A Systematic Framework for Reconciling Solar Energy Production Models with Operational Data: A Literature Synthesis and Methodology Development

Sahil Shah[1,*], Aravind Reddy Boozula[2], Radhika Girish Lampuse[3], and Jigar Janakbhai Thakkar[4]

ABSTRACT

The growing deployment of utility-scale photovoltaic (PV) systems has heightened the importance of accurate energy yield predictions for financial planning and operational management. However, a persistent gap exists between predicted and actual performance, creating significant challenges for project developers, investors, and operators. This paper presents a comprehensive analysis of the performance gap phenomenon in solar PV systems through a systematic review of current literature and industry practices. We examine the primary sources of discrepancy between energy production models and operational data, including resource assessment uncertainties, system loss characterization, and operational factors. Building on this analysis, we develop a structured framework for reconciling predictions with real-world performance that integrates statistical approaches for uncertainty quantification with methodologies for model calibration. The proposed framework enables more accurate performance assessments through the systematic identification and correction of modeling biases across different temporal scales and operating conditions. This research contributes to improved risk assessment in PV project development and advances modeling methodologies by providing practitioners with a practical approach to align predictive models with operational realities, ultimately enhancing the reliability of financial models and performance guarantees in the solar energy sector.

Keywords: Energy yield prediction, model reconciliation, performance gap analysis, Solar photovoltaic.

1. INTRODUCTION

Photovoltaic (PV) systems have emerged as a cornerstone of global renewable energy strategies, with utility-scale installations experiencing particularly rapid growth over the past decade. The International Energy Agency reports that solar PV capacity has increased more than 20-fold since 2010, with utility-scale systems making up approximately 70% of new installations [1], [2]. This accelerated deployment has intensified the need for accurate energy yield predictions, as these forecasts directly influence investment decisions, power purchase agreements, and operational strategies [3].

Despite significant advancements in modeling techniques, a persistent gap remains between predicted and actual performance of PV systems. This "performance gap"—defined as the discrepancy between modeled energy predictions and operational results—represents a critical challenge for the solar industry [4]. The concept of performance gaps has been recognized across energy systems, with early investigations by Bordass *et al*. [5] establishing that significant disparities between predicted and actual energy performance are common in complex systems. Recent studies have documented mean absolute errors ranging from 3% to 8% in annual energy predictions, with seasonal variations often exceeding 10% [6]. These deviations occur despite increasingly sophisticated modeling tools, suggesting underlying challenges in accurately capturing the complex interactions between equipment

performance, environmental conditions, and operational realities.

The consequences of these prediction errors extend beyond simple accounting discrepancies. Project financing relies heavily on energy yield assessments, with debt service coverage ratios and investment returns calculated based on predicted generation profiles [3]. Reich *et al.* [7] demonstrated that the target performance ratios used in financial models often exceed 90%, setting expectations that may be unrealistic given real-world operational conditions. Financial stakeholders typically apply confidence intervals (P50/P90 values) to manage uncertainty, but these statistical approaches assume underlying modeling methodologies are free from systematic biases—an assumption increasingly questioned by operational data [8]. As the industry continues to mature and profit margins tighten, improving forecast accuracy has become essential to reducing financing costs and ensuring long-term project viability.

Several factors contribute to the observed performance gaps. Solar resource assessment—the foundation of energy yield predictions—remains subject to uncertainty from both measurement limitations and the inherent variability of solar radiation [9]. Mondol *et al.* [10] identified that even high-quality solar radiation models can introduce errors of 4%–6% in energy yield calculations due to spatial and temporal resolution limitations. Meng *et al.* [11] analyzed 246 identical rooftop PV systems and found surprising variability in performance even among systems with nearly identical designs and locations, highlighting the challenges in accounting for all influential factors. Loss factors such as soiling, shading, and system availability are frequently estimated based on rules of thumb rather than site-specific assessments [12]. Moreover, long-term degradation rates, which significantly impact lifetime project performance, are often based on limited field data that may not represent current technology or installation practices [13]. Jordan and Kurtz [14] conducted a landmark analytical review of PV degradation rates, finding median rates of 0.5%–0.6% per year but with substantial variation across technologies and climates that is often overlooked in modeling.

The solar industry has responded with increasingly standardized approaches to performance modeling and validation. Measurement and reporting protocols have been developed to ensure consistency in data collection and analysis [15]. Technical due diligence methodologies have evolved to incorporate more rigorous validation procedures and uncertainty assessments [9]. King *et al.* [4] developed one of the most influential array performance models that continues to underpin many modern simulation tools, though adaptations for newer technologies remain ongoing. However, these advances have primarily focused on procedural standardization rather than addressing the underlying causes of systematic modeling errors or developing robust frameworks for reconciling predictions with operational reality.

Recent research has begun exploring more sophisticated approaches to aligning modeled and operational performance. Yagli *et al.* [16] proposed sequential reconciliation techniques to improve forecasting accuracy by systematically adjusting predictions based on observed

patterns. Zuhaib *et al.* [17] conducted comprehensive performance analysis of utility-scale solar farms that incorporated both physical and environmental factors, demonstrating the importance of holistic approaches to performance assessment. These advances suggest opportunities to move beyond static modeling approaches toward dynamic frameworks that evolve throughout a project's operational life.

This paper aims to address critical gaps in current practice by developing a comprehensive framework for analyzing and reconciling differences between modeled predictions and operational performance of PV systems. Through a systematic review of current literature and industry practices, we examine the primary sources of performance discrepancies, evaluate existing methodologies for model validation and calibration, and propose a structured approach for improving prediction accuracy. The research focuses specifically on utility-scale PV installations, where performance gaps have the most significant financial implications and where operational data collection is typically more robust.

The remainder of this paper is structured as follows: Section 2 presents a comprehensive literature review exploring the historical development of PV performance modeling, current understanding of performance gaps, and previous reconciliation attempts. Section 3 outlines our methodology for analyzing performance gaps and developing a reconciliation framework. Section 4 examines key sources of discrepancy between predicted and actual performance, categorizing them into physical, modeling, and operational factors. Section 5 introduces our proposed framework for reconciling models with operational data, including validation approaches and implementation considerations. Finally, Section 6 summarizes our conclusions and identifies directions for future research.

2. Literature Review

2.1. Historical Development of PV Performance Modeling

The evolution of photovoltaic (PV) performance modeling has paralleled the industry's growth from specialized niche applications to mainstream power generation. Early modeling approaches in the 1980s and 1990s primarily focused on simple correlations between solar irradiance and power output, with limited consideration of environmental factors or system losses [9]. These rudimentary models provided reasonable estimates for small, well-maintained systems but proved inadequate as commercial and utility-scale installations expanded.

The early 2000s saw significant advancements in modeling sophistication. Stein and Klise [9] categorized PV performance models into three generations: first-generation models using constant efficiency values, second-generation models incorporating temperature corrections and basic loss factors, and third-generation models featuring detailed component-level simulations with comprehensive loss analyses. This evolution reflected growing recognition of the complex interactions between equipment specifications, environmental conditions, and

system configuration that influence actual performance. De Soto *et al*. [18] made significant contributions to second-generation models by improving the five-parameter model and validating it against measured data, demonstrating substantial accuracy improvements over earlier approaches.

Current industry-standard modeling tools integrate multiple physical models with extensive databases of component specifications, historical weather data, and empirical loss factors. According to Kurtz *et al*. [3], these sophisticated platforms have substantially improved baseline accuracy but still struggle with site-specific variations and dynamic operational conditions. The fundamental challenge has shifted from basic modeling capability to accurately capturing the myriad factors that create disparities between modeled and measured performance. King *et al*. [4] established a comprehensive photovoltaic array performance model at Sandia National Laboratories that became foundational for modern simulation tools, particularly in accounting for spectral and angular effects often overlooked in simplified models.

2.2. Performance Gap Characterization and Measurement

The "performance gap" in PV systems refers to discrepancies between predicted and actual energy production. While the concept appears straightforward, methodologies for quantifying and characterizing this gap vary considerably across the literature. Van Dronkelaar *et al*. [19] established a theoretical framework for performance gap analysis in building energy systems that has since been adapted to PV applications, emphasizing the distinction between model inadequacy (systematic errors in the model structure), specification gap (differences between design and as-built conditions), and operational deviation (differences between assumed and actual operation). Burman *et al*. [20] further advanced this framework by proposing measurement and verification protocols specifically designed to address the credibility gap in energy performance, providing a methodological approach applicable to PV systems.

Harrison and Jiang [6] conducted one of the most comprehensive investigations of PV performance gaps using dynamic simulation modeling. Their case study demonstrated annual performance deviations of 5%–7% between predicted and measured output, with seasonal variations exceeding 12% during winter months. The study identified irradiance data quality, snow losses, and inverter performance modeling as primary contributors to the observed discrepancies. Reich *et al*. [7] questioned whether performance ratios above 90% are realistic in practice, finding that even well-designed systems rarely achieve such values consistently due to unavoidable real-world losses. Importantly, they found that conventional adjustment factors applied in commercial modeling tools failed to adequately account for these site-specific variations.

