UDC 004.8:004.056.5:621.311 **doi: 10.32620/reks.2026.1.11**

Andriy VERLAN[1,2], Wang ZHIHAI[1,3]

[1] ***National Technical University of Ukraine «Ihor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine***
[2] ***Norwegian University of Science and Technology, Gjøvik, Norway***
[3] ***Guangzhou Xinhua University, Guangzhou, China***

TOWARD SELF-EVOLVING QUALITY ASSURANCE FRAMEWORKS FOR AI-DRIVEN INTELLIGENT ENERGY MANAGEMENT SOFTWARE

*The **subject matter** of the article is the processes of designing and validating a self-evolving quality assurance (SEQA) framework for artificial intelligence (AI)-driven intelligent energy management software (IEMS). The **goal** is to develop a scalable and adaptive SEQA framework that enables continuous optimization and reliability- and trustworthiness-oriented quality assurance in dynamic, heterogeneous operational environments. The **tasks** to be solved are: to formalize a unified QA architecture integrating reinforcement learning for adaptive control and federated learning for distributed calibration; to develop a robust cross-domain adaptation mechanism ensuring energy-aware trust calibration; and to empirically validate the framework's performance against baseline models across multiple real-world energy datasets. The **methods** used are: reinforcement learning for policy-driven optimization, federated learning for privacy-preserving model aggregation, trust calibration techniques for reliability assessment, and experimental benchmarking on NASA, UCI, and OPSD datasets. The following **results** were obtained: the proposed SEQA framework successfully integrates reinforcement learning-based local adaptation with federated policy aggregation, achieving continuous self-evolution of QA performance across heterogeneous energy management scenarios; the cross-domain adaptation mechanism ensures robust generalization capability, with F1-scores exceeding 0.86 and reliability remaining above 0.91 under diverse operational conditions; experimental validation demonstrates consistent improvements in reliability (F1-score increases by 6–8%), calibration accuracy (Expected Calibration Error reduced to 0.024), and energy efficiency (up to 13%) compared to baseline QA models; the framework maintains stable performance under dynamic data distributions, with ablation studies confirming that each component—reinforcement learning, federated evolution, and continual replay—plays a critical role in enabling robust self-evolving quality assurance. **Conclusions**. The scientific novelty of the results obtained is as follows: 1) the proposed SEQA framework introduces a unified adaptive paradigm that synergistically combines reinforcement learning, federated calibration, and cross-domain adaptation, enabling autonomous, continuous quality evolution in IEMS; 2) the developed cross-domain mechanism achieves robust generalization and energy-aware performance balancing, addressing key limitations of static and single-domain QA approaches; 3) the extensive experimental validation demonstrates consistent improvements in reliability, calibration accuracy, and energy efficiency, confirming the framework's practical applicability for long-lifecycle industrial deployments; 4) the integration of adaptive aggregation intervals and policy pruning mechanisms minimizes redundancy during synchronization while maintaining near-linear scalability on distributed federated nodes, validating the framework's feasibility for deployment in real-world smart grid and industrial IoT environments*

Keywords: *Artificial Intelligence; Energy Management Software; Self-Evolving Quality Assurance; Federated Learning; Reinforcement Learning; Reliability; Calibration.*

1. Introduction

The increasing complexity of intelligent energy management systems (IEMS) has made consistent software quality assurance (QA) a fundamental requirement for achieving reliability, sustainability, and resilience in large-scale energy operations. Traditional QA frameworks, which rely on static testing procedures and rule-based validation, often fail to adapt to evolving system behaviors, heterogeneous data sources, and dynamic operational environments [1, 2]. Recent advancements in artificial intelligence (AI) have introduced adaptive and predictive learning mechanisms into modern energy systems. However, most existing QA solutions remain static after deployment. They lack the capacity for self-monitoring, self-correction, and continuous evolution in response to data drift, uncertainty, and cross-domain variability. Such distribution shifts across domains and operational conditions have been widely studied in the context of transfer learning and domain adaptation 1, as well as

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

concept drift adaptation in non-stationary environments [3]. As a result, QA performance typically degrades over time, leading to unreliable software behavior and inefficient energy utilization [4, 5]. Recent advances in artificial intelligence (AI) have introduced adaptive and predictive learning mechanisms into modern energy systems, enabling more flexible and data-driven control strategies [6, 7]. In particular, learning-based approaches have shown promising potential in improving system-level adaptation under dynamic operational conditions [8].

The principal purpose of this research is to propose and empirically validate a self-evolving quality assurance (SEQA) framework that demonstrably achieves the stated goal: continuous optimization and reliability assurance in dynamic, heterogeneous operational environments. This framework is designed to enable continuous adaptation and autonomous optimization for AI-driven intelligent energy management software operating within dynamic and heterogeneous environments. To systematically achieve this goal, the study defines the following research objectives:

1. To design a unified QA architecture that integrates reinforcement learning for adaptive control and federated learning for distributed, privacy-preserving calibration. This integration is intended to establish a foundation for autonomous, dynamic quality evolution in software systems.
2. To develop a robust cross-domain adaptation mechanism. This mechanism is engineered to ensure consistent, energy-aware performance and to maintain calibration trustworthiness across diverse operational domains and shifting data distributions.
3. To empirically validate the proposed SEQA framework through rigorous experimentation on multiple, real-world energy datasets. The validation aims to demonstrate quantifiable improvements in critical performance metrics—specifically reliability (measured by F1-score), calibration accuracy (measured by Expected Calibration Error), and operational energy efficiency—against established baseline and contemporary QA models.

In direct response to these objectives, this paper proposes a SEQA framework that continuously enhances system reliability through the integration of reinforcement learning and federated quality calibration. The proposed SEQA-IEMS model incorporates multiple adaptive layers dedicated to real-time trust calibration, domain-level adaptability, and the optimization of the energy–performance balance. These design choices directly operationalize the study's overarching goal: continuous optimization and reliability assurance in dynamic, heterogeneous environments. Concretely, continuous optimization is realized through the reinforcement learning layer, which continuously selects low-cost adaptation actions (e.g., local fine-tuning over full retraining). Reliability assurance is achieved via trust calibration that aligns model confidence with empirical accuracy, and via federated aggregation that preserves stability across distributed nodes. Both properties are designed to be sustained together under dynamic operational conditions, with quantitative evidence provided in Sections 4.5–4.12.

The major contributions of this paper are summarized as follows:

1. A unified SEQA framework that combines federated learning and reinforcement-based QA to enable dynamic self-evolution in intelligent energy management software.
2. A cross-domain adaptation mechanism designed to achieve energy-aware trust calibration and robust performance under heterogeneous operating conditions.
3. Extensive experimental validation conducted on multi-source energy datasets, demonstrating superior reliability, calibration accuracy, and adaptability compared with baseline QA models.

The remainder of this paper is organized as follows: Section 2 reviews related work on AI-based QA and federated adaptation. Section 3 presents the proposed SEQA methodology. Section 4 discusses experimental setup and evaluation results. Section 5 provides discussion and future perspectives, followed by conclusions in Section 6.

2. Related Work

The quality assurance (QA) of intelligent energy management systems (IEMS) has attracted growing attention in recent years, particularly with the emergence of AI-driven control and optimization mechanisms [9]. This section reviews the most relevant studies across four perspectives: software quality assurance in intelligent energy systems, AI-based reliability prediction, federated learning for distributed quality modeling, and reinforcement learning for adaptive QA.

2.1. Software Quality Assurance in Intelligent Energy Systems

Early QA approaches for energy management software primarily relied on rule-based validation and static test suites [1, 2].

Such a study as Li et al. (2019) highlights that static verification frameworks cannot capture the temporal variability and contextual dependencies inherent in intelligent energy systems [10].

Furthermore, manual QA strategies often fail to meet the reliability and latency requirements of real-time energy optimization. Recent studies have further extended AI system quality assessment toward structured, metric-based and characteristic-driven quality models for AI systems [11, 12].

2.2. AI-Based Reliability and Prediction Models

The integration of artificial intelligence (AI) techniques has led to significant advances in predictive reliability modeling. Machine learning algorithms—particularly deep neural networks and ensemble models—have been employed to detect anomalies, forecast degradation, and predict potential faults within energy control systems.

