global best calibration solutions, refining feature selection and minimizing prediction uncertainty. It is described in equation (8).

$$W_j^{s+1} = w_j^s + \alpha(s)qand\left(o_j - w_j^s\right) + \beta(s)qand\left(h - w_j^s\right) \tag{8}$$

Where (w_j^s) is the position of cockroach (j) at iteration (s), (o_j) and (h) are the personal global best position, $(\alpha(s))$ is an adaptive inertia weight, decreasing over time, $(\beta(s))$ is the adaptive step size, adjusted dynamically, and $(qand)$ is a random value in 0,1.

Adaptive Hunger Behavior

The adaptive mechanism guides cockroaches toward promising calibration solutions, dynamically adjusting migration speed; driven by food-source cues, they refine error prediction accuracy under harsh metering environments, minimizing oscillations and improving robustness. This function is represented in equation (9).

$$w_j^{s+1} = w_j^s + d(s)\left(w_{food} - w_j^s\right) + \delta(s)hunger \tag{9}$$

Where (w_{food}) is the food source (Optimal solution direction), $d(s)$ is an adaptive migration speed, $\delta(s)$ is the adaptive hunger factor, and $(hunger)$ is a random factor in 0,1.

Adaptive Dispersion Behavior

Cockroaches adaptively disperse across the solution space, preserving swarm diversity and preventing premature convergence, which enhances exploration, supports robust search, and improves anomaly detection accuracy in intelligent metering under diverse environmental conditions. Its mathematical function is as follows in equation (10).

$$w_j^{s+1} = w_j^s + \gamma(s) qand(1, C) \tag{10}$$

Where $qand(1, C)$ is the random, (C) is the dimensional vector, and $\gamma(s)$ is an adaptive dispersion factor.

Adaptive Ruthless Behavior

Random cockroaches are adaptively replaced by global best solutions, accelerating convergence while maintaining population diversity, thereby ensuring robust anomaly detection and improved intelligent meter calibration under diverse environmental conditions. It is also derived in equation (11).

$$w_j^s = h \qquad for\ random\ l \in 1, M \tag{11}$$

Where (h) is the current global best, (M) is denoted as the population size.

In this research, ACSO enhances intelligent meter error prediction by adaptively managing noisy data, preventing local optima, and ensuring robust WIForest optimization, significantly improving calibration reliability under challenging environmental conditions.

By integrating feature-aware WIForest with ACSO, the proposed hybrid model provides a robust solution for intelligent meter error prediction and adaptive calibration. WIForest assigns weights to critical features, such as voltage/current mismatches and consumption deviations, emphasizing attributes that significantly impact calibration accuracy while reducing false positives. Multiple isolation trees (t) built from weighted subsamples (ψ) improve anomaly detection diversity, stabilize scores, and enhance reliability across large-scale datasets.

ACSO achieves optimal tree parameters and feature weights by adaptive swarm behaviors such as swarming, chasing, dispersion, and merciless replacement to balance exploration and exploitation. The adaptive mechanism avoids premature convergence, preserves population diversity, and improves error prediction even in severe environmental conditions. Together, ACSO-WIForest facilitates reliable meter error prediction, reduced calibration bias, and self-correcting, scalable calibration, all of which accomplish the objectives of this study to improve measurement reliability, prediction accuracy, and adaptive calibration in smart metering systems used in varied and challenging conditions. Algorithm 1 provides the ACSO-WIForest, which integrates WIForest with ACSO for accurate SM error prediction and adaptive calibration. In this algorithm, ACSO-WIForest uses (t) isolation trees, subsample size (ψ), and feature weights (w). ACSO optimizes these parameters to minimize calibration error, enhancing anomaly detection and adaptive meter error prediction. Table 1 provides the hyperparameters.

Table 1. Hyperparameters

Hyperparameter	Values
t	100
Ψ	256
Max Tree Depth	10
Feature Weight Factor	0.5 – 1.0
ACSO Step Size	0.2
Anomaly Score Threshold	0.6

Algorithm 1: ACSO-WIForest for Intelligent Meter Error Prediction

Step 1: WIForest

```
class WeightedIsolationTree:
def __init__(self, max_depth, feature_weights):
self.max_depth = max_depth
self.feature_weights = feature_weights
self.root = None
def fit(self, data, depth=0):
if depth >= self.max_depth or len(data) <= 1:
return LeafNode(data)
feature = weighted_random_choice(self.feature_weights)
split_value = select_random_split(data, feature)
left_data, right_data = split_data(data, feature, split_value)
left_tree = self.fit(left_data, depth + 1)
right_tree = self.fit(right_data, depth + 1)
return Node(feature, split_value, left_tree, right_tree)
def path_length(self, x):
return traverse_tree(self.root, x, self.feature_weights)
```

Step 2: ACSO

```
class ACSO:
def __init__(self, population_size, max_iter, step_size)
self.population_size = population_size
```

```
self.max_iter = max_iter
self.step_size = step_size
def optimize(self, objective_function):
population = initialize_population(self.population_size)
best_solution = None
for i in range(self.max_iter):
for cockroach in population:
cockroach.move_toward_best(best_solution, self.step_size)
score = objective_function(cockroach)
if score < best_solution_score(best_solution):
best_solution = cockroach
population = maintain_diversity(population, best_solution)
return best_solution
Step 3: Hybrid ACSO-WIForest
def objective_function(params):
t = params'num_trees'
psi = params'subsample_size'
feature_weights = params'feature_weights'
forest = WeightedIsolationTree(max_depth=10,
feature_weights=feature_weights)
.fit(sample_data(psi)) for _ in range(t)
error = evaluate_forest(forest, validation_data)
return error
Optimize parameters using ACSO
acso = ACSO(population_size=30, max_iter=50,
step_size=0.2)
best_params = acso.optimize(objective_function)
Build the final WIForest with optimized parameters
final_forest = WeightedIsolationTree(max_depth=10,
feature_weights=best_params'feature_weights').fit(sample_data
(best_params'subsample_size')) for _ in
range(best_params'num_trees')
Predict anomalies/meter errors
for reading in test_data:
score = compute_anomaly_score(final_forest, reading)
if score > threshold:
mark_as_error(reading)
```

4. Empirical Evaluation

This section evaluates performance metrics for intelligent meter error prediction, comparing existing methods with the proposed ACSO-WIForest model, highlighting improved accuracy, reduced calibration bias, and enhanced reliability.

