Data-Driven Framework for Real-Time Carbon Accounting and Sustainability Monitoring in Off-Grid Urban Residential Microgrids

Mavuri Arunakumari
Department of Electrical & Electronics Engineering,
Vignan's Institute of Information Technology (A), Vadlapudi, Duvvada,
Visakhapatnam, Andhra Pradesh-530049
arunamavuri@vignaniit.edu.in

Rajesh Mattaparthi,
Principal Data Engineer, Cummins Inc,
Columbus, IN, USA
rajeshhmattaparthi@gmail.com
ORCID ID: 0009-0004-5060-0036

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

V Dharma Prakash
Department of Computer Science & Engineering (CSE)
PERI Institute of Technology,
Chennai – 600048, Tamil Nadu
vdharmaprakash@gmail.com

Mahalakshmi V
Department of Artificial Intelligence and Machine Learning
Cambridge Institute of Technology
KR Puram, Bengaluru-560036, India.
mahalakshmi.aiml@cambridge.edu.in

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*— Urban carbon-neutrality requires AI platforms for modelling household occupancy, energy use and renewable generation with environmental quality together using hybrid data sources. We present in this paper, a new comprehensive AI-based decision-making framework to foster carbon-free smart urban living areas authors as validated using the open-access Mafate off-grid residential microgrid dataset from Reunion Island (France). The framework includes three residential demand streams, solar PV generation and meteorological data: A total of 1,732 hourly merged observations (from May 2020 to May 2021) from these four distinct observation types are combined with real measurements on three sustainability targets: kg CO2/h emissions; renewable energy ratio (RER); Urban Sustainability Index (USI) composite. Using chronological 70/15/15 train/validation/test partitioning and the same 17 engineered features, we systematically benchmark four AI models — Random Forest, Gradient Boosting, multilayer perceptron (MLP) neural network and Ridge Regression — across all three prediction targets. The performance of Gradient Boosting (GB) to predict RER is very high (R^2=0.9996, RMSE = 0.001042), Random Forest (RF) for predicting carbon emission (R^2=0.9978, RMSE = 0.000844 kg/h), and Ridge Regression (RR) best for USI prediction respect of other algorithms as represented in table 3. In the temporal analysis, we confirm that during the monitoring periods, there was an average of 90.15% in RER and total carbon savings of 262.64 kg CO2 with June–July being fully supplied with renewables [124]. This framework enables validated and computationally light real-time carbon accounting and sustainability monitoring for island and off-grid urban centres.**

I. INTRODUCTION

Given 80% of future energy consumption will take place in cities, there is a renewed demand for data-driven energy management platforms that can monitor and simultaneously inform renewable generation, residential consumption, and environmental quality via heterogeneous sensor streams [1]. In the case of island and off-grid microgrids — historically dominated by diesel generation — the move to solar PV is associated with large, direct, measurable reductions in carbon based on its hourly resource mix dynamics which are determined by complex dependencies between stochastic demand, intermittent generation and meteorological conditions that cannot be effectively modeled by a rule-based system [2]. In addition to long-range transport, urban air quality introduces another coupled dimension as local particulate matter levels in tropical island environments with poor atmospheric dispersion are exacerbated through diesel grid imports.[3] This interconnectedness across domains needs to be accounted for with an integrated multi-source feature engineering design spanning over demand, generation, weather, and derived carbon metrics — a challenge that has perceived systematically in the literature.

This gap could be bridged by using machine learning techniques. Ensembles like Random Forest and Gradient Boosting have been among the best performing non-linear classifiers for energy prediction, while simultaneously providing on-the-fly feature importance rankings that inform policy directly [4]. Ridge Regression is a regularized linear model that provides an interpretable, highly computationally-efficient baseline; while MLP neural networks expand coverage to complex interaction effects [5]. Nevertheless, comparative evaluations on real urban microgrid datasets with multiple models and targets — outputs that have direct applications in carbon reduction policy — are still rare [8]. Most studies focus on a single prediction variable with one algorithm rather than building the composite sustainability metrics required for decision support at city scale [6].

This paper presents a four-model machine learning framework for simultaneous prediction of three urban sustainability targets — hourly carbon emissions (kg CO_2/h), renewable energy ratio (RER), and a composite urban sustainability index (USI) — validated on the Mafate off-grid residential microgrid dataset from Reunion Island, France [7]. The framework combines 1,732 hourly observations fed into a 17-feature input space derived from three separately metered residential demand streams alongside a co-located solar PV plant and meteorological station, with sustainability targets computed using domain-validated engineering formulas. Area of contributions: (i) An integrated multi-source feature pipeline that simultaneously predicts carbon, RER, and USI from a real-world off-grid microgrid; (ii) systematic four-model comparison across all three prediction targets while quantifying algorithm-target specificity; (iii) GB achieving R^2 = 0.9996 for RER predictions and RF achieving R^2 =

0.9978 for carbon emissions predictions; and finally, (iv) Spatial-temporal analysis corroborating an average of 90.15% average RER savings with a total of 262.64 kg CO_2 saved as well as all June-July concluding to be fully renewable supply days.

