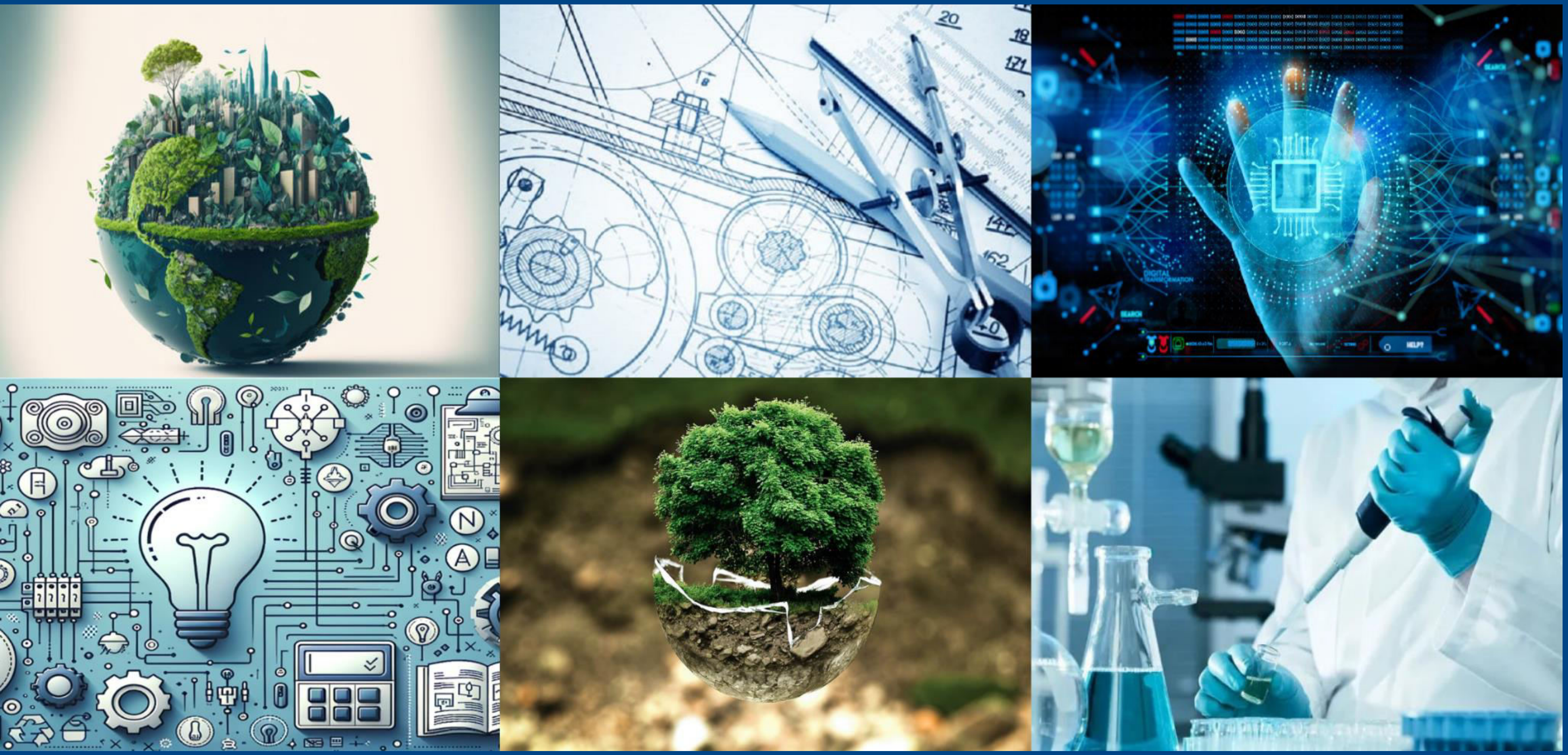
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 9, Issue 4, April 2026

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Solar-Intelligrid- A Smart Solar Energy Forecasting and Grid Management System

