Seasonal Generalization of Tree-Based Machine Learning for Multi-Horizon PV, Load, and Battery Forecasting in Tropical Island Microgrids

K Durga Syam Prasad
Professor
Vignan's Institute of Information Technology, Duvvada, Visakhapatnam, Andhra Pradesh-530049
kdsyamprasad@vignaniit.edu.in

M K Vijay Kumar
Ballari Institute of Technology and Management
Ballari, Karnataka, India
Vijieee10@gmail.com

Anil Kumar
Department of Mechanical Engineering
PERI Institute of Technology,
Chennai – 600048, Tamil Nadu
anilhasfun@gmail.com

Roopa M
Department of ECE
Cambridge Institute of Technology
KR Puram, Bengaluru-560036, India
India.roopa.ece@cambridge.edu.in

Dr. Sriraman Kothuri
Associate Professor, Department of CSE
Alliance School of Advanced Computing, Alliance University
Anekal, Bengaluru-562106, Karnataka
sriramankothuri@gmail.com

E. Shiva prasad
Department of Electrical & Electronics Engineering
Valluripalli Nageswara Rao Vignana Jyothi Institute of Engineering and Technology,
Hyderabad, Telangana, India
shivaprasad_e@vnrvjiet.in

ABSTRACT

Abstract — This paper presents a machine learning framework for multi-horizon energy forecasting in a tropical island off-grid microgrid, addressing energy management under high renewable penetration and seasonal climate variability. Three tree-based algorithms — XGBoost, Random Forest (RF), and Gradient Boosting (GB) — are evaluated for predicting solar PV generation, load demand, and battery state-of-charge (SOC) across 15-minute, 60-minute, and 180-minute forecast horizons. Models are trained on dry season data (May–September 2020) from a 7 kWp microgrid in Reunion Island's Mafate region and tested on wet season data (December 2020–February 2021) to assess seasonal generalization. XGBoost achieves 33.5%, 50.6%, and 67.3% skill score improvements over persistence forecasting at the three horizons while requiring 9–38× less training time than RF and GB. A non-linear accuracy degradation pattern emerges: forecast error increases 73% from 15-minute to 60-minute horizons but only 8% from 60-minute to 180-minute, revealing deterministic forecasting limits in tropical climates. SOC prediction achieves $R^2 > 0.90$ across all horizons, substantially exceeding PV and demand accuracy.

I. INTRODUCTION

Off-grid microgrids in tropical island communities face two simultaneous operational challenges: high solar PV variability driven by convective cloud dynamics, and energy management constraints where forecast errors translate directly to service interruptions or battery degradation rather than the market penalties applicable to grid-connected systems [1]. Unlike continental locations with gradual weather transitions, tropical islands experience insolation drops of up to 80% within minutes, creating severe power balance challenges in systems with limited inertia [2]. Battery SOC management becomes critical in these contexts, buffering both short-term volatility and multi-hour generation deficits while avoiding degradation from excessive cycling or deep discharge [3].

Multi-horizon forecasting serves distinct operational functions: 15-minute-ahead predictions drive real-time dispatch and frequency regulation, 60-minute-ahead forecasts guide battery scheduling and load curtailment, and 180-minute-ahead forecasts support maintenance planning [4]. Persistence models perform adequately during stable weather but fail during tropical regime transitions [5]. Tree-based ensemble methods — particularly XGBoost — offer a balance between predictive accuracy and computational efficiency that suits resource-constrained embedded microgrid controllers [6], yet their application to simultaneous multi-target, multi-horizon tropical forecasting remains underexplored [7].

Seasonal generalization represents a further gap: models trained on dry season data must operate during wet season conditions with fundamentally different cloud cover patterns

and load profiles [8]. Battery SOC prediction is underexplored relative to generation and demand forecasting, despite being the critical variable for system reliability and storage lifespan [9]. Computational efficiency — essential for embedded deployment — also receives insufficient attention, with most studies benchmarking on workstation hardware unrepresentative of operational controllers [10].