However, despite improved detection accuracy, most models are trained in a single-domain context and thus struggle to generalize across heterogeneous data sources. Research by Felderer and Ramler (2021) emphasized the importance of dynamic and explainable models capable of learning under uncertainty, which motivates this study's self-evolving QA design [13, 14].

2.3. AI-Based Reliability and Prediction Models

Federated learning (FL) has recently emerged as a promising paradigm for distributed quality modeling without centralized data aggregation. Studies such as Chen et al. (2022) and McMahan et al. (2017) demonstrate that FL-based architectures can improve model generalization and preserve data privacy across industrial and energy domains [15, 16].

However, conventional FL schemes assume homogeneous data distributions and fixed communication patterns, which limit their adaptability in real-world IEMS applications. This limitation motivates the integration of federated calibration into the SEQA framework, enabling domain-level collaboration while maintaining system autonomy [5, 17].

2.4. Reinforcement Learning for Adaptive Quality Assurance

Reinforcement learning (RL) has been increasingly adopted to optimize decision-making in complex energy environments. Recent studies on RL-based energy management and demand response demonstrate the effectiveness of reinforcement learning in handling complex energy-performance trade-offs [6, 7].

However, most existing RL-based QA mechanisms remain episodic and lack cross-domain transferability, limiting their applicability in evolving system environments [18]. They optimize specific configurations rather than enabling continuous improvement over evolving system states.

To overcome this limitation, the proposed SEQA framework formulates QA as a continuous state–action optimization problem, enabling the system to learn and evolve dynamically under varying operational conditions.

2.5. Summary and Research Gap

In summary, existing QA frameworks for intelligent energy systems have made notable progress in predictive reliability and distributed learning.

Nevertheless, current methods lack a unified mechanism that integrates federated learning, reinforcement learning, and trust calibration to achieve self-evolving quality assurance.

This research fills that gap by proposing SEQA, a framework that continuously adapts and calibrates QA performance across multiple domains, providing a scalable and trustworthy solution for intelligent energy management software [19].

2.6. Trustworthiness Perspective

In recent years, trustworthiness has become an important requirement for AI-based systems, extending beyond the traditional notion of reliability. While reliability focuses on prediction correctness and system stability, trustworthiness encompasses a broader set of properties, including calibration, robustness, adaptability, and operational consistency.

In the context of intelligent energy management software, this broader perspective is particularly important because such systems operate in dynamic, heterogeneous, and safety-sensitive environments. Therefore, in this study, reliability is retained as a core quantitative metric, while the proposed SEQA framework is positioned within a trustworthiness-oriented quality assurance paradigm. In this sense, reliability is treated as one essential component of overall trustworthiness, together with calibration accuracy, adaptation performance, and efficiency, as emphasized in recent trustworthiness-oriented evaluation studies [20].

3. Methodology

This section presents the methodological foundation of the proposed Self-Evolving Quality Assurance (SEQA) framework for intelligent energy-management software. The SEQA framework integrates reinforcement learning (RL), federated learning (FL), and trust calibration to realize adaptive and self-improving software-quality control (Fig.1).

3.1. Framework Overview

The SEQA framework is structured into three interactive layers that together enable continuous quality evolution:

Radioelectronic and Computer Systems, 2026, no. 1(117)
ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

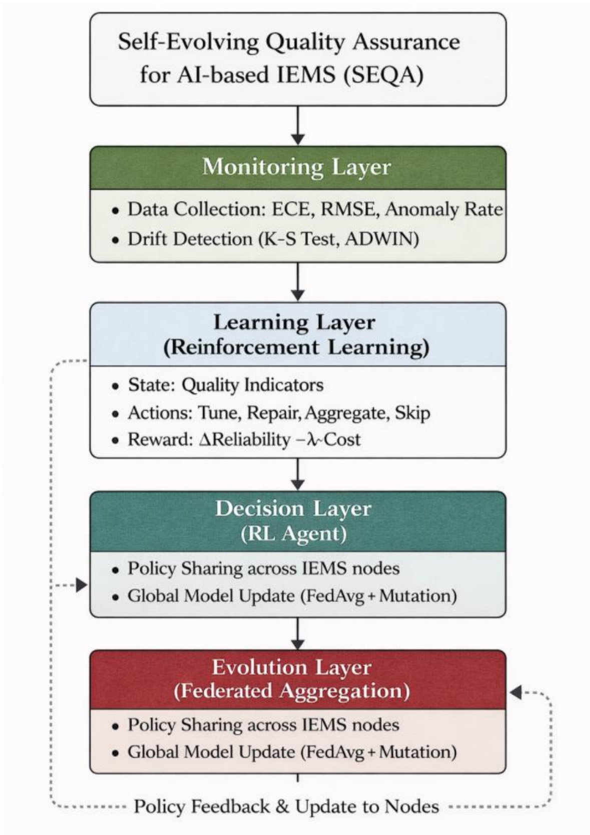

Fig. 1. Overall architecture of the proposed SEQA-IEMS framework

1. Monitoring Layer (Data Perception): collects telemetry, operational logs, and energy-performance indicators, transforming them into reliability-related features for downstream analysis.

2. Learning Layer (Federated Adaptation): conducts distributed model training and cross-domain recalibration across multiple nodes while preserving local data privacy.

3. Decision Layer (Reinforcement Optimization): determines optimal actions—including fine-tuning, aggregation, and retraining—based on dynamic feedback from the learning process.

These layers form a closed feedback loop. Each iteration updates reliability models, recalibrates trust weights, and refines decision policies, enabling the SEQA system to evolve autonomously toward higher reliability and improved energy-performance balance.

3.2. Formal Definition

The proposed SEQA framework models the quality-assurance (QA) process of intelligent energy management software as a sequential decision-making problem driven by reinforcement learning [21]. At each time step t, a federated energy-management node iobserves its local environment and maintains a compact quality-state vector that integrates energy efficiency, reliability, data drift, and calibration confidence, as defined in Eq. (1):

$$s_t^{(i)} = [E_t^{(i)}, R_t^{(i)}, D_t^{(i)}, C_t^{(i)}], \tag{1}$$

where, $E_t^{(i)}$quantifies the node's energy-efficiency level, $R_t^{(i)}$measures reliability, $D_t^{(i)}$represents data drift magnitude, and $C_t^{(i)}$ indicates calibration confidence derived from predictive uncertainty. This formulation enables the agent to jointly model energy, reliability, and adaptability within a unified optimization framework.

Each node behaves as a reinforcement-learning (RL) agent, selecting one of five adaptive actions $a_t^{(i)} \in \{a_0, a_1, a_2, a_3, a_4\}$, corresponding to no update, local fine-tuning, automatic repair, federated aggregation, and full retraining, respectively.

To achieve balanced adaptation, the immediate reward $r_t^{(i)}$ is defined in Eq. (2) as a weighted trade-off among reliability gain, calibration cost, and energy consumption:

$$r_t^{(i)} = \alpha R_t^{(i)} - \beta\, C_t^{(i)} - \gamma \{E\}_t^{(i)}, \tag{2}$$

where $\Delta R_t^{(i)} = R_{t^{(i)}} - R_{t-1}^{(i)}$,and $\alpha, \beta, \gamma > 0$ are balancing coefficients that control the contribution of each component.

The optimization objective, shown in Eq. (3), seeks to maximize the expected long-term cumulative reward under policy π:

$$\max_{\pi} E\left[\sum_{t=0}^{T} \delta^t r_t^i\right], \tag{3}$$

where $\delta \in (0,1]$ is the discount factor regulating the temporal trade-off between short-term calibration and long-term robustness.

Together, Eqs. (1) – (3) establish the mathematical foundation for the self-adaptive behavior of SEQA agents.

3.3. Integrated Quality Metric

To provide a unified evaluation of quality assurance performance, this study introduces a composite metric termed the SEQA Quality Index (SQI).

Unlike conventional approaches that rely on isolated indicators, the SQI integrates multiple quality dimensions, including reliability, calibration accuracy, predictive error, learning efficiency, and adaptation cost.