More recently, Meng *et al*. [11] analyzed an unprecedented dataset of 246 identical rooftop PV systems, providing unique insights into performance variability even under nominally identical conditions. Advanced performance analysis techniques [21] have further enhanced the ability to diagnose system-specific issues through statistical pattern recognition. Their findings revealed production differences exceeding 8% among systems with identical components, orientations, and geographic proximity, highlighting the challenges in predicting individual system performance even with sophisticated models and high-quality inputs. The study emphasized the intrinsic variability in real-world systems that may remain unaccounted for in even the most advanced deterministic models.

The IEA Report [8] on PV system yield predictions established standardized metrics for performance gap assessment, recommending the decomposition of overall deviation into specific categories: resource assessment error, loss factor estimation error, and operational deviation. This structured approach enables more systematic identification of error sources and facilitates targeted improvement of modeling methodologies. The conceptual framework for performance gap analysis is illustrated in Fig. 1, which shows the hierarchical breakdown of factors contributing to discrepancies between predicted and actual performance across three primary domains.

2.3. Factors Affecting Performance Gap

The literature identifies numerous factors contributing to the observed performance gap in PV systems. Sepúlveda-Oviedo [13] provided the most comprehensive review of operational factors affecting performance, categorizing them into environmental factors (soiling, shading, irradiance variation, temperature), equipment factors (module degradation, inverter efficiency, mismatch losses), and operational factors (availability, maintenance practices, grid interaction).

Solar resource assessment uncertainty remains a fundamental contributor to performance gaps. Traditional modeling approaches rely heavily on typical meteorological year (TMY) datasets, which by definition cannot capture the inter-annual variability of solar resources. Mondol *et al*. [10] demonstrated that the choice of solar radiation model significantly impacts simulation accuracy, with errors varying by geographic location and temporal resolution of the data. Stein and Klise [9] demonstrated that even high-quality TMY data can introduce uncertainties of 3%–5% in annual energy predictions due to natural climate variations. Mondol *et al*. [10] demonstrated that the choice of solar radiation model significantly impacts simulation accuracy, with errors varying by geographic location and temporal resolution of the data. More recent satellite-derived irradiance datasets have improved spatial resolution but may still contain systematic biases in specific regions or climatic conditions.

System losses represent another major source of performance discrepancy. Deceglie *et al*. [12] highlighted how soiling losses alone can account for 2%–6% annual energy reduction in many environments, with significant seasonal variations that standard models often fail to capture. The impact of environmental factors such as dust, humidity, and air velocity on soiling accumulation [22] creates complex interactions that are difficult to predict using simplified models. The study presented a novel methodology for extracting soiling loss profiles directly from operational
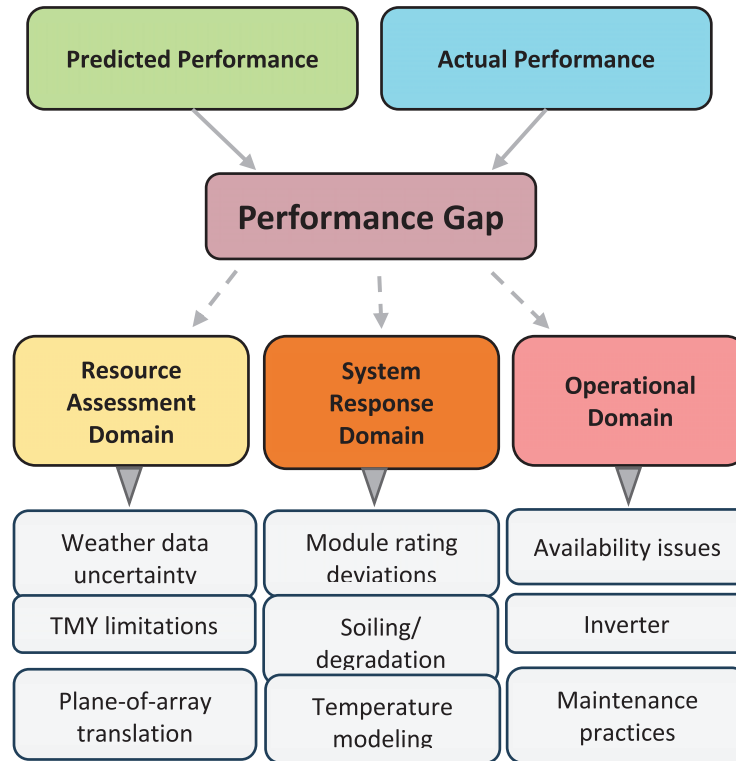
Fig. 1. Conceptual framework for peformance gap analysis.

data, demonstrating the potential for improved modeling through systematic analysis of performance patterns.

Equipment performance deviations from manufacturers' specifications contribute significantly to the performance gap. Kurtz *et al*. [3] documented systematic differences between nameplate ratings and actual field performance of PV modules, with measured power output frequently 2%–3% below rated values even before degradation effects. Driesse *et al*. [23] analyzed numerous commercial inverters and found that actual efficiency curves can diverge significantly from manufacturer specifications, particularly under non-ideal input conditions. Similarly, inverter efficiency profiles in real-world conditions often deviate from the idealized curves used in performance models, particularly during low-irradiance operation.

Degradation effects represent a particularly challenging aspect of long-term performance prediction. While most models incorporate linear degradation assumptions (typically 0.5%–0.7% annually), actual degradation patterns exhibit significant variability. Zuhaib *et al*. [17] demonstrated that environmental factors such as high temperatures, humidity, and dust exposure can accelerate degradation rates, leading to performance ratios declining faster than predicted in harsh environments. Lindig *et al*. [1] established an international collaboration framework for calculating performance loss rates, revealing significant variations in degradation patterns across different climates and technologies that are rarely captured in standard modeling approaches. Table I summarizes the typical magnitudes of performance gap contributions reported in recent literature, categorized by source of discrepancy and providing quantitative ranges with supporting references for each factor.

2.4. Reconciliation Approaches

Recent literature has proposed various approaches to reconcile the gap between modeled and measured performance. Yagli *et al*. [16] introduced the concept of "sequential reconciliation" for solar forecasts, demonstrating how hierarchical adjustments to predictions based on observed patterns could significantly improve accuracy. This approach builds on earlier work by Yang *et al*. [24] on geographical hierarchy in solar forecasting, which demonstrated that reconciling predictions across spatial scales improves overall accuracy. While primarily focused on short-term forecasting, their methodological framework offers valuable insights for longer-term yield reconciliation, particularly the importance of preserving relationships between different temporal aggregation levels.

Yang *et al*. [25] explored operational solar forecasting for real-time markets, developing techniques to continuously update predictions based on observed performance patterns. The statistical foundations for these hierarchical reconciliation approaches were established by Hyndman *et al*. [26], who developed computational methods for reconciling forecasts across grouped time series that have since been adapted to solar applications. Their work highlighted the potential for machine learning approaches to capture complex, non-linear relationships between environmental conditions and system performance that may be overlooked in physics-based models.

The industry has increasingly adopted empirical correction factors to bridge the gap between models and measurements. Pless *et al*. [15] established protocols for measuring and reporting PV performance that included methodologies for developing site-specific correction factors based on operational data. Similarly, Kurtz *et al*.

TABLE I: Summary of Performance Gap Magnitudes from Literature

| Source of discrepancy	Typical range (% of annual production)	Notes	Key references
Solar resource assessment	3%–5%	Higher uncertainty in regions with variable climate	[9], [10]
Soiling losses	2%–6%	Location and climate dependent	[12]
Temperature modeling	1%–3%	Larger in hot climates	[17]
Module rating deviation	2%–3%	Initial deviation from nameplate	[3]
Degradation rate	0.5%–2% per year	Cumulative effect increases with time	[14]
Inverter performance	1%–2%	Larger at partial loading	[23]
Mismatch & wiring	1%–3%	Higher in non-optimal layouts	[11]
Availability & downtime	0.5%–3%	Dependent on O&M practices	[8]
Snow & shading	1%–8%	Highly location specific	[6]

Evolution of PV Performance Modeling
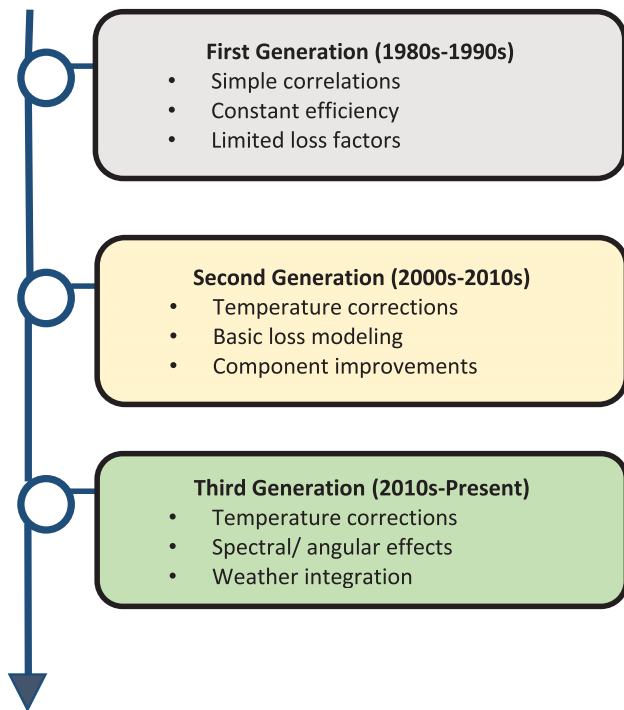
Fig. 2. Evolution of PV performance modeling approaches.