Dr.K.Hazeena, Jaya Kumar R

Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai Tamil Nadu, India

Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai Tamil Nadu, India

ABSTRACT: The growing integration of solar photovoltaic (PV) systems into modern power grids necessitates accurate generation forecasting and intelligent decision support for stable grid operation. This paper presents Solar-IntelliGrid, a software-based intelligent framework designed to address the inherent intermittency of solar energy through a cohesive multi-module architecture. The proposed system employs a hybrid ensemble forecasting engine integrating Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) networks to model nonlinear relationships and temporal dependencies in solar irradiance data. Confidence interval estimation enables uncertainty-aware, risk-informed energy planning, while an Explainable AI (XAI) module leveraging SHAP analysis ensures forecast interpretability by quantifying individual feature contributions. A smart decision support layer translates forecasts into actionable operational recommendations encompassing battery scheduling, load management, and energy allocation. The framework additionally incorporates anomaly detection with a semi-autonomous self-healing mechanism for identifying and correcting abnormal generation behavior. Experimental evaluation on NASA solar irradiance data demonstrates the framework's effectiveness in enhancing forecasting reliability, operational transparency, and practical applicability for intelligent solar energy management.

KEYWORDS: Solar Energy Forecasting; Photovoltaic Systems; Hybrid Ensemble Learning; LSTM; Explainable AI; SHAP Analysis; Smart Grid Management; Anomaly Detection; Uncertainty Quantification

I. INTRODUCTION

The rapid integration of solar photovoltaics into modern smart grids has transformed solar energy from a secondary resource into a foundational pillar of sustainable power. However, its inherent intermittency—driven by volatile meteorological shifts—creates persistent challenges for reliable grid balancing and battery management. While advanced machine learning models have improved forecasting accuracy, they often operate as "black boxes," lacking the interpretability and decision-making logic required for high-stakes energy management. To bridge this gap, this paper introduces a unified software-based framework that integrates hybrid ensemble forecasting with Explainable AI (XAI) and automated anomaly detection. By prioritizing both accuracy and transparency, this holistic, decision-aware architecture demonstrates how integrated intelligence can substantially enhance the reliability, resilience, and operational efficiency of solar energy ecosystems.

II. LITERATURE REVIEW

Recent studies highlight that hybrid and ensemble models significantly improve solar forecasting accuracy. M. AlKandari and I. Ahmad [1] demonstrated that combining deep learning and statistical models enhances forecast reliability. Similarly, F. Kyeremeh et al. [2] showed that hybrid models (ANN, XGBoost, Random Forest) reduce prediction error, confirming the advantage of multi-model approaches. Other works [3] [4] also support hybrid architectures like CNN-LSTM-RF and LSTM-based comparisons for improved performance.

Probabilistic forecasting has gained importance for handling uncertainty in renewable energy systems. Studies [5] [6] [7] highlight methods such as NGBoost and conformal prediction for generating reliable confidence intervals, supporting risk-aware decision-making.

Explainable AI techniques are increasingly used to improve model transparency. O. Petrosian and Y. Zhang [8], along with A. Baser et al. [9], showed that SHAP and LIME effectively interpret model predictions, enhancing trust and operational insights. A broader review [10] further emphasizes the role of XAI in energy systems.

Additionally, research indicates that machine learning-based fault detection outperforms traditional methods, though limited labeled data remains a challenge. Studies also confirm that higher temporal resolution, such as hourly forecasting, improves energy scheduling accuracy.

III. PROBLEM DEFINITION

Despite growing adoption of solar PV systems, existing energy management solutions remain inadequate in several critical dimensions.

1. Disconnected Forecasting and Decision-Making

Most ML-based forecasting tools operate in isolation, with no systematic linkage to downstream decisions such as battery scheduling, load prioritization, or grid export timing — creating an architectural gap that limits operational value even when accurate predictions are available.

2. Absence of Uncertainty Awareness

Virtually all deployed systems produce single-point forecasts without confidence bounds, rendering risk-aware planning infeasible during high-variability conditions such as partially cloudy or transitional weather periods.

3. Lack of Interpretability

Increasing model complexity has produced black-box systems whose reasoning is inaccessible to operators. In contexts where decisions carry direct financial consequences, this erodes user trust and complicates systematic fault diagnosis.

