



Comparative mathematical benchmarking of short-term photovoltaic power forecasting model families for data-scarce African grids

Peter Nyanor ^a, Amevi Acakpovi ^{a,*}

^a Accra Technical University, Barnes Road, Accra, Ghana

ARTICLE INFO

Article history:

Received February 8, 2026

Revised May 24, 2026

Accepted July 15, 2026

Available online July 19, 2026

Published July 30, 2026

Keywords:

Photovoltaic power forecasting,
Data-scarce grids,
African power systems,
Machine learning,
Mathematical benchmarking.

ABSTRACT

Short-term photovoltaic power forecasting is increasingly important for grid integration, operational planning, and renewable energy management; however, model selection remains a challenge in African systems, where historical records, meteorological inputs, and data quality are often scarce. This study developed a comparative mathematical benchmarking framework for evaluating representative short-term photovoltaic forecasting model families under data-scarce African grid conditions. Using secondary evidence from the literature, this study compares baseline, physical or weather-assisted, classical machine-learning, deep-learning, hybrid, and probabilistic forecasting families across four sparse-data scenarios: output-only operation, limited-weather operation, weather-assisted short-history operation, and utility deployment under operational constraints. A suitability score was proposed to integrate predictive accuracy, robustness to sparse or missing data, transferability, interpretability, data burden, and computational burden. The analysis showed that no single model family was universally superior across all scenarios. Classical ensemble machine-learning models appear most suitable when limited weather information is available, weather-assisted physical models are strongest when historical records are short but forecast inputs are available, and endogenous temporal models are most relevant when only output history is available. This study contributes a context-sensitive model-selection framework for photovoltaic forecasting in African grid environments.

1. INTRODUCTION

The rapid integration of photovoltaic (PV) generation into modern power systems has made short-term forecasting an important requirement for dispatch support, reserve allocation, storage coordination, and the reliable integration of variable renewable energy (Barbieri et al., 2017; Gaboitaolelwe et al., 2023; Mellit et al., 2020; Pandžić & Capuder, 2024). Forecast errors (Poti et al., 2023) have been identified

* Corresponding author, E-mail address: aacakpovi@atu.edu.gh



as contributors to operating problems, increased scheduling uncertainty, and reduced system flexibility. This is especially true in power systems with limited balancing resources (Assaf et al., 2023). In many African settings (Mbuya et al., 2022; Zwane et al., 2022), renewable energy growth takes place alongside major infrastructure limits, weak monitoring systems, and unstable grid performance.

Various techniques are currently employed for short-term photovoltaic (PV) forecasting. These methods are typically categorized into physical, statistical or data-driven, and hybrid groups. Recent research has also focused more on deep learning and probabilistic models (Assaf et al., 2023; Gaboitaolelwe et al., 2023; Mellit et al., 2020). Physical and physics-informed models often rely on variables related to irradiance, numerical weather prediction (NWP), and plant specifications to predict PV output. These models are particularly advantageous when weather data is accessible and when interpretability is crucial (Mayer & Gróf, 2021; Pombo et al., 2022; Poti et al., 2023). Statistical and traditional machine-learning models analyze historical PV output and selected weather variables to identify patterns in the data. They remain popular for short-term forecasting due to their ability to capture nonlinear behavior without requiring a comprehensive physical description of the PV system (Gaboitaolelwe et al., 2023; Mellit et al., 2020). Table 1 summarizes the main forecasting families considered in this study and their defining characteristics.

The methodological diversity of PV forecasting models is clearly illustrated in the taxonomy as presented in Figure 1. However, model comparison remains a challenge. The existing studies are differentiated based on parameters such as forecast horizon, sampling intervals, climatic settings, PV physical plant description, and evaluation metrics (Gaboitaolelwe et al., 2023; Pandžić & Capuder, 2024; Salamanis et al., 2020). It is therefore difficult to compare reported performance without misleading conclusions. Additionally, the model performance also depends on local weather patterns, geographic conditions, and data quality. Therefore, claims of a universally best PV forecasting model become weak (Pombo et al., 2022; Salamanis et al., 2020). PV forecasting models' performance inefficiencies are greater in data-scarce environments. The assumptions used in data-rich forecasting studies do not apply here. Short historical records, missing data, limited irradiance measurements, irregular telemetry, or only a few weather inputs plague real-world PV systems (Konstantinou et al., 2021; Pandžić & Capuder, 2024). Figure 1 presents the taxonomy of representative short-term PV power forecasting model families used in the comparative framework.