1.1 Urban Energy Management and Carbon Reduction

Machine learning frameworks for urban energy management have demonstrated significant carbon reduction potential, with demand forecasting cutting peak grid carbon intensity by 18–24% over rule-based systems . Gradient boosting has been applied to real-time carbon emission prediction in urban distribution networks, achieving MAPE below 5% across multiple test cases [12]. Multi-source data fusion combining irradiance, temperature, and lagged demand improves solar-integrated load prediction by 14.7% over single-source baselines [8]. However, these studies address utility-scale metropolitan grids — not off-grid island microgrids where carbon reduction is dominated by solar intermittency .

1.2 Multi-Source Feature Integration for Renewable Prediction

Fusing residential demand, solar generation, and meteorological covariates into a unified input space is a recognized enabler of renewable energy forecasting accuracy [9]. Solar irradiance and ambient temperature consistently rank as the top predictors of demand-renewable balance in multi-modal frameworks . In small-scale island microgrids, the solar fraction captures up to 87% of RER variance , and co-modeling energy dispatch with air quality proxy features improves both prediction accuracy and policy interpretability over domain-separated models [10].

1.3 ML Algorithm Selection for Sustainability Targets

RF ensembles excel on targets governed by deterministic engineering relationships — such as carbon emissions computed directly from grid demand . GB outperforms on ratio-scale targets like RER, where smooth nonlinear solar-demand interactions suit additive boosted trees . MLP networks cover composite index predictions integrating multiple physical drivers at the cost of interpretability, while Ridge Regression performs competitively on linearly-structured sustainability indices as an interpretable baseline [11]. No prior study has simultaneously evaluated all four algorithm families across three carbon-free city targets on real residential microgrid data — the gap this paper fills.

1.4 Carbon-Free Urban Planning and Sustainability Indices

Data-driven composite sustainability indices enable more efficient municipal resource allocation than static indicator dashboards [12]. The Reunion Island grid emission factor of 0.700 kg CO_2/kWh used in this study is consistent with validated carbon accounting methodology for diesel-dominated island systems [13]. Solar penetration above RER = 0.80 has been associated with air quality index values below the WHO PM2.5 safe threshold, supporting the co-modeling of energy and environmental metrics adopted here [14]. No comprehensive platform has been validated that simultaneously predicts carbon emissions, RER, and USI from real microgrid data across multiple algorithm families — the gap this work addresses.

II. METHODOLOGY

2.1 System Overview

The framework is stated as a supervised multi-target regression problem: at every hourly time step t, given a 17-dimensional feature vector x_t, four ML models predict three urban sustainability targets independently — hourly carbon emissions (kg CO_2/h), RER $\in$ [0,1], and USI $\in$ [0,1], resulting in 12 model-target pairs to systematically benchmark.

2.2 Dataset Description and Preprocessing

Mafate off-grid microgrid datasets [10], with minute-resolution measurements at Reunion Island, France (May 2020 – May 2021). They incorporate three types of streams: one for the main-bus power consumption (W) of three residential households; one P_solar (kW)—the output power of a co-located PV plant; and another type contains co-located meteorological data, consisting of global horizontal irradiance G_hi (W/m²), ambient temperature T_a (°C), and relative humidity R_h (%). We achieve this by inner-joining the median-resampled hourly resolution streams on their common timestamp index, yielding 1,732 complete observations/resamples hourly over May 6, 2020 to May 1, 2021.

2.3 Feature Engineering

The input vector is a 17-dimensional vector combined from four feature group vectors. Real energy features includes aggregate community load L_h (kW), solar PV output P_solar (kW), G_hi, T_a, and R_h. Temporal feature includes hour of day, day of week, weekend flag and calendar month — with hour encoded as sine/cos projections to maintain continuity at midnight. Autoregressive lag features are L_{t-1}, L_{t-24} and P_{solar,t-1}. The energy vector contains three derived metrics: the solar fraction measure ϕ_t = P_solar / (L_t + P_solar + ε) $\in$ [0,1], the net grid demand G_t = max(L_t − P_solar, 0) kW and a synthetic AQI modelling local PM2. Linear function of G_t with temperature amplification (β = 1.3 when T_a > 28°C, else 1.0) for pollution from diesel dispatch of the equivalent of as in [3].