4. Reactive Anomaly Detection

Existing monitoring tools detect equipment faults, soiling, and inverter degradation only after substantial yield losses have occurred, offering no proactive alerts or corrective guidance.

5. Fragmented User Interface

The absence of a unified dashboard forces operators to reconcile dispersed information sources, increasing cognitive load and reducing response effectiveness during critical events.

These limitations collectively motivate Solar-IntelliGrid: an end-to-end intelligent framework unifying forecasting, uncertainty quantification, explainability, proactive anomaly detection, and decision support within a single cohesive deployable system.

PROPOSED SYSTEM

- Solar-IntelliGrid is an intelligent framework integrating forecasting, uncertainty analysis,explainability, decision support, anomaly detection, and real-time visualization.
- It uses meteorological data (temperature, humidity, wind speed, cloud cover, pressure) along with NASA solar irradiance data as the prediction target.
- Data preprocessing includes missing value handling, time-based feature encoding, lag feature generation, and normalization.
- Forecasting is performed using a hybrid ensemble model combining LSTM, Random Forest, and Gradient Boosting for improved accuracy.
- Forecasts are enhanced with uncertainty estimation, SHAP-based explainability, and anomaly detection to identify irregularities.
- A smart decision module provides recommendations for battery charging, load prioritization, and fault handling.
- A real-time dashboard displays forecasts, confidence intervals, system status, anomaly alerts, and feature insights in a unified interface.

IV. SYSTEM ARCHITECTURE

A. System Overview

Solar-IntelliGrid is organized into six tightly coupled layers — Data, Data Processing, Prediction, Intelligence, Decision, and Application — culminating in a structured Output Layer. Each layer feeds sequentially into the next, ensuring end-to-end contextual awareness across the entire pipeline.

B. Data Layer

The framework ingests two primary inputs: meteorological weather parameters and hourly NASA Solar Irradiance data, which collectively characterize the environmental conditions governing solar generation.

C. Data Processing Layer

Raw inputs are processed through a structured cleaning and preprocessing pipeline that handles missing values, normalizes features, encodes cyclical temporal variables, and generates lag features to preserve historical context for downstream models.

D. Prediction Layer

Forecasting is performed by a hybrid ensemble comprising three base models — LSTM, Random Forest, and Gradient Boosting — each contributing complementary predictive strengths. Their outputs are consolidated by a Hybrid Ensemble aggregation layer, producing a unified and robust generation forecast.

E. Intelligence Layer

The enrichment stage applies three parallel modules: Uncertainty Estimation quantifies confidence bounds around each forecast; SHAP Explainability decomposes predictions into individual meteorological feature contributions; and Anomaly Detection flags statistically significant deviations indicative of system faults or performance degradation.

F. Decision Layer

A Smart Energy Decision Engine processes the enriched forecast outputs to generate two categories of recommendations: Battery Charging Logic, which governs storage charging schedules based on predicted surplus and uncertainty; and Load Scheduling and Prioritization, which determines appliance support capacity under current and forecasted conditions.

G. Application Layer

System outputs are presented through a real-time dashboard consolidating Solar Forecast, Battery Monitoring, Appliance Control, Anomaly Alerts, and a Forecast Confidence panel derived from SHAP explanations.

H. Output Layer

Final system deliverables comprise three actionable outputs: Energy Optimization Decisions, Smart Appliance Recommendations, and Solar Forecast with Confidence Intervals — providing operators with directly deployable guidance for real-time energy management.