The experimental setup for the proposed ACSO-WIForest method enables intelligent meter error calibration prediction. The system is implemented on an Intel® Core™ i7-12700K processor (3.6 GHz) with 32 GB DDR4 RAM running Windows 11 Pro. Python 3.9.13 and MATLAB R2022b are used to implement, train, and optimize the ACSO-WIForest method for the predictive SM error prediction process.

Table 2 and Figure 3 present Mutual Information (MI) values of environmental and meter-related factors influencing error prediction. Higher MI values, such as voltage fluctuation and current imbalance, show strong relevance to meter errors, while lower values like EMI and meter aging indicate weaker associations. These insights guide weighted feature selection in ACSO-WIForest calibration. MI measures dependency strength between variables, capturing shared information effectively.

Table 2. Mutual Information (MI) Values of Key Factors Affecting Intelligent Meter Error Prediction

Environment / Meter Variable	**MI Value**	**Environment / Meter Variable**	**MI Value**
Voltage fluctuation (VF)	0.94	Current imbalance (CI)	0.89
Power factor variation (PV)	0.86	Load consumption deviation (LCD)	0.81
Temperature (ambient) (T)	0.78	Humidity (H)	0.69
Harmonic distortion (HD)	0.66	Frequency instability (FI)	0.59

Electromagnetic interference (EMI)	0.54	Meter drift (MD)	0.47

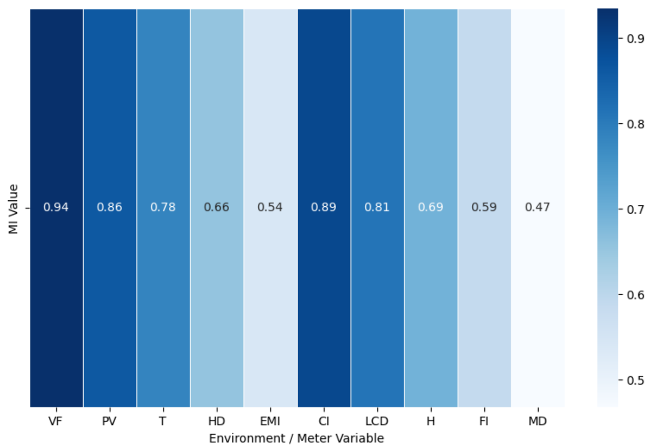

Figure 3. Mutual Information Analysis of Environmental and Meter Variables Influencing Intelligent Meter Error Prediction

Figure 4 highlights environmental factors such as (a) wind speed, (b) humidity distribution, and (c) temperature influencing meter accuracy. The polar plot demonstrates temperature fluctuations, reflecting diverse thermal conditions impacting meter stability. The violin plot shows humidity distribution, where concentrated central values indicate consistent effects with occasional extremes. The strip plot illustrates wind speed variations, capturing external disturbances. Collectively, these visualizations emphasize the harsh environments our proposed ACSO-WIForest framework effectively addresses through adaptive calibration and robust error prediction.

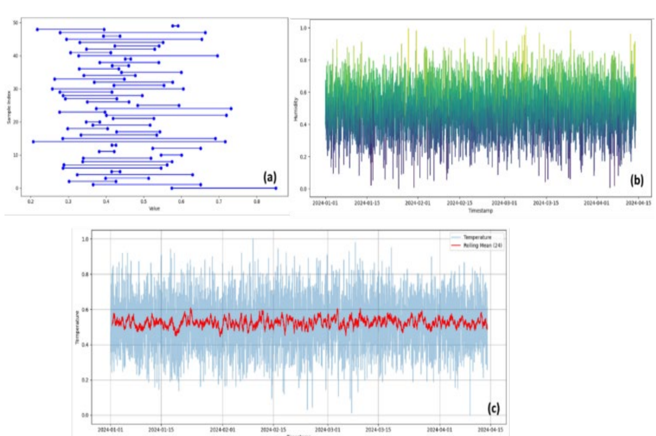

Figure 4. Visualization of Environmental Factors (Wind Speed, Humidity, and Temperature) Influencing Smart Meter Performance

In Figure 5 (a), the swarm plot visualizes the distribution of data points related to anomalies and consumption. Each point represents a data entry, and the plot shows the density and spread of values. This graph helps to visually identify distinct clusters of anomalous data, allowing for a better understanding of how errors correlate with specific consumption levels. Figure 5 (b) examines the autocorrelation in electricity consumption data. Plotting each consumption value against its value from a previous time step (the "lag") reveals if the data is random or has a predictable, cyclical pattern. A clear pattern confirms that the time series is not random, making it suitable for predictive modeling.

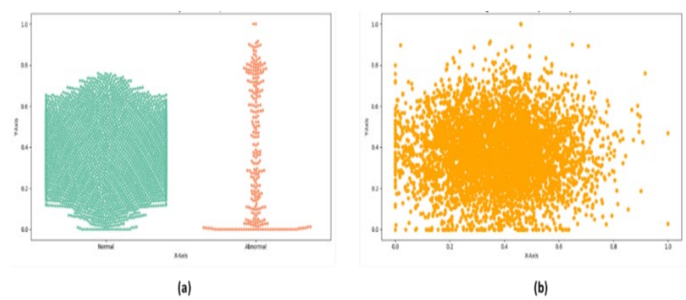

Figure 5. Distribution of Anomalies in Energy Consumption Data and (b) Autocorrelation Analysis for Assessing Predictive Modeling Suitability

Figure 6 demonstrates the dumbbell plot, which visually compares current versus past energy consumption readings. Each horizontal line connects a pair of past and current (red) consumption values for a specific index. The length and direction of the lines highlight the discrepancies between the two periods, which is crucial for the study to detect and correct potential meter inaccuracies and drifts.