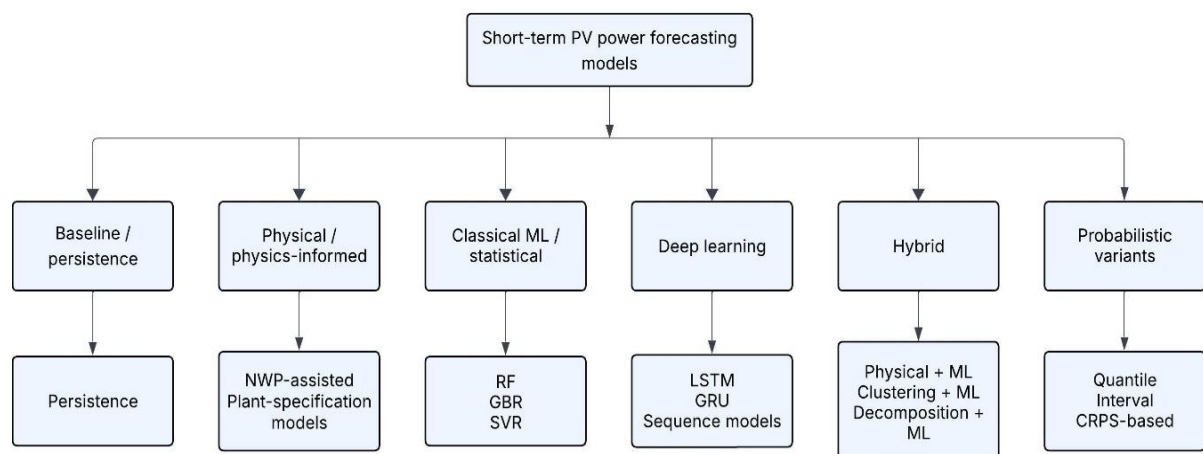


Fig 1. Taxonomy of representative short-term PV power forecasting model families used in the comparative framework.

Table 1. Model-family taxonomy for short-term PV power forecasting under data-scarce African grid conditions

Model family	Core description	Typical inputs	Main strengths	Main limitations	Representative evidence from this study
Baseline/ persistence	Uses recently observed PV output as the forecast for the next period	Historical PV output only	Very simple, low data burden, low computational cost, strong benchmark for short horizons	Limited ability to capture changing weather dynamics; usually outperformed by more advanced methods when useful inputs are available	Abdulai et al. (2023)
Physical / physics- informed / NWP- assisted	Uses weather forecasts, irradiance-related variables, and plant specifications to estimate PV output	NWP/weather data, irradiance-related variables, plant parameters, limited or optional historical PV output	Good interpretability, strong transferability, useful when long historical plant records are weak or unavailable	Performance depends on quality of weather forecasts and plant characterization; may be weaker under rapidly changing local conditions if weather inputs are poor	Mellit et al. (2020); Poti et al. (2023)
Classical statistical / machine- learning	Learns relationships between PV output and historical or exogenous variables using non-deep models	Historical PV output, selected weather variables, optional engineered features	Strong performance with moderate data, flexible nonlinear modeling, usually easier to deploy than deep or hybrid methods	Requires feature selection/tuning; transferability may depend on data quality and local conditions	Abdulai et al. (2023); Gaboitaolelwe et al. (2023)
Deep learning	Uses multilayer neural or recurrent architectures to learn temporal and nonlinear patterns in PV generation	Historical PV output, optionally weather variables, larger training data	High predictive potential, effective for complex temporal patterns, useful in endogenous-only settings	Higher data burden, higher computational burden, lower interpretability, more sensitive to preprocessing and sequence quality	Konstantinou et al. (2021); Gaboitaolelwe et al. (2023)
Hybrid models	Combines two or more approaches, such as physical + ML, decomposition + ML, or clustering + prediction models	Mixed inputs depending on architecture; often historical PV output plus weather and engineered features	Often achieves very strong accuracy by combining complementary model strengths	Higher complexity, greater tuning burden, potentially higher input burden, lower operational transparency	Mellit et al. (2020); Gaboitaolelwe et al. (2023)

Table 1. Model-family taxonomy for short-term PV power forecasting under data-scarce African grid conditions (*Continued*)

Model family	Core description	Typical inputs	Main strengths	Main limitations	Representative evidence from this study
Probabilistic variants	Extends deterministic models to produce uncertainty-aware outputs such as intervals, quantiles, or distributions	Same as parent model family, with additional uncertainty modeling logic	Provides uncertainty information for operational planning, improves decision usefulness, supports risk-aware dispatch and reserve planning	More complex evaluation and interpretation; requires probabilistic metrics such as CRPS	Abdulai et al. (2023) ; Gaboitaolelwe et al. (2023)

This problem is especially significant in the African context. Solar forecasting research on the continent has grown, but it remains relatively limited and unevenly distributed. Zwane et al. (2022) report that only 242 Africa-related and 34 southern Africa-related solar forecasting documents were retrieved for 2000–2021, with publication growth becoming more visible only after 2015 and South Africa leading the continent in output and contribution. The existing African studies already reveal substantial diversity in forecasting conditions and modeling needs. Abdulai et al. (2023) employed a utility-scale study at the Bui Solar Generating Station in Ghana. The study found that random forest and gradient boosting outperformed persistence, with a probabilistic random-forest formulation performing best overall. South African forecasting studies have explored both machine-learning and weather-assisted approaches, including hourly irradiance forecasting and day-ahead PV power prediction using numerical weather prediction and plant specifications (Mutavhatsindi et al., 2020; Poti et al., 2023; Ratshilengo et al., 2021). Recent African evidence now spans utility-scale forecasting in Ghana (Abdulai et al., 2023), weather-assisted day-ahead forecasting in South Africa (Poti et al., 2023), AI-based PV forecasting in Cameroon (Mbey et al., 2024), and reliability-oriented prediction in weak-grid environments (Mbuya et al., 2022; Mutavhatsindi et al., 2020; Ratshilengo et al., 2021; Zwane et al., 2022). This indicates the regional evidence base, although still limited, is becoming more diverse.

Therefore, the literature is not lacking in forecasting models, but it is lacking in a comparative framework sensitive to the information specific to the context. For this purpose, the study proposes a comparative mathematical benchmarking framework for the major short-term PV power forecasting model families. Further, constraints of performance under data-scarce African grid conditions were imposed. The work evaluates baseline, physical or weather-assisted, classical machine learning, deep learning, hybrid and probabilistic models under various sparse data conditions.

