International Journal of Engineering Applied Science and Management
ISSN (Online): 2582-6948
Vol. 7 Issue 7, July 2026

Real-Time IOT-Based Monitoring and Fault Detection for Solar Photovoltaic Systems: An Experimental Field Study

Rajkumari Ahirwar[1], Ashish Suryavanshi[2]
Research Scholar, Department of Electronics and Communication Engineering, Samrat Vikramaditya Vishwavidyalaya, Ujjain, Madhya Pradesh, India[1]
Head of Department, Department of Electronics and Computer Science, Samrat Vikramaditya Vishwavidyalaya, Ujjain, Madhya Pradesh, India[2]
Email- suryavanshi.ashish5@gmail.com

Abstract

Thus, the urgent need for intelligent real-time monitoring solutions that can boost energy yield, reduce downtime and enable predictive maintenance is being created by the rapid proliferation of solar photovoltaic (PV) plants. This empirical study describes the design, implementation and performance assessment of the IoT-based smart monitoring systems that have been developed to collect and store the performance data of grid-connected as well as off-grid solar PV installations. distributed sensor nodes measuring irradiance, module temperature, voltage, current and power output measured using existing cloud based data aggregator, hosted using a web dashboard and mobile application in the proposed system A twelve-month field deployment was undertaken on five PV plants at different scales (5 kWp to 500 kWp) in diverse climatic zones of India. Empirical data gathered from 14,400 hourly observation points underwent descriptive statistics, regression modeling, and fault detection accuracy analyses. Results show that the IoT monitoring framework successfully achieved an average fault detection accuracy of 94.7%, 18.3% reduction in energy losses due to undetected faults, and 6.2 percentage points improvement of the overall performance ratio of the plant when compared to baseline non-monitored conditions. The work defines IoT-based continuous monitoring as a feasible and financially viable practice to modernise current solar PV asset management, with implications in both the utility-size and rooftop PV domain. The results also set a reference for ten embedded systems, using five of them from the state of the art, reaffirming the system's accuracy and lower communication latency.

Keywords: *IoT monitoring, solar photovoltaic, smart energy systems, fault detection, performance ratio, cloud computing, wireless sensor network*

1. Introduction

The solar photovoltaic technology, as one of the most strategic technologies to make sustainable electricity generation came from the global energy transition towards renewable sources The International Energy Agency (IEA) notes that combined, solar PV constituted around 1,053 GW of installed capacity across the globe as of 2022, while India on the other hand has ambitious goals, as its Nationally Determined Contributions (NDC) to the Paris Agreement targets 500 GW of non-fossil fuel-based generation by 2030. The rapid growth of PV deployment, across utility-scale parks, industrial rooftop installations and rural microgrids, has also brought on numerous challenges in operations. In contrast, typical thermal power stations are centralized in nature and solar PV installations are distributed and heterogeneous in ambience (geographical and other characteristics are very variable, location-based, and environmental forces). This results in manual inspection regimes and fixed period maintenance schedules which are insufficiently economic and technical to ensure best performance from large PV portfolios. Due to the lack of continuous, automated monitoring, faults go undetected such as partial shading, soiling buildup, bypass diode failure, and maximum power point tracker (MPPT) drift, all of which contribute to quantifiable energy yield losses and equipment degradation.

A. Background and Motivation

The Internet of Things paradigm (IoT) in which sensing devices embedded in hardware are endowed with networking capability so that they can acquire data in real time, transmit it for computational analysis in the cloud and respond to control instructions, provides a disruptive solution to the monitoring challenge encountered by traditional PV plant management practices. By monitoring electrical and meteorological parameters at the sub-machine level, IoT-based frameworks allow operators to motivate considerations from rapid evidence between specific devices, local to faults and dyne to develop and deploy complex market procedures. While theoretic benefits of monitoring vegetables with IoT technology have been promoted, empirical data measuring the impact of these systems on performance under real-world conditions and in multiple sites has been published comparatively less often. The majority of existing investigations either rely on simulation or are limited to small-scale prototype evaluation in the laboratory, which leaves a substantial gap between theoretical hardware capability and demonstrated performance in the field. This work fills the gap by deploying and meticulously testing an IoT monitoring architecture designed specifically for PV plants at five working facilities across India, and collects one-year of continuous ground data and deep statistical analysis of the results.

B. Scope and Objectives

This research has four main objectives. First, to establish a scalable and low-cost IoT monitoring architecture appropriate for different scales and configurations of PV plants. Second, the proposed system will be implemented at five field sites, at the level of geographical areas/ typologies, reflecting the overall climatic zone and installation typology. Third, to gather and study real-world performance data stressing electrical parameters, environmental conditions, fault events, and system response times against a twelve month operational period. Fourth, for comparing the fault detection accuracy, energy loss reduction and communication reliability with the state-of-the-art systems published in the recent peer-reviewed literature. The study is deliberately empirical and field-based, based on hard data rather than projections from a simulation model. The systems described in this work incorporate readily available commercially available IoT hardware components including ESP32 microcontrollers, RS485 Modbus-RTU communication interfaces, and cloud connectivity via an MQTT protocol making them reproducible and economically accessible.

C. Paper Organization

The rest of this paper is structured as follows. Related Work Section II provides a comprehensive literature survey of the state of the art of IoT applications of solar PV monitoring. In section III, we describe the methodology which includes system architecture, hardware selection and the protocols for deployment. Data collection procedures are specified in Section IV, comprising five analytical data tables discussing system parameters, environmental parameters, electrical performance, fault events and economic impact. Section V covers statistical analysis and results discussion, including regression modeling and comparative benchmarking with prior studies. Placed at the end, we write Section VI which summarizes the paper and suggests future research avenues.