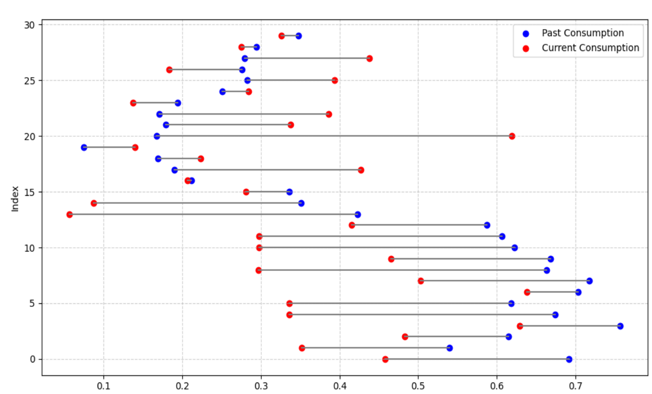

Figure 6. Comparison of Past and Current Energy Consumption Values Using a Dumbbell Plot to Highlight Measurement Variations

Figure 7 graphically visualizes the relationships between selected features: Electricity Consumed, Temperature, Humidity, and Wind Speed. The diagonal plots show the distribution of each variable. The scatterplots show their pairwise correlations, which help the researcher understand how environmental factors (Temperature, Humidity, Wind Speed) influence meter readings and energy consumption.

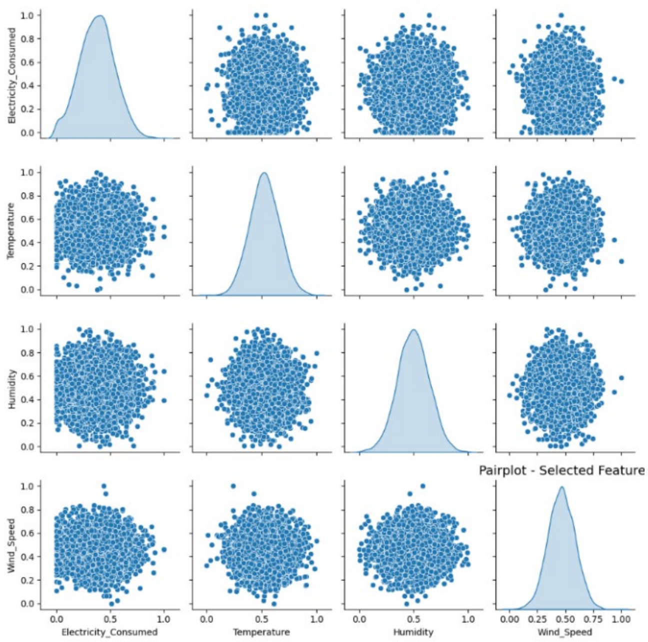

Figure 7: Pairwise Relationships Between Environmental Factors and Energy Consumption Variables for Analyzing Their Influence on Meter Performance

The research proposes an Adaptive Calibration Framework for Intelligent Meter Error Prediction using Machine Learning (Figure 8). It addresses metering inaccuracies caused by environmental factors and hardware issues. The framework uses RPCA and an ACSO-WIForest model to predict and correct errors. Validated with real-world data, the approach significantly improves accuracy and reliability, even with limited data, ensuring more trustworthy energy monitoring.

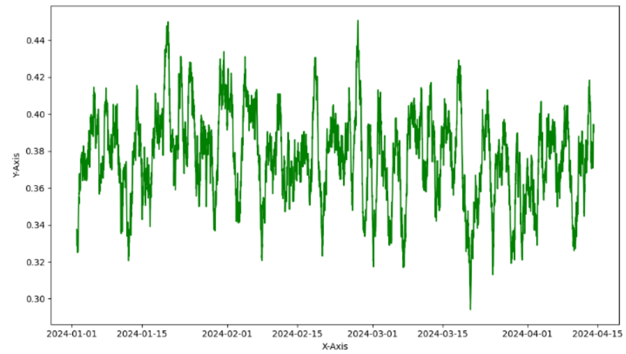

Figure 8: Proposed Adaptive Calibration Framework for Intelligent Meter Error Prediction Using the ACSO-WIForest Model

Figure 9 (a), "KDE - Temperature," is a Kernel Density Estimate (KDE) plot showing the distribution of temperature data. The bell-shaped curve indicates that the temperatures are most frequently clustered around a central value, suggesting a normal distribution. Figure 9 (b), "Line Gradient - Temperature," is a line graph showing the fluctuations in temperature over a period from January 1st to January 11th, 2024. The dashed pink line highlights the rapid and frequent changes in temperature, indicating high volatility.