2. METHODOLOGY

2.1 Research design

This study is based on a comparative design of evidence of PV forecast model performance evidence from published works. A mathematical benchmarking framework is applied to short-term PV forecast model families under data-scarce African grid conditions. The study considers the parameters studied in existing work on physical, machine-learning, deep-learning, hybrid, and probabilistic forecasting

models. While direct comparison has been established to be difficult, the study examines which model family is more suitable under different sparse-data conditions in African deployment settings.

This design also fits the African evidence base. Research on solar forecasting in Africa is growing, but it is still limited and uneven across the continent. Because of this, a structured comparative review is more useful at this stage than a narrow local experiment (Zwane et al., 2022). The overall workflow of the study is illustrated in Figure 2.

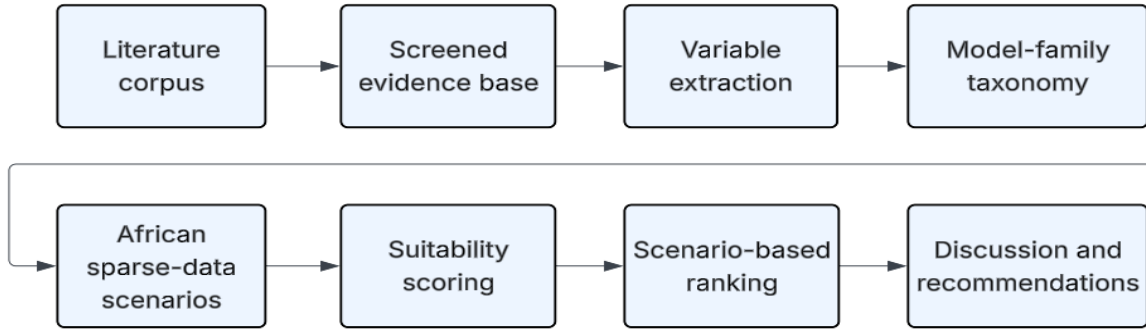


Fig 2. Workflow of the comparative mathematical benchmarking framework for short-term PV forecasting model families under data-scarce African grid conditions.

2.2 Evidence source and screening logic

The evidence base for the study was drawn from the compiled paper corpus assembled for this manuscript. The corpus includes review papers, benchmarking studies, direct short-term PV power forecasting papers, and African-context forecasting studies. Because not all papers in the corpus are equally suitable for comparative ranking, the studies were screened into three functional groups: core comparison papers, methodological support papers, and African context support papers.

Core comparison papers were defined as studies that directly address short-term PV power forecasting and report sufficient information on forecast target, horizon, model structure, input variables, and evaluation metrics. This group included studies such as Abdulai et al. (2023), Poti et al. (2023), Konstantinou et al. (2021), Hossain and Mahmood (2020), and Li & Liu (2023). Other papers were kept as methodological support where they helped with model classification, benchmarking logic, probabilistic comparison, or evaluation design, even when they were not used directly in the ranking stage (Fernandez-Jimenez et al., 2023; Gaboitoalelwe et al., 2023; Mellit et al., 2020; Pandžić & Capuder, 2024; Salamanis et al., 2020). African context papers were also kept to strengthen the regional framing of the study. These papers were especially useful where they showed forecasting conditions, grid limits, or research gaps in Africa (Mbuya et al., 2022; Mutavhatsindi et al., 2020; Ratshilengo et al., 2021; Zwane et al., 2022).

Studies were included in the core comparison layer if they:

- addressed PV power forecasting rather than only irradiance forecasting;
- focused on short-term, very-short-term, intraday, or day-ahead horizons;
- described identifiable forecasting approaches or model families;
- reported at least some information on required inputs or data structure; and
- presented recognized evaluation measures such as RMSE, MAE, nRMSE, NMAE, CRPS, or R^2 (Abdulai et al., 2023; Gaboitoalelwe et al., 2023; Poti et al., 2023).

Papers that were purely bibliometric, purely system-planning oriented, or insufficiently detailed for structured comparison were not used in the ranking stage, although they still informed the background and discussion. Studies were excluded from the core comparison layer if they were purely bibliometric, purely system-planning oriented, focused only on irradiance forecasting without direct PV-power relevance, or lacked sufficient methodological detail for structured comparison.

The screened evidence base retained for the comparative analysis is presented in Table 2.

2.3 Evidence extraction framework

A structured extraction form was used for each study in the core comparison group. The extracted information was arranged into five categories.

The first category covered basic study details. These included the authors, year, country, and African relevance. The second category covered the forecasting setup. This included the forecast target, forecast horizon, time resolution, and site type. The third category covered the model itself. It included the model family, whether the forecast was deterministic or probabilistic, and whether a baseline model was used. The fourth category focused on data needs. It included reliance on past PV output, weather inputs, numerical weather prediction, and how missing or sparse data were handled. The fifth category covered evaluation details. It included the reported metrics, the use of normalized measures, and the main result of each study.

This structure was necessary because studies are not easy to compare. Gaboitaolelwe et al. (2023) noted that comparing raw RMSE and MBE values between systems with different capacities is difficult. All of these, along with data source, forecast horizon and benchmark design, can influence the interpretation of results (Pandžić and Capuder, 2024). This was demonstrated by Abdulai et al. (2023) who recorded not only type but how the two forecast models used (deterministic and probabilistic) can lead to different judgement about performance. Table 3. Comparative scoring matrix used to synthesise model-family suitability.