[3] proposed standardized approaches for energy performance evaluation that incorporate empirical adjustments to theoretical models based on measured performance ratios.

Despite these advances, the literature reveals a persistent need for more systematic, integrated approaches to performance gap reconciliation. While many studies have identified specific contributing factors or proposed corrections for individual aspects of performance modeling, comprehensive frameworks that address multiple sources of discrepancy within a coherent methodology remain limited. This gap in the literature provides the foundation for the present study's contribution in developing a structured reconciliation framework. Fig. 2 demonstrates the evolution of PV performance modeling approaches over three generations, showing the progression from simple correlations to comprehensive analytical frameworks currently used in industry practice.

3. Methodology

3.1. Research Design

This study employs a systematic approach to analyze performance gaps in utility-scale PV systems and develop a comprehensive framework for reconciling modeled predictions with operational data. The research methodology consists of four interconnected components:

- a structured literature review to establish the current state of knowledge
- categorization and analysis of performance gap sources
- development of a reconciliation framework, and
- case-based validation of the proposed approach

The literature review focused on peer-reviewed journal articles, technical reports from recognized research institutions, and industry standards published between 2005 and 2025. Key search terms included "photovoltaic performance gap," "solar energy yield prediction," "PV model validation," and "reconciling solar forecasts." We prioritized studies with quantitative analysis of utility-scale systems and those providing methodological frameworks rather than purely case-specific findings. This approach ensured a comprehensive foundation for understanding both the technical and methodological aspects of the performance gap phenomenon.

To assess the relative significance of different factors contributing to performance gaps, we compiled quantitative data from multiple studies, normalizing findings to consistent metrics (typically percentage of annual energy yield) where possible. This meta-analysis allowed us to identify patterns across diverse geographical locations, system configurations, and operational conditions, revealing which factors consistently contribute most significantly to performance discrepancies.

3.2. Performance Gap Analysis Framework

Our methodology for analyzing performance gaps builds on the theoretical framework established by Van Dronkelaar *et al.* [19] but extends it specifically for PV applications. We categorize performance gap sources into three main domains:

1. *Resource Assessment Domain:* Factors related to the characterization and prediction of solar resource,

including measurement uncertainties, spatial interpolation errors, temporal resolution limitations, and stochastic variability.

2. *System Response Domain:* Factors related to how the PV system converts available solar resource into electrical energy, including equipment specifications, physical models, loss characterization, and degradation patterns.

3. *Operational Domain:* Factors related to system operation and maintenance, including availability, curtailment, control strategies, and maintenance practices.

This structured categorization enables more systematic identification of error sources and facilitates targeted improvement strategies. For each domain, we analyze both systematic biases (persistent directional errors) and random variations (stochastic uncertainties), as these require different reconciliation approaches.

For quantitative analysis, we adopt the normalized metrics recommended by the IEA Report [8], including the Performance Ratio (PR) for overall system performance evaluation and specific decomposition metrics for individual loss factors. This standardized approach enables consistent comparison across different studies and system configurations.

3.3. Reconciliation Framework Development

The development of our reconciliation framework follows a modular structure that addresses each performance gap domain sequentially while preserving the interdependencies between domains. The framework integrates three key methodological approaches:

1. *Statistical Reconciliation:* Adapting the sequential reconciliation methodology proposed by Yagli *et al.* [16] to address temporal hierarchies in performance data, with extensions to incorporate both short-term variations and long-term trends.

2. *Model Calibration:* Systematically adjusting model parameters based on operational data to minimize prediction errors, following a Bayesian approach that updates prior distributions of model parameters as new data becomes available.

3. *Loss Factor Decomposition:* Isolating and quantifying specific loss mechanisms from aggregate performance data, building on the methodology developed by Deceglie *et al.* [12] for soiling loss extraction and extending it to other loss factors.

The framework development process incorporated an iterative design approach, with successive refinements based on theoretical considerations and practical constraints identified in the literature. Particular attention was given to ensuring the framework's adaptability to different system configurations, data availability scenarios, and operational contexts.

3.4. Validation Methodology

While full empirical validation of the proposed framework is beyond the scope of this study, we employ a case-based validation approach using publicly available
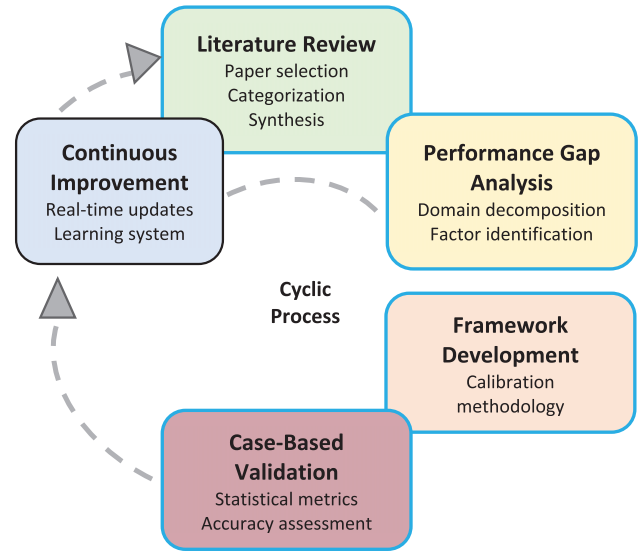

Fig. 3. Methodological framework overview.

datasets and published case studies. This allows us to evaluate the framework's effectiveness across diverse scenarios without requiring extensive new field data collection.

The validation methodology involves applying the reconciliation framework to published case studies with well-documented performance gaps, then comparing the reconciled predictions with actual measured performance. We assess improvement using standard statistical metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and bias metrics such as Mean Bias Error (MBE). This approach provides a preliminary validation of the framework while acknowledging the limitations of using secondary data. Fig. 3 provides an overview of the methodological framework, showing the integration of literature review, performance gap analysis, framework development, and validation components in a systematic approach to developing the reconciliation methodology.

4. PERFORMANCE GAP ANALYSIS

4.1. Resource Assessment Domain

The resource assessment domain encompasses factors related to the characterization and prediction of the solar resource available to a PV system. Analysis of the literature reveals several key challenges in this domain that contribute significantly to performance gaps.

4.1.1. Data Source Uncertainties

The choice of solar resource data represents a fundamental source of uncertainty in performance prediction. Stein and Klise [9] demonstrated that different data sources can produce variations of 3%–5% in annual energy predictions for the same location. These variations stem from different measurement methodologies, spatial resolution, and underlying models used to derive irradiance values. Table II provides a comprehensive comparison of solar resource data sources, showing the trade-offs between accuracy and spatial coverage for different measurement approaches.

TABLE II: COMPARISON OF SOLAR RESOURCE DATA SOURCES

Data source type	Typical GHI uncertainty (Hourly)	Typical GHI uncertainty (Annual)	Key limitations	Best applications
Ground measurements	2%–3%	1%–2%	Limited spatial coverage, maintenance requirements	Reference validation, high-accuracy applications
Satellite models	4%–8%	3%–5%	Biases in cloudy/high-aerosol regions	Utility-scale development, locations without ground data
Reanalysis datasets	8%–15%	5%–8%	Coarse resolution, systematic biases	Regional studies, long-term analysis
Typical meteorological year	N/A	3%–5%	Cannot capture inter-annual variability	System design, financial modeling

Ground-measured data from high-quality meteorological stations historically provided the reference standard for resource assessment. However, as highlighted by Mondol *et al*. [10], even well-maintained ground stations introduce measurement uncertainties of 2%–3% for global horizontal irradiance (GHI) and 3%–5% for direct normal irradiance (DNI). Moreover, the limited geographical coverage of ground stations necessitates spatial interpolation, introducing additional uncertainty for sites distant from measurement locations.