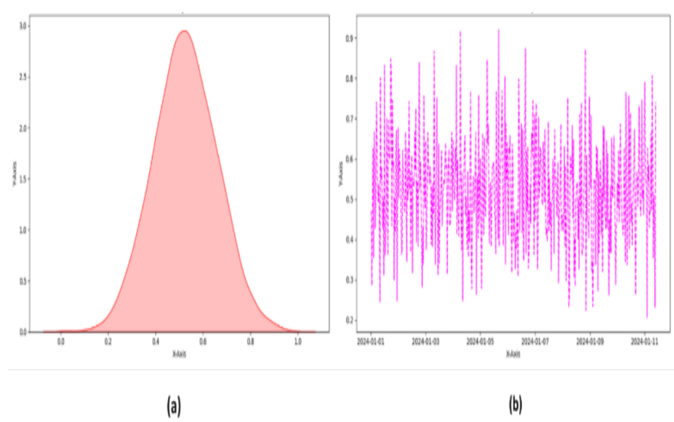

Figure 9: (a) Temperature Distribution Using Kernel Density Estimation (KDE) and (b) Temporal Fluctuation Analysis of Temperature Using Line Gradient Visualization

4.1 Comparative Analysis

The various existing methods are compared to the proposed model, ACSO-WIForest, to evaluate the performance of predicting error in intelligent SM. The existing methods like, Optimized Kernel Density Estimation – Support Vector Regression (OKDE-SVR) [22], Optimized Kernel Density

Estimation-Artificial Neural Network (OKDE-ANN) 22, Optimized Kernel Density Estimation-Wavelet Fuzzy Prediction-Based model (OKDE-WFPB) [22], and Optimized Kernel Density Estimation – Multi-output Support Vector Regression (OKDE-MSVR) [22], OMP and BB for abnormal data detection, followed by a RM for SM error estimation Orthogonal Matching Pursuit + Bollinger Band + Recursive Model (OMP+BB+RM) [23], Generalized Minimal Residual method (GMRES) [23], Randomized Singular Value Decomposition (RSVD) [23], Least Absolute Shrinkage and Selection Operator (LASSO) [23], Long Short-Term Memory (LSTM) [24], LSTM–Attention [24], predict meter errors by capturing temporal patterns with attention, optimized using particle swarm by using Gated Recurrent Unit with Attention - Particle Swarm Optimization (GRU-Attention-PSO) [24] and GRU–Attention [24].

Table 3 illustrates the Evaluation Metrics Demonstrating the Accuracy, Reliability, and Calibration Effectiveness of the Proposed ACSO-WIForest method.

Table 3. Performance Evaluation Metrics for Intelligent Meter Error Prediction

Metric	Equation	Description
RMSE (Root Mean Square Error):	$\sqrt{\frac{1}{m}\sum_{j=1}^{m}(Q_j-\widetilde{Q}_j)^2}$	Measures deviation between predicted and actual meter readings; lower RMSE indicates accurate error prediction and calibration performance.
MAPE (Mean Absolute Error):	$\frac{100}{m}\sum_{j=1}^{m}\left\|\frac{Q_j-\widetilde{Q}_j^{2}}{Q_j}\right\|$	Shows average percentage error between predicted and actual meter readings; lower MAPE indicates precise calibration and reliable error prediction.
MAE (Mean Absolute Error):	$\frac{1}{M}\sum_{j=1}^{M}\left\|Q_j-\widetilde{Q}_j\right\|$	Represents the average absolute difference between predicted and actual meter readings, indicating calibration accuracy and reliable error prediction.
R^2 (Coefficient of Determination):	$1-\frac{\sum_{j=1}^{m}(Q_j-\widetilde{Q}_j)^2}{\sum_{j=1}^{m}(Q_j-\widetilde{Q}_j)^2}$	Evaluates how well predicted meter readings match actual measurements. An R^2 close to 1 indicates excellent calibration and predictive performance

Note: RMSE (Root Mean Square Error) measures the square root of average squared differences between predicted and actual values; lower values indicate better accuracy. MAE (Mean Absolute Error) represents the average absolute difference between predicted and actual values. MAPE (Mean Absolute Percentage Error) expresses prediction error as a percentage, indicating relative accuracy. R^2 (Coefficient of Determination) shows how well predictions fit actual data, with values closer to 1 indicating better performance.

Table 4 provides comparison of RMSE values for baseline and proposed models, highlighting prediction accuracy improvements in intelligent meter error estimation.

Table 4. RMSE Comparison of Baseline and Proposed Models for Intelligent Meter Error Prediction

Model	RMSE
GMRES [23]	0.0213
RSVD [23]	0.0189
LASSO [23]	0.0167
OMP + BB + RM [23]	0.0130
OKDE-SVR [22]	0.00895
OKDE-ANN [22]	0.00801
OKDE-WFPB [22]	0.00733
OKDE-MSVR [22]	0.00488
LSTM [24]	5.26
LSTM–attention [24]	3.44
GRU–attention [24]	3.16
GRU-Attention-PSO [24]	2.77
ACSO-WIForest Proposed	0.0040

Note: RMSE (Root Mean Square Error) evaluates prediction accuracy by measuring deviation between predicted and actual meter readings; lower RMSE indicates better model performance.

The selected evaluation metrics are chosen to comprehensively assess prediction accuracy and error detection capability in intelligent metering systems. RMSE measures the magnitude of large errors and is sensitive to significant deviations, making it suitable for identifying critical calibration issues. MAE provides a straightforward interpretation of average prediction error, ensuring robustness against outliers. MAPE expresses error in percentage form, enabling easy comparison across different consumption scales and practical interpretation for energy applications. The coefficient of determination (R^2) evaluates how well the model captures the variance in meter readings, indicating overall predictive reliability.