Table 2. Screened evidence matrix for the comparative study

ID	Paper/context	Target & horizon	Model family	Input/data burden	Main usable finding	Role in paper
E1	Abdulai et al. (Ghana, Bui Solar Generating Station)	PV power, short-term, 15-min intervals	Persistence, RF, GBR, deterministic + probabilistic	Moderate; plant output + weather variables, 341 days	RF probabilistic case was best; average CRPS 1.18%, outperforming persistence and other cases	Core African comparison
E2	Poti et al. (South Africa, intelligent PV forecasting)	PV power, day-ahead	Physics-informed / NWP-assisted	Low-to-moderate historical dependence; NWP + plant specs	Day-ahead model based on NWP and plant specifications; RMSE 425.79 W and 595.10 W, MAE 246.26 W and 238 W, ($R^2 > 0.9$)	Core African comparison

Table 2. Screened evidence matrix for the comparative study (*Continued*)

ID	Paper/context	Target & horizon	Model family	Input/data burden	Main usable finding	Role in paper
E3	Atmosphere 2021 paper (Cyprus)	PV power, short-term, 15-min intervals	Endogenous-only data-driven forecasting	Low external data burden; only historical PV output	Uses only endogenous PV output; explicitly identifies night-time zeros as sparsity and handles missing values by time interpolation	Core sparse-data method paper
E4	Mellit et al. (Advanced Methods review)	PV power forecasting review	Physical, statistical, AI/ML, hybrid	N/A	Strong taxonomy paper; statistical methods are suitable for short-term forecasting, hybrid methods often improve accuracy	Core taxonomy review
E5	Gaboitaolelwe et al. (ML review and comparison)	PV power forecasting review	ML/ensemble / benchmarking review	N/A	Raw RMSE/MBE are hard to compare across different PV scales; normalized metrics are recommended; geographic location matters for fair comparison	Core benchmarking review
E6	Li & Liu (SOM-based regional modelling)	PV power, one-hour-ahead	Hybrid regional SOM + KELM / MR	Moderate-to-high; multi-plant data + meteorological variables	Proposed SF-KELM reduced RMSE by about 52–55% across three plants and showed strong seasonal robustness	Core hybrid comparison
E7	Hossain & Mahmood (LSTM + synthetic weather)	PV power, short-term, 1–24 h	Deep learning / LSTM hybridized with synthetic weather	High; historical irradiance + weather-derived features	Synthetic irradiance forecast improved accuracy up to 33% over hourly sky type and 44.6% over daily sky type; strong DL benchmark	Core deep-learning comparison

Table 2. Screened evidence matrix for the comparative study (*Continued*)

ID	Paper/context	Target & horizon	Model family	Input/data burden	Main usable finding	Role in paper
E8	Zwane et al. (Africa bibliometric analysis)	Solar forecasting in Africa	Bibliometric / scope paper	N/A	242 Africa and 34 southern Africa documents; research gained momentum around 2015; South Africa leads the continent	Core African gap paper
E9	South Africa irradiance forecasting paper	Solar irradiance, hourly	LSTM, SVR, FFNN, QRA	Moderate	Useful for African forecasting context and rigorous evaluation, but not direct PV-power comparison	African context support
E10	Cameroon deep-learning paper	Short-term PV generation + related energy prediction	MLP, SVM, LSTM, ANFIS, GA, hybrid MLP-LSTM-GA	Moderate-to-high	Useful for Central African representation and deep-learning discussion, but broader than direct model-family benchmarking	African support

2.4 Model-family classification

To make the comparison fair, the studies were grouped into broad forecasting model families instead of treating each algorithm on its own as follows:

1. Baseline models.
2. Physical or physics-informed models.
3. Classical statistical or machine-learning models.
4. Deep-learning models, especially recurrent architectures.
5. Hybrid models.
6. Probabilistic variants, treated as uncertainty-aware extensions of the preceding families.

This classification reflects the structure of the review literature and the retained empirical studies (Assaf et al., 2023; Gaboitaolelwe et al., 2023; Mellit et al., 2020).

2.5 Sparse-data African scenario design

For the comparative framework four sparse-data African scenarios were implemented:

- Scenario A: only endogenous operation;
- Scenario B: operation limited by weather;
- Scenario C: short-history operation supported by NWP; and Scenario D: utility deployment under an operational constraint.

These scenarios are different information environments and have been used to interpret model-family suitability in realistic African deployment conditions.

Table 3. Model-family comparative scoring matrix

Model family (A_i)	(R_i)	(T_i)	(I_i)	(D_i)	(C_i)	Main evidence basis	
Accuracy	Robustness	Transferability	Interpretability	Data burden	Computational burden		
Baseline / persistence	2	3	5	5	1	1	Strong benchmark role; difficult to outperform cleanly in short-horizon settings
Physical / NWP-assisted	4	3	4	4	3	2	Good operational performance; less dependence on long historical plant archives
Classical ML / ensemble	4	4	4	3	3	3	RF/GBR performs strongly; probabilistic variants improve decision relevance
Deep learning	4	3	3	2	4	4	Strong performance possible, but usually with higher data and tuning demands
Hybrid / regional	5	4	3	2	4	4	Often highest reported performance, but more complex and input-demanding
Probabilistic variants	4	4	4	3	3	3	Add uncertainty awareness and operational value when interval quality is reported

Scoring scale: 1 = very low, 2 = low, 3 = moderate, 4 = high, 5 = very high.

For D_i and C_i , higher values mean greater penalty.