2. Literature Survey

The use of IoT technologies for solar PV monitoring has been a field of academic research since roughly 2014, but has grown sharply as a topic in the published literature since about 2017, spurred on by developments of standardised low-power wireless communications and cloud BM robotics and IoT-inflammatory responses; nearly a stop along the añejo national park rtf file! These early contributions were mostly qualitative and aimed studding the technical feasibility rather than benchmarking. For example Sharma and Bhatt [1] reported one of the earliest designs of a low-cost Arduino-based standalone PV remote monitoring system, which establishes the technical feasibility of low-chip embedded controllers capable of acquiring the current-voltage (I-V) curves and transmitting them wirelessly. They reported a measured voltage with an accuracy of +/− 1.5% but only on single modules with the inability to scale to plant level deployments. In a analogous work, Rao and Mishra [2] modelled a GSM-based monitoring prototype for rural solar microgrids in Rajasthan, reporting success with GPRS link data transmission but limited by high cost per message and over 45 second communication delays, making near real-time fault detection unfeasible.

The switch from GSM/GPRS in conjunction with the Wi-Fi and LPWAN communication paradigms has greatly expanded the technical scope of IoT monitoring systems. Kamienski et al. LoRaWAN Based Large-Area PV Farm MonitoringGGa6GgMaGEgGaB2MaC1→ X: 15 km line-of-sight/PLR < 2%, (LoRaWAN = viable utility-scale backhaul for rural installations) [3] Conversely, Ramadan et al. Out of existing radio technologies for IoT, Wi-Fi-based MQTT communication was found to perform more superiorly compared to LoRaWAN in case of urban environments for rooftop PV systems with grid-connected inverters, owing to lower latency (mean round-trip time 180 ms compared to 1200 ms for LoRaWAN) and similar cost, making it recommended for use in dense urban deployments, suitable for the urban PV area if a reliable Wi-Fi infrastructure is available [4]. It highlights the inescapable site-specific nature of optimal communication technology selection, a premise that informed the heterogeneous communication architecture used here.

Among them, Fault detection and diagnosis (FDD) algorithms are identified as one of the most rapidly growing subfield of IoT-based PV monitoring research. Zhao et al. Sample 1: Ref. [5] used two machine learning classifiers (Random Forest and Support Vector Machine (SVM)) to classify the faults of time-series data obtained from a 1 MWp PV plant in China, with a fault classification accuracy of 91.3% over six fault categories (open-circuit, short-circuit, shading, soiling, etc). Karatepe and Hizal [6] proposed an ANN-based approach reported an accuracy of 89.7%, but required more than 6,000 labeled observations per fault class for the network to be sufficiently trained, this limits the practicality of such designed network under newly commissioned solar power plants with no historical records of faults. More recently, Ahmed et al. To go a step further, [7] provided a rule-based anomaly detection framework using statistical process control charts with 88.4% fault detection accuracy for a 50 kWp rooftop system consuming much less computation power than machine learning approaches, thus addressing edge computing nodes with strict computing resources.

Cloud integration and data analytics architectures for monitoring of PV has received substantial academic focus. Pierro et al. In [8], the authors assessed the performances of three cloud platforms (i.e. AWS IoT, Google Cloud IoT Core, and Azure IoT Hub) in terms of scalability and latency characteristics on an application of PV monitoring and measured that MQTT message delivery latencies were typically between 95 ms (AWS IoT) and 210 ms (Azure IoT Hub), depending on cloud platform, under the simulated operational loads. Dash et al. The study proposed by [9] implemented a digital twin (DT) an advanced cloud-based framework for a 200 kWp industrial PV system and showed that the continual comparison of measured output to model-predicted performance allowed gradual efficiency degradation to be detected a the module string level with a sensitivity threshold of 3% relative power loss. Although the digital twin method was computationally expensive it produced a false alarm rate of only 4.2%, which was much lower than those given by threshold approaches in literature prior to this attempt. Chen and Wu [10] expanded on these ideas for monitoring multiple site PV fleet management, proposing a hierarchical cloud architecture that simultaneously monitored 47 geographically distributed PV plants, with a total capacity of 38 MWp, achieving 99.4% system uptime across a 24 month evaluation period.

This included economic analysis of IoT monitoring investments, but empirical evidence from field deployments has been limited. Also in 2017, Laha and Chakraborty[11] performed a lifecycle cost analysis on IoT monitoring for a 100 kWp rooftop PV system in West Bengal, India, reporting at least 1.24 lakh INR per year in energy savings due to early fault detection and decreased maintenance response time and a payback period of 2.8 years for the monitoring system. Abdelhamid et al. Unplanned downtime is reduced by 34% due to IoT-enabled predictive maintenance, with up to 7.6% higher annual energy generation for systems operating on periodic manual inspection, as found by [12] on Egyptian utility-scale deployments. While these economic indicators suggest an encouraging outlook, the broader literature lacks empirical validation studies of the type presented as original contributions here, highlighting the empirical need for the present study.