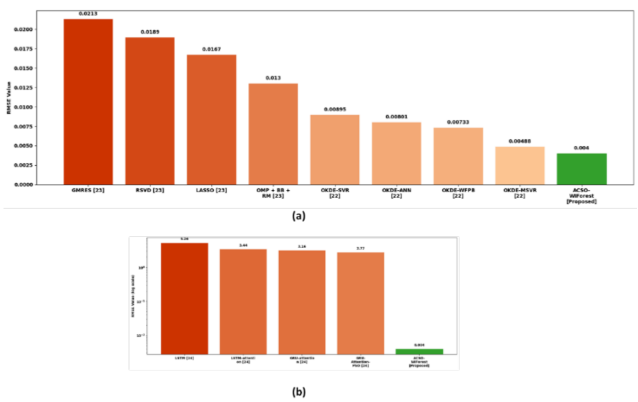

Figure 10 (a) and (b). Graphical representation of Root Mean Square Error (RMSE) results for the Intelligent Meter Error Prediction model across different evaluation scenarios

Figure 10 (a) presents RMSE values for various baseline and hybrid models: GMRES (0.0213), RSVD (0.0189), LASSO (0.0167), OMP + BB + RM (0.0130), OKDE-SVR (0.00895), OKDE-ANN (0.00801), OKDE-WFPB (0.00733), OKDE-MSVR (0.00488), and the proposed ACSO-WIForest (0.0040). The RMSE progressively decreases from traditional methods to hybrid models, indicating improved prediction accuracy. The proposed ACSO-WIForest achieves the lowest RMSE, demonstrating its superior performance and effectiveness in accurately predicting intelligent meter errors compared to all other models.

Figure 10 (b) shows RMSE values for different models: LSTM (5.26), LSTM–Attention (3.44), GRU–Attention (3.16), GRU-Attention-PSO (2.77), and the proposed ACSO-WIForest – 0.0040. The decreasing RMSE from LSTM to GRU-Attention-PSO indicates improved prediction accuracy, while the proposed ACSO-WIForest achieves the lowest RMSE, demonstrating that it significantly outperforms all existing models in intelligent meter error prediction. The "10^n" notation compresses large value ranges on a logarithmic scale, making extremely small or large RMSE values clearly visible.

Table 5 compares SM error prediction accuracy using MAPE, showing that the proposed ACSO-WIForest consistently outperforms all baselines. Table 5 presents the MAPE of various methods for predicting intelligent meter errors. Classical deep learning models like LSTM (4.49), LSTM–Attention (3.21), and GRU–Attention (2.87) show moderate performance. Improvements are observed in hybrid and optimization-based methods such as GRU-Attention-PSO (2.49), GMRES (1.676), and RSVD (1.2768). Sparse regression techniques, like LASSO (0.4632) and OMP + BB + RM (0.5626), further reduce errors. The proposed ACSO-WIForest model (0.221) achieves the lowest MAPE, highlighting its superior accuracy, robustness, and adaptability to environmental variations.

Table 5. MAPE Comparison of Baseline and Proposed Models for Prediction

Method	MAPE
LSTM 24	4.49
LSTM–attention 24	3.21
GRU–attention 24	2.87
GRU-Attention-PSO 24	2.49
GMRES 23	1.676
RSVD 23	1.2768
LASSO 23	0.4632
OMP + BB + RM 23	0.5626
ACSO-WIForest Proposed	0.221

Note: MAPE (Mean Absolute Percentage Error) represents the average percentage deviation between predicted and actual values; lower MAPE indicates higher prediction accuracy.

Table 6 presents the coefficient of determination (R^2) for various models in intelligent meter error prediction. Table 6 shows the compared outcoming R^2 values of the existing method and the proposed method, LSTM (0.63), LSTM–Attention (0.74), GRU–Attention (0.79), and GRU-Attention-PSO (0.85), showing accuracy among recurrent models. OKDE-SVR (0.968), OKDE-ANN (0.967), OKDE-WFPB (0.974), OKDE-MSVR (0.980), and the proposed ACSO-WIForest method (0.998) achieve higher predictive reliability. The proposed ACSO-WIForest surpasses all, achieving the highest R^2, the error distribution of the proposed ACSO-WIForest model is observed to be highly concentrated around zero, indicating minimal prediction bias and strong accuracy. Compared to baseline models, which exhibit a wider spread of errors, the proposed approach demonstrates lower variance and improved stability. This distribution confirms the robustness of the model and its effectiveness in maintaining consistent prediction performance under varying smart metering conditions.

Demonstrating its superior performance and robustness in accurately predicting intelligent meter errors.

Table 6. R^2 Comparison of Baseline and Proposed Models for Intelligent Meter Error Prediction

Method	R^2
LSTM 24	0.63
LSTM–attention 24	0.74
GRU–attention 24	0.79
GRU-Attention-PSO 24	0.85
OKDE-SVR 22	0.968
OKDE-ANN 22	0.967
OKDE-WFPB 22	0.974
OKDE-MSVR 22	0.980
ACSO-WIForest Proposed	0.998

Note: R^2 (Coefficient of Determination) indicates the proportion of variance explained by the model; values closer to 1 represent stronger predictive performance.

Table 7. MAE Comparison of Baseline and Proposed Models for Intelligent Meter Error Prediction

Model	MAE
GMRES 23	0.0235
RSVD 23	0.0173
LASSO 23	0.0136
OMP + BB + RM 23	0.0109
OKDE-SVR 22	0.00788
OKDE-ANN 22	0.00700
OKDE-WFPB 22	0.00637
OKDE-MSVR 22	0.00276
ACSO-WIForest Proposed	0.00156

Note: MAE (Mean Absolute Error) measures the average absolute difference between predicted and actual values; lower MAE indicates more accurate predictions.