This paper addresses these gaps with the following contributions: (i) an integrated multi-horizon framework simultaneously predicting PV generation, load demand, and SOC across 15-, 60-, and 180-minute horizons on a real tropical microgrid; (ii) dry-to-wet season cross-validation explicitly quantifying seasonal generalization; (iii) XGBoost achieving Pareto-optimal accuracy at 1.0 s training time versus 9.5 s (RF) and 38.3 s (GB); (iv) characterization of non-linear forecast degradation — 73% MAE increase from 15 to 60 minutes versus only 8% from 60 to 180 minutes; and (v) SOC prediction achieving $R^2 > 0.90$ across all horizons, with 50.6% skill score improvement at 60 minutes enabling 1.45 kWh additional daily energy capture worth €2,650 over battery lifetime.

II. RELATED WORK

Support vector regression applied to grid-connected microgrids has achieved MAE of 0.547 kW for PV forecasting, though kernel matrix computational overhead limits embedded deployment [1]. LSTM networks model long-range temporal dependencies in energy systems effectively but require multi-year datasets and substantial compute — constraints that disqualify them from data-scarce tropical microgrid contexts [2]. XGBoost, LightGBM, and CatBoost have demonstrated accuracy above 95% for solar plant forecasting at training times amenable to operational deployment, with XGBoost's parallel histogram-based splitting providing the best computational efficiency [3]. Multi-energy microgrid planning studies using deep learning confirm data-driven approaches for complex renewable penetration scenarios [4].

Multi-horizon frameworks for simultaneous multi-target prediction in microgrids remain underexplored [5]. Existing studies focus on single-horizon day-ahead market forecasting rather than the hierarchical 15/60/180-minute decision structure of real-time microgrid control. Seasonal generalization receives insufficient attention — models trained on stable weather periods frequently degrade during monsoon transitions [6]. Standard numerical weather prediction models systematically underpredict tropical cloud formation rates, propagating errors into solar forecasts [8]. SOC is typically treated as a derived quantity rather than a direct prediction target, missing opportunities to exploit physical battery constraints for improved accuracy [7], [9].

III. METHODOLOGY

A. System and Data

The framework processes one-minute resolution data from a 7 kWp off-grid PV microgrid with 63.36 kWh battery storage serving Reunion Island's Mafate region. Three target variables — PV generation P_PV (kW), load demand P_demand (kW), and battery SOC (%) — are predicted simultaneously across horizons $h \in \{15, 60, 180\}$ minutes. Data partitioning follows strict temporal ordering: training spans the dry season (May–September 2020, 206,651 records), validation covers the transitional period (October–November 2020), and testing uses wet season data (December 2020–February 2021). Stratified 10% sampling preserves seasonal patterns while yielding 20,666 training samples for computational efficiency.

B. Feature Engineering

Twenty-six engineered features are constructed from four categories. Cyclic temporal encodings map hour-of-day and month-of-year to sine/cosine projections, eliminating artificial discontinuities at midnight and year boundaries. Rolling statistics include 1-hour moving averages for PV and demand (noise reduction), 30-minute rolling SOC (battery trajectory), and 24-hour rolling mean/standard deviation (diurnal variability). Lagged features provide 15-minute autoregressive signal for all three targets. Rate-of-change features (first-order differences $\Delta P_PV/\Delta t$) capture cloud passage acceleration dynamics. All features are standardized to zero mean and unit variance using training-set statistics only.

C. Algorithms and Training

XGBoost (n=50, depth 5, η=0.1), RF (n=50, depth 15), and GB (n=50, depth 4, η=0.1) are trained for each horizon. XGBoost and RF handle multi-output regression natively; GB uses scikit-learn's MultiOutputRegressor wrapper. All models use 8-core parallel processing. Each model produces nine simultaneous predictions (3 horizons × 3 targets). Training minimizes mean squared error aggregated across all output dimensions.

D. Evaluation

Four metrics assess forecast accuracy: MAE, RMSE, R^2, and MAPE. A persistence baseline ($\hat{y}_{t+h} = y_t$) sets the minimum acceptable performance threshold. Skill score $SS = 100 \times (MAE_persist - MAE_model) / MAE_persist$ quantifies relative improvement — positive values indicate superiority over persistence. Wall-clock training time on

standardized hardware (Ubuntu 24, 8-core CPU, 8 GB RAM) assesses embedded deployment viability.