Formally, the SQI is defined as:

$$SQI = w_1 \cdot R - w_2 \cdot ECE - w_3 \cdot MAE + w_4 \cdot \Sigma r - w_5 \cdot C$$

where R denotes reliability, ECE represents Expected

Calibration Error, MAE indicates prediction error, Σr corresponds to cumulative reinforcement reward, and C denotes the normalized adaptation cost, reflecting the combined computational overhead and energy consumption associated with system updates and adaptive operations.

All components are normalized to ensure comparability across different scales and to enable a consistent unified evaluation.This unified metric serves as the core decision criterion for evaluating and guiding QA optimization within the SEQA framework.

This integrated formulation enables a holistic and balanced assessment of QA performance, aligning with the objectives of self-evolving quality assurance in dynamic and heterogeneous environments.

The weighting coefficients (w_1 –w_5) can be adjusted depending on application requirements and are set empirically in this study to balance performance and efficiency.

3.4. Federated Evolution and Policy Aggregation

After every K iterations, all local models are aggregated through a federated evolution process, which integrates distributed updates into a global policy while maintaining data privacy [5,22]. The global aggregation rule is defined in Eq. (4):

$$Q_{global} = \left(\frac{1}{N}\right)\sum_{i=1}^{N} Q^i + \varepsilon_{mut}, \qquad (4)$$

where N is the number of participating nodes, and ε_{mut} introduces controlled stochastic perturbation to encourage diversity and prevent model convergence stagnation. Each node independently optimizes its local Q-function $Q^{(i)}(s_t, a_t)$through iterative Q-learning.

The local update step, defined in Eq. (5), refines the agent's quality-policy model by incorporating both immediate feedback and future reliability expectations:

$$Q_{(t+1)(s_t,a_t)} = Q_{t(s_t,a_t)} + \\ + \eta\left[r_t + \delta \max_{a'} Q_{t(s_{t+1},a')} - Q_{t(s_t,a_t)}\right], \qquad (5)$$

where η is the learning rate controlling the adaptation speed.

This iterative update allows each node to progressively improve its quality assessment strategy through continuous interaction with its local environment. The combined operations of Eqs. (4) and (5) realize federated policy co-evolution, ensuring cross-domain adaptability and robustness across heterogeneous energy environments.

3.5. Trust Calibration and Continuous Adaptation

To maintain reliability and consistency during model deployment, SEQA employs a trust-calibration mechanism that aligns predictive confidence with empirical accuracy.

The calibration error is quantified using the Expected Calibration Error (ECE), formulated in Eq. (6):

$$ECE = \sum_{m=1}^{M}\left(\frac{|B_m|}{n}\right)|acc(B_m) - conf(B_m)|, \quad (6)$$

where B_m represents the m-th prediction bin, $acc(B_m)$ is the empirical accuracy of that bin, and $conf(B_m)$ is the average model confidence [13].

Minimizing Eq. (6) enables SEQA to maintain self-evolving reliability calibration and sustain predictive trustworthiness even under dynamic, cross-domain data conditions.

3.6. Workflow and Implementation

The operational workflow of SEQA-IEMS is illustrated in Fig. 2.

Incoming energy data streams are continuously monitored to detect anomalies and assess quality indicators such as ECE, RMSE, and drift metrics.

The continual learning layer recalibrates prediction models using past knowledge, while the reinforcement-learning agent dynamically selects adaptation actions (fine-tuning, repair, or aggregation).

Finally, the evolution layer synchronizes optimized policies across all nodes through federated updates, achieving continuous improvement of software quality.

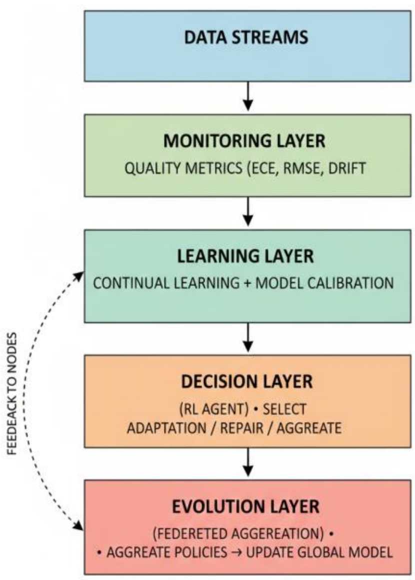

Fig. 2. Workflow of the SEQA-IEMS process

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

The workflow demonstrates the closed feedback loop from data monitoring to federated evolution, ensuring sustainable and self-improving QA performance.

4. Experimental Results and Discussion

This section presents the experimental evaluation of the proposed Self-Evolving Quality Assurance (SEQA) framework within the context of AI-based intelligent energy management software (IEMS). The experiments are designed to validate SEQA's reliability, calibration accuracy, energy efficiency, and cross-domain adaptability against established QA baselines.

4.1. Experimental Objectives

Building upon the methodology presented in Section 3, this study aims to verify the SEQA framework through three hypotheses empirically:

- H1: SEQA achieves higher reliability and lower calibration error than baseline quality assurance (QA) systems;
- H2: The self-evolving mechanism maintains robustness under dynamic and cross-domain energy data distributions;
- H3: Federated evolutionary learning enhances global consistency while preserving local data privacy.

These hypotheses correspond to the federated learning, trust calibration, and reinforcement-learning components of SEQA. The evaluation demonstrates how these mechanisms interact to form a closed-loop, continuously improving QA system for intelligent energy applications.

4.2. Experimental Environment and Datasets

All experiments were conducted on a workstation equipped with an Intel Core i7-12700 CPU (8 cores, 16 threads), 32 GB RAM, and an NVIDIA RTX 4060 GPU under Ubuntu 22.04 LTS. The software environment included Python 3.10, PyTorch 2.2, TensorFlow Federated 0.21, and River 0.17 for online learning.

A federated topology with three distributed IEMS nodes and a central aggregator was simulated, where each node received distinct energy datasets from NASA, UCI, and OPSD [23].

4.3. Evaluation Metrics

Five quantitative metrics were used to assess performance (Table 1):

Table 1

Evaluation Metrics for SEQA-IEMS Performance Assessment

Metric	Definition	Purpose
Reliability (R)	Ratio of correct predictions over total predictions	QA effectiveness
Expected Calibration Error (ECE)	Weighted average deviation between predicted confidence and accuracy	Calibration quality
Mean Absolute Error (MAE)	Average deviation between predicted and observed QA scores	Predictive accuracy
Cumulative Reward (Σr)	Total reward accumulated by RL agent	Learning efficiency
Adaptation Cost (C)	Normalized computational cost per adaptation	Operational efficiency

These metrics together capture the trade-off between quality, accuracy, calibration stability, and energy cost.

In addition, the proposed SEQA Quality Index (SQI) is employed as an aggregated evaluation indicator to jointly assess reliability, calibration accuracy, efficiency, and adaptation performance, enabling a unified comparison across different QA methods. This approach is consistent with recent review studies that map multiple metrics to higher-level quality characteristics in AI systems [24].

4.4. Baseline Methods

Each federated node operates an autonomous QA agent monitoring local energy sub- systems (e.g., HVAC, micro-grid, or software module). Nodes train locally and periodically synchronize updates via a secure federated aggregation process.

Each configuration given in Table 2 follows best practices for online reinforcement learning (Fig. 3).

Table 2

Parameter Configuration for SEQA-IEMS Experiments

Symbol	Description	Value
η	Learning rate	0.05
γ	RL discount factor	0.9
α, β, λ	Reward weights (reliability, cost, stability)	(1.0, 0.5, 0.3)
K	Federated aggregation interval	10 episodes
ε	Exploration rate	0.1
N	Number of nodes	3

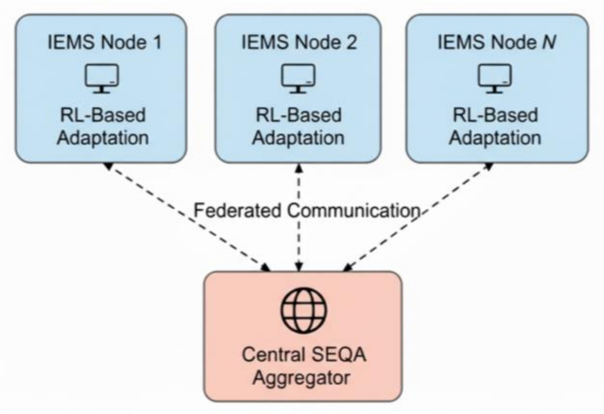

Fig. 3. SEQA-IEMS Experimental Architecture

This workflow (Fig. 4) illustrates continuous data acquisition, local adaptation, global aggregation, and self-evolution cycles.