Satellite-derived datasets have become increasingly prevalent for utility-scale project development due to their global coverage and improving accuracy. Comprehensive spatial assessment of solar energy potential [27] has demonstrated both the opportunities and challenges in resource characterization across diverse geographical regions. Yang *et al*. [25] found that modern satellite models achieve typical uncertainties of 4%–8% for hourly GHI and 10%–15% for hourly DNI, with improved performance at larger temporal aggregations. However, systematic biases persist in regions with frequent cloud cover, high aerosol content, or snow cover, which can significantly impact annual energy predictions.

4.1.2. Temporal Representaion Challenges

The temporal representation of solar resource data introduces another significant source of discrepancy. Most performance modeling relies on Typical Meteorological Year (TMY) datasets that synthesize representative conditions rather than predicting actual weather for a specific period. The IEA Report [8] noted that TMY data introduces inherent uncertainty in annual predictions due to natural climate variability, typically ranging from 3%–5% for most locations but exceeding 8% in regions with high inter-annual variability.

Harrison and Jiang [6] found that the choice of reference year for resource data explained approximately 40% of the observed performance gap in their case study. Systems designed using TMY data may consistently under or over-perform depending on whether the actual meteorological conditions diverge from the typical pattern. This effect is particularly pronounced for short operational periods (1–3 years) before regression to long-term means occurs.

Temporal resolution of resource data represents another important consideration. Traditional hourly data fails to capture high-frequency irradiance variations that impact system performance, particularly for systems with high DC/AC ratios where clipping losses may be underestimated. Meng *et al*. [11] demonstrated that sub-hourly

variations can impact annual energy predictions by 1%–3%, with greater impact for systems employing single-axis tracking.

4.1.3. Plane-of-Array Translation

Translating horizontal irradiance data to the plane-of-array (POA) represents another significant source of uncertainty. Most resource datasets provide global horizontal irradiance, which must be decomposed into direct and diffuse components and then transposed to the tilted plane of the PV array. Each step in this process introduces additional uncertainty.

According to Stein and Klise [9], decomposition models introduce uncertainties of 5%–15% for hourly diffuse estimates, while transposition models add 2%–5% uncertainty for fixed-tilt systems. For tracking systems, the uncertainty increases due to the complexity of modeling the changing orientation throughout the day. Yang *et al*. [25] found that different transposition models can produce variations of 2%–4% in annual energy predictions for the same horizontal irradiance dataset, with larger variations in locations dominated by diffuse irradiance.

4.2. System Response Domain

The system response domain encompasses factors related to how the PV system converts the available solar resource into electrical energy. Analysis of the literature reveals several critical factors in this domain that contribute to the performance gap.

4.2.1. Module Performance Characterization

Accurate characterization of PV module performance under varying operating conditions represents a fundamental challenge in performance modeling. Standard Test Conditions (STC) provide a reference point for module specifications, but real-world operation rarely matches these idealized conditions.

Kurtz *et al*. [10] documented systematic differences between nameplate ratings and actual field performance of PV modules, with measured power output frequently 2%–3% below rated values even before degradation effects. This initial deviation creates a baseline offset in performance projections that persists throughout system operation.

Temperature modeling introduces additional uncertainty. While the temperature coefficient of power (typically -0.3% to -0.5% per °C for crystalline silicon modules) is well-established, accurately predicting module temperature based on ambient conditions and mounting configuration remains challenging. Zuhaib *et al*. [17]

found that standard temperature models could introduce errors of 1%–3% in annual energy predictions, with larger deviations in locations with extreme temperatures or high wind variability.

Spectral response represents another source of discrepancy, particularly for thin-film technologies with narrower spectral sensitivity. Most performance models use simplified spectral corrections or assume an average Air Mass modifier. King *et al*. [4] developed more sophisticated spectral corrections that improved prediction accuracy, but these require detailed spectral data rarely available for commercial projects.

4.2.2. System Losses Characterization

Accurate characterization of system losses represents a significant challenge in performance modeling. While some losses can be calculated from first principles (e.g., DC wiring losses), others rely heavily on empirical assumptions.

Soiling losses represent a particularly challenging aspect of performance modeling. Deceglie *et al*. [12] demonstrated that soiling can reduce annual energy production by 2%–6% in many environments, with significant seasonal and geographic variations. Traditional models typically apply a constant soiling loss factor based on location type, failing to capture the dynamic nature of soiling accumulation and removal.

Snow losses present similar challenges, with high geographic and seasonal variability that is difficult to predict accurately. Harrison and Jiang [6] found that snow-related losses explained approximately 25% of the winter performance gap in their northern climate case study, highlighting the importance of location-specific modeling for accurate predictions.

Mismatch losses due to manufacturing tolerance, non-uniform soiling, or partial shading represent another source of uncertainty. Meng *et al*. [12] found that even among identical systems installed side-by-side, mismatch effects could explain production differences of 1%–3%, suggesting intrinsic variability that deterministic models struggle to capture.

4.2.3. Degradation Patterns

Long-term degradation represents a critical factor for lifetime energy predictions. Most performance models assume linear degradation rates based on industry averages, typically 0.5%–0.7% per year for crystalline silicon. However, actual degradation patterns show significant variability.

Jordan and Kurtz [14] conducted an extensive review of degradation rates, finding median values of 0.5%–0.6% per year but with a wide distribution ranging from 0.2% to over 1% depending on technology, climate, and installation quality. Importantly, they found that degradation is often non-linear, with higher rates in the first year (light-induced degradation) followed by a slower long-term trend.

Lindig *et al*. [1] established an international framework for calculating performance loss rates from operational data, revealing significant variations across different climates and technologies. Their work demonstrated that standard degradation assumptions often fail to capture the complex interactions between technology, climate, and installation quality that determine actual performance loss rates.

4.3. Operational Domain

The operational domain encompasses factors related to system operation, maintenance, and interaction with the broader electrical infrastructure. Literature analysis reveals several key factors in this domain that contribute to performance gaps.

4.3.1. System Availability and Curtailment

System availability represents a significant source of discrepancy between modeled and actual performance. While performance models typically assume high availability (98%–99%), actual systems may experience more frequent or longer outages due to component failures, grid issues, or maintenance activities.

The IEA Report [8] noted that availability losses in commercial systems typically range from 0.5% to 3% annually, with higher values for systems with less robust monitoring and maintenance programs. Importantly, availability losses often cluster in high-irradiance periods due to inverter thermal shutdown or grid-related curtailment, magnifying their impact on annual energy production.

Curtailment due to grid constraints represents an increasing challenge as PV penetration grows. Yang *et al*. [25] highlighted that grid curtailment typically appears as a performance gap in standard analysis, though it represents an external constraint rather than a modeling error. Proper accounting for curtailment events requires detailed operational logs often unavailable in simplified performance assessments.

4.3.2. Inverter Performance

Inverter performance modeling presents unique challenges due to the non-linear efficiency curve and sensitivity to operating conditions. Driesse *et al*. [23] demonstrated that actual inverter efficiency curves can deviate significantly from manufacturer specifications, particularly under low-load conditions or extreme temperatures.

Most performance models use a weighted efficiency metric (e.g., CEC or Euro efficiency) that may not accurately reflect the actual operating profile at a specific location. This simplified approach can introduce errors of 1%–2% in annual energy predictions, particularly for systems with atypical DC/AC ratios or operating in extreme climates.

Clipping losses due to inverter power limitations represent another source of discrepancy, particularly for systems with high DC/AC ratios designed to maximize energy harvest in non-peak conditions. Accurate prediction of clipping losses requires high-resolution irradiance data to capture brief high-irradiance periods that may be averaged out in hourly data.

4.3.3. Maintenance Practices

Maintenance practices significantly impact system performance but are rarely incorporated explicitly in performance models. Regular cleaning, vegetation management, and proactive repairs can substantially
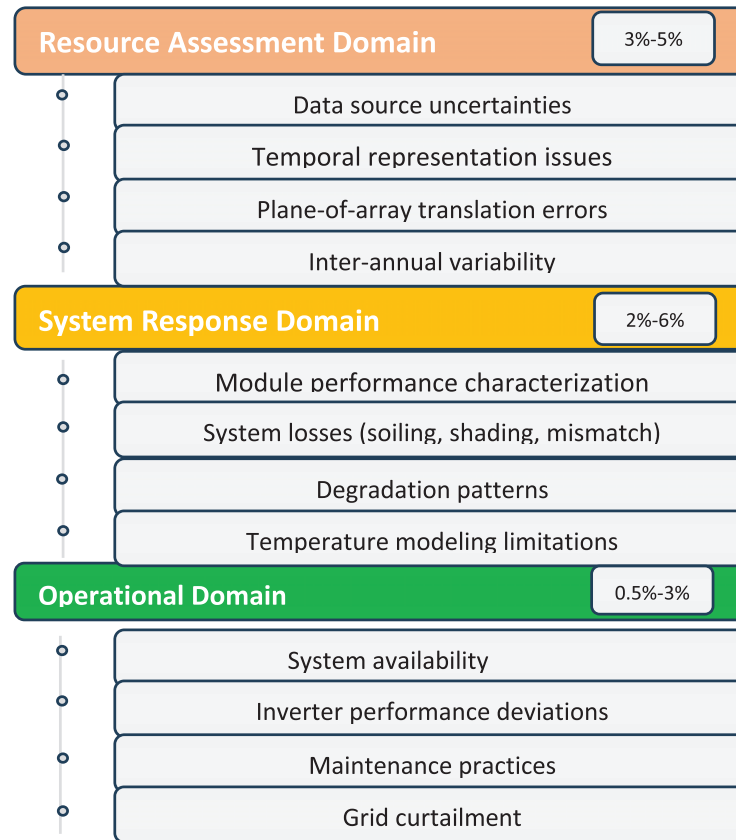
Fig. 4. Categorization of performance gap contributing factors.

reduce performance gaps, while deferred maintenance exacerbates them.