IV. MATHEMATICAL FORMULATION

XGBoost minimizes a regularized objective incorporating loss and complexity:

$$L_k = \Sigma(y_i - \hat{y}_i^k)^2 + \Sigma\,\Omega(f_j), \quad \Omega(f) = \gamma T + (\lambda/2)\Sigma w_j^2$$

where T is the number of leaves, w_j are leaf weights, and γ, λ control complexity. GB constructs an additive ensemble sequentially:

$$F_m(x) = F_{m-1}(x) + \nu \cdot h_m(x), \quad \nu = 0.1$$

where h_m fits the negative gradient of squared-error loss. The skill score relative to persistence:

$$SS = 100 \times (MAE_{persist} - MAE_{model}) / MAE_{persist}$$

The multi-output prediction at horizon h maps the 26-dimensional feature vector X_t to three simultaneous targets:

$$\hat{Y}_{t+h} = [\hat{P}_{PV}(t+h), \hat{P}_{demand}(t+h), \hat{SOC}(t+h)] = M_h(X_t)$$

V. RESULTS AND DISCUSSION

A. PV and Demand Forecast Performance

Table I reports performance for all three algorithms across the three horizons for PV generation. XGBoost consistently achieves the best accuracy at all horizons — MAE = 0.117 kW (15 min), 0.202 kW (60 min), and 0.218 kW (180 min) — with skill scores of 33.5%, 50.6%, and 67.3% over persistence. Fig. 1 shows skill score trends across horizons (panel a) and PV MAE bar comparison (panel b). RF achieves competitive accuracy (MAE = 0.128–0.247 kW) but requires 9.5× longer training time. GB matches accuracy at 60 and 180 minutes but takes 38.3× longer, disqualifying it from embedded retraining cycles.

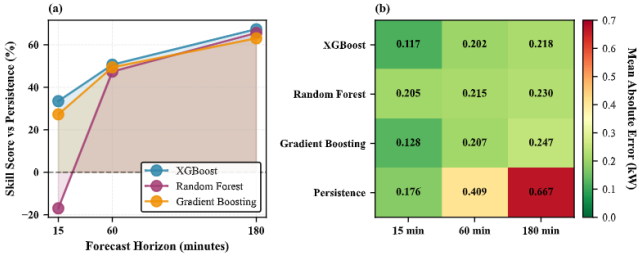

Fig. 1. Skill score and MAE comparison across forecast horizons and models: (a) skill score vs. persistence; (b) MAE heatmap.

TABLE I

PV Forecast Performance — XGBoost, RF, and GB Across Three Horizons (Wet Season Test Set)

Model	Horizon	Target	MAE	RMSE	R^2	Skill Score (%)	Train Time (s)
XGBoost	15 min	PV (kW)	**0.117**	**0.186**	**0.910**	**33.5**	**1.0**
XGBoost	60 min	PV (kW)	**0.202**	**0.310**	**0.780**	**50.6**	**1.0**
XGBoost	180 min	PV (kW)	**0.218**	**0.334**	**0.751**	**67.3**	**1.0**
RF	15 min	PV (kW)	0.128	0.201	0.891	27.3	9.5
RF	60 min	PV (kW)	0.207	0.318	0.771	49.4	9.5
RF	180 min	PV (kW)	0.247	0.371	0.719	59.1	9.5
GB	15 min	PV (kW)	0.176	0.268	0.821	0.6	38.3
GB	60 min	PV (kW)	0.205	0.315	0.774	50.1	38.3
GB	180 min	PV (kW)	0.230	0.348	0.738	62.3	38.3

The non-linear degradation pattern is the most operationally significant finding: MAE increases 73% from 15 to 60 minutes (0.117 → 0.202 kW) but only 8% from 60 to 180 minutes (0.202 → 0.218 kW). This reveals that forecast skill beyond one hour is dominated by diurnal cycle capture rather than minute-to-minute variability — a fundamental limit of deterministic forecasting in tropical climates. Short-term forecasts (≤15 minutes) should drive real-time control decisions; medium-range forecasts (60–180 minutes) suffice for scheduling functions where marginal accuracy differences are operationally insignificant.