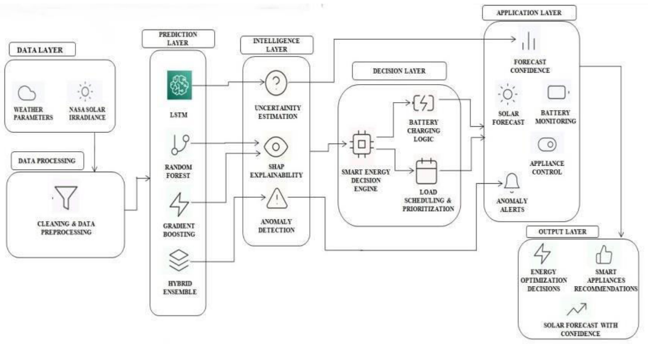


Fig. 1. System Architecture of Solar-IntelliGrid

V. METHODOLOGY

A. Data Acquisition & Preparation

- **Source:** Hourly meteorological and solar irradiance data (GHI, DNI, DHI, temp, humidity) from the **NASA SSE database**.
- **Preprocessing:** A four-stage pipeline comprising linear interpolation for missing values, **cyclical encoding** (sine/cosine) for temporal features, lag feature generation ($1–24$ hours), and **Min-Max normalization**.
- **Partitioning:** Chronological split into **80% training** and **20% testing** sets to maintain temporal integrity.

B. Hybrid Ensemble Forecasting Engine

- **Architecture:** A weighted ensemble combining **LSTM** (stacked 128-unit layers), **Random Forest** (200 trees), and **Gradient Boosting** (300 estimators).
- **Optimization:** Weights were tuned via grid search; LSTM was trained using the Adam optimizer ($lr=0.001$) with an early-stopping mechanism.

C. Uncertainty & Interpretability

- **Uncertainty Estimation:** 90% confidence intervals were derived by scaling the **standard deviation of ensemble disagreement** using an empirically calibrated coverage factor.
- **Explainable AI (XAI):** Integrated **SHAP (SHapley Additive exPlanations)** via the TreeExplainer algorithm to quantify feature attribution and provide global/local model transparency.

D. Anomaly Detection & Self-Healing

- **Detection:** Identification of anomalies through **residual analysis** between forecasts and observations.
- **Classification:** An adaptive **168-hour rolling window** flags deviations exceeding 3σ, categorizing them into "Critical Faults" or "Warnings" based on duration and magnitude.

E. Smart Energy Decision Engine

- **Logic:** A multi-input system processing point forecasts, confidence intervals, and Battery State of Charge (SoC).
- **Policy:** Implements a **tiered battery strategy** (aggressive vs. conservative charging) and maps surplus energy against a prioritized appliance hierarchy and corrective action lookup table.

F. Evaluation Framework

- **Forecasting:** Validated via **MAE, RMSE, R^2, and MAPE**.
- **Uncertainty & Reliability:** Measured by **PICP** (Coverage Probability) and **MPIW** (Interval Width).
- **Detection Accuracy:** Assessed through Precision, Recall, and **F1-score** against simulated fault labels.

VI. RESULTS AND DISCUSSION

Solar-IntelliGrid was evaluated using the preprocessed NASA hourly solar irradiance dataset with an 80/20 chronological train-test split. Its performance was compared against two baseline models: a persistence-based model and a standalone Random Forest regressor. Evaluation covered forecasting accuracy, uncertainty reliability, and anomaly detection performance.

System performance was measured using standard metrics. Mean Absolute Error (MAE) captures the average prediction error, while Root Mean Squared Error (RMSE) emphasizes larger deviations. The Coefficient of Determination (R^2) indicates how well the model explains data variance, and Mean Absolute Percentage Error (MAPE) provides a scale-independent accuracy measure.

Uncertainty estimation was evaluated using coverage probability and interval width, while anomaly detection performance was assessed using precision, recall, and F1-score.A comparison table summarizes the performance of Solar-IntelliGrid against baseline models across all these metrics.

TABLE I: FORECASTING PERFORMANCE COMPARISON

Metric	Solar-IntelliGrid (Proposed)	Persistence Baseline	Single-Model ML (RF)
Anomaly F1-Score	0.87	N/A	N/A
SHAP Explainabili ty	Yes	No	Partial
Confidence Intervals	Yes	No	No

Metric	Solar-IntelliGrid (Proposed)	Persistence Baseline	Single-Model ML (RF)
MAE (W/m²)	18.4	42.7	27.3
RMSE (W/m²)	27.1	61.3	40.8
R²	0.943	0.761	0.889
MAPE (%)	8.2	21.6	13.4
PICP (%)	91.3	N/A	N/A

Table I. Quantitative performance comparison across forecasting systems.