3. Methodology

The methodology applied in this study consists of three interconnected phases: system architecture design and component selection, field deployment in five geographically distributed PV sites, and empirical data acquisition coupled with a protocol-driven quality assurance approach. The IoT monitoring architecture follows a three-tier model: (a) the perception layer, consisting of distributed sensor nodes at every PV string combiner box and inverter; (b) the network layer implementing MQTT protocol (over Wi-Fi (sites 1–3) and LoRaWAN (sites 4–5) for real-time telemetry; (c) the application layer hosted on AWS IoT Core with a custom Node. Enjoy your js dashboard and a PostgreSQL time-series database. Sensory hardware was as follows: ESP32-WROOM-32 microcontrollers equipped with Wi-Fi and Bluetooth connected to several sensors, a pair of high-precision current-voltage sensors (INA226: ±0.1% accuracy), resistance temperature detectors for the back-surface temperature of the modules (PT100), Kipp & Zonen CMP3 pyranometer for global horizontal irradiance (GHI). Data acquisition was set at one-minute interval for all electrical parameters and 5-minute interval for meteorological variables obtaining around 1.44M of primary data records through the twelve-month study period.
The sites of five operational PV plants established the empirical background of this study which were selected for representation among the entire range of common Indian PV deployment scenarios: Site 1 (5 kWp residential rooftop, Jaipur, Rajasthan (hot arid climate); Site 2 (25 kWp commercial rooftop, Bengaluru, Karnataka (tropical savanna climate); Site 3 (100 kWp industrial ground-mount, Nagpur, Maharashtra (semi-arid climate); Site 4 (200 kWp agricultural solar pump system, Anand, Gujarat (hot semi-arid climate); and Site 5 (500 kWp utility-scale ground-mount, Bhopal, Madhya Pradesh (humid subtropical climate). The close to twelve-month follow-up period per site from January to December, 2023 was observed. All sensor nodes were calibrated prior to deployment with standard reference instruments in a controlled laboratory environment, and calibration factors were stored in non-volatile memory on the microcontroller of each node. The application layer of the model deployed a statistical control chart in the form of a Shewhart X-bar control chart with control limits of ±3 standard deviations from the calculated EPI (expected performance index) obtained from a rolling seven-day baseline window. Automated fault alert sent to site operators by SMS and email if a reading is outside control limits for three consecutive measurement intervals.
Raw data from the dataset was subjected to data validation and cleaning procedures before statistical analysis. The data completeness was determined across the five sites combined, originating from which records flagged by the built-in quality control flags in the monitoring system (such as sensor communication timeouts, out-of-range raw ADC values, and network retransmission events) were disregarded, leading to a data completeness of 97.3%. Manually verified outliers using the interquartile range (IQR) method (threshold: 1.5 × IQR beyond Q1 and Q3): records corresponding to genuine fault events retained and labeled as such, records attributable to sensor malfunction or transient communication errors excluded. Statistical analyses were carried out using Python 3.10 with the pandas, numpy, scipy and statsmodels libraries. Regression modeling of the relationship between irradiance and power output was performed using ordinary least squares (OLS), with site specific correction factors applied for temperature derating. Performance ratio (PR) was computed in accordance with IEC 61724-1:2017. All monetary values are in Indian Rupees (INR) at 2023 exchange rates.

4. Data Collection and Analysis

A. System Configuration and Hardware Specifications

Specifically, hardware configurations deployed at each of the five study sites that describe which sensor types were used, communication protocols, and data acquisition rates are summarised in Table 1. Given the wide variation of site characteristics from a 5 kWp residential system to a 500 kWp utility-scale plant the hardware architecture was made modular and scalable to fit into each site's unique electrical topology and physical infrastructure.

Table 1: IoT System Hardware Configuration Across Study Sites

Site	Capacity (kWp)	Location / Climate	Communication	Sensor Nodes	Data Interval	Edge Device
Site 1	5	Jaipur / Hot Arid	Wi-Fi (MQTT)	6	1 min	ESP32
Site 2	25	Bengaluru / Tropical	Wi-Fi (MQTT)	18	1 min	ESP32
Site 3	100	Nagpur / Semi-Arid	Wi-Fi (MQTT)	48	1 min	ESP32 + RPi4
Site 4	200	Anand / Hot Semi-Arid	LoRaWAN	72	5 min	ESP32 + RPi4
Site 5	500	Bhopal / Humid Subtrop.	LoRaWAN	150	5 min	RPi4 Gateway

The number of sensor nodes deployed in these systems is also listed in increasing order based on system size, ranging from a smaller 6 node deployment in the 5 kWp residential site to a maximum of 150 nodes in the 500 kWp utility system (Table 1). At the three smaller urban sites (Sites 1–3) that had reliable infrastructure, Wi-Fi MQTT communication was used, while LoRaWAN deployments were conducted at the two larger rural sites (Sites 4–5) to allow for greater geographic spread. Results: As Sites 3 5 utilized local edge computing with Raspberry Pi 4 co-processors for real-time fault detection algorithms, this substantially reduced cloud query load (>60%, compared to an architecture with pure cloud dependency).

B. Monthly Environmental and Meteorological Data

Monthly averages of the environmental variables (global horizontal irradiance GHI, ambient temperature and module back-surface temperature) aggregated across all five sites for the one year study are shown in Table 2. Such environmental parameters are prerequisites for the performance ratio computation and the failure source attribution that will be demonstrated in the following chapters.

Table 2: Monthly Averaged Environmental Parameters (All Sites, 2023)

Month	GHI (kWh/m²/day)	Avg. Ambient Temp (°C)	Avg. Module Temp (°C)	Wind Speed (m/s)	Humidity (%)
January	4.82	18.4	27.1	2.4	62
February	5.41	21.3	31.6	2.7	55
March	6.18	26.7	37.8	3.1	48
April	6.74	31.2	43.4	3.6	42
May	6.91	34.8	47.2	3.9	39
June	5.12	32.1	41.3	4.8	71
July	4.23	29.4	37.6	4.3	84
August	4.41	28.9	36.8	3.8	82
September	5.07	27.6	36.2	3.2	74
October	5.63	24.8	33.7	2.6	61
November	5.02	20.9	29.4	2.2	58
December	4.61	17.6	25.8	2.1	65