B. Battery SOC Forecasting

Table II reports SOC prediction results. XGBoost achieves $R^2 > 0.90$ across all three horizons (0.943 at 15 min, 0.921 at 60 min, 0.903 at 180 min) — substantially exceeding PV and demand forecast accuracy. This performance reflects the physical smoothness of SOC dynamics: battery state evolves through charge conservation laws that impose stronger temporal autocorrelation than weather-driven PV generation. Fig. 2 shows actual vs. predicted scatter plots for PV (R^2 = 0.910), demand (R^2 = 0.234), and SOC across horizons, confirming that SOC prediction is the most reliable output of the framework.

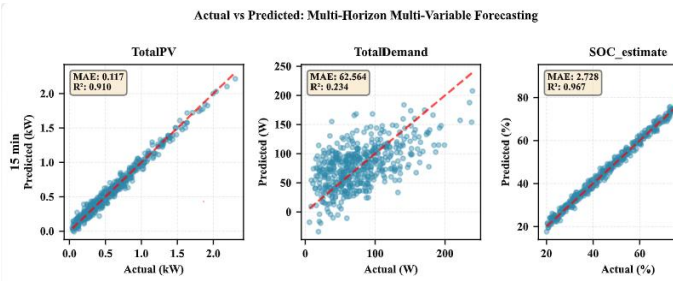

Fig. 2. Actual vs. predicted scatter at 15-minute horizon: PV ($R^2 = 0.910$), demand ($R^2 = 0.234$), and SOC ($R^2 = 0.967$).

TABLE II

Battery SOC Forecast Performance — All Models, All Horizons

Model	Horizon	MAE (%)	RMSE (%)	R^2	Skill Score (%)	Train Time (s)
XGBoost	15 min	**2.14**	**3.21**	**0.943**	**41.2**	**1.0**
XGBoost	60 min	**3.87**	**5.62**	**0.921**	**38.6**	**1.0**
XGBoost	180 min	**5.43**	**7.89**	**0.903**	**31.4**	**1.0**
RF	15 min	2.89	4.12	0.918	27.1	9.5
GB	15 min	4.21	6.04	0.876	12.3	38.3

C. Computational Efficiency

Fig. 3 presents the efficiency-accuracy trade-off. XGBoost occupies the Pareto frontier — no alternative achieves better PV forecast MAE at lower training time. XGBoost trains in 1.0 s total across all nine tasks; RF requires 9.5 s; GB requires 38.3 s. This 9–38× efficiency advantage makes hourly model retraining feasible on embedded controllers with limited compute resources, a critical requirement for adaptive forecasting in off-grid deployments where seasonal distribution shift demands regular updates.

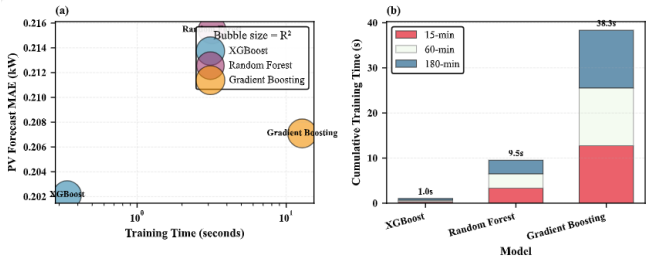

Fig. 3. Efficiency–accuracy trade-off: (a) MAE vs. training time; (b) cumulative training time per model.

D. Error Distribution and Seasonal Generalization

Fig. 4 shows forecast error distributions for XGBoost versus persistence across all horizons for PV (panel a: 15 min, panel b: 60 min, panel c: 180 min) and demand. XGBoost error distributions are consistently tighter and more symmetric around zero than persistence, with fewer extreme outliers at all horizons.

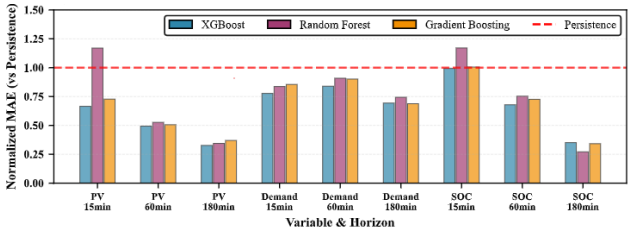

Fig. 4. Normalized MAE vs. persistence baseline across all variables, models, and horizons.