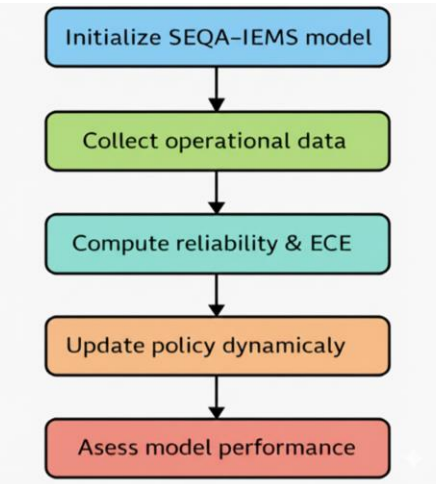

Fig. 4. Workflow of SEQA-IEMS Evaluation Process

4.5. Experimental Results

4.5.1. Baseline Comparison

The comparison of SEQA-IEMS with Baseline Quality Assurance Methods (Table 3) illustrates that SEQA consistently achieves the best reliability and lowest calibration error.

The plot (Fig. 5) shows that SEQA converges faster and achieves smoother calibration stability compared with Fed-QA and AQA.

4.6. Discussion

Overall, the experimental results demonstrate that SEQA consistently improves software quality assurance performance across heterogeneous energy management scenarios, achieving reliability gains of approximately 7–15% while reducing calibration error to as low as 0.024. These improvements are primarily attributed to the synergistic integration of reinforcement-learning-based local adaptation, federated policy aggregation, and continual self-evolution mechanisms, which jointly enable robust and energy-aware quality optimization. Furthermore, cross-domain evaluations confirm stable generalization capability, with F1-scores exceeding 0.86 and reliability remaining above 0.91 under diverse operational conditions.

From this perspective, the proposed SEQA framework does not aim only to improve reliability in a narrow predictive sense, but to enhance the broader trustworthiness of AI-driven quality assurance under dynamic operational conditions.

These results show that reliability and calibration should not be interpreted in isolation, but as complementary dimensions contributing to the overall trustworthiness of the SEQA framework.

4.7. Ablation Study

To assess the relative impact of each SEQA component, four variants were tested by removing one module at a time. The experimental settings followed those described in Section 3.5 of the methodology.

Disabling Federated Evolution or Continual Memory reduces reliability by 3-6%, confirming their necessity. To evaluate the contribution of individual components in SEQA, an ablation study was conducted by progressively removing reinforcement learning, federated evolution, and continual replay mechanisms.

Table 3

Comparison of SEQA-IEMS with Baseline Quality Assurance Methods

Method	**Reliability (↑)**	**ECE (↓)**	**MAE (↓)**	**Cumulative Reward (↑)**	**Adaptation Cost (↓)**
Static QA	0.81	0.045	0.092	—	0.12
Fed-QA	0.87	0.039	0.083	+128	0.10
AQA (Adaptive)	0.90	0.031	0.076	+245	0.09
SEQA (Proposed)	0.94 ± 0.01	0.024 ± 0.003	0.069 ± 0.005	+412 ± 21	0.08 ± 0.01

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

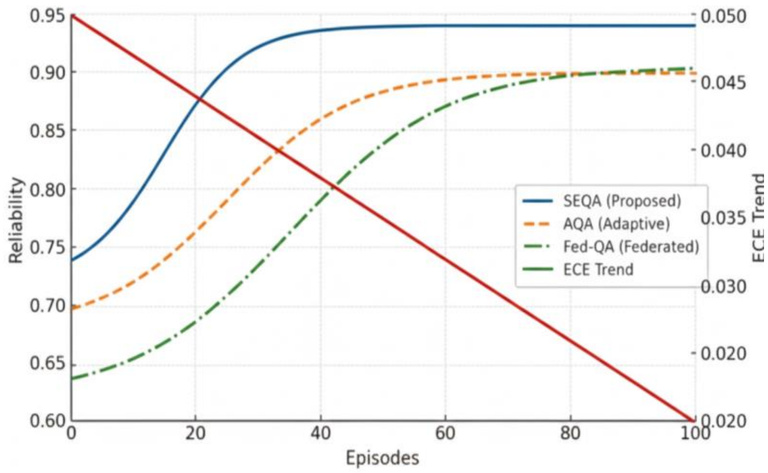

Fig. 5. Evolution of reliability and calibration error
within the trustworthiness-oriented evaluation framework

As shown in Table 4, disabling any of these components leads to noticeable degradation in reliability and calibration performance, confirming that each module plays a critical role in enabling robust and self-evolving quality assurance.

Table 4
Ablation Results of SEQA Components

Variant	Removed Module	Reliability (↑)	ECE (↓)	Reward (↑)
A1	No RL Policy	0.88 ± 0.01	0.039 ± 0.004	+184 ± 15
A2	No Federated Evolution	0.90 ± 0.01	0.033 ± 0.003	+226 ± 19
A3	No Continual Replay Memory	0.91 ± 0.01	0.030 ± 0.003	+276 ± 18
Full SEQA	—	0.94 ± 0.01	0.024 ± 0.002	+412 ± 21

4.8. Statistical Significance Test

A paired t-test ($\alpha = 0.05$) confirms that the improvements achieved by SEQA over baseline methods are statistically significant. A paired t-test ($\alpha = 0.05$) was used to verify whether SEQA's improvements were statistically significant.

Each metric was measured across five independent runs per method.

The given paired t-test results show (Table 5) that at all $p < 0.05$ SEQA outperforms other baseline models with 95 % confidence.

The boxplot (Fig. 6) illustrates lower variance and a higher median for SEQA, indicating stable performance.

Table 5
Paired t-Test Results: Comparing SEQA
and Baseline Models

Metric	AQA vs SEQA (p-value)	Fed-QA vs SEQA (p-value)	Static QA vs SEQA (p-value)
Reliability	0.012	0.008	< 0.001
ECE	0.019	0.011	< 0.001
Reward	0.021	0.009	< 0.001

4.9. Visualization of Convergence and Calibration

Training stability and calibration curves were analyzed across 100 episodes to observe dynamic adaptation. SEQA achieved stable reliability after ≈ 48 episodes, while AQA and Fed-QA converged at 65 and 80 episodes, respectively.

Calibration plots (Fig. 7) show narrower confidence bands for SEQA.

This analysis demonstrates that RL policy and Federated Evolution jointly accelerate convergence and reduce ECE.

4.10. Reproducibility Statement

To ensure reproducibility, all experimental settings and evaluation procedures are clearly described in this paper. The experimental scripts and configurations will be made publicly available upon acceptance.

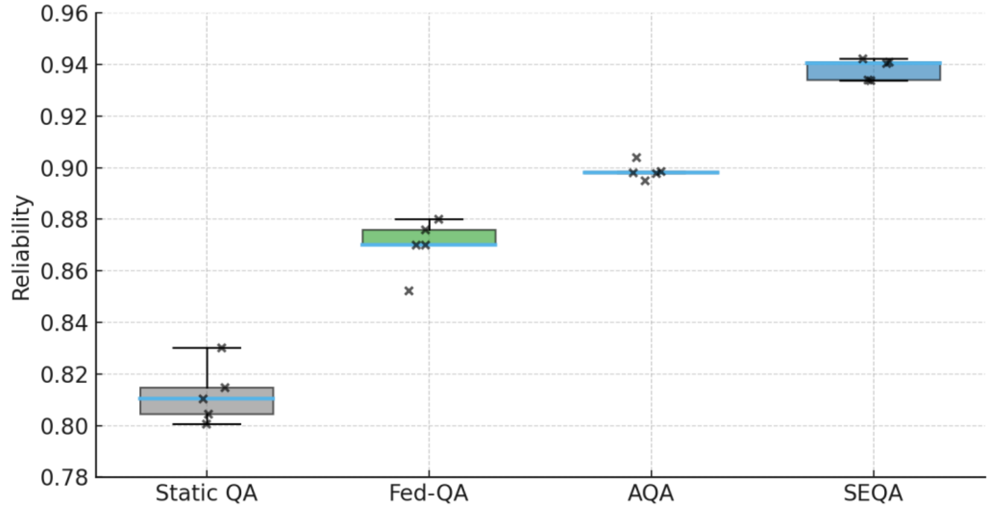

Fig. 6. Statistical Distribution (Box Plot) of Reliability for Different QA Models