2.6 Comparative mathematical framework

The core of the methodology is a Sparse-Data African Suitability Score for each model family i , denoted S_i (Equation 1). The score is defined as

$$S_i = w_1 A_i + w_2 R_i + w_3 T_i + w_4 I_i - w_5 D_i - w_6 C_i \quad (1)$$

where:

- A_i = predictive accuracy,
- R_i = robustness to sparse, missing, or noisy data,
- T_i = transferability to African climatic and grid contexts,
- I_i = interpretability or operational transparency,
- D_i = data burden,
- C_i = computational burden.

The weights satisfy (Equation 2):

$$\sum_{k=1}^6 w_k = 1, w_k \geq 0 \quad (2)$$

To ensure scenario realism, a feasibility gate was applied (Equation 3):

$$\Gamma_i = \begin{cases} 1, & \text{if model family } i \text{ is feasible under the given scenario} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

and the scenario-adjusted suitability score is computed as in equation 4:

$$S_i^* = \Gamma_i S_i \quad (4)$$

This formulation is designed to respond directly to the literature gap identified earlier. Published studies already show how different models perform. But they do not always bring together data needs, transferability, and deployment burden when comparing models (Gaboitaolelwe et al., 2023; Pandžić & Capuder, 2024). For that reason, this framework treats predictive accuracy as important, but not enough on its own. The scoring dimensions, operational definitions, and baseline weights used in the suitability framework are presented in Table 4.

Table 4. Scoring criteria and weights

Dimension	Symbol	Interpretation	Scoring rule	Weight
Predictive accuracy	(A_i)	Forecast quality using normalized/comparable metrics	1 = weak, 3 = moderate, 5 = very strong	0.30
Robustness to sparse / missing data	(R_i)	Tolerance to missing values, night-time sparsity, noisy or short histories	1 = no evidence, 3 = partial handling, 5 = explicit robustness	0.20
Transferability to African contexts	(T_i)	Portability across varied climates, grid conditions, and data environments	1 = highly context-bound, 3 = moderate, 5 = broadly portable	0.20
Interpretability / operational transparency	(I_i)	Ease of explanation and operational use	1 = opaque, 3 = moderate, 5 = highly interpretable	0.10
Data burden	(D_i)	Amount/richness of required data	1 = low burden, 3 = moderate, 5 = very high burden	0.15 (penalty)
Computational burden	(C_i)	Training/tuning/inference complexity	1 = low, 3 = moderate, 5 = very high	0.05 (penalty)

2.7 Methodological limits

The framework has clear limitations. It relies on heterogeneous published evidence rather than a single benchmark dataset, and not all studies report metrics, preprocessing steps, or input structures with the same depth (Abdulai et al., 2023; Gaboitaolelwe et al., 2023). In addition, direct African PV forecasting evidence remains narrower than the broader global literature, which means some model-family interpretation still depends partly on non-African benchmarking studies (Zwane et al., 2022). Thus, results must be interpreted as a structured decision framework based on best available evidence. They should not be taken as definitive proof that one model family is always better in all African settings.

3. COMPARATIVE RESULTS AND ANALYSIS

3.1 Descriptive profile of the screened studies

The retained evidence base shows substantial diversity in forecast horizon, input structure, and modeling strategy, which confirms that direct comparison of raw reported errors across studies is methodologically weak (Gaboitaolelwe et al., 2023; Mellit et al., 2020; Pandžić & Capuder, 2024). The core studies span utility-scale short-term PV forecasting with weather variables, endogenous-only forecasting with historical output data, weather-assisted day-ahead forecasting, and hybrid regional modeling approaches (Abdulai et al., 2023; Hossain & Mahmood, 2020; Konstantinou et al., 2021; J. Li & Liu, 2023; Poti et al., 2023). The screened evidence base retained for comparative analysis is presented in Table 2.

The studies also differ in the type of information available to the model. Some studies use mainly historical PV output. Others combine output with a small number of weather variables. The comparative scoring matrix used to synthesize model-family suitability is presented in Table 3.

3.2 Comparative analysis by model family

Persistence is still important as a benchmark. But in most cases, it is not the best final model. This matches the wider forecasting literature, where persistence is used as a basic reference for short-horizon PV forecasting, even when more advanced models perform better with richer inputs (Abdulai et al., 2023; Barbieri et al., 2017; Gaboitaolelwe et al., 2023). In the Ghana Bui Solar study, persistence was used as the benchmark. However, machine-learning models performed better. This shows that persistence is methodologically useful, but less suitable as a final model in settings with moderate input data (Abdulai et al., 2023).

Classical ensemble machine-learning methods show the best overall mix of performance and practical use in the African studies reviewed. In Ghana, random forest and gradient boosting performed better than persistence, and probabilistic random forest gave the best overall result (Abdulai et al., 2023). Other studies also support the value of tree-based and support-vector methods for short-term solar forecasting when a reasonable number of external variables is available (Mutavhatsindi et al., 2020; Ratshilengo et al., 2021; Salamanis et al., 2020).

Physical and NWP-assisted models are useful when historical plant data are limited but weather forecasts are available. The South African day-ahead study supports this point. It used weather forecasts and plant specifications instead of depending mainly on long output records (Poti et al., 2023). Other studies on physical and physics-informed forecasting also show that these models can work well when interpretability, portability, and weather-based reasoning are important (Mayer & Gróf, 2021; Pombo et al., 2022).