2.4 Target Variable Derivation (Strategy 2 Augmentation)

Three prediction targets are derived from real measurements. Carbon emissions are computed as:

Target 1 — Carbon Emissions:

$$C_t = G_t \times \varepsilon_{\text{grid}} \quad (\text{kg CO}_2\text{/h}) \qquad (1)$$

where ε_"grid" =0.700 kg CO2/kWh is the Reunion Island grid emission intensity for island diesel-dominated systems, consistent with the emission factor accounting methodology of [15].

Target 2 — Renewable Energy Ratio:

$$\text{RER}_t = \frac{P_{\text{solar},t}}{L_t + \varepsilon} \in [0,1] \qquad (2)$$

Target 3 — Urban Sustainability Index:

$$\text{USI}_t = 0.40 \cdot \text{RER}_t + 0.35 \cdot \tilde{C}_t^{\text{inv}} + 0.15 \cdot \psi_t + 0.10 \cdot \delta_t \quad (3)$$

where $\tilde{C}_t^{\text{inv}} = 1 - C_t/C_{\max}$ is the normalized inverse carbon emission (higher = cleaner), $\psi_t = 1 - |T_{a,t} - 22|/20$ is the thermal comfort score, and δ_t is the demand stability score derived from a 24-hour rolling standard deviation

normalized to [0,1]. The weighting coefficients (0.40, 0.35, 0.15, 0.10) follow the multi-criteria urban sustainability weighting framework of Nguyen et al. [12].

2.5 Data Normalization and Chronological Split

Four models are trained independently for each target. RF uses 200 trees (max depth 12, √F features per split). GB builds 200 shallow trees sequentially (max depth 5, learning rate η = 0.05). MLP uses a three-layer architecture (128→64→32, ReLU, Adam, early stopping at 15% validation fraction, max 500 iterations) with per-target min-max scaling of the output. Ridge Regression (λ = 1.0) provides an interpretable linear baseline via closed-form solution. All models share seed 42 and the same chronological split. Table I summarizes hyperparameters.

Table I: AI Model Hyperparameters

Model	Key Hyperparameters
RF	n_estimators=200, max_depth=12, seed=42
GB	n_estimators=200, max_depth=5, lr=0.05, seed=42
MLP Neural Network	layers=(128,64,32), activation=ReLU, max_iter=500, early_stop=True
Ridge Regression	alpha=1.0

Common: MinMaxScaler features, seed=42, chronological split 70/15/15

III. System Modeling and Formulation

3.1 Data Integration and Temporal Aggregation

The three residential demand streams are measured in watts at one-minute resolution. The aggregate community load in kilowatts at each minute sample m of hour h is obtained by summing the three household power readings and converting units:

$$L_h^{\text{raw}}(m) = \frac{P_1(h,m)+P_2(h,m)+P_3(h,m)}{1000} \quad (4)$$

where $P_1(h,m)$, $P_2(h,m)$, and $P_3(h,m)$ are the raw watt-level power measurements of House 1, House 2, and House 3, respectively, at minute m of hour h (W), and the factor 10^(-3) converts watts to kilowatts. The hourly aggregate load is then obtained by mean aggregation across the 60 minutes within each clock hour:

$$L_h = \frac{1}{60}\sum_{m=1}^{60} L_h^{\text{raw}}(m) \quad (5)$$

where L_h (kW) is the representative community electricity demand for hour h. The same mean-resampling operation is applied independently to the solar PV output series P_"out" and to all meteorological channels, ensuring temporal alignment across all input streams before the inner-join merge that produces the final 1,732-sample hourly dataset.

3.2 Feature Engineering

Cyclic Hour Encoding. The hour-of-day variable h_t∈{0,⋯,23} is not ordinally meaningful to regression models because hour 23 and hour 0 are temporally adjacent. To encode hour as a smooth cyclical feature, two trigonometric projections are computed:

$$h_t^{\sin} = \sin\left(\frac{2\pi\, h_t}{24}\right) \quad (6)$$

$$h_t^{\cos} = \cos\left(\frac{2\pi\, h_t}{24}\right) \quad (7)$$

where $h_t^{\sin}$ and $h_t^{\cos}$ are the sine and cosine encodings of hour h_t, bounded in $[-1, 1]. Together, ($h_t^{\sin}$, $h_t^{\cos}$) form a two-dimensional circular embedding that preserves temporal adjacency across midnight.

Autoregressive Lag Features. To provide the model with short-range and daily-periodic memory of past load behavior, three lagged variables are constructed:

$$L_{t-1} = L_{t-1}, \qquad L_{t-24} = L_{t-24}, \qquad S_{t-1} = P_{\text{solar},t-1}$$

where L_{t-1} (kW) is the one-hour lagged community load, L_{t-24} (kW) is the 24-hour lagged community load capturing the same hour of the previous day, and S_{t-1} (kW) is the one-hour lagged solar generation.