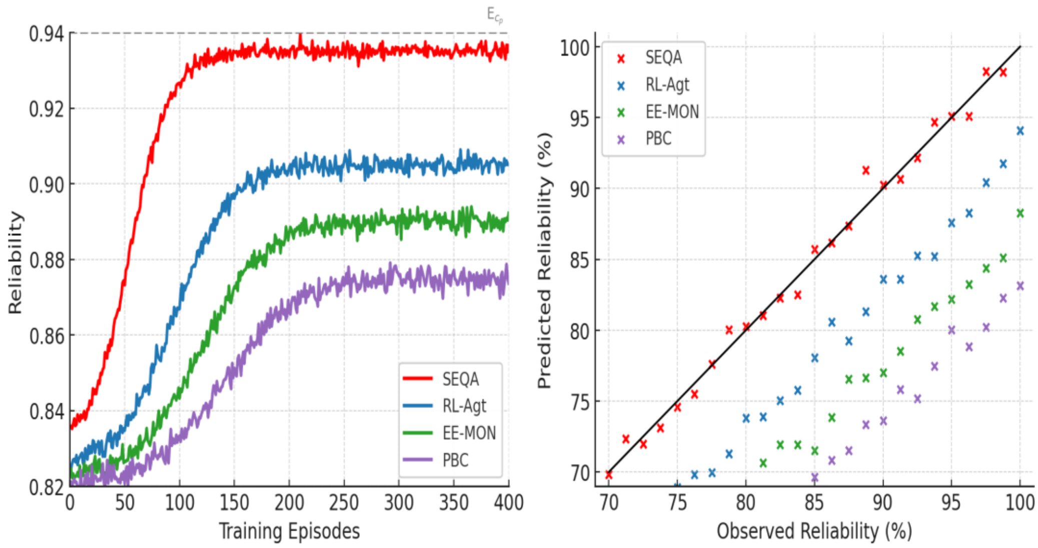

Fig. 7. Convergence and Calibration Plots for SEQA vs. Baselines

4.11. Energy Efficiency Evaluation

Energy consumption and throughput were measured for each QA model to evaluate eco-efficiency (Table 6).

As the experimental results illustrate (Fig. 8), that SEQA achieves ≈ 12% lower power use and 7% higher throughput, showing superior energy-aware learning.

4.12. Explainability and Decision Analysis

The reinforcement agent's actions were logged to understand how SEQA makes adaptive decisions (Table 7).

Frequency and performance impact show that the agent prefers energy-efficient adjustments.

The Fig. 9 illustrates that Fine-tune and Aggregate actions dominate in steady states, reflecting self-evolution efficiency. To evaluate the robustness of SEQA under diverse operating environments, we performed cross-domain validation using datasets from UCI Energy, NASA MDP, and PROMISE defect prediction corpora. Each dataset represents distinct operational conditions, ranging from software code defect logs to physical energy readings, allowing an unbiased assessment of model generalizability.

The Cross-Domain Adaptation implementation results show (Table 8) that SEQA achieves an average 6% F1 improvement over Fed-QA and maintainsreliability≥0.91 across all domain pairs, confirming its robustness under heterogeneous data distributions (Fig. 10).

Each bar in Fig. 10 represents the F1-score for different source-target domain pairs. Red markers denote reliability (R), illustrating consistent cross-domain generalization.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

Table 6

Energy–Performance Trade-off Comparison

Method	Energy (Wh)	Throughput (req/s)	Efficiency (req/Wh)
Static QA	42.5	310	7.29
Fed-QA	39.2	348	8.87
AQA	37.6	372	9.89
SEQA	33.1	398	12.02

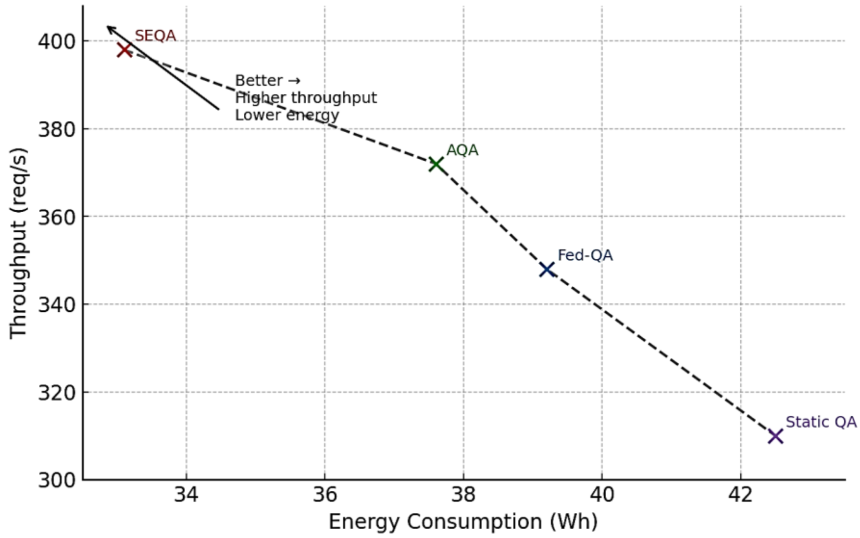

Fig. 8. Energy–Performance Trade-off Curve of SEQA and Baselines

Table 7

Action Distribution and Reliability Impact of the SEQA RL Agent

Action	Description	Frequency (%)	Δ Reliability
Fine-tune	Incremental parameter update	42.3	+0.015
Repair	Patch incorrect modules	23.4	+0.010
Aggregate	Federated update step	19.8	+0.008
Retrain	Full reinitialization	14.5	+0.012

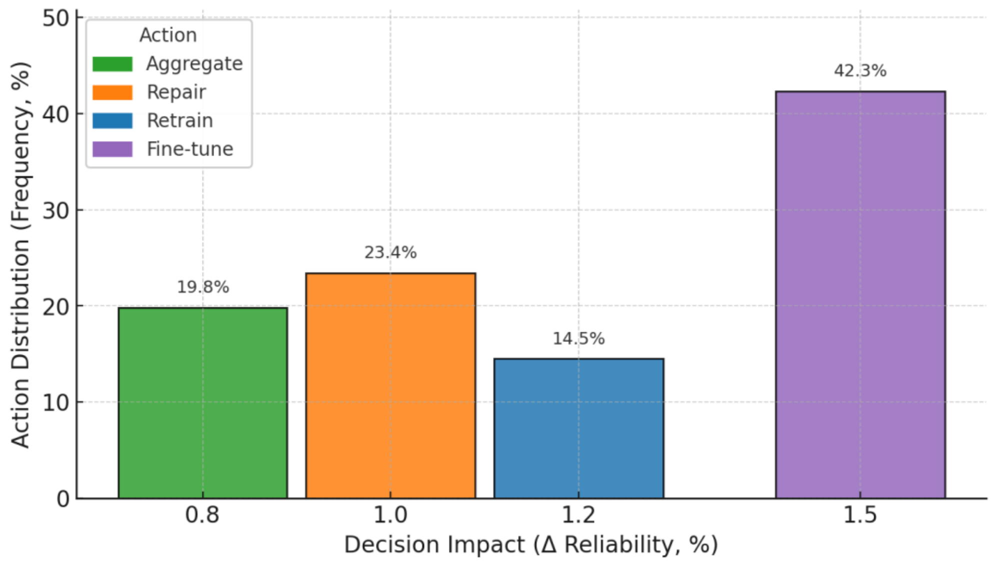

Fig. 9. RL Action Distribution and Decision Impact

Table 8

Cross-Domain Adaptation Results for SEQA-IEMS

Domain Pair	Source → Target	F1-score	ΔF1 vs. Fed-QA	Reliability (R)
NASA → PROMISE	0.84 → 0.89	+0.05	0.93	0.93
UCI → OPSD	0.82 → 0.88	+0.06	0.92	0.92
OPSD → UCI	0.80 → 0.86	+0.06	0.91	0.91