Deceglie *et al.* [12] demonstrated that cleaning frequency significantly impacts annual soiling losses, with optimized cleaning schedules potentially reducing these losses by 30%–50% compared to calendar-based approaches. Zuhaib *et al.* [17] found that proactive maintenance could reduce annual performance losses by 1%–3%, with greater impact in harsh environments where degradation accelerates without proper intervention.

The challenge for performance modeling lies in predicting future maintenance practices at the design stage. Most models incorporate simplified assumptions about maintenance frequency and effectiveness, which may not reflect actual operational decisions driven by resource constraints or changing economic conditions over the project lifetime. Industry best practices for operation and maintenance [28] provide guidance for optimizing these decisions, but site-specific factors often require customized approaches. As shown in Fig. 4, the categorization of performance gap contributing factors reveals the relative distribution of error sources across the three analytical domains, with operational factors contributing approximately 0.5%–3% of total discrepancies.

4.4. Temporal Patterns in Performance Gaps

Temporal patterns in performance gaps provide valuable insights for reconciliation approaches. Analysis of the literature reveals distinct patterns across different timescales, from diurnal to inter-annual. The temporal patterns in performance gaps are summarized in Table III, which categorizes typical patterns across different timescales and identifies the primary contributing factors for each temporal domain.

4.4.1. Diurnal Patterns

At the diurnal scale, performance gaps typically show systematic patterns related to solar angle and temperature effects. Harrison and Jiang [6] observed larger morning/evening discrepancies compared to mid-day, with models consistently overestimating performance at low sun angles. This pattern suggests limitations in handling incidence angle effects and low-irradiance inverter performance.

Temperature-related performance gaps show the opposite pattern, with models typically underestimating temperature derating during peak irradiance hours. Zuhaib *et al.* [17] found that standard temperature models consistently underestimated module temperatures during periods of high irradiance with low wind speeds, resulting in overprediction of mid-day performance.

4.4.2. Seasonal Patterns

Seasonal patterns in performance gaps are particularly pronounced in regions with distinct seasonal variations in climate. Harrison and Jiang [6] found that winter performance gaps in northern climates were 2–3 times larger than summer gaps, primarily due to snow losses and low-irradiance effects not adequately captured in models.

Soiling-related seasonal patterns vary by climate type. In regions with distinct dry seasons, soiling effects progressively increase until the onset of rains, creating a characteristic sawtooth pattern in performance ratios. Deceglie *et al.* [12] developed a methodology to isolate

TABLE III: Temporal Patterns in Performance Gaps

Temporal scale	Typical pattern	Primary contributing factors	Reconciliation implications
Diurnal	Large gaps at low sun angle	Incidence angle effects, low-irradiance inverter performance	Improve angle-of-incidence modeling, refine inverter low-irradiance models
Seasonal	Large winter gaps in temperature climates	Snow losses, spectral effects, temperature model limitations	Develop climate-specific seasonal correction factor
Annual cycle	Progressive decline in performance ratio until rainfall/ cleaning	Soiling accumulation, module degradation	Implement dynamic soiling models with climate triggers
Inter-Annual	Deviation from linear degradation assumptions	Non-linear degradation, climate variability	Replace linear degradation with technology-specific models

TABLE IV: Framework Implementation Stages and Requirments

Implementation stage	Key activities	Data requirements	Expected outcomes
Initial assessment	Performance gap quantification, preliminary decomposition	Minimum 3 months operational data, original model predictions	Baseline performance gap metrics, preliminary diagnosis
Basic calibration	Resource and system response calibration for major factors	6–12 months operational data with meteorological context	First-order corrected model, improved short-term predictions
Comprehensive calibration	Complete domain decomposition, detailed parameter calibration	12+ months operational data, detailed operational logs	Fully calibrated model with domain-specific adjustments
Operational integration	Implementation of continuous monitoring and adaptive calibration	Ongoing performance and meteorological data streams	Self-improving prediction system with scenario-based forecasts

these patterns from performance data, enabling more accurate characterization of seasonal soiling dynamics.

4.4.3. Inter-Annual Patterns

Inter-annual patterns in performance gaps provide insights into long-term degradation and climate variability effects. Jordan and Kurtz [14] found that actual degradation rates often deviate from the linear assumptions used in most performance models, with higher initial rates followed by more gradual decline.

Climate variability introduces another source of inter-annual variation. The IEA Report [8] noted that natural climate cycles can cause year-to-year variations in solar resource of 3%–8%, translating directly to energy yield variations that may appear as performance gaps when compared to TMY-based predictions.

4.5. Temporal Patterns in Performance Gaps

The financial implications of performance gaps extend beyond simple energy production shortfalls. Detailed analysis of the literature reveals several key impact pathways that affect project economics. Table IV outlines the framework implementation stages and requirements, showing how the reconciliation approach can be deployed with varying levels of data availability and analytical sophistication.

4.5.1. Revenue Impact

The direct revenue impact of performance gaps depends on the project's revenue structure. For projects with fixed Power Purchase Agreements (PPAs), each percentage point of energy underperformance translates directly to an equivalent percentage of revenue reduction. Kurtz *et al*. [3] estimated that a 5% performance gap in a typical utility-scale PV project represents approximately 3%–4% reduction in project Internal Rate of Return (IRR), potentially threatening project viability in competitive markets.

Projects with merchant revenue models or capacity-based payment structures experience more complex financial impacts. Yang *et al*. [25] noted that performance gaps during high-price periods have disproportionate revenue impacts, highlighting the importance of time-of-delivery performance accuracy for financial projections.

4.5.2. Financing Impact

Performance gaps significantly impact project financing through risk perception and required returns. The IEA Report [8] highlighted that financing entities typically apply risk adjustments to energy yield predictions based on perceived uncertainty, with higher uncertainty translating to larger contingency reserves and higher financing costs.

Traditional P90 statistical approaches attempt to quantify this uncertainty but typically assume normally distributed, independent performance variations. Reich *et al*. [7] demonstrated that these assumptions often underestimate actual performance risk, particularly for systematic biases that persist across multiple years.

4.5.3. Long-term Valuation Impact

Performance gaps impact long-term asset valuation and secondary market transactions. System underperformance relative to original projections typically triggers valuation adjustments during refinancing or acquisition, potentially creating significant financial losses for original investors.

The IEA Report [8] noted that secondary market transactions increasingly incorporate performance reconciliation assessments to establish new baseline expectations, with systems demonstrating consistent underperformance experiencing valuation discounts of 5%–15% depending on the severity and understood causes of the performance gap. The financial impact pathways of performance gaps are illustrated in Fig. 5, demonstrating how technical discrepancies translate into primary, secondary, and tertiary economic consequences for project stakeholders.
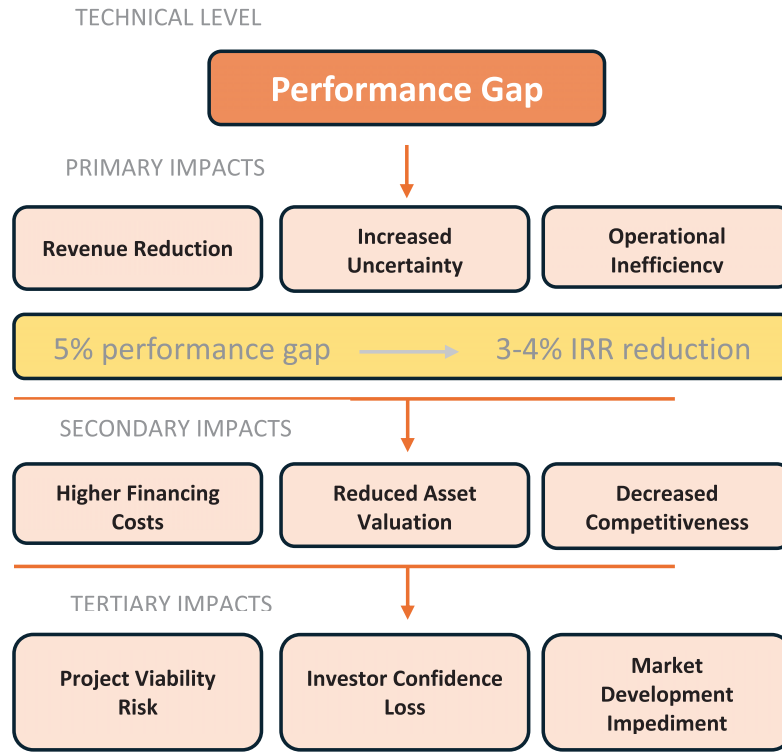
Fig. 5. Financial impact pathways of performance gaps.