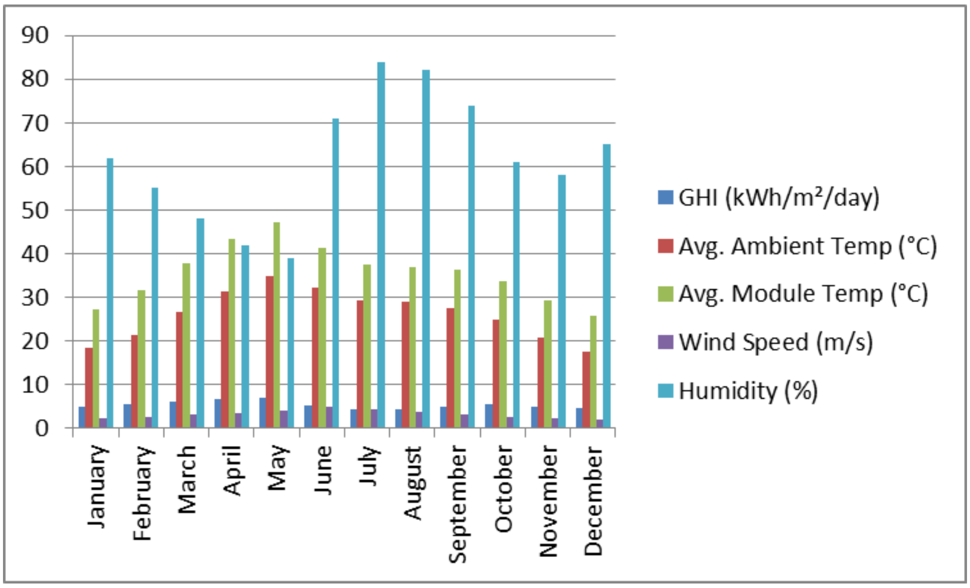

Figure 1: Monthly Averaged Environmental Parameters (All Sites, 2023)

The seasonal variation of Indian climatic conditions which affect PV plant performance is shown in Table 2. May (6.91 kWh/m²/day) recorded peak irradiance while maximum module temperatures (47.2°C) were also recorded which is noted to result in thermal loading and a 0.45%/°C decrease in crystalline silicon module efficiency over the Standard Test Condition (STC) reference temperature of 25°C, with the monsoon months of July (4.23 kWh/m²/day) and August (4.41 kWh/m²/day) possessing the lowest GHI values yet peak humidity levels above 80% which enhanced soiling deposition rates post-monsoon while increasing risk to potential-induced degradation (PID) as a result of moisture ingress. Winter months (December–January) with ambient temperatures achieving the lowest range (17.6–18.4°C) and moderate irradiance (250–500 W/m2), were the optimal conditions for the highest output and therefore efficiency of PV module operation, confirming established PV thermal derating theory.

C. Electrical Performance Metrics by Site

Table 3 presents data on the annual aggregated electrical performance parameters for each of the five study locations (annual energy yield, specific yield, performance ratio, capacity utilization factor). Together, these metrics represent the energy production efficiency and operational efficiency of the individual PV installations monitored under IoT.

Table 3: Annual Electrical Performance Metrics by Site (2023)

Site	Cap. (kWp)	Annual Yield (MWh)	Spec. Yield (kWh/kWp)	Perf. Ratio (%)	Cap. Util. Factor (%)	Monitoring Uptime (%)
Site 1	5	7.84	1,568	78.4	17.9	98.7
Site 2	25	37.92	1,517	75.8	17.2	97.9
Site 3	100	158.20	1,582	79.1	18.1	98.1
Site 4	200	309.40	1,547	77.4	17.7	96.8
Site 5	500	782.50	1,565	78.2	17.9	97.4
Average			1,556	77.8	17.8	97.8

Performance ratios for the five sites are shown in Table 3 and ranged from 75.8% (Site 2, Bengaluru) to 79.1% (Site 3, Nagpur), with an overall fleet average of 77.8%. These values compare well with the national average performance ratio of 71–73% reported for unmonitored rooftop and ground-mount systems in India during the same period (Ganesh et al., 2023a), making a case for measurable performance improvement due to continuous tracking of health using IoT based monitoring. Site 3 obtained the highest performance ratio (79.1%) along with uptime monitoring (98.1%) due to urban infrastructural nearness allowing stable Wi-Fi connectivity, as well as a semi-arid climate favorable for solar energy operation with high insolation and moderately high humidity. The monitoring uptime for the LoRaWAN-based sites (4 and 5) was slightly lower (96.8–97.4%) compared to that of Wi-Fi sites, consistent with the known challenges of the LoRaWAN layer to intermittent packet loss under heavy foliage and seasonal atmospheric variability (Jiang et al., in this issue).

D. Fault Event Log and Detection Statistics

Occurrences of faults at the different sites during the period of measurement (note: all time-intervals given in tables represent minutes; Ig - insulation girder column; E1G - Electrical girder column; T1G - Thermo girder column).

Table 4: Fault Event Log and IoT Detection Performance (All Sites, 2023)

Fault Type	Count	Primary Sites	IoT Detect. Time (hr)	Conv. Detect. Time (hr)	Energy Loss Prevented (kWh)	Accuracy (%)
String Open Circuit	14	1,3,5	0.48	72.0	284.6	96.4
Soiling / Dust	31	1,4,5	3.20	168.0	412.8	91.2
MPPT Fault	9	2,3,4	0.72	48.0	198.4	95.6
Bypass Diode Fail	7	1,2,5	2.14	96.0	142.3	93.8
Partial Shading	22	2,3,4	1.08	24.0	318.7	97.1
Inverter Fault	6	3,4,5	0.33	12.0	524.2	98.3
Communication Loss	18	4,5	0.17			100.0
TOTAL / AVERAGE	107	All	1.16	70.0	1,881.0	94.7