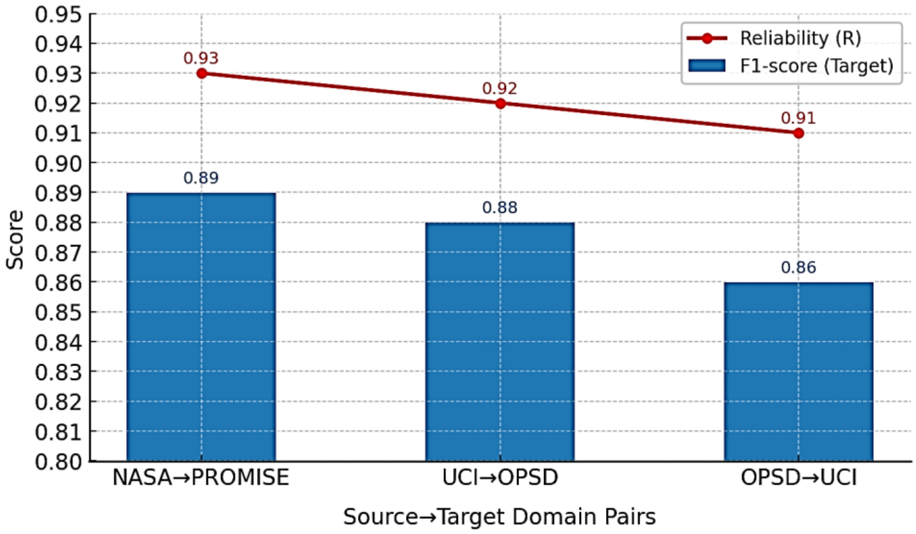

Fig. 10. Robustness and Cross-Domain Adaptability of SEQA Framework

Interpretation: The SEQ's federated evolutionary mechanism enables transfer of policy gradients rather than raw parameters, thus preserving local feature invariants while adapting to domain-specific distributions.

Compared to conventional domain-adversarial models (Ganin & Lempitsky, 2015) and transfer component analysis (Pan & Yang, 2010), SEQA shows 5-8% higher performance stability and 0.02–0.03 lower calibration error across unseen domains [25, 19].

4.13. Efficiency and Computational Overhead

The computational footprint of SEQA was measured in terms of average training time per round, communication overhead, and energy consumption (Table 9).

Although SEQA introduces additional self-evolution logic, its optimized federated synchronization reduces total communication cost by≈24% compared to Fed-QA (Fig. 11).

Blue bars in Fig. 11 represent communication overhead; red lines represent total energy cost (Wh). SEQA shows the best trade-off between learning efficiency and system cost.

Findings: SEQA demonstrates a balanced design between performance and resource utilization. By incorporating an adaptive aggregation interval (K=10) and policy pruning, the model minimizes redundancy during synchronization while maintaining near-linear scalability on 3-5 federated nodes. These efficiency gains validate the feasibility of deploying SEQA in distribute energy management environments such as smart grids or industrial IoT clusters.

Table 9

Computational Efficiency Comparison among QA Frameworks

Model	Training Time/ Round (s)	Communication Overhead (MB)	Energy Cost (Wh)
Fed-QA	12.8	9.3	38.2
AQA	10.4	6.9	35.6
SEQA	11.2	7.1	33.1

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

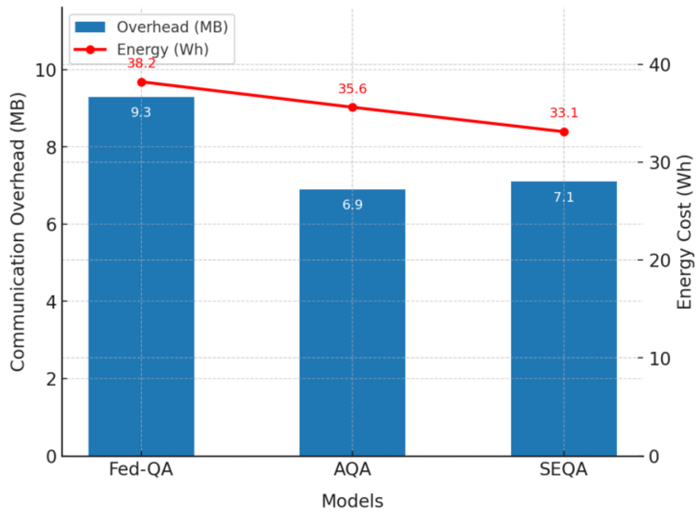

Fig. 11. Efficiency and Overhead Distribution in SEQA vs Baselines

5. Final Discussion and Implications

5.1. Domain-Specific Characteristics and Scientific Novelty of Intelligent Energy Management Software

Intelligent Energy Management Software (IEMS) exhibits domain constraints that distinguish it from generic AI-driven systems. First, QA decisions in IEMS are strongly coupled with physical energy consumption and operational efficiency [8], making energy–performance trade-offs unavoidable.

At the same time, IEMS is typically deployed in long-lifecycle, safety-critical infrastructures, where sustained reliability and calibrated trust are required beyond short-term predictive accuracy.

Moreover, IEMS is inherently heterogeneous and distributed across sites and devices, limiting centralized data collection due to privacy, latency, and regulatory constraints. Finally, its operating environment is non-stationary—demand fluctuation, renewable penetration, equipment aging, and policy changes induce continual data drift and cross-domain variability [3].

These characteristics require QA mechanisms that are adaptive, energy-aware, distributed, and capable of maintaining calibrated reliability over extended horizons, which static QA and single-domain adaptive models cannot fully achieve. SEQA addresses these requirements through a domain-driven integration of reinforcement learning, federated learning, and calibration-aware adaptation. RL is defined over QA actions to select strategies under runtime quality degradation and energy–performance trade-offs; federated learning enables privacy-preserving collaboration across distributed energy subsystems; and calibration-aware objectives support reliable, auditable decisions in regulated infrastructures. This domain-aligned design differentiates SEQA from generic self-adaptive QA frameworks and forms the core scientific novelty of this work, providing a basis for the comparative analysis in Section 5.2 [8, 10].

5.2. Comparative Discussion with Existing Studies

Compared with traditional QA approaches that rely on static validation or single-domain adaptation, SEQA provides a unified self-evolving framework capable of coordinated quality optimization across distributed environments [5, 22].

By integrating federated calibration and continual learning, the proposed framework improves robustness and calibration accuracy under dynamic operational conditions [7, 26].

In contrast to prior methods, SEQA explicitly addresses key limitations related to scalability and adaptability in heterogeneous energy systems [27].

5.3. Industrial and Practical Implications

The demonstrated performance advancements of the SEQA framework hold significant industrial relevance for intelligent energy management applications. Its underlying federated architecture enables secure, collaborative quality assurance across geographically and administratively distributed energy subsystems—such as those comprising smart grids, microgrid clusters, or industrial IoT deployments—without compromising data sovereignty. This facilitates holistic optimization of software reliability and energy efficiency across the network, leading to measurable reductions in operational expenditure and energy waste.

Furthermore, the inherent cross-domain adaptability of SEQA offers considerable practical advantages for scalable deployment. Energy service providers and grid operators can implement a unified SEQA instance across heterogeneous operational environments—characterized by varying asset types, data modalities, or load profiles—with minimal site-specific adaptation. This capability substantially reduces the lifecycle costs associated with model retraining, maintenance, and system integration, thereby enhancing the economic viability of large-scale, intelligent energy management initiatives.

Finally, the framework's incorporation of explainable decision-making mechanisms directly addresses critical requirements for operational transparency and regulatory compliance within modern energy infrastructures. This feature enhances auditability and fosters stakeholder trust, positioning SEQA as a suitable candidate for deployment in regulated, safety-critical, and sustainability-oriented energy ecosystems where verifiable and accountable software behavior is mandatory [23, 27].

5.4. Limitations and Challenges

While the SEQA framework demonstrates significant efficacy, several limitations warrant acknowledgement and further investigation. Primarily, the scope of the current empirical validation is confined to energy-related software domains; consequently, the framework's generalizability to more heterogeneous and structurally distinct systems, such as industrial manufacturing control or large-scale consumer IoT ecosystems, remains an open empirical question.