5. PROPOSED RECONCILIATION FRAMEWORK

5.1. Framework Overview

Based on our analysis of performance gap factors and existing reconciliation approaches, we propose a comprehensive framework for reconciling modeled predictions with operational data. The framework incorporates three key components: (1) structured performance gap decomposition, (2) model calibration methodology, and (3) forward prediction adjustment.

The proposed framework is designed to address the limitations of existing approaches while incorporating their strengths. Unlike single-factor correction approaches, our framework provides a systematic methodology for identifying and addressing multiple sources of discrepancy simultaneously. The modular structure allows implementation with varying levels of data availability, making it applicable across diverse operational contexts. Fig. 6 provides an overview of the proposed reconciliation framework, showing the integration of performance data, meteorological inputs, and operational logs through structured decomposition and model calibration processes.

5.2. Structured Performance Gap Decomposition

The first component of our framework involves structured decomposition of observed performance gaps into domain-specific components using statistical pattern recognition and physical insights. This decomposition enables targeted corrections rather than generic adjustments.

5.2.1. Resource Domain Decomposition

For the resource assessment domain, we adapt the methodology proposed by Yagli *et al*. [16] to identify systematic biases in resource data. The approach involves comparing observed performance patterns with expected response to standardized weather variables, isolating resource-related discrepancies from system-specific effects.

Key metrics for resource domain decomposition include:

- *Irradiance Bias Index (IBI):* Quantifies the systematic bias in irradiance estimates by comparing the ratio of measured to modeled performance across different irradiance bins, normalizing for other effects.

$$IBI = (1/N) \times \Sigma_i [(Pmeasured, i / Pmodeled, i)$$
$$\times (Gmodeled, i / Gmeasured, i)]$$

where
N–number of data points
Pmeasured, i–measured power at time i
Pmodeled, i–modeled power at time i
Gmodeled, i–modeled irradiance at time i
Gmeasured, i–measured irradiance at time i

- *Temporal Pattern Correlation (TPC):* Assesses the alignment between predicted and actual performance patterns at different timescales (hourly, daily, monthly), helping identify temporal resolution limitations.

For each timescale τ (hourly, daily, monthly):
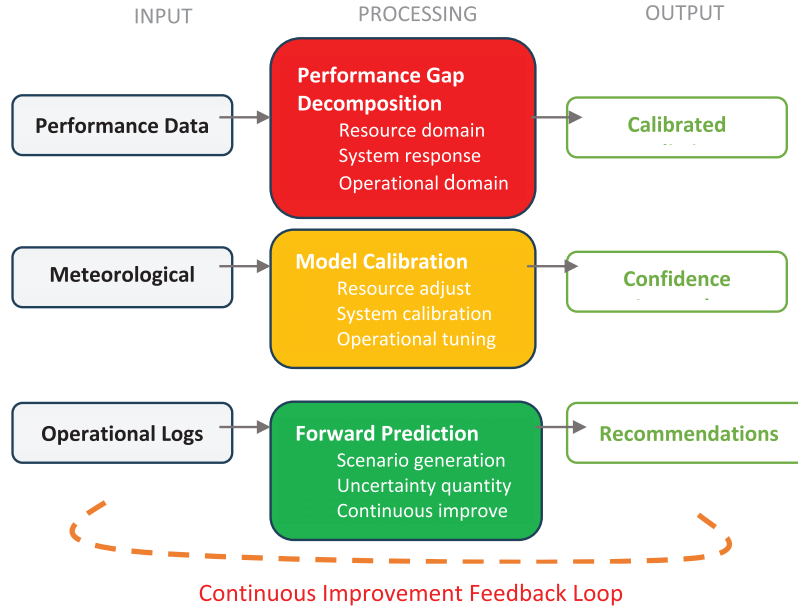
$$TPC\tau = corr(Pmeasured, \tau, Pmodeled, \tau)$$

Fig. 6. Reconciliation framework overview.

where

Pmeasured, τ–aggregated measured performance at timescale τ

Pmodeled, τ–aggregated modeled performance at timescale τ

- *Clear Sky Deviation Metric (CSDM):* Isolates resource assessment errors by examining performance only during clear sky conditions when system response is most predictable:

$$CSDM = median[(Pmeasured, cs/Pmodeled, cs)$$
$$|kt > 0.65]$$

where

Pmeasured, cs–measured power during clear sky conditions

Pmodeled, cs–modeled power during clear sky conditions

kt–clearness index threshold for clear sky identification

5.2.2. System Response Domain Decomposition

For the system response domain, we build on the methodology developed by Deceglie *et al*. [12] for soiling loss extraction, extending it to other system-specific factors. The approach uses pattern recognition to identify characteristic signatures of different loss mechanisms.

Key techniques for system response decomposition include:

- *Temperature Response Analysis:* Isolates temperature model errors by examining the relationship between ambient temperature and performance ratio deviations, controlling for irradiance effects:

$$TRE = \beta_1 - \beta model$$

where β_1 is derived from regression:

$$PR = \beta_0 + \beta_1 (Tmodule - TSTC) + \varepsilon$$

- *Soiling Signature Extraction:* Identifies soiling patterns through analysis of performance trends during dry periods and step changes following precipitation events.

For dry periods between cleaning events:

$$SRdaily = -dPR/dt|precipitation = 0$$

where

PR–Performance Ratio

t–time in days

- *Degradation Trend Analysis:* Separates long-term degradation from seasonal and operational variations using statistical filtering techniques.

5.2.3. Operational Domain Decomposition

For the operational domain, we focus on identifying discrete events and operational patterns that impact performance. This component relies heavily on operational metadata and performance time series analysis.

Key approaches for operational domain decomposition include:

- *Availability Event Detection:* Identifies system downtime and partial availability events through statistical analysis of performance time series, flagging periods with anomalous performance patterns.
- *Inverter Clipping Detection:* Quantifies inverter clipping losses by identifying characteristic plateaus in power output during high-irradiance periods.

- *Curtailment Pattern Recognition:* Distinguishes grid curtailment from performance issues by identifying characteristic patterns and correlating with grid data when available.

5.3. Model Calibration Methodology

The second component of our framework involves systematic calibration of model parameters based on the decomposed performance gap analysis. This calibration process adjusts specific model components rather than applying generic correction factors.

5.3.1. Resource Model Calibration

For resource model calibration, we propose a site-specific correction approach that addresses systematic biases while preserving the statistical properties of the resource data. The methodology includes:

- *Irradiance Scaling Factors:* Adjusts irradiance estimates based on the observed Irradiance Bias Index, with separate factors for different sky conditions (clear, partly cloudy, overcast).
- *Decomposition Model Adjustment:* Calibrates the diffuse fraction model based on observed performance patterns at different sun angles, improving plane-of-array translation accuracy.
- *Temporal Pattern Alignment:* Adjusts the temporal distribution of irradiance to better match observed performance patterns, addressing temporal resolution limitations.

5.3.2. System Response Model Calibration

For system response model calibration, we focus on adjusting specific loss factors and performance parameters based on operational data. Key calibration targets include:

- *Module Temperature Model:* Adjusts temperature coefficients and thermal parameters based on observed temperature response analysis, improving accuracy across different operating conditions.
- *Soiling Loss Model:* Develops site-specific soiling accumulation and removal models based on extracted soiling signatures, capturing seasonal and weather-dependent patterns.
- *Degradation Model:* Replaces simplified linear degradation assumptions with calibrated models based on observed degradation trends, potentially incorporating non-linear components.

5.3.3. Operational Model Calibration

For operational model calibration, we focus on improving the representation of operational factors based on observed patterns. Key aspects include:

- *Availability Model:* Calibrates availability assumptions based on historical patterns, potentially incorporating seasonal variations or correlation with weather events.

- *Inverter Performance Model:* Refines inverter efficiency curves and clipping behavior based on operational data, improving accuracy across the operating range.
- *Maintenance Impact Model:* Develops predictive models for maintenance effects based on historical patterns, enabling more accurate long-term performance projections.

5.4. Forward Prediction Adjustment

The third component of our framework involves applying calibrated models to improve forward predictions while accounting for inherent uncertainties. This component addresses the critical need for accurate future performance projections rather than simply explaining historical discrepancies.