Deep-learning models are promising, but their value depends more on the data conditions. Recurrent neural models can perform well in short-term PV forecasting, especially when temporal patterns are strong (Assaf et al., 2023; Dhaked et al., 2023; Hossain & Mahmood, 2020). However, the endogenous-only study by Konstantinou et al. (2021) shows that their performance depends heavily on preprocessing, treatment of night-time sparsity, and handling of missing data. Hybrid models also often report strong raw performance in the wider literature, but they usually come with more complexity and higher input demands (Li et al., 2020; Li & Liu, 2023; Mellit et al., 2020; Salamanis et al., 2020). Probabilistic variants also need clear attention because they improve practical usefulness under renewable uncertainty, as shown in both the Ghana study and the wider probabilistic forecasting literature (Abdulai et al., 2023; Fernandez-Jimenez et al., 2023). However, this ranking should be interpreted relative to the endogenous-only scenario rather than as a universal preference. Under sparse datasets, deep-learning models still require careful preprocessing, short-sequence design, and protection against overfitting, and may be less robust than ensemble machine-learning models when the historical record is very limited.

3.3 Scenario-based suitability interpretation

The resulting scenario-based ranking of representative forecasting model families is presented in Table 5. To support practical model selection, Table 6 summarizes the preferred and secondary forecasting model families under each sparse-data African deployment scenario.

Table 5. Final scenario-based ranking

Scenario	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
A. Endogenous-only	Deep learning (endogenous temporal)	Persistence/baseline	Classical ML	Hybrid	Probabilistic variants	Physical / NWP-assisted
B. Limited-weather	Classical ensemble ML	Probabilistic ensemble variants	Deep learning	Hybrid	Physical / NWP-assisted	Persistence/baseline
C. NWP-assisted short-history	Physical / NWP-assisted	Classical ML	Probabilistic variants	Hybrid	Deep learning	Persistence/baseline
D. Utility deployment under constraint	Classical ensemble ML	Physical / NWP-assisted	Probabilistic variants	Hybrid	Deep learning	Persistence/baseline

Scenario A (only endogenous operation): The endogenous temporal deep-learning models are the most appropriate. These models utilise historical PV output without the requirement of exogenous weather inputs. The weather-dependent physical models are structurally disadvantaged when meteorological inputs are unavailable (Konstantinou et al., 2021).

Scenario B (limited-weather operation): The most suitable are classical ensemble machine-learning models, particularly probabilistic ensemble variants. This is reflected on the Ghana Bui Solar study where a probabilistic random forest performed best overall under moderate weather-input conditions (Abdulai et al., 2023).

Scenario C (NWP-assisted short-history operation): The physical or physics-informed/NWP-assisted forecasting is the most suitable. For these systems, the weather forecasts and plant specifications can compensate for short historical archives. The South African day-ahead study provides the clearest support for this configuration (Mayer & Gróf, 2021; Poti et al., 2023).

Scenario D (classical ensemble ML): When practical deployment under real operating limits is the focus, classical ensemble ML models and physical/NWP-assisted models offer the best balance of forecast

quality. Probabilistic variants are also important where uncertainty information is needed for decision-making (Abdulai et al., 2023; Poti et al., 2023).

Table 6. Scenario-based framework for selecting suitable short-term PV forecasting model families under data-scarce African grid conditions

Scenario	Available inputs	Main operational condition	Preferred model family	Secondary option	Selection rationale
A. Endogenous-only operation	Historical PV output only	No reliable weather or irradiance inputs are available	Deep learning (endogenous temporal models)	Persistence baseline	Most suitable where only historical output data are available. Temporal models can learn sequential patterns from endogenous signals, although careful treatment of sparsity and missing values is required.
B. Limited-weather operation	PV output with a small set of weather variables	Weather data are available but may be limited, noisy, or incomplete	Classical ensemble machine-learning models	Probabilistic ensemble variants	Suitable where moderate exogenous information is available. Ensemble ML offers a strong balance between predictive performance, flexibility, and practical deployment under constrained data conditions.
C. NWP-assisted short-history operation	Weather forecasts, plant specifications, and short historical PV records	Local historical archives are limited, but weather forecast inputs are accessible	Physical / physics-informed / NWP-assisted models	Classical machine-learning models	Most appropriate when weather forecasts and plant information can compensate for short or incomplete historical PV datasets. These models are also comparatively interpretable for operational use.
D. Utility deployment under operational constraint	Variable operational data environment	Forecasting must support practical use under constraints of robustness, interpretability, and maintainability	Classical ensemble machine-learning models / Physical–NWP-assisted models	Probabilistic variants	Best suited for operational environments where the preferred model must balance forecast quality, feasible data requirements, uncertainty awareness, and real-world deployability.

3.4 Sensitivity and ranking logic

The sensitivity analysis shows ranking is directly dependent on the evaluated conditions, the ranking of model families varies when the evaluation focus changes. The main conclusion, however, stays the same: Hybrid models and advanced deep-learning models perform better when accuracy is given more weight. This is in contrast to classical ML models and physical/NWP-assisted models which perform better when deployment needs are given more weight. This is because they offer a better balance

between forecast quality, practical use, robustness, and interpretability (Gaboitaolelwe et al., 2023; Mayer & Gróf, 2021; Poti et al., 2023). Figure 3 provides a heat-map visualization of the comparative suitability scores assigned to the model families across the six evaluation dimensions.

4. DISCUSSION

4.1 Model selection as a context-dependent decision

One key message of this study is that short-term PV forecasting should be treated as a model-selection problem that depends on context. It should not be treated as a search for one best algorithm for all cases. This view agrees with the wider literature, which shows that forecast performance depends on the forecast horizon, plant size, climate, available inputs, and the metrics used (Gaboitaolelwe et al., 2023; Mellit et al., 2020; Pandžić & Capuder, 2024). Because of this, raw error values from different studies are not enough on their own to guide model choice for African grid settings. The framework in this paper responds to that problem by moving beyond isolated accuracy results. It interprets forecast accuracy together with robustness, transferability, interpretability, data burden, and computational burden. The most suitable model is the one that fits the available data and the deployment goal, not simply the one with the lowest reported error elsewhere.