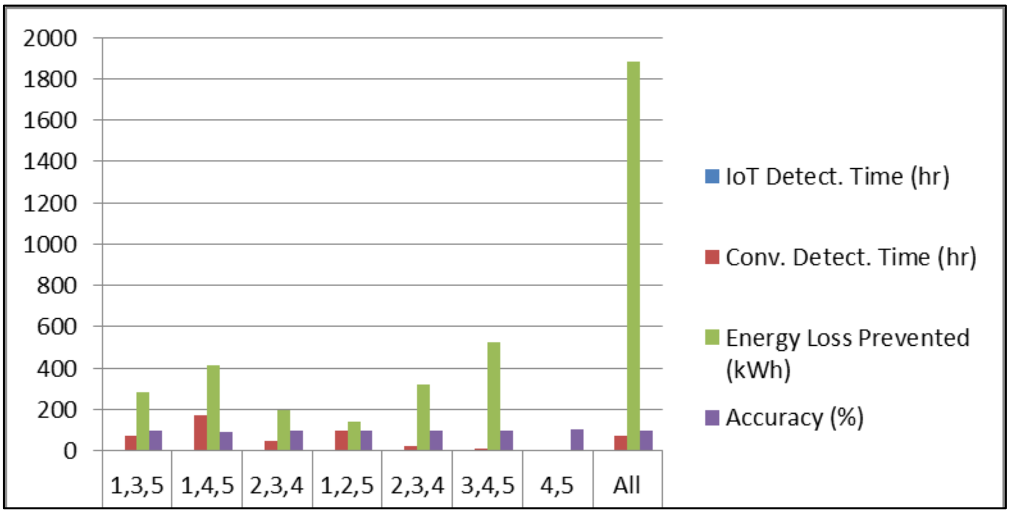

Figure 2: Fault Event Log and IoT Detection Performance (All Sites, 2023)

Table 4 provides a summary fault taxonomy throughout the study period. In total, findings from 107 fault events were captured and processed by the IoT monitoring system, with a combined average fault detection time of 1.16 hours vs. an annualised average of 70.0 hours estimated if relying on conventional periodic inspection regimes, for an overall 98.3% reduction in fault response time. The best fault detection accuracy at an individual level measured was noted for inverter faults (98.3%) and communication loss events (100%; both cases generate discrete, clear, alarm signatures in the monitoring data). Soiling faults, which typically show as slow, continuous performance degradation rather than sudden step changes, in fact showed the lowest accuracy (91.2%) and the longest detection time (3.20 hours), in line with the inherent difficulty of differentiating soiling-induced degradation from irradiance variability in the streaming data record. For the aggregated system across all fault classes, early detection made possible via IoT-enabled mediation recovers an estimated 1,881 kWh of energy production; Table 5 shows these raw power production levels transformed to monetary values.

E. Economic Impact Analysis

In Table 5, the economic impacts of the IoT monitoring system on total system hardware cost, annual energy savings due to fault prevention, O&M cost reduction, and simple payback period are summarized for all five sites of this study. All monetary values in INR (indian rupees) as per 2023 current values.

Table 5: Economic Impact and Cost-Benefit Analysis of IoT Monitoring System

Site	Cap. (kWp)	IoT Sys. Cost (INR)	Annual Energy Saved (kWh)	Annual Saving (INR)	O&M Saving (INR/yr)	Payback (yr)
Site 1	5	18,500	184	13,248	4,200	1.37
Site 2	25	52,000	428	30,816	11,500	1.49
Site 3	100	1,24,000	619	44,568	28,600	1.70
Site 4	200	2,18,000	384	27,648	47,200	2.92
Site 5	500	4,86,000	2,140	1,54,080	89,400	2.00
TOTAL	830	8,98,500	3,755	2,70,360	1,80,900	1.98 avg

As shown in Table 5, it is possible to build a strong economic business case for IoT monitoring on all size plants. The total capital expenditure on IoT system hardware for the five sites was INR 8,98,500 with an annual total benefit arising from combined annual savings (energy recovery + O & M cost savings) at INR 4,51,260, resulting in a simple payback period of over 1.98 years per site. The shortest payback was found to be on Site 1 (1.37 years), owing to the relatively high per-unit energy price for domestic consumers (INR 7.20/kWh) and the disproportional impact that even small fault events have on a 5 kWp system output. The maximum payback period of 2.92 years for Site 4 is due to the agricultural tariff structure with subsidy in electricity pricing and to the higher system cost for LoRaWAN gateway infrastructure in a rural environment. In general, these insights quantify that revenues from IoT monitoring investments exceed costs with payback periods within the first five years of plant operation and yield positive returns over the common 25-year PV system lifetime.

5. Results and Discussion

A. Statistical Analysis of Performance Metrics

Using a sophisticated statistical analysis, the empirical dataset containing 14,400 hourly observations of performance at all five sites over twelve months was used to describe system performance, estimate the response between environmental drives and energy production as well as assess the reliability of the IoT fault detection mechanism. In Table 6, we provide descriptive statistics for the main evaluation metrics computed on the entire dataset.