5.4.1. Scenario-Based Prediction

Rather than producing deterministic point forecasts, our framework employs a scenario-based approach that captures the range of plausible outcomes. This methodology includes:

- *Scenario Definition:* Develops multiple scenarios representing different combinations of resource conditions, system response, and operational factors.
- *Probability Assignment:* Assigns probability weights to each scenario based on historical patterns and calibrated model uncertainties.
- Aggregated Prediction: Combines scenario results into probabilistic forecasts with explicit confidence intervals, providing more nuanced information for decision-making.

5.4.2. Uncertainty Quantification

Our framework incorporates explicit uncertainty quantification throughout the prediction process, addressing the limitations of traditional P50/P90 approaches. Key aspects include:

- *Domain-Specific Uncertainty:* Quantifies uncertainty components for each domain (resource, system, operational) separately before combining into aggregate metrics.
- *Correlation Modeling:* Accounts for correlations between different uncertainty sources, avoiding the underestimation that occurs when assuming independence.
- *Confidence Interval Calculation:* Provides more accurate confidence intervals based on calibrated uncertainty models rather than generic assumptions.

5.4.3. Continuous Improvement Process

The framework incorporates a continuous improvement process that updates calibrations as new operational data

TABLE V: Framework Capability Under Different Data Availability Scenarios

Data availability scenario	Resource domain capability	System response domain capability	Operational domain capability
Comprehensive data	Full calibration with uncertainty quantification	Detailed loss factor decomposition and calibration	Complete operational model calibration
Limited meteorological data	Basic bias correction with higher uncertainty	Simplified loss decomposition with combined factors	Limited operational pattern recognition
Missing operational logs	Full resource calibration	Detailed system response calibration	Basic availability detection only
Production data only	Simplified aggregate calibration without domain-specific adjustments	Combined calibration without factor isolation	Minimal operational insights

becomes available. This aspect is particularly important for long-term performance optimization:

- *Performance Monitoring:* Continuously compares actual performance with predictions, identifying emerging discrepancies.
- *Incremental Calibration:* Updates model parameters based on new data without requiring complete recalibration.
- *Learning System:* Accumulates knowledge about site-specific patterns, progressively improving prediction accuracy over time.

5.5. Practical Implementation Considerations

Practical implementation of the reconciliation framework requires consideration of data availability, computational requirements, and integration with existing workflows. Our analysis identifies several key considerations for effective implementation.

5.5.1. Data Requirements

The framework's effectiveness depends critically on data quality and completeness. Minimum data requirements include:

- *Performance Data:* Revenue-grade production measurements at hourly or sub-hourly resolution, ideally with inverter-level granularity.
- *Meteorological Data:* On-site measurements of global horizontal irradiance, plane-of-array irradiance (where available), ambient temperature, and wind speed.
- *Operational Logs:* Records of outages, maintenance activities, and grid curtailment events.

The framework is designed to function with varying levels of data completeness, with degraded but still valuable functionality when certain elements are unavailable. As detailed in Table V, the framework's capability adapts to different data availability scenarios, maintaining core functionality even when certain data streams are unavailable.

5.5.2. Integration with Existing Systems

Integration with existing performance monitoring and asset management systems represents a critical consideration for practical implementation. The framework is designed with a modular architecture that facilitates integration through several pathways:

- *Data Connectors:* Standardized interfaces for extracting performance and meteorological data from common monitoring platforms.
- *Calibration Services:* API-based services that perform specific calibration functions without requiring full system implementation.
- *Visualization Components:* Standardized outputs for integration with existing dashboards and reporting tools.

This approach enables gradual implementation without requiring wholesale replacement of existing systems, reducing adoption barriers while still providing substantial benefits.

5.5.3. Cost-Benefit Considerations

The economic value of implementing the reconciliation framework varies depending on project scale, performance issues, and financial structure. Our analysis suggests several key considerations for cost-benefit assessment:

- *Project Scale Effects:* Implementation costs scale sub-linearly with project size, making the framework more economically viable for larger systems or portfolios.
- *Performance Gap Magnitude:* Economic benefits increase with the magnitude of existing performance gaps, as larger improvements yield greater financial returns.
- *Revenue Structure Impacts:* Projects with performance-based revenue structures (e.g., merchant plants or performance guarantees) typically realize greater economic benefits from improved predictions.

Initial implementation typically requires 2–4 weeks of analyst time plus ongoing monitoring costs, with payback periods ranging from 6–24 months depending on project specifics. For large portfolios, economies of scale can significantly improve the cost-benefit ratio through shared implementation resources.

5.5.4. Illustrative Example: 50 MW Utility-Scale PV System

To demonstrate the practical application of our reconciliation framework, we present an illustrative example based on typical performance patterns observed in utility-scale PV systems:

System Characteristics

- **Capacity:** 50 MW DC (40 MW AC)

- **Location:** Southwestern United States (high desert climate)
- **Technology:** Crystalline silicon modules with single-axis tracking
- **Operational Period:** 24 months
- **Initial Performance Gap:** -6.2% (annual basis)

Step 1: Performance Gap Decomposition

Using the framework's decomposition methodology, the initial 6.2% gap was attributed to:

Resource Assessment Domain (2.1%)

- Irradiance Bias Index: 0.97 (3% underestimation)
- Clear Sky Deviation Metric: 1.02 (2% overestimation during clear conditions)
- Net resource contribution: -2.1%

System Response Domain (2.8%)

- Temperature model error: -1.2% (underestimated operating temperatures)
- Soiling losses: -1.1% (constant 2% assumed vs. 3.1% measured average)
- Initial module deviation: -0.5% (flash test vs. nameplate)

Operational Domain (1.3%)

- Availability: -0.8% (97.5% actual vs. 98.5% modeled)
- Inverter clipping: -0.3% (underestimated due to hourly data resolution)
- Grid curtailment: -0.2% (not originally modeled)

Step 2: Model Calibration

Based on the decomposition results, the following calibrations were applied:

1. **Resource Calibration:**

   - Adjusted diffuse fraction model with site-specific coefficients
   - Implemented clearness index-dependent scaling factors

2. **System Response Calibration:**

   - Updated temperature model parameters (Uc $= 25$ W/m^{2}K, Uv $= 1.5$ W·s/m^{3}K)
   - Developed monthly soiling factors based on observed patterns
   - Applied -0.5% adjustment to module power rating

3. **Operational Calibration:**

   - Revised availability to 97.5% with seasonal variation
   - Refined inverter model with sub-hourly clipping estimation

Step 3: Results

After calibration:

- **Year 1 Retrospective:** MAE reduced from 6.2% to 1.8%
- **Year 2 Prediction:** Initial gap of 5.8% reduced to 2.1%

- **Uncertainty Bounds:** P90 confidence interval narrowed from $\pm4.5\%$ to $\pm2.8\%$

Key Insights

1. **Dominant Factors:** System response factors (particularly temperature and soiling) contributed most significantly to the performance gap in this desert environment.
2. **Temporal Patterns:** Soiling showed strong seasonal accumulation with 4–5 discrete cleaning events annually from rainfall.
3. **Calibration Impact:** The framework reduced prediction error by 71% while providing actionable insights for O&M optimization.

This example illustrates how the framework systematically identifies, quantifies, and corrects performance gap sources, leading to substantially improved prediction accuracy and operational insights.

6. CONCLUSION AND FUTURE RESEARCH

6.1. Summary of Key Findings

This study has examined the persistent gap between predicted and actual performance in photovoltaic systems, developing a comprehensive framework for analyzing and reconciling these discrepancies. Our analysis reveals several key findings that advance the understanding of performance gaps and provide pathways for improvement.

First, performance gaps are multi-factorial and domain-specific, requiring structured approaches to decomposition and reconciliation. Our analysis demonstrates that resource assessment, system response, and operational factors all contribute significantly to observed discrepancies, with their relative importance varying by location, technology, and operational context. This complexity explains why simplified correction approaches often fail to provide lasting improvements.

Second, temporal patterns in performance gaps provide valuable diagnostic insights that can guide targeted reconciliation efforts. Diurnal, seasonal, and inter-annual patterns reflect different underlying mechanisms that require specific analytical approaches and correction methodologies. The temporal signature of performance gaps often reveals more about their causes than aggregate metrics alone.

Third, existing modeling approaches incorporate simplifications and assumptions that systematically contribute to performance gaps. These include idealized soiling patterns, linear degradation assumptions, simplified temperature models, and generalized loss factors that fail to capture site-specific conditions. While these simplifications are often necessary for tractable modeling, they must be recognized and addressed through calibration processes.

Finally, effective reconciliation requires a structured methodology that addresses multiple domains simultaneously while preserving their interdependencies. Our proposed framework demonstrates that systematic decomposition, model calibration, and forward prediction adjustment can significantly improve alignment between

modeled and actual performance, enhancing both technical accuracy and financial reliability.

6.2. Practical Implications

The practical implications of this research extend across the PV project lifecycle, from initial development through long-term operation and financial transactions.