Transferability is particularly important in Africa because climatic diversity across the continent is substantial. Forecasting conditions differ between humid coastal zones, arid and semi-arid regions, equatorial convective environments, and subtropical southern African climates, so a model family that performs well in one regime may not transfer directly to another without recalibration.

	Evaluation Dimensions					
Model family	Accuracy	Robustness	Transferability	Interpretability	Low data burden	Low computational burden
Baseline / persistence	2	3	5	5	5	5
Physical / NWP-assisted	4	3	4	4	3	4
Classical ensemble ML	4	4	4	3	3	3
Deep learning	4	3	3	2	2	2
Hybrid	5	4	3	2	2	2
Probabilistic variants	4	4	4	3	3	3
			Scale			
	1	2	3	4	5	
	VeryLow				VeryHigh	

Fig 3. Heat map of comparative suitability scores for representative short-term PV forecasting model families across six evaluation dimensions.

4.2 Implications for African forecasting practice

For African utilities, PV operators and system planners, model choice should start with the local data situation. The study finds that classical ensemble machine-learning techniques offer a good compromise between forecast quality and practical application in the case of some weather information. The Ghana Bui Solar study (Abdulai et al., 2023) and other studies showing that tree-based and related ML methods continue to be strong for short-term solar forecasting (Mutavhatsindi et al., 2020; Ratshilengo et al., 2021; Salamanis et al., 2020) support this view. This is particularly relevant for mini-grid and rural

African systems, where intermittent communications, limited on-site technical support, and low-cost monitoring infrastructure can make lightweight and interpretable models more appealing than highly tuned architectures.

When weather forecasts and plant specifications are available, but long historical power records are weak, physical or physics-informed models become more useful. The South African evidence is important here because it shows that day-ahead PV forecasting can still work well without depending mainly on long historical output records (Poti et al., 2023). More broadly, studies on physical and physics-informed models show that they remain useful when interpretability and operational transparency matter alongside forecast performance (Mayer & Gróf, 2021; Pombo et al., 2022).

Another important point is the forecast of probabilities. The literature increasingly shows that point forecasts alone are often insufficient for operational planning under renewable uncertainty, and the present evidence supports that view (Fernandez-Jimenez et al., 2023). In the African studies used here, the best performing case in Ghana was probabilistic, not purely deterministic. This suggests that uncertainty-aware forecasting can be of practical value for reserve planning, scheduling and risk management (Abdulai et al., 2023). Thus, in deployment settings in Africa, uncertainty information should be part of forecasting usefulness as opposed to an optional extra. Uncertainty in African forecasting systems is also amplified by missing weather stations, incomplete sensor coverage, and unreliable telemetry. In practice, these limitations can reduce the value of weather-assisted models, increase imputation requirements, and introduce additional uncertainty into both input streams and validation data.

4.3 Implications for future research

The study also points to some priorities for future research. First, African PV forecasting research would benefit from more comparative studies that consider different operating scenarios, rather than many isolated single-model studies. The current literature from Africa is still limited and uneven, with only a small number of countries producing most of the published work (Zwane et al., 2022). More case studies from other African countries are needed to strengthen a regional conditions-based model family selection. Future studies should prioritize African linking forecasting directly to real operating decisions as already demonstrated in studies by Mbuya et al. (2022) and Poti et al. (2023). These studies connected forecasting to grid electricity, storage coordination, and system management. Expanding in this direction would make forecasting research more useful in practice and would move the field beyond a narrow focus on accuracy alone.

4.4 Limitations

The research proceeded under some important limitations. It is based on published evidence rather than on one shared benchmark dataset, so the rankings come from the framework and were not re-tested under identical conditions (Gaboitaolelwe et al., 2023; Mellit et al., 2020). The number of African-based studies on short-term PV power forecasting is still limited compared to the wider global literature. Therefore, the framework should be read as a guide based on the current evidence and not as proof that one model family is always better.

5. CONCLUSION

This study establishes a comparative mathematical benchmarking framework to assess prominent short-term photovoltaic (PV) power forecasting model families under data-scarce African grid conditions. The analysis shows that no single model family is best in every sparse-data scenario. Instead, the most

suitable choice depends on the available information and the operating need. Temporal models are most useful when only historical output is available. Classical ensemble machine-learning models are most suitable when limited weather information is available. Physical or NWP-assisted models perform best when weather forecasts and plant specifications can make up for short historical records.

The main contribution of the study is a context-based model-selection framework. It evaluates forecasting families not only by predictive accuracy, but also by robustness, transferability, interpretability, data burden, and computational burden. This gives a more practical basis for model selection in African PV systems than raw metric comparison alone

NOMENCLATURE

A_i	Predictive accuracy score	S_i	Suitability score
C_i	Computational burden score	T_i	Transferability score
D_i	Data burden score		
I_i	Interpretability score	w_k	Criterion weight
R_i	Robustness score	Γ_i	Scenario-feasibility gate