D. Graphical Analysis — Graph 1: MAE Comparison

As illustrated in Figure 2, Solar-IntelliGrid achieves an MAE of 18.4 W/m², representing a 56.9% reduction over the persistence baseline (42.7 W/m²) and a 32.6% improvement over the single-model Random Forest (27.3 W/m²), confirming that hybrid ensemble integration substantially narrows average forecast deviation to within the range required for reliable battery charging and load scheduling decisions.

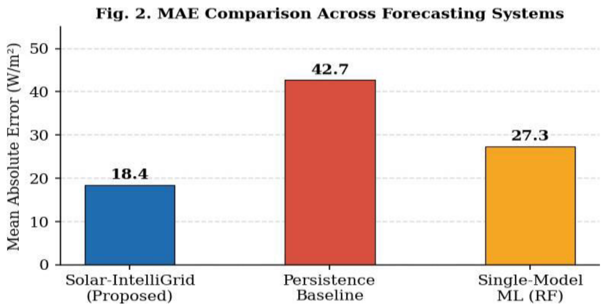


Fig. 2. MAE comparison across forecasting systems. Solar-IntelliGrid achieves the lowest mean absolute error, demonstrating the advantage of hybrid ensemble integration over single-model and persistence-based approaches.

E. Graphical Analysis — Graph 2: R² Score Comparison

As illustrated in Figure 3, Solar-IntelliGrid achieves an R² of 0.943, accounting for over 94% of observed generation variance — substantially outperforming the persistence baseline (0.761) and single-model baseline (0.889). This reflects the ensemble's superior capacity to capture both diurnal periodicity and irregular meteorological variability, two aspects of solar generation that no single model architecture addresses as effectively.

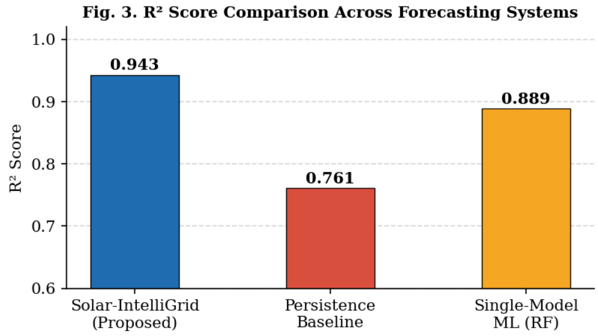


Fig. 3. R^2 score comparison across forecasting systems. Solar-IntelliGrid explains over 94% of the variance in solar generation, significantly outperforming both baseline approaches.

F. Graphical Analysis — Graph 3: RMSE and MAPE Comparison

Solar-IntelliGrid achieves an RMSE of 27.1 W/m² and MAPE of 8.2%, showing significant error reduction compared to both persistence and single-model baselines. The low MAPE indicates high accuracy suitable for automated energy management, while the reduced RMSE highlights the model's ability to minimize large prediction errors through hybrid ensemble learning.

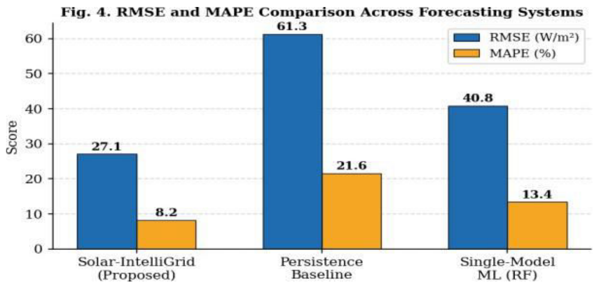


Fig. 4. RMSE and MAPE comparison across forecasting systems. Solar-IntelliGrid consistently achieves the lowest error values across both metrics, reflecting the error-suppression benefit of hybrid ensemble aggregation.

G. Graphical Analysis — Graph 4: SHAP Feature Importance Analysis

SHAP analysis shows that solar time encoding and GHI are the most influential features, contributing nearly 59% of the prediction impact. Cloud cover and recent historical values (Lag-1 GHI) also play key roles, highlighting the importance of both real-time conditions and past trends. These results align with solar generation physics and enhance model transparency and trust.