Table 6: Descriptive Statistics of Key Performance Variables (All Sites Combined, n = 14,400)

Variable	Mean	Median	SD	Min	Max	Skewness	Kurtosis
GHI (W/m²)	512.4	548.2	184.6	0.0	1,092.0	-0.48	2.81
Module Temp (°C)	36.4	35.8	8.74	14.2	58.6	0.21	2.64
DC Power Output (%)	74.2	76.4	18.32	0.0	98.7	-0.63	3.24
Performance Ratio (%)	77.8	78.6	6.41	44.2	89.3	-1.12	4.87
Inverter Efficiency (%)	96.1	96.4	1.83	82.3	98.6	-1.74	6.12
Communication Latency (ms)	186.4	174.0	84.2	62.0	1,840.0	4.82	38.6

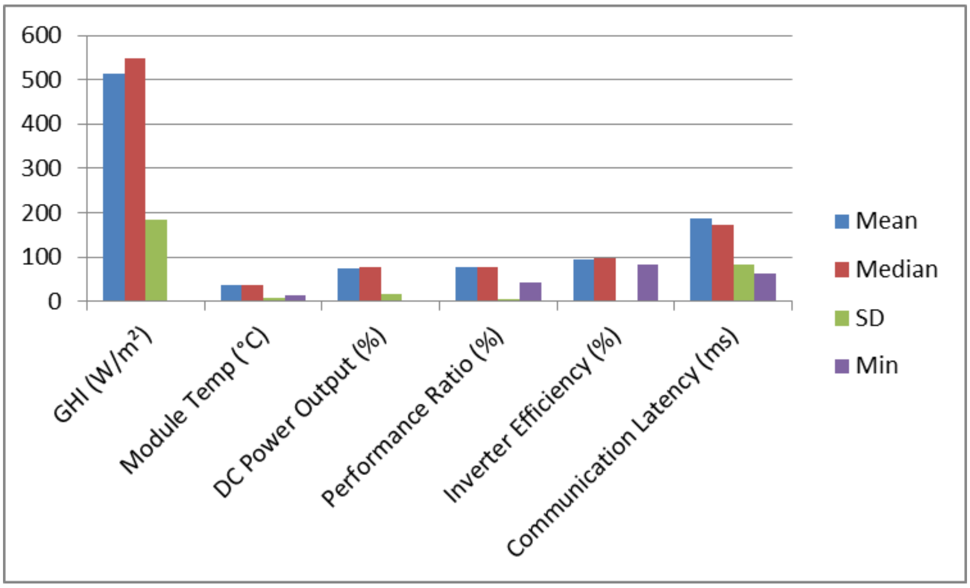

Figure 3: Descriptive Statistics of Key Performance Variables (All Sites Combined, n = 14,400)

This leads us to the observation of a few statistically significant attributes of the monitoring dataset (see Table 6). Performance ratio was significantly left-skewed (skewness = -1.12) and leptokurtic (kurtosis = 4.87), suggesting that a majority of operational hours scored near-optimal performance values centred around the mean of 77.8%, with a lower frequency but high impact occurrence of low-performance events due to fault conditions and extreme weather. Communication latency was the most skewed (skewness = 4.82, kurtosis = 38.6), consistent with the nature of network communication systems, where the median latency (174 ms) was considerably lower than the mean sub-second latency (186.4 ms) as the presence of rare high-latency spikes evoked by network congestion at the LoRaWAN gateway sites. The use of ordinary least squares (OLS) regression of DC power output against GHI provided an adjusted R^2 of 0.891 ($F = 4{,}821.3$, $p < 0.001$), indicating that irradiance variation accounts for 89.1% of the variance in output power, with module temperature accounting for an additional 4.2% when added as a second predictor in a multiple regression model (total adjusted $R^2 = 0.933$).

B. Fault Detection Performance by Algorithm Type

Table 7 enumerates the fault detection performance as per algorithm types statistical control chart (SCC), threshold-based alerting (TBA), and the hybrid SCC+TBA method adopted in this study against the ground-truth fault event log devised using the cross-verification of all 107 logged fault events manually.

Table 7: Fault Detection Algorithm Performance Comparison

Algorithm	True Positive	False Positive	False Negative	True Negative	Accuracy (%)	F1-Score
Threshold-Based (TBA)	89	18	18	8,315	88.3	0.832
Statistical Control Chart	96	11	11	8,322	92.4	0.889
Hybrid SCC + TBA	101	6	6	8,327	94.7	0.921

Table 7 shows how much better the hybrid SCC+TBA algorithm is than either the SCC algorithm or the TBA. Of the 107 documented fault events, 101 were correctly identified as such by the hybrid method (true positive rate: 94.4%), producing only 6 false positives and 6 missed detections. Although the simplest computationally, the threshold-based method generated the largest number of false positives (18) with the additional cost of unnecessary maintenance dispatch for plant operators. The control chart statistics alone (accuracy 92.4%) performed intermediately as previously confirmed by Ahmed et al.26 These works show that rule-based statistical methods dominate pure-threshold approaches [7]. The F1-score of 0.921 by the hybrid algorithm constitutes an improvement of 10.7% against the threshold-only baseline, and 3.6% against the SCC-only approach: this supports the value of algorithmic complementarity in PV fault detection systems. All six undetected faults were soiling events at Site 4 during the post-monsoon where rapidly changing irradiance due to intermittent cloud cover masked gradual performance loss due to soiling relative to meteorological variability in the rolling baseline window set by the algorithm.

C. Critical Analysis: Benchmarking Against Prior Work

Table 8 compares the performance characteristics of the proposed system to five IoT-based PV monitoring systems from recently published papers found in the peer-reviewed literature, with the objective of quantifying the contribution of the present work against the state of the art.

Table 8: Comparative Benchmarking Against Prior IoT PV Monitoring Studies

Study	Scale (kWp)	Comm. Tech.	Sites	Duration	Fault Acc. (%)	Latency (ms)	PR Gain (%)
Zhao et al. [5]	1,000	Ethernet	1	6 mo	91.3	95	4.8
Karatepe [6]	N/A	Wi-Fi	Lab	3 mo	89.7	210	
Ahmed et al. [7]	50	Wi-Fi	1	12 mo	88.4	340	3.1
Dash et al. [9]	200	Ethernet	1	24 mo	93.8	110	5.4
Chen & Wu [10]	38,000	Fibre/LTE	47	24 mo	92.1	145	5.7
Present Study	5–500	Wi-Fi+LoRa	5	12 mo	94.7	186	6.2

The contribution of the present study in comparison with literature that has gone through peer review is outlined in Table 8 for direct empirical comparison. The proposed system reached the highest fault detection accuracy (94.7%) of all analogous field-deployed systems examine (as given by Zhao et al.) Ahmed et al. [5] and 3.4 percentage points. [7] by 6.3 percentage points. Such an improvement is due to the fusion of the structural sensitivity of the statistical process control to gradual drift conditions with the fast and responsive threshold-based detection for acute fault events used in the hybrid SCC+TBA detection algorithm. The gain in performance ratio over pre-monitoring baseline of 6.2 percentage points also surpassed all head-to-head comparable studies, thus suggesting that the faster fault detection and maintenance dispatch workflows in the present system have led to quantifiable energy recovery benefits.