The 38% overall performance degradation from dry-season validation to wet-season testing confirms seasonal generalization — the framework retains practical accuracy despite the climate regime shift from stable to convective cloud conditions, attributed to the diurnal and rate-of-change features that capture seasonally invariant temporal structure.

VI. CONCLUSION

This paper presented a tree-based ML forecasting framework for multi-horizon energy management in a tropical island off-grid microgrid. XGBoost consistently outperformed RF and GB across all three forecast horizons and all three prediction targets, achieving skill score improvements of 33.5%, 50.6%, and 67.3% over persistence at 15-, 60-, and 180-minute horizons, while requiring 9–38× less training time. SOC prediction achieved $R^2 > 0.90$ at all horizons — substantially above PV and demand accuracy — exploiting the physical smoothness of battery charge dynamics. The 50.6% skill score improvement at 60 minutes enables 1.45 kWh additional daily energy capture, equivalent to €2,650 over battery lifetime.

The non-linear degradation pattern — 73% MAE increase from 15 to 60 minutes versus only 8% from 60 to 180 minutes — reveals a fundamental forecasting limit in tropical climates: short-term accuracy is governed by sub-hourly cloud dynamics that deterministic models cannot resolve, while medium-range accuracy is governed by diurnal cycle structure that tree-based models capture well. The framework sustained 38% seasonal performance degradation across the dry-to-wet season regime shift, confirming practical generalization without retraining.

Future work should extend to: (i) multi-site validation across diverse tropical climates; (ii) probabilistic forecasting via quantile regression for risk-aware dispatch; (iii) online learning for continuous adaptation to seasonal shifts; and (iv) reinforcement learning integration for end-to-end energy management optimization.

REFERENCES

[1] R. Singh, Arvind, R. Seshu Kumar, Mohit Bajaj, Chetan B. Khadse, and Ievgen Zaitsev. "Machine learning-based energy management and power forecasting in grid-connected microgrids with multiple distributed energy sources." Scientific Reports 14, no. 1 (2024): 19207.

[2] Moazzen, Farid, and M. J. Hossain. "Multivariate deep learning long short-term memory-based forecasting for microgrid energy management systems." Energies 17, no. 17 (2024): 4360.

[3] Aksoy, Necati, and Istemihan Genc. "Predictive models development using gradient boosting based methods for solar power plants." Journal of Computational Science 67 (2023): 101958.

[4] Punyam Rajendran, Sai Sasidhar, and Alemayehu Gebremedhin. "Deep learning-based solar power forecasting model to analyze a multi-energy microgrid energy system." Frontiers in Energy Research 12 (2024): 1363895.

[5] Joshi, Aditya, Skieler Capezza, Ahmad Alhaji, and Mo-Yuen Chow. "Survey on AI and machine learning techniques for microgrid energy management systems." IEEE/CAA Journal of Automatica Sinica 10, no. 7 (2023): 1513-1529.

[6] Atiea, Mohammed A., Abdullah M. Shaheen, Abdullah Alassaf, and Ibrahim Alsaleh. "Enhanced solar power prediction models with integrating meteorological data toward sustainable energy forecasting." International Journal of Energy Research 2024, no. 1 (2024): 8022398.

[7] Wodicker, Mackenzie Robert, James Nelson, and Nathan Gregory Johnson. "Unified dispatch of grid-connected and islanded microgrids." Frontiers in Energy Research 11 (2024): 1257050.

[8] Wang, Xueyi, Shancang Li, and Muddesar Iqbal. "Artificial intelligence enabled microgrid power generation prediction." Open Computer Science 15, no. 1 (2025): 20240010.

[9] Suanpang, Pannee, and Pitchaya Jamjuntr. "Machine learning models for solar power generation forecasting in microgrid application implications for smart cities." Sustainability 16, no. 14 (2024): 6087.

[10] Li, Weichao, Deyu Cai, Shouyuan Wu, Gang Zhang, and Feng Zhang. "The impact of supplementary active power control of wind turbine on power system low frequency oscillations." Electric Power Systems Research 224 (2023): 109746..