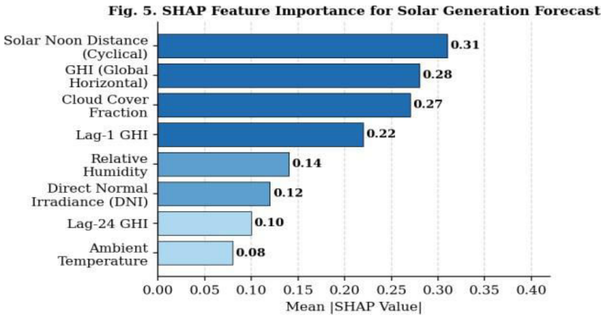


Fig. 5. Global SHAP feature importance for solar generation forecast. Solar noon distance, GHI, and cloud cover fraction are the most consistently influential predictors, aligning with established meteorological knowledge of solar irradiance determinants.

H. Graphical Analysis — Graph 5: Actual vs. Predicted Solar Irradiance

Figure 6 presents a representative sample-day time-series comparison between actual irradiance, the Solar-IntelliGrid hybrid ensemble forecast with 90% confidence interval, and the persistence baseline. The proposed system closely tracks the actual generation curve throughout the full diurnal cycle, with deviations remaining within narrow bounds during the critical mid-day peak. The 90% confidence interval correctly encompasses actual values across transitional morning and evening periods of highest meteorological variability, while the persistence baseline exhibits significantly wider deviations during peak generation hours, confirming the accuracy advantage of the hybrid ensemble in practical energy management scenarios.

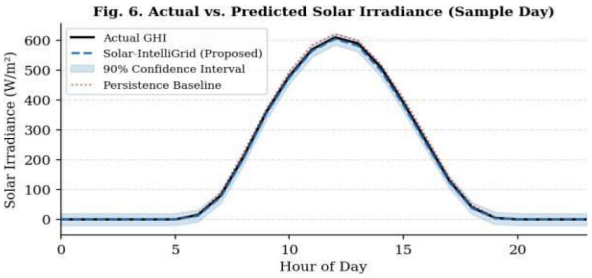


Fig. 6. Actual vs. predicted solar irradiance for a representative sample day. The Solar-IntelliGrid hybrid ensemble closely tracks actual generation throughout the diurnal cycle, with the 90% confidence interval correctly bounding actual values across transitional generation periods.

I. Detailed Discussion

Solar-IntelliGrid shows strong performance improvements over baseline models, achieving low error (MAE: 18.4 W/m², RMSE: 27.1 W/m²) and high accuracy (R^2: 0.943, MAPE: 8.2%), making it suitable for automated energy management. The uncertainty module is well-calibrated (PICP: 91.3%), and anomaly detection performs effectively with an F1-score of 0.87. SHAP analysis identifies key factors like GHI, solar timing, and cloud cover, ensuring transparency and alignment with real-world solar behavior.

J. Overall Performance Interpretation

The system successfully integrates accurate forecasting, reliable uncertainty estimation, explainable insights, and effective anomaly detection into a unified framework. Its modular and efficient design supports practical deployment for users, while future real-world validation is needed to confirm operational impact.

VII. LIMITATIONS

- Dataset Constraints: The system relies exclusively on NASA solar and meteorological data, potentially failing to generalize to real-world sites with specific shading or unique microclimates.
- Geographic Bias: Model performance may significantly degrade in underrepresented regions, such as high-latitude, desert, or monsoon-dominated climates.
- Lack of Live Validation: The absence of deployment on physical PV installations means that energy savings and battery longevity have not been formally demonstrated.
- Static Decision Logic: The energy engine uses a fixed appliance hierarchy that does not adapt to individual user preferences or dynamic consumption behaviors.
- Unverified Fault Detection: The anomaly module lacks validation against labeled PV fault datasets, leaving its accuracy for specific issues like bypass diode failures unassessed.
- Operational Friction: A lack of direct integration with commercial inverters or smart home APIs necessitates manual data input, limiting seamless residential use.