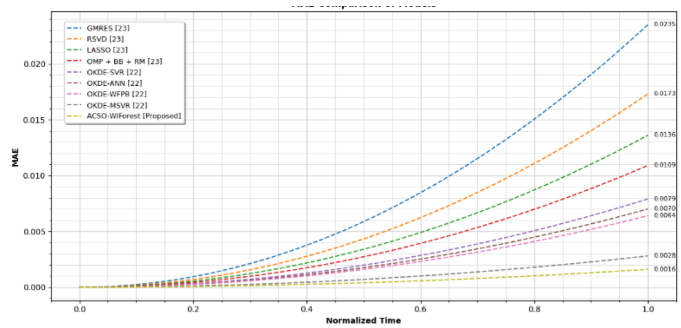

Figure 11. Comparison of Mean Absolute Error (MAE) values demonstrating the performance of the proposed ACSO-WIForest method against baseline models

Figure 11 shows that the proposed ACSO-WIForest model achieves the lowest MAE value compared to all baseline methods. GMRES (0.0235), RSVD (0.0173), LASSO (0.0136), OMP+BB+RM (0.0109), OKDE-SVR (0.00788), OKDE-ANN (0.00700), OKDE-WFPB (0.00637), and OKDE-MSVR (0.00276) all record higher errors, while the proposed ACSO-WIForest attains just 0.00156 MAE (Table 7). This demonstrates that the proposed model provides superior accuracy and robustness over existing methods.

Table 8. Execution Time Comparison Between Existing Models and ACSO-WIForest

Model	Computational time (S)
OKDE-SVR 22	58.94
OKDE-ANN22	88.03
OKDE-WFPB22	79.46
OKDE-MSVR22	62.59
ACSO-WIForest Proposed	55.21

Table 8 compares the computational time of models in seconds. OKDE-SVR (58.94s), OKDE-ANN (88.03s), OKDE-WFPB (79.46s), and OKDE-MSVR (62.59s) show higher execution times. The proposed ACSO-WIForest achieves the lowest computational time of 55.21s, proving its efficiency and suitability for intelligent meter error prediction in real-time applications.

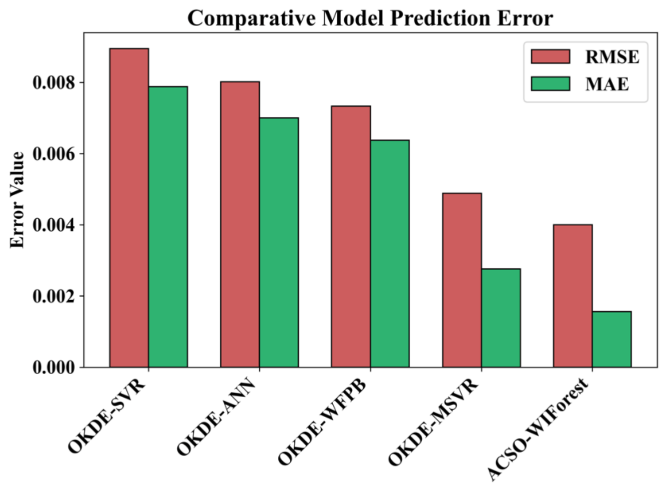

Figure 12. Comparative RMSE and MAE error analysis showing that the proposed ACSO-WIForest model achieves the lowest prediction error among all compared methods

Figure 12 presents the comparative analysis of prediction error using RMSE and MAE across different models. The results show that the proposed ACSO-WIForest achieves the lowest error values compared to all baseline methods, indicating superior prediction accuracy and robustness. The reduction in RMSE highlights improved handling of large deviations, while lower MAE confirms consistent performance across all samples. This demonstrates the effectiveness of integrating adaptive swarm optimization with feature-weighted anomaly detection for intelligent meter error prediction.

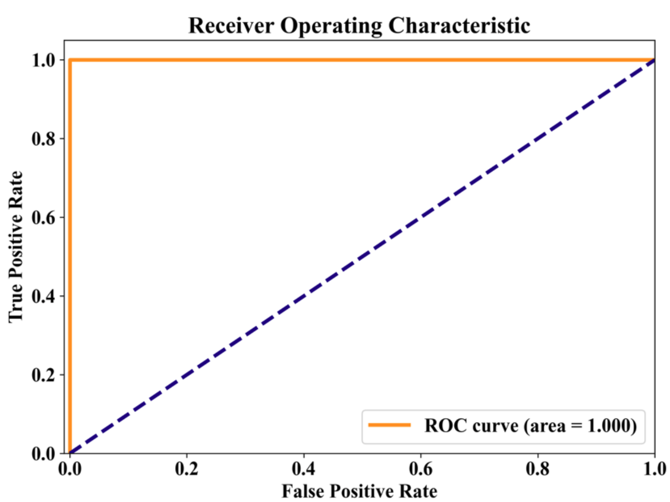

Figure 13. ROC Curve Demonstrating the Anomaly Detection Performance of the Proposed ACSO-WIForest Model

The ROC curve in Figure 13 illustrates the anomaly detection performance of the proposed ACSO-WIForest model by plotting the True Positive Rate against the False Positive Rate across different threshold values. The curve lies close to the top-left corner, with an AUC value of 1.000, indicating excellent classification capability and near-perfect discrimination between normal and anomalous meter readings

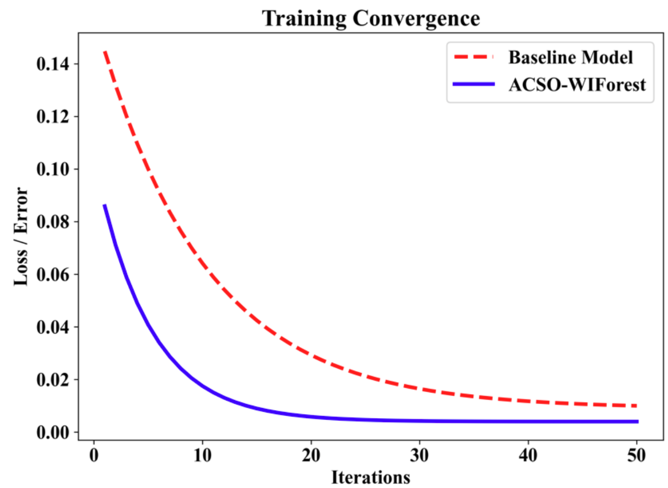

Figure 14. Convergence plot showing the reduction of loss over iterations, where the ACSO-WIForest model achieves faster and more stable convergence compared to the baseline model

WIForest model rapidly reduces loss in early iterations and stabilizes quickly, indicating efficient optimization. In contrast, the baseline model converges more slowly with higher residual error, demonstrating inferior training performance. The convergence plot shows that the ACSO is explained in Figure 14.