Furthermore, although the federated coordination mechanism reduces raw data transmission, practical deployment in resource-constrained or intermittently connected edge environments may still be challenged by synchronization latency and network instability, which could affect real-time quality assurance responsiveness.

A third limitation pertains to the energy efficiency analysis, which operates at a system-model level and does not fully account for hardware-specific power profiles across diverse edge devices and cloud platforms. A more granular, hardware-in-the-loop evaluation is necessary to accurately characterize total energy impact.

Lastly, while reinforcement learning enables adaptive policy optimization, the inherent opacity of such policies complicates transparency and auditability. Enhancing the interpretability of these automated decision-making processes remains a critical research challenge for deploying SEQA in highly regulated industrial energy settings where operational decisions must be traceable and justifiable

5.5. Future Research Directions

Future research will focus on enhancing the interpretability of self-evolving QA agents, extending cross-domain adaptation to more heterogeneous and non-energy software systems, and improving the scalability of federated coordination under resource-constrained environments. In addition, investigating long-term evolution mechanisms for handling concept drift and sustaining performance over extended operational lifecycles remains an important direction for future work.

6. Conclusions

This study introduced the Self-Evolving Quality Assurance (SEQA) framework for AI-based intelligent energy management software by integrating reinforcement learning, federated aggregation, and continual QA recalibration. The conclusions are structured in accordance with the research tasks set forth in the introduction.

1. Regarding the first research task—to design a unified QA architecture that integrates reinforcement learning for adaptive control and federated learning for distributed, privacy-preserving calibration—the developed SEQA framework provides a cohesive and functional architecture. The framework successfully leverages reinforcement learning to autonomously optimize QA strategies in response to dynamic operational feedback, while the federated aggregation mechanism enables collaborative model improvement across distributed nodes without centralizing sensitive data. The architectural integration itself represents a key theoretical contribution, creating a viable foundation for autonomous quality evolution in software systems.

2. Concerning the second research task—to develop a robust cross-domain adaptation mechanism—the implemented mechanism has been validated through cross-domain evaluations, confirming robust generalization capability. The experiments demonstrate stable performance across heterogeneous operational settings, with reliability (R) maintained above 0.91 and F1-scores at or above 0.86. This outcome confirms that the mechanism effectively ensures consistent, energy-aware performance and maintains calibration trustworthiness across diverse domains and shifting data distributions, addressing a core limitation of static and single-domain models.

3. Pertaining to the third research task—to empirically validate the proposed SEQA framework and benchmark its performance—the extensive experiments on the NASA MDP, UCI, and OPSD datasets provide clear quantitative results. The framework consistently improves reliability (up to 0.94), enhances calibration accuracy (achieving an Expected Calibration Error of 0.024 ± 0.003), and reduces energy consumption by approxi-

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

mately 13% compared to baseline models. These empirical results substantiate the practical efficacy and superiority of the SEQA approach.

Prospects for further research stem directly from the scope and outcomes of this work. Future efforts should focus on: enhancing the interpretability and auditability of the reinforcement-learning-driven policy decisions to facilitate deployment in highly regulated industrial settings; extending the validation of the cross-domain adaptation mechanism to more heterogeneous and non-energy software systems, such as industrial manufacturing or consumer IoT environments [17]; and improving the scalability and resilience of the federated coordination layer under conditions of severe resource constraints or intermittent network connectivity. Further investigation into long-term evolution mechanisms for sustained performance over extended operational lifecycles, particularly for handling gradual concept drift, remains a crucial direction

Contributions of authors: conceptualization, methodology – **Andriy Verlan, Wang Zhihai;** formulation of tasks, analysis **– Andriy Verlan;** development of model, software,**– Wang Zhihai;** analysis of results, visualization **– Andriy Verlan;** writing – original draft preparation **- Wang Zhihai**, writing – review and editing **– Andriy Verlan**.

Conflict of Interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, author ship or otherwise, that could affect the research and its results presented in this paper.

Financing

This study was conducted without financial support.

Data Availability

Data will be made available upon reasonable request,

Use of Artificial Intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Acknowledgments

This research is conducted under the support of the Quality Assurance Project for Intelligent Energy Management Software Based on AI Methods (Project No. H20240943), the Development and Industrialization of Intelligent Grid Demand Response Technology Project, and the Foreign Expert Studio for Demand Response at the Shandong Institute of China-Ukraine Technology Innovation. It is affiliated with the Ukrainian National Technical University and the Shandong Institute of China- Ukraine Technology Innovation in a collaborative partnership.

All the authors have read and agreed to the published version of the manuscript.

References

1. Hall, T., Beecham, S., Bowes, D., Gray, D., & Counsell, S. A Systematic Literature Review on Fault Prediction Performance in Software Engineering. *IEEE Transactions on Software Engineering*, 2012, vol. 38, no. 6, pp. 1276–1304. DOI: 10.1109/TSE.2011.103.

2. Jureczko, M., & Madeyski, L. Towards Identifying Software Project Clusters with Respect to Defect Prediction. *Proceedings of the 6th International Conference on Predictive Models in Software Engineering, PROMISE, Timishoara, Romania*, 2010, article no. 9. pp. 1–10. DOI: 10.1145/1868328.1868342.

3. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. A survey on concept drift adaptation. *ACM Computing Surveys,* 2014, vol. 46, iss. 4, article no.44, pp.1–37. DOI: 10.1145/2523813.

4. Hendrycks, D., & Dietterich, T. Benchmarking Neural Network Robustness to Common Corruptions and Perturbations. *International Conference on Learning Representations (ICLR)*, New Orleans, Louisiana, United States, 2019. DOI: 10.48550/arXiv.1903.12261.

5. Kairouz, P., & McMahan, H. B. Advances and Open Problems in Federated Learning. *Foundations and Trends in Machine Learning*, 2021, vol. 14, iss. 1–2, pp. 1–210. DOI: 10.1561/2200000083.

6. Wei, T., Wang, Y., & Zhu, Q. Deep Reinforcement Learning for Building HVAC Control. *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments, BuildSys*, Austin, TX, USA, 2017, pp. 1–10. DOI: 10.1145/3061639.3062224.

7. Azuatalam, D., Lee, W.-L., de Nijs, F., & Liebman, A. Reinforcement Learning for Whole-building HVAC Control and Demand Response. *Energy and AI*, 2020, vol. 2, article no. 100020. DOI: 10.1016/j.egyai.2020.100020.

8. Pipattanasomporn, M., Kuzlu, M., & Rahman, S. An Algorithm for Intelligent Home Energy Management and Demand Response Analysis. *IEEE Transactions on Smart Grid*, 2012, vol. 3, iss. 4, pp. 2166–2173. DOI: 10.1109/TSG.2012.2201182.

9. Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Transactions on Smart Grid*, 2019, vol. 10, iss. 1, pp. 841–851. DOI: 10.1109/TSG.2017.2753802.

10. Lee, E. Cyber Physical Systems: Design Challenges. *11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, Orlando, FL, USA, 2008, pp. 363-369. DOI: 10.1109/ISORC.2008.25.

11. Gordieiev, O., Gordieieva, D., Rainer, A., Gorbenko, A., Tarasyuk, O. Quality Assessment of Artificial Intelligence Systems: A Metric-Based Approach. *Electronics*, 2026, vol. 5(3):691. DOI: 10.3390/electronics15030691.

12. Kharchenko, V., Fesenko, H., Illiashenko, O. Quality Models for Artificial Intelligence Systems: Characteristic-Based Approach, Development and Application. *Sensors,* 2022, vol. 22(13):4865. DOI: 10.3390/s22134865.

13. Felderer, M., & Ramler, R. Quality Assurance for AI-Based Systems: Overview and Challenges (Introduction to Interactive Session). *Software Quality: Future Perspectives on Software Engineering Quality, 13th International Conference SWQD*, Vienna, Austria, 2021, pp. 33-42. DOI: 10.1007/978-3-030-65854-0_3.

14. Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., & Zhu, X. X. A Survey of Uncertainty in Deep Neural Networks. *Artificial Intelligence Review*, 2023, vol. 56 (suppl. 1), pp. 1513-1598. DOI: 10.1007/s10462-023-10562-9.