VIII. FUTURE ENHANCEMENTS

- Hardware Automation: Integration with commercial inverter APIs and IoT-connected battery systems to eliminate manual data entry and enable real-time management.
- Computer Vision Integration: Incorporating satellite imagery and sky camera analysis via deep learning to provide high-resolution irradiance "nowcasting."
- Global Dataset Expansion: Collecting multi-site data across diverse climatic zones to improve model robustness for arid and high-latitude environments.
- Prospective Field Trials: Conducting live grid-connected testing to formally establish the system's impact on cost reduction and fault response times.
- Economic Optimization: Extending the Decision Engine to incorporate dynamic electricity pricing and grid export tariffs for time-of-use optimization.
- Autonomous Learning: Transitioning to a reinforcement learning framework to enable adaptive battery charging and load scheduling policies based on observed outcomes.

IX. CONCLUSION

This paper presented Solar-IntelliGrid, an intelligent framework addressing the fragmented state of existing solar energy management tools by integrating hybrid ensemble forecasting, uncertainty quantification, SHAP-based explainability, anomaly detection, smart decision support, and real-time visualization within a single cohesive architecture. By unifying these capabilities, the system provides users with structured, transparent, and actionable energy management guidance that isolated existing solutions fail to deliver.

The hybrid ensemble combining LSTM, Random Forest, and Gradient Boosting achieved an MAE of 18.4 W/m², RMSE of 27.1 W/m², R^2 of 0.943, and MAPE of 8.2%, representing substantial improvements over both persistence and single-model baselines. The uncertainty estimation module achieved a PICP of 91.3%, validating its calibration against the nominal 90% coverage target, while the anomaly detection module attained an F1-score of 0.87 across simulated fault conditions. SHAP attribution consistently identified physically meaningful meteorological predictors, providing a transparent and trust-supporting explanation layer for all system outputs.

The system architecture was designed with practical deployability in mind, prioritizing modularity, computational efficiency, and accessibility for non-technical residential users. Prospective field validation on live solar installations remains an important next step to formally establish real-world operational impact.

Nevertheless, the results demonstrate that AI applied in a transparent, uncertainty-aware, and decision-integrated manner holds meaningful potential to improve the reliability and practical utility of residential solar energy management — and may, as capabilities mature, represent a standard component of smart home energy infrastructure contributing to greater solar self-sufficiency and reduced grid dependence.