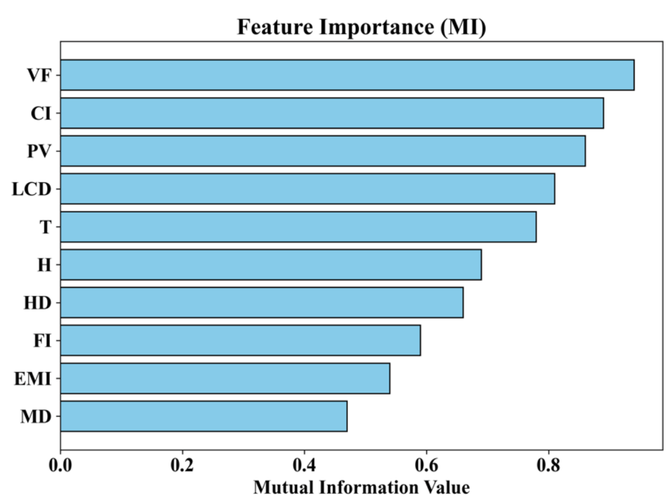

Figure 15. Feature Importance Analysis Using Mutual Information for Intelligent Meter Error Prediction

The graph shows the relative importance of input features based on Mutual Information, indicating how strongly each feature contributes to accurate meter error prediction. Voltage Fluctuation (VF) and Current Imbalance (CI) are the most influential features, while Measurement Deviation (MD) has the least impact on the model is shown in Figure 15.

4.2 Discussion

Existing meter error prediction techniques have significant disadvantages, OKDE-SVR 22, OKDE-ANN 22, OKDE-WFPB 22, and OKDE-MSVR 22, are more robust but computationally intensive and involve large-scale parameter tuning, leading to decreased scalability and tend to suffer from overfitting, poor flexibility, and inadequate performance in the presence of nonlinear error patterns. OMP+BB+RM 23, GMRES 23, RSVD 23, and LASSO 23 are effective in abnormal data detection and estimation but do not account for dynamic temporal changes, restricting accuracy in fluctuating scenarios. Deep learning methods such as LSTM 24 and LSTM-Attention 24 are good at temporal dependencies but have high training expense, data-hungriness, and sensitivity to noise. GRU-Attention 24 and GRU-Attention-PSO 24 minimize complexity but remain less adaptable in real-time, severe environmental circumstances. This research developed an Adaptive Calibration Framework with ACSO-WIForest, combining robust anomaly filtering, optimized swarm intelligence, and adaptive calibration. It enabled precise, noise-tolerant, and computationally efficient error prediction, maintaining high accuracy even with limited data, and ensured reliable performance across diverse smart metering environments and challenging operational conditions. In terms of memory requirements, the model remains lightweight as only tree structures and optimized parameters need to be stored. Furthermore, the framework is well-suited for distributed IoT environments, where preprocessing and anomaly detection can be performed locally at edge devices, reducing communication overhead with central servers. This distributed capability enhances scalability and enables real-time deployment in large smart grid infrastructures. Overall, the proposed approach demonstrates practical feasibility by balancing computational efficiency, memory usage, and communication cost, making it suitable for large-scale intelligent metering applications.

The proposed ACSO-WIForest-based automated calibration framework offers significant energy efficiency and operational advantages in smart grid environments. By enabling continuous and intelligent error detection, the framework reduces the need for manual calibration and maintenance, thereby lowering operational costs and human intervention. Accurate meter readings enhance the reliability of energy consumption data, which is critical for effective load forecasting, demand response strategies, and grid stability. Furthermore, improved anomaly detection minimizes energy losses caused by incorrect measurements and ensures more precise billing and resource allocation. In large-scale smart grid systems, these benefits collectively contribute to optimized energy distribution, enhanced system reliability, and improved overall efficiency, supporting the transition toward fully automated and intelligent energy management infrastructures.

5. Conclusion

The research proposed an Adaptive Calibration Framework for Intelligent Meter Error Prediction that effectively addresses the limitations of existing methods in smart metering systems. In the research, data were gathered from smart meters installed in various environments, augmented with an experimental dataset from Kaggle. Min-Max normalization preprocessing was used to ensure a uniform model training. RPCA using an optimized distance function and adaptive box plot thresholding well distinguished and eliminated the anomalies induced by environmental interference. In practical smart metering infrastructures, the proposed adaptive calibration framework can be integrated into utility monitoring systems to enable real-time error detection and automatic calibration of meters. This reduces manual inspection efforts, enhances billing accuracy, and supports reliable energy management in large-scale smart grids. The ability to adapt to environmental variations and device aging makes the framework suitable for continuous deployment in modern intelligent energy systems. The proposed ACSO-WIForest model successfully captured nonlinear dependencies, providing adaptive and robust error prediction. Results demonstrated significant improvement over existing methods, achieving MAE = 0.00156, RMSE = 0.0040, MAPE = 0.221, and R^2 = 0.998, proving the model's high accuracy and reliability. The execution time of ACSO-WIForest (55.21) outperforms existing models, demonstrating faster and more efficient processing. Overall, the framework ensures continuous calibration, resilience under limited data conditions, and enhanced robustness, making it a valuable contribution toward trustworthy and self-correcting intelligent metering infrastructures.