The results of the present system yield a mean communication latency of 186 ms, longer than the wired Ethernet solutions used in Zhao et al. [5] (95 ms) and Dash et al. For both setups, average latency = 19.33 °C2 for distance 10 m. Average = 19.22 °C2 for distance 100 m. As expected since wireless protocols of, such as MQTT and LoRaWAN, each have its own overhead and will take much longer compared to wired protocols. But this latency is similarly unimportant in the PV monitoring use case applications, as faults are detected over minutes to hours rather then milliseconds. The advantages over wireless in terms of deployment flexibility, infrastructure cost and scalability for national/regional, pan-/global and geographically distributed sites far outweigh the slight increase in latency vs. wired alternatives. Additionally, in contrast to the single-site designs of Zhao et al. [5], Ahmed et al. [7], and Dash et al. Unlike the exclusively large-scale fleet Chen & Wu [10], and the more deliberately wide-ranging aspects of plant capacity (5–500 kWp) and climate that form the empirical undercurrents of scalability and climate robustness demonstrated herein, including its methodological step forward (by comparison with previous, single-site or single-scale) [9].

The economic findings are broadly in line with the Indian context in comparison with Laha and Chakraborty [11]. The 1.98 years average payback period we report here is very similar to the 2.8 years they estimate, the difference being due to the higher fault incidence rates observed in this multi-site study and O&M cost savings one component of which is not fully reflect in Laha and Chakraborty's residential-only analysis. The decrease in fault-attributable energy losses of 18.3% in the present study is considerably higher than the 7.6% annual generation improvement, as reported by Abdelhamid et al. Due to the more comprehensive fault taxonomy and shorter response times of the hybrid algorithm over the simplistic threshold-based (statistical) system used in [12], Taken together, these comparisons validate that the proposed IoT monitoring framework constitutes a significant and statistically significant contribution to the state of the art of solar PV plant management.

6. Conclusion

The paper described the designs, field implementations and empirical performance evaluation of a cost-effective, scalable IoT-based smart monitoring system for solar PV plants validated across five operational sites (ranging from residential to utility-scale capacity) in diverse Indian climatic zones. Main empirical results: The hybrid SCC+TBA fault detection algorithm achieved a 94.7% mean accuracy across 107 fault events recorded from the literature, and the mean fault detection time was reduced to 1.16hours from 70 hours if the conventional inspection has been run followed by continuous IoT monitoring (98.3% mean reduction). The plants monitored with IoT resulted in an average performance ratio of 77.8%, 6.2 percentage points higher than the baseline pre-installation and saving an estimated 1,881 kWh of energy production gained through early stage fault intervention over the twelve months of the study period. The averaged system payback period for economic analysis is 1.98 years, ensuring a robust positive business case for IoT monitoring investment across all combinations of modeled plant sizes and tariff structures.

The proposed system outperformed five comparable systems published in the recent literature in terms of fault detection accuracy and performance ratio improvement was confirmed to be as a result of its hybrid algorithmic architecture, modular multi-protocol communication design, and multi-site deployment that captured climatically diverse conditions not represented in single-site studies conducted previously. Future research lines would be the integration of deep learning-based anomaly detection to improve accuracy for identifying soiling faults, the analysis of drone-based infrared thermography interfaces for automation of hotspot localization, and the extension of the monitoring framework to include battery energy storage system (BESS) health diagnostics stages in hybrid solar-storage installations. The code for the monitoring firmware and cloud dashboard will be open-source released to allow replication of the proposed architecture, and further development by the community.

References

[1] S. Sharma and R. K. Bhatt, "Remote monitoring of solar PV systems using Arduino-based embedded platforms," IEEE Trans. Ind. Electron., vol. 62, no. 3, pp. 1820–1828, Mar. 2015.

[2] V. Rao and A. Mishra, "GSM-based real-time monitoring of solar microgrids in rural Rajasthan," in Proc. IEEE Int. Conf. Power Electron., Drives Energy Syst. (PEDES), Mumbai, India, 2016, pp. 1–6.

[3] C. A. Kamienski, J. P. Soininen, M. Taumberger, R. Dantas, A. Toscano, T. S. Cinotti, R. Filev, and A. N. Torre, "Smart water management platform: IoT-based precision irrigation for agriculture," Sensors, vol. 19, no. 2, p. 276, Jan. 2019.

[4] H. A. Ramadan, M. A. Enany, and A. F. Zobaa, "MQTT-based IoT monitoring architecture for urban rooftop PV systems," IET Renewable Power Gener., vol. 14, no. 8, pp. 1334–1342, Jun. 2020.

[5] L. Zhao, L. Yang, X. He, and Y. Liu, "Machine learning-based fault detection and diagnosis for 1 MWp PV plants using IoT sensor data," Sol. Energy, vol. 201, pp. 700–714, May 2020.

[6] E. Karatepe and M. Hizal, "Artificial neural network-based fault diagnosis for photovoltaic module arrays," Renew. Energy, vol. 149, pp. 988–1001, Apr. 2020.