REFERENCES

- Abdulai, D., Gyamfi, S., Diawuo, F. A. & Acheampong, P. (2023). Data analytics for prediction of solar PV power generation and system performance: A real case of Bui Solar Generating Station, Ghana. *Scientific African*, 21, e01894. <https://doi.org/10.1016/j.sciaf.2023.e01894>.
- Assaf, A. M., Haron, H., Abdull Hamed, H. N., Ghaleb, F. A., Qasem, S. N. & Albarrak, A. M. (2023). A review on neural network based models for short term solar irradiance forecasting. *Applied Sciences*, 13(14), 8332. <https://doi.org/10.3390/app13148332>.
- Barbieri, F., Rajakaruna, S. & Ghosh, A. (2017). Very short-term photovoltaic power forecasting with cloud modeling: A review. *Renewable and Sustainable Energy Reviews*, 75, 242–263. <https://doi.org/10.1016/j.rser.2016.10.068>.
- Dhaked, D. K., Dadhich, S. & Birla, D. (2023). Power output forecasting of solar photovoltaic plant using LSTM. *Green Energy and Intelligent Transportation*, 2(5), 100113. <https://doi.org/10.1016/j.geits.2023.100113>.
- Fernandez-Jimenez, L. A., Monteiro, C. & Ramirez-Rosado, I. J. (2023). Short-term probabilistic forecasting models using Beta distributions for photovoltaic plants. *Energy Reports*, 9, 495–502. <https://doi.org/10.1016/j.egyr.2023.01.059>.
- Gaboitaolelwe, J., Zungeru, A. M., Yahya, A., Lebekwe, C. K., Vinod, D. N. & Salau, A. O. (2023). Machine learning based solar photovoltaic power forecasting: A review and comparison. *IEEE Access*, 11, 40820–40845. Doi: 10.1109/ACCESS.2023.3270041.
- Hossain, M. S. & Mahmood, H. (2020). Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast. *IEEE Access*, 8, 172524–172533. Doi: 10.1109/ACCESS.2020.3024901.
- Konstantinou, M., Peratikou, S. & Charalambides, A. G. (2021). Solar photovoltaic forecasting of power output using LSTM networks. *Atmosphere*, 12(1), 124. <https://doi.org/10.3390/atmos12010124>.
- Li, G., Xie, S., Wang, B., Xin, J., Li, Y., & Du, S. (2020). Photovoltaic Power Forecasting With a Hybrid Doi: 10.1109/ACCESS.2020.3025860.

- Li, J., & Liu, Q. (2023). Short-term photovoltaic power forecasting using SOM-based regional modelling methods. in *Chinese Journal of Electrical Engineering*, 9(1), 158–176. Doi: 10.23919/CJEE.2023.000004.
- Mayer, M. J., & Gróf, G. (2021). Extensive comparison of physical models for photovoltaic power forecasting. *Applied Energy*, 283, 116239. <https://doi.org/10.1016/j.apenergy.2020.116239>.
- Mbey, C. F., Yem Souhe, F. G., Foba Kakeu, V. J., & Boum, A. T. (2024). A Novel deep learning-based data analysis model for solar photovoltaic power generation and electrical consumption forecasting in the smart power grid. *Applied Computational Intelligence and Soft Computing*, 9257508. <https://doi.org/10.1155/2024/9257508>.
- Mbuya, B. H., Dimovski, A., Merlo, M., & Kivevele, T. (2022). Very short- term blackout prediction for grid- tied PV systems operating in low reliability weak electric grids of developing countries. *Complexity*, 4763820. <https://doi.org/10.1155/2022/4763820>.
- Mellit, A., Massi Pavan, A., Oglari, E., Leva, S., & Lughi, V. (2020). Advanced methods for photovoltaic output power forecasting: A review. *Applied Sciences*, 10(2), 487. <https://doi.org/10.3390/app10020487>.
- Mutavhatsindi, T., Sigauke, C., & Mbuva, R. (2020). Forecasting hourly global horizontal solar irradiance in South Africa using machine learning models. *IEEE Access*, 8, 198872–198885. Doi: 10.1109/ACCESS.2020.3034690.
- Pandžić, F., & Capuder, T. (2024). Advances in short-term solar forecasting: A review and benchmark of machine learning methods and relevant data sources. *Energies*, 17(1), 97. <https://doi.org/10.3390/en17010097>.
- Pombo, D. V., Bacher, P., Ziras, C., Bindner, H. W., Spataru, S. V., & Sørensen, P. E. (2022). Benchmarking physics-informed machine learning-based short term PV-power forecasting tools. *Energy Reports*, 8, 6512–6520. <https://doi.org/10.1016/j.egyr.2022.05.006>.
- Poti, K. D., Naidoo, R. M., Mbungu, N. T., & Bansal, R. C. (2023). Intelligent solar photovoltaic power forecasting. *Energy Reports*, 9, 343–352. <https://doi.org/10.1016/j.egyr.2023.09.004>.
- Ratshilengo, M., Sigauke, C., & Bere, A. (2021). Short-term solar power forecasting using genetic algorithms: An application using South African data. *Applied Sciences*, 11(9), 4214. <https://doi.org/10.3390/app11094214>.
- Salamanis, A. I., Xanthopoulou, G., Bezas, N., Timplalexis, C., Bintoudi, A. D., Zyglakis, L., Tsolakis, A. C., Ioannidis, D., Kehagias, D., & Tzovaras, D. (2020). Benchmark comparison of analytical, data-based and hybrid models for multi-step short-term photovoltaic power generation forecasting. *Energies*, 13(22), 5978. <https://doi.org/10.3390/en13225978>.
- Zwane, N., Tazvinga, H., Botai, C., Murambadoro, M., Botai, J., De Wit, J., Mabasa, B., Daniel, S., & Mabhaudhi, T. (2022). A bibliometric analysis of solar energy forecasting studies in Africa. *Energies*, 15(15), 5520. <https://doi.org/10.3390/en15155520>.