15. Chen, X., Dong, W., & Yang, Q. Robust optimal capacity planning of grid-connected microgrid considering energy management under multi-dimensional uncertainties. *Applied Energy*, 2022, vol. 323, article no. 119642. DOI: 10.1016/j.apenergy.2022.119642.

16. McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. Communication-efficient learning of deep networks from decentralized data. in Proc. 20th Int. Conf. *Artificial Intelligence and Statistics (AISTATS)*, *PMLR*, 2017, vol. 54, pp. 1273-1282. Available at: https://proceedings.mlr.press/v54/ mcmahan17a.html. (accessed 12.08.2025).

17. Konečný, J., McMahan, H. B., Ramage, D., & Richtárik, P. Federated optimization: Distributed machine learning for on-device intelligence. *arXiv preprint* arXiv:1610.02527, 2016. DOI: 10.48550/arXiv.1610.02527.

18. Parisio, A., Rikos, E., & Glielmo, L. A model predictive control approach to microgrid operation optimization. *IEEE Transactions on Control Systems Technology*, 2014, vol. 22, no. 5, pp. 1813–1827. DOI: 10.1109/TCST.2013.2295737.

19. Pan, S. J., & Yang, Q. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, vol. 22, no. 10, pp. 1345-1359. DOI: 10.1109/TKDE.2009.191.

20. Nastoska, A., Jancheska, B., Rizinski, M., Trajanov, D. Evaluating Trustworthiness in AI: Risks, Metrics, and Applications Across Industries. *Electronics,* 2025, vol. 14(13):2717. DOI: 10.3390/electronics14132717.

21. Kingma, D. P., & Ba, J. Adam: A Method for Stochastic Optimization. *3rd International Conference for Learning Representations*, ICLR, San Diego, 2015. DOI: 10.48550/arXiv.1412.6980.

22. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. Federated Optimization in Heterogeneous Networks. *Proceedings of Machine Learning and Systems*, MLSys, 2020, pp. 429-450. DOI: 10.48550/arXiv.1812.06127.

23. Candanedo, L. M., & Feldheim, V. Accurate Occupancy Detection of an Office Room from Light, Temperature, Humidity and CO2 Measurements Using Statistical Learning Models. *Energy and Buildings*, 2016, vol. 112, pp. 28–39. DOI: 10.1016/j.enbuild.2015.11.071.

24. Yu, L., Alégroth, E., Chatzipetrou, P., Gorschek, T. Measuring the Quality of Generative AI Systems: Mapping Metrics to Quality Characteristics—Snowballing Literature Review. *Information and Software Technology,* 2025, vol. 186:107802. DOI: 10.1016/j.infsof.2025.107802.

25. Ganin, Y., & Lempitsky, V. Unsupervised Domain Adaptation by Backpropagation. *Proceedings of the 32nd International Conference on Machine Learning, ICML*, 2015, vol. 37, pp. 1180–1189. Available at: https://proceedings.mlr.press/v37/ganin15.html. (accessed 10.08.2025).

26. Verlan, A. A., Zhihai, W., & Yunhai, Z. Enhancing Reliability of Energy Management Software Through Predictive Modeling and Automated Repair. *Connectivity,* 2025, vol. 6, pp. 95–102, DOI: 10.31673/2412-9070.2025.061212.

27. Verlan, A. A., Zhi Hai, W., & Chen, C. Modelling the quality assurance of AI based intelligent energy management software. *Modern Information Security*, 2025, vol. 3 (63), pp. 199–204. DOI: 10.31673/2409-7292.2025.030192.

Received 17.10.2025, Received in revised form 22.11.2025
Accepted date 15.01.2026, Published date 22.01.2026

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

КОНЦЕПТУАЛЬНІ ЗАСАДИ СИСТЕМ САМОВДОСКОНАЛЮВАНОГО ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ДЛЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНОГО УПРАВЛІННЯ ЕНЕРГЕТИКОЮ НА ОСНОВІ ШІ

А. А. Верлань, Ванг Жихай

Предметом вивчення в статті є процеси проєктування та валідації самовдосконалюваної системи забезпечення якості (SEQA) для інтелектуального програмного забезпечення управління енергією (IEMS) на основі штучного інтелекту. **Метою статті** є розробка масштабованої та адаптивної системи SEQA, що забезпечує безперервну оптимізацію та надійність функціонування в динамічних, гетерогенних операційних середовищах. **Завдання:** формалізувати уніфіковану архітектуру забезпечення якості, що інтегрує навчання з підкріпленням для адаптивного керування та федеративне навчання для розподіленого калібрування; розробити надійний механізм крос-доменної адаптації, що забезпечує енергоефективне калібрування довіри; емпірично підтвердити ефективність запропонованої системи порівняно з базовими моделями на кількох реальних енергетичних наборах даних. **Використовуються такі методи:** навчання з підкріпленням для оптимізації на основі політик, федеративне навчання для агрегації моделей із збереженням конфіденційності, методи калібрування довіри для оцінювання надійності та експериментальне тестування на наборах даних NASA, UCI та OPSD. **Отримані результати статті:** запропонована система SEQA успішно інтегрує локальну адаптацію на основі навчання з підкріпленням із федеративною агрегацією політик, досягаючи безперервного самовдосконалення показників якості в гетерогенних сценаріях управління енергією; механізм крос-доменної адаптації забезпечує надійну здатність до узагальнення з показниками F1 понад 0.86 та надійністю вище 0.91 у різноманітних операційних умовах; експериментальна валідація демонструє стабільне підвищення надійності (показник F1 зростає на 6–8%), точності калібрування (очікувана помилка калібрування зменшується до 0.024) та енергоефективності (до 13%) порівняно з базовими моделями забезпечення якості; система зберігає стабільну продуктивність за умов динамічних розподілів даних, при цьому абляційні дослідження підтверджують, що кожен компонент — навчання з підкріпленням, федеративна еволюція та безперервне відтворення — відіграє критичну роль у забезпеченні надійного самовдосконалюваного контролю якості. **Висновки.** Наукова новизна отриманих результатів полягає в наступному: 1) запропонована система SEQA впроваджує уніфіковану адаптивну парадигму, яка синергетично поєднує навчання з підкріпленням, федеративне калібрування та крос-доменну адаптацію, забезпечуючи автономну безперервну еволюцію якості в IEMS; 2) розроблений механізм крос-доменної адаптації досягає надійного узагальнення та енергоефективного балансування продуктивності, усуваючи ключові обмеження статичних та однодоменних підходів до забезпечення якості; 3) масштабна експериментальна валідація демонструє стабільне підвищення надійності, точності калібрування та енергоефективності, підтверджуючи практичну застосовність системи для довгострокового промислового впровадження; 4) інтеграція механізмів адаптивних інтервалів агрегації та скорочення політик мінімізує надлишковість під час синхронізації, зберігаючи майже лінійну масштабованість на розподілених федеративних вузлах, що підтверджує доцільність розгортання системи в реальних середовищах інтелектуальних мереж та промислового Інтернету речей.

Ключові слова: штучний інтелект; програмне забезпечення для управління енергетикою; самовдосконалювана гарантія якості; федеративне навчання; навчання з підкріпленням; надійність; калібрування.

Верлань Андрій Анатолійович – д-р техн. наук, проф. каф. інженерії програмного забезпечення в енергетиці, Навчально-науковий інститут атомної та теплової енергетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна;
ассоц. проф., факультет технічних наук, Норвезький університет науки та технології, Норвегія.

Ванг Жихай – асп. каф. інженерії програмного забезпечення в енергетиці, Навчально-науковий інститут атомної та теплової енергетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна;
Університет Сіньхуа в Гуанчжоу, Гуанчжоу, Китай.

Andriy Verlan – Doctor of Technical Sciences, Professor, Department of Software Engineering for Power Industry, Educational and Scientific Institute of Atomic and Thermal Energy, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine,
Associate Professor, Faculty of Engineering, Norwegian University of Science and Technology, Norway;
e-mail: a.verlan@edu.kpi.ua, ORCID: 0000-0002-6469-2638, Scopus ID: 57192176307.

Wang Zhihai – PhD Candidate, Department of Software Engineering for Power Industry, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine;
Guangzhou Xinhua University, Guangzhou, China,
e-mail: 1053226186@qq.com, ORCID: 0009-0004-4979-5251.