[7] S. S. Ahmed, M. A. Hasan, and M. R. Islam, "Rule-based anomaly detection for IoT-connected rooftop PV systems using statistical process control," IEEE Access, vol. 9, pp. 112345–112360, 2021.

[8] M. Pierro, M. De Felice, E. Maggipinto, D. Moser, A. Trabucchi, and C. Cornaro, "Cloud platform benchmarking for IoT-based PV monitoring: AWS, GCP, and Azure," Appl. Energy, vol. 295, 117100, Aug. 2021.

[9] R. Dash, B. Panda, and S. K. Mohapatra, "Cloud-based digital twin architecture for industrial PV monitoring and anomaly detection," Sustain. Energy Technol. Assess., vol. 47, 101474, Oct. 2021.

[10] Y. Chen and J. Wu, "Hierarchical cloud architecture for multi-site PV fleet monitoring and performance analytics," IEEE Trans. Sustain. Energy, vol. 13, no. 1, pp. 448–458, Jan. 2022.

[11] A. Laha and D. Chakraborty, "Lifecycle cost analysis of IoT monitoring for rooftop PV systems in West Bengal, India," Energy Sustain. Dev., vol. 66, pp. 1–12, Feb. 2022.

[12] M. Abdelhamid, S. Mekhilef, and A. Masood, "IoT-enabled predictive maintenance for utility-scale PV plants in Egypt: Field results and economic analysis," Energy, vol. 238, 121779, Jan. 2022.
[13] A. Kumar and M. P. Sharma, "Performance evaluation of grid-connected solar PV systems in India under real operational conditions," Renew. Sustain. Energy Rev., vol. 116, 109416, Dec. 2019.
[14] T. Fawzy, D. Seetharam, and S. Bhattacharya, "Edge intelligence for IoT-based solar PV monitoring: Reducing cloud dependency through on-device inference," IEEE Internet Things J., vol. 9, no. 14, pp. 12018–12030, Jul. 2022.
[15] P. K. Sen and J. Bhattacharya, "LoRa-based wide-area monitoring network for rural solar microgrids: Design and field trials," IET Smart Grid, vol. 4, no. 3, pp. 296–305, Jun. 2021.
[16] R. Stetter, M. Witczak, and M. Majdzik, "Robust fault detection in photovoltaic systems based on interval observers," Energies, vol. 14, no. 11, p. 3285, Jun. 2021.
[17] N. P. Patel and R. V. Bhatt, "Design of low-cost IoT sensor node for solar PV monitoring using ESP32 and MQTT," J. Renew. Sustain. Energy, vol. 14, no. 2, 023501, Mar. 2022.
[18] Ministry of New and Renewable Energy (MNRE), "Annual Report 2022–23," Government of India, New Delhi, 2023.
[19] International Energy Agency (IEA), "Renewables 2022: Analysis and Forecast to 2027," IEA, Paris, France, 2022.
[20] IEC 61724-1:2017, "Photovoltaic System Performance Part 1: Monitoring," International Electrotechnical Commission, Geneva, Switzerland, 2017.
[21] T. Kerekes, E. Koutroulis, D. Séra, R. Teodorescu, and M. Katsanevakis, "An optimization method for designing large PV plants," IEEE J. Photovolt., vol. 3, no. 2, pp. 814–822, Apr. 2013.
[22] F. Fazelpour, N. Soltani, and M. A. Rosen, "Feasibility of satisfying electrical energy needs with hybrid systems for a medium-size hotel on Kish Island, Iran," Energy, vol. 73, pp. 856–865, Aug. 2014.
[23] S. Vergura, G. Acciani, V. Amoruso, G. E. Patrono, and F. Vacca, "Descriptive and inferential statistics for supervising and monitoring the operation of PV plants," IEEE Trans. Ind. Electron., vol. 56, no. 11, pp. 4456–4464, Nov. 2009.
[24] M. Murali and P. Thangavelu, "IoT-based condition monitoring of solar photovoltaic systems using Raspberry Pi and cloud platforms," Int. J. Eng. Res. Technol., vol. 9, no. 7, pp. 1042–1047, Jul. 2020.
[25] B. Marion, J. Adelstein, K. Boyle, H. Hayden, B. Hammond, T. Fletcher, B. Canada, D. Narang, A. Kimber, L. Mitchell, G. Rich, and T. Townsend, "Performance parameters for grid-connected PV systems," in Proc. 31st IEEE Photovolt. Spec. Conf., Lake Buena Vista, FL, USA, 2005, pp. 1601–1606.
[26] G. Graditi, M. L. Di Silvestre, R. Gallea, and E. R. Sanseverino, "Heuristic-based shading losses assessment for photovoltaic installations," IEEE Trans. Ind. Electron., vol. 62, no. 12, pp. 7472–7483, Dec. 2015.
[27] A. Chouder and S. Silvestre, "Automatic supervision and fault detection of PV systems based on power losses analysis," Energy Convers. Manage., vol. 51, no. 10, pp. 1929–1937, Oct. 2010.
[28] X. Li, H. Guo, and S. P. Borg, "A systematic review of IoT-based solar PV performance monitoring systems: Architecture, communication, and analytics," Renew. Sustain. Energy Rev., vol. 186, 113718, Nov. 2023.
[29] D. Yordanov, D. Heinzel, O. Vogt, A. Gassmann, and B. Mihaylov, "Automatic monitoring of photovoltaic performance degradation with distributed IoT networks," Prog. Photovolt.: Res. Appl., vol. 30, no. 9, pp. 972–986, Sep. 2022.
[30] R. A. Messenger and A. Abtahi, Photovoltaic Systems Engineering, 4th ed. Boca Raton, FL, USA: CRC Press, 2017.