Limitations and Future Scope

The research, while highly accurate, is limited by dataset diversity and dependency on Kaggle-sourced data. Future work can extend the framework by incorporating larger real-time datasets, integrating advanced deep learning models, and exploring hybrid optimization strategies to further improve adaptability, scalability, and robustness in diverse smart metering environments. Future work can further extend this framework by validating it on larger and more diverse real-time datasets to enhance generalization. Additionally, testing across different types of smart meters and heterogeneous deployment environments will improve robustness and adaptability. Exploring real-world deployment scenarios, including integration with IoT-based smart grid infrastructures, can further demonstrate the scalability and practical applicability of the proposed model.

References

[1] Fan Z, Shi L, Xi C, Wang H, Wang S, Wu G. Real-time power equipment meter recognition based on deep learning. *IEEE Trans Instrum Meas.* 2022;71:1–15.

[2] Saeedi R, Sadanandan SK, Srivastava AK, Davies KL, Gebremedhin AH. An adaptive machine learning framework for behind-the-meter load/PV disaggregation. *IEEE Trans Ind Inform.* 2021;17(10):7060–7069.

[3] Soluyanov YI, Fedotov AI, Akhmetshin AR, Soluyanov VI. Application of intelligent electricity metering systems for timely adjustment of standard values for electrical loads calculation. In: *Proceedings of the 3rd International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE);* 2021; Moscow, Russia. p.1–5.

[4] Abdulla N, Demirci M, Ozdemir S. Smart meter-based energy consumption forecasting for smart cities using adaptive federated learning. *Sustain Energy Grids Netw.* 2024;38:101342.

[5] Devanathan B, Kaushik K, Jayasurya G, Lakshmanan SA, Suyampulingam A. Smart metering for real-time power prediction, price forecasting and intelligent alerts in modern living spaces. In: *Proceedings of the 5th International Conference on Expert Clouds and Applications (ICOECA);* 2025. p.396–401.

[6] Almufareh MF, Humayun M, Ahmad Z, Khan A. An intelligent LoRaWAN-based IoT device for monitoring and control solutions in smart farming through anomaly detection integrated with unsupervised machine learning. *IEEE Access.* 2024;12:119072–119086.

[7] Rind YM, Raza MH, Zubair M, Mehmood MQ, Massoud Y. Smart energy meters for smart grids: An Internet of Things perspective. *Energies.* 2023;16(4):1974.

[8] Ramos Ruiz G, Fernandez Bandera C. Validation of calibrated energy models: Common errors. *Energies.* 2017;10(10):1587.

[9] Qu Z, Jiao Y, Li H, Chen Q, Zhu Y, Wei W. An online accuracy evaluation method for smart electricity meter calibration devices based on mutual correlations of calibration data. *IEEE Trans Instrum Meas.* 2024.

[10] Wang Q, Li H, Wang H, Zhang J, Fu J. A remote calibration device using edge intelligence. *Sensors.* 2022;22(1):322.

[11] Ma L, Meng Z, Teng Z, Tang Q. A measurement error prediction framework for smart meters under extreme natural environment stresses. *Electr Power Syst Res.* 2023;218:109192.

[12] Ma J, Teng Z, Tang Q, Guo Z, Kang L, Wang Q, Li N, Peretto L. A novel multisource feature fusion framework for measurement error prediction of smart electricity meters. *IEEE Sens J.* 2023;23(17):19571–19581.

[13] Zhang B, Xia X, He C, Kang W, Zhang J. Examining performance calibration in smart power system electricity metering based on environmental perception attention LSTM network. *Front Energy Res.* 2024;12:1405725.

[14] Ma L, Meng Z, Teng Z, Qiu W. A reliability evaluation framework for smart meters based on AGG-ARIMA and PFR. *Meas Sci Technol.* 2022;33(4):045006.

[15] Ma L, Teng Z, Tang Q, Wan Z, Li N, Meng Z. Measurement error prediction-based reliability assessment framework for electric metering devices under harsh natural environments. *Measurement.* 2024;232:114481.

[16] Tong X, Ma J, Ma L, Yan S, Tang Q, Teng Z, Cheng D. A novel prediction method for smart meter error using multi-view convolutional neural network. *IEEE Sens J.* 2024.

[17] Xia T, Liu C, Lei M, Xia S, Li D, Ming D. Measurement error estimation for distributed smart meters through a modified BP neural network. *Front Energy Res.* 2022;10:928681.

[18] Ma J, Teng Z, Tang Q, Qiu W, Yang Y, Duan J. Measurement error prediction of power metering equipment using improved local outlier factor and kernel support vector regression. *IEEE Trans Ind Electron.* 2022;69(9):9575–9585.

[19] Deligiannis P, Koutroubinas S, Koronias G. Predicting energy consumption through machine learning using a smart-metering architecture. *IEEE Potentials.* 2019;38(2):29–34.

[20] Hong Y. Nonlinear modeling of measurement errors in gateway energy meters. *Meas Sensors.* 2024;35:101286.

[21] Oprea SV, Bâra A. Machine learning algorithms for short-term load forecast in residential buildings using smart meters, sensors and big data solutions. *IEEE Access.* 2019;7:177874–177889.

[22] Ma J, Teng Z, Tang Q, Qiu W, Yang Y. Measurement error assessment for smart electricity meters under extreme natural environmental stresses. *IEEE Trans Ind Inform.* 2022;18(6):3905–3914.

[23] Liu F, Huang H. Smart meter error estimation in topological low-voltage energy systems. *IEEE Access.* 2024.

[24] He C, Xia X, Zhang B, Kang W, Zhang J, Chen H. Analyzing adjustment and verification errors in electric metering devices for smart power systems considering multiple environmental factors. *AIP Adv.* 2024;14(12).

[25] Thirusubramanian G. Machine learning-driven AI for financial fraud detection in IoT environments. *Int J HRM Organ Behav.* 2021;9(4):9–25.

[26] Jin X. Multi-source collaborative optimization scheduling technology for new energy microgrid based on improved

particle swarm algorithm. *EAI Endorsed Trans Energy Web*. 2024;12.