For project developers, the reconciliation framework provides a methodology for more accurate energy yield predictions that reduce financial risk and improve competitiveness. By identifying and addressing systematic biases in performance models, developers can avoid the costly consequences of energy shortfalls and strengthen stakeholder confidence in project projections.

For system operators, the framework offers a structured approach to performance optimization that goes beyond simple monitoring. By decomposing performance gaps into specific contributing factors, operators can prioritize interventions that deliver the greatest improvement in energy production and financial returns. The continuous calibration aspect ensures that models remain aligned with actual performance as systems age and operating conditions evolve.

For financial stakeholders, the improved prediction accuracy and explicit uncertainty quantification enhance risk assessment and valuation processes. More reliable performance projections reduce the need for conservative risk adjustments, potentially lowering financing costs and improving project economics. The structured nature of the reconciliation process also provides greater transparency for due diligence and asset transactions.

For the broader industry, standardizing approaches to performance gap analysis and reconciliation creates opportunities for improved benchmarking and knowledge sharing. As more projects implement structured reconciliation methodologies, collective understanding of performance factors will improve, driving further refinements in modeling approaches and industry standards.

6.3. Limitations and Future Research

While the proposed framework addresses many limitations of existing approaches, several areas require further research and development.

First, the framework relies on statistical decomposition techniques that have inherent limitations in distinguishing between overlapping effects. Future research should explore advanced pattern recognition approaches, potentially incorporating machine learning techniques, to improve the separation of concurrent factors affecting performance.

Second, the current approach to uncertainty quantification focuses primarily on technical factors rather than operational decisions. Future work should develop more sophisticated models for predicting maintenance practices, replacement schedules, and other operational interventions that significantly impact long-term performance.

Third, the validation methodology employed in this study relies primarily on published case studies rather than comprehensive empirical testing. Future research should implement the framework across diverse utility-scale systems to assess its effectiveness in different contexts and refine the methodology based on practical experience.

Fourth, the increasing deployment of bifacial modules, tracking systems, and storage-coupled PV presents new modeling challenges not fully addressed in the current framework. Future extensions should incorporate these technologies, developing specific decomposition and calibration techniques for their unique performance characteristics.

Finally, the framework currently focuses on reconciling predictions with historical performance rather than optimizing future decisions. An important direction for future research is the integration of reconciliation methodologies with operational optimization algorithms to maximize energy production and financial returns throughout the project lifecycle.

In conclusion, this research contributes to bridging the gap between predicted and actual performance in PV systems, providing both analytical insights and practical methodologies. By systematically addressing the complex factors that create performance discrepancies, the proposed framework advances the industry's capacity for accurate energy yield prediction and long-term performance optimization. As utility-scale PV deployment continues to accelerate globally, improved reconciliation approaches will play an increasingly important role in ensuring that these systems deliver their expected energy and financial benefits.

CONFLICT OF INTEREST

The authors declare that they do not have any conflict of interest.

REFERENCES

[1] Lindig S, Moser D, Curran AJ, Rath K, Khalilnejad A, French RH, *et al.* International collaboration framework for the calculation of performance loss rates: data quality, benchmarks, and trends (towards a uniform methodology). *Prog Photovolt*. 2021;29(6):573–602. doi: 10.1002/pip.3397.

[2] International Renewable Energy Agency. *Renewable energy statistics 2023*. IRENA; 2023.

[3] Kurtz S, Newmiller J, Kimber A, Flottemesch R, Riley E, Dierauf T, *et al*. *Analysis of Photovoltaic System Energy Performance Evaluation Method*. 2013. doi: 10.2172/1111193.

[4] Kratochvil J, Boyson W, King D . *Photovoltaic Array Performance Model*. 2004. doi: 10.2172/919131.

[5] Bordass B, Cohen R, Field J. Energy performance of non-domestic buildings: closing the credibility gap. *Proceedings of the 2004 Improving Energy Efficiency in Commercial Buildings Conference*, pp. 1–10, Frankfurt, Germany, 2004.

[6] Harrison S, Jiang L. An investigation into the energy performance gap between the predicted and measured output of photovoltaic systems using dynamic simulation modelling software—a case study. *Int J Low Carbon Technol*. 2017b;13(1):23–9. doi: 10.1093/ijlct/ctx016.

[7] Reich NH, Mueller B, Armbruster A, van Sark WGJHM, Kiefer K, Reise C. Performance ratio revisited: is PR > 90% realistic? *Prog Photovolt: Res Applicat*. 2012;20(6):717–26. doi: 10.1002/pip.1219.

[8] IEA PVPS Task 13. *Uncertainties in PV System Yield Predictions and Assessments*. Report IEA-PVPS T13-18:2023. International Energy Agency Photovoltaic Power Systems Programme. 2023.

[9] Stein J, Klise G. *Models Used to Assess the Performance of Photovoltaic Systems*. 2009. doi: 10.2172/974415.

[10] Mondol JD, Yohanis YG, Norton B. Solar radiation modelling for the simulation of photovoltaic systems. *Renew Energy*. 2007;33(5):1109–20. doi: 10.1016/j.renene.2007.06.005.

[11] Meng B, Loonen R, Hensen J. Performance variability and implications for yield prediction of rooftop PV systems–Analysis

of 246 identical systems. *Appl Energy*. 2022b;322:119550. doi: 10.1016/j.apenergy.2022.119550.

[12] Deceglie MG, Micheli L, Muller M. Quantifying soiling loss directly from PV yield. *IEEE J Photovolt*. 2018;8(2):547–51. doi: 10.1109/jphotov.2017.2784682.

[13] Sepúlveda-Oviedo EH. A review of operational factors affecting photovoltaic system performance. *Energy Conv Manag: X*. 2023;20:100442. doi: 10.1016/j.ecmx.2023.100442.

[14] Jordan DC, Kurtz SR. Photovoltaic Degradation Rates—An analytical review. *Prog Photovolt*. 2011;21(1):12–29. doi: 10.1002/pip.1182.

[15] Pless S, Deru M, Torcellini P, Hayter S. *Procedure for Measuring and Reporting the Performance of Photovoltaic Systems in Buildings*. 2005. doi: 10.2172/859414.

[16] Yagli GM, Yang D, Srinivasan D. Reconciling solar forecasts: sequential reconciliation. *Sol Energy*. 2019b;179:391–7. doi: 10.1016/j.solener.2018.12.075.

[17] Zuhaib M, Khan HA, Rihan M. Performance analysis of a Utility-Scale Grid integrated solar farm considering physical and environmental factors. *J Inst Eng (India) Ser B*. 2020b;102(2):363–75. doi: 10.1007/s40031-020-00500-6.

[18] De Soto W, Klein S, Beckman W. Improvement and validation of a model for photovoltaic array performance. *Sol Energy*. 2005;80(1):78–88. doi: 10.1016/j.solener.2005.06.010.

[19] Van Dronkelaar C, Dowson M, Spataru C, Mumovic D. A review of the regulatory energy performance gap and its underlying causes in non-domestic buildings. *Front Mech Eng*. 2016;1. doi: 10.3389/fmech.2015.00017.

[20] Burman E, Mumovic D, Kimpian J. Towards measurement and verification of energy performance under the framework of the European directive for energy performance of buildings. *Energy*. 2014;77:153–63. doi: 10.1016/j.energy.2014.05.102.

[21] Livera A, Theristis M, Makrides G, Georghiou GE. Recent advances in failure diagnosis techniques based on performance ratio analysis and partial shading detection. *Sol Energy*. 2019;179:398–405.

[22] Mekhilef S, Saidur R, Kamalisarvestani M. Effect of dust, humidity and air velocity on efficiency of photovoltaic cells. *Renew Sustain Energ Rev*. 2012;16(5):2920–5.

[23] Driesse A, Jain P, Harrison S. Beyond the curves: modeling the electrical efficiency of photovoltaic inverters. *Conference Record of the IEEE Photovoltaic Specialists Conference*, 2008. doi: 10.1109/pvsc.2008.4922827.

[24] Yang D, Quan H, Disfani VR, Liu L. Reconciling solar forecasts: geographical hierarchy. *Sol Energy*. 2017;146:276–86. doi: 10.1016/j.solener.2017.02.010.

[25] Yang D, Wu E, Kleissl J. Operational solar forecasting for the real-time market. *Int J Forecast*. 2019c;35(4):1499–1519. doi: 10.1016/j.ijforecast.2019.03.009.

[26] Hyndman RJ, Lee AJ, Wang E. Fast computation of reconciled forecasts for hierarchical and grouped time series. *Comput Statist Data Anal*. 2015;97:16–32. doi: 10.1016/j.csda.2015.11.007.

[27] Prăvălie R, Patriche C, Bandoc G. Spatial assessment of solar energy potential at global scale. A geographical approach. *J Clean Prod*. 2019;209:692–721.

[28] National Renewable Energy Laboratory. *Best Practices for Operation and Maintenance of Photovoltaic and Energy Storage Systems*. NREL; 2023.