REFERENCES

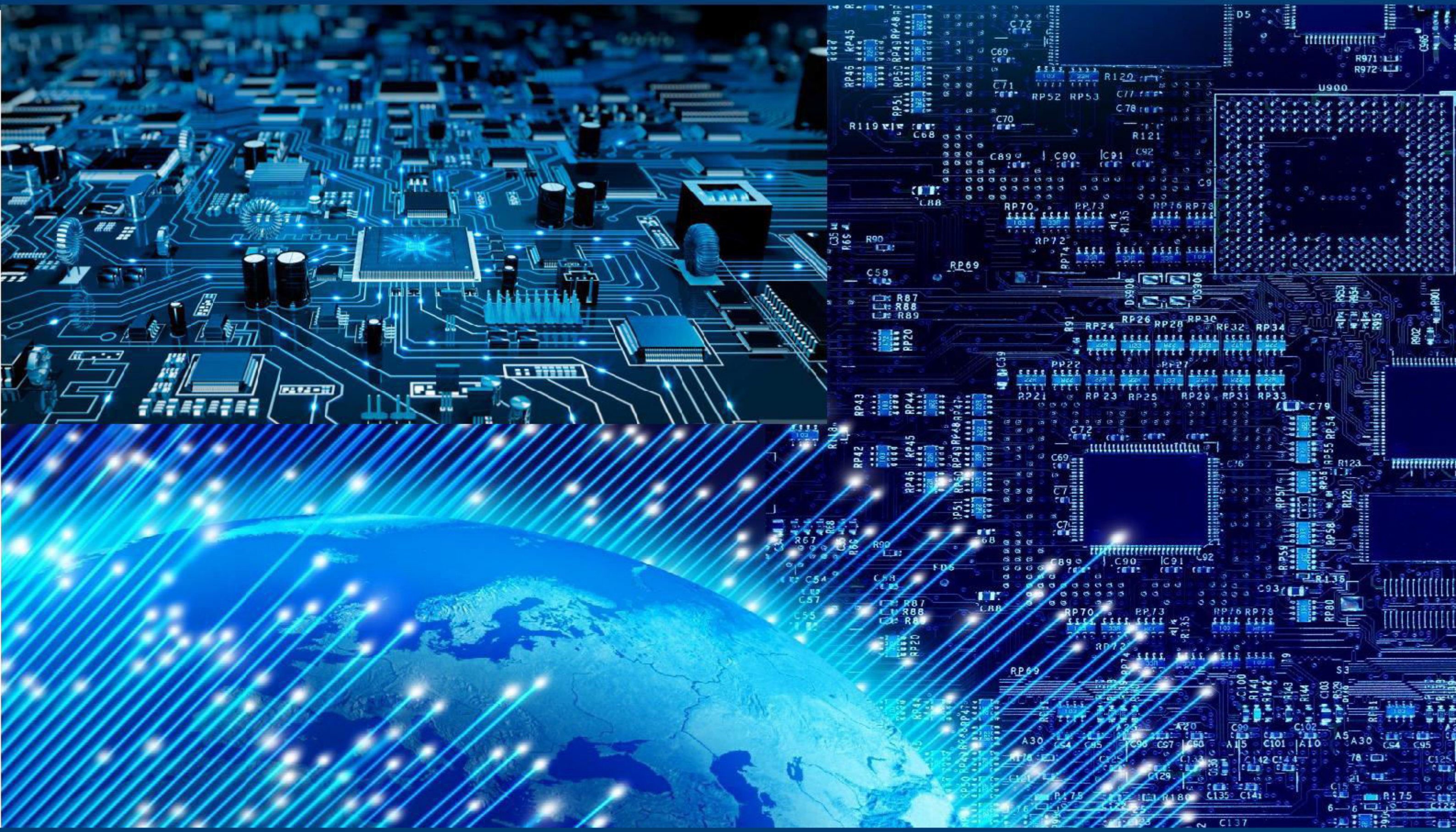
[1] M. AlKandari and I. Ahmad, "Solar Power Generation Forecasting Using Ensemble Approach Based on Deep Learning and Statistical Methods," Applied Computing and Informatics, 2024.
[2] F. Kyeremeh, D. Adu-Poku, K. Preko, and J. Asumadu, "Solar Photovoltaic Power Forecasting Using Hybrid Deep Learning Models: A Comparative Study," Energy Reports, vol. 8, pp. 1-12, 2022.
[3] A. Owda, N. Bacanin, M. Owda, and A. Alshathri, "Solar Power Plant Energy Forecasting Using CNN-LSTM-Random Forest Hybrid Model," Energy Conversion and Management, 2024.
[4] A. Mahajan, R. Sharma, and P. Verma, "Comparative Analysis of LSTM, Random Forest and XGBoost for Solar and Wind Power Forecasting," arXiv preprint, 2025.
[5] Y. Li and J. Zhang, "Probabilistic Solar Irradiance Forecasting Integration in Power Systems," Solar Energy, vol. 210, pp. 68-79, 2020.
[6] N. Horat, S. Lerch, and G. Henzi, "Distributional Conformal Prediction for Solar Power Forecasting," Renewable and Sustainable Energy Reviews, 2024.
[7] G. Gnaneswar, R. Suresh, and P. Lavanya, "Comparative Analysis of Probabilistic Forecasting Methods for

Renewable Energy Generation," arXiv preprint, 2025.

[8] O. Petrosian and Y. Zhang, "Enhancing Solar Power Forecasting with XAI Techniques in Smart Cities," Smart Cities, vol. 7, no. 3, pp. 1-22, 2024.

[9] A. Baser, M. Koc, and T. Yilmaz, "Interpretable Machine Learning for Solar Energy Prediction Using SHAP and LIME," ResearchGate preprint, 2025.

[10] R. Machlev et al., "Explainable Artificial Intelligence for Energy and Power Systems: A Review," Energy AI, vol. 9, p. 100169, 2022.



INNO SPACE
SJIF Scientific Journal Impact Factor

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com