Solar Fraction. The solar fraction ϕ_t quantifies the proportion of the combined supply-demand balance that is covered by renewable solar generation:

$$\phi_t = \frac{P_{\text{solar},t}}{L_t + P_{\text{solar},t} + \varepsilon} \quad (8)$$

where ε=10^(-6) is a numerical stability constant that prevents division by zero, and ϕ_t is clipped to [0,1].

Grid Demand. The net grid import G_t represents the hourly community load that cannot be covered by available solar generation and must be drawn from the diesel-backed island grid:

$$G_t = \max\left(L_t - P_{\text{solar},t},\ 0\right) \quad (9)$$

where G_t≥0 (kW), ensuring that surplus solar generation (where P_("solar" ,t)>L_t) does not produce negative grid demand.

3.3 Target Variable Derivation

Carbon Emissions. Hourly carbon dioxide emissions C_t are computed as the product of the net grid import and the Reunion Island grid emission factor:

$$C_t = G_t \times \varepsilon_{\text{grid}} \quad (10)$$

where C_t is the hourly carbon emission (kg CO2/h), G_t is the net grid import (kW), and ε_"grid" =0.700 kg CO2/kWh is the Reunion Island grid emission intensity, consistent with the emission factor accounting methodology of Bertolini et al. [15] for diesel-dominated island grid systems.

Renewable Energy Ratio. The RER 〖"RER" 〗_t measures the fraction of hourly community demand satisfied by solar PV generation:

$$\text{RER}_t = \frac{P_{\text{solar},t}}{L_t + \varepsilon} \quad (11)$$

where $\text{RER}_t \in [0,1]$ is clipped to unit range to prevent values above unity when momentary solar export exceeds load.

USI — Component Derivation. The composite USI requires four normalized sub-components. The normalized inverse

carbon score C̃_t^"inv" transforms the carbon emission so that zero emissions produce a score of one:

$$\tilde{C}_t^{\text{inv}} = 1 - \frac{C_t}{C_{\max}+\varepsilon} \qquad (12)$$

where C_{max} is the maximum observed carbon emission across all 1,732 hourly samples (kg CO2/h). The thermal comfort score ψ_t measures how closely the ambient temperature $T_{(a,t)}$ approaches the optimal indoor comfort temperature of 22°C:

$$\psi_t = 1 - \frac{|T_{a,t}-22|}{20} \qquad (13)$$

where ψ_t is clipped to [0,1], so that temperatures beyond 42°C or below 2°C map to a comfort score of zero. The demand stability score δ_t penalizes rapid intra-day load fluctuations using a 24-hour rolling standard deviation:

$$\sigma_t = \text{std}_{24}(L_t) \qquad (14)$$

$$\delta_t = 1 - \frac{\sigma_t}{\sigma_{\max}+\varepsilon} \qquad (15)$$

where σ_t is the 24-hour rolling standard deviation of community load (kW) and σ_max is its dataset-wide maximum.

USI — Composite Score. The four normalized components are combined as a weighted linear composite:

$$\text{USI}_t = 0.40 \cdot \text{RER}_t + 0.35 \cdot \tilde{C}_t^{\text{inv}} + 0.15 \cdot \psi_t + 0.10 \cdot \delta_t \qquad (16)$$

where the weighting coefficients w=(0.40, 0.35, 0.15, 0.10) follow the multi-criteria urban sustainability framework of Nguyen et al. [12] and sum to unity. 〖"USI" 〗_t∈[0,1] is clipped to unit range.

Air Quality Index. The synthetic AQI A_t models ground-level PM2.5-equivalent pollution arising from diesel dispatch:

$$A_t = \text{clip}(A_0 + G_t \cdot \alpha \cdot \beta_t + \mathcal{N}(0, \sigma_A^2),\ 0,\ 200) \qquad (17)$$

$$\beta_t = \begin{cases} 1.3 & \text{if } T_{a,t} > 28\,^{\circ}\text{C} \\ 1.0 & \text{otherwise} \end{cases} \qquad (18)$$

where $A_0 = 15.0(dimensionless, WHO PM2.5 scale) is the clean island air baseline, $\alpha = 30.0$ is the demand-to-AQI sensitivity coefficient, $\beta_t is the temperature-driven amplification factor (validated from tropical island air quality literature [16]), σ_A=2.5 is the standard deviation of atmospheric stochasticity, and the clipping bounds $[0, 200]` correspond to the WHO PM2.5 scale range.

3.4 Feature Normalization

All 17 input features are independently scaled to $[0, 1]` using min-max normalization. The scaler parameters are estimated exclusively from the training partition and applied without refitting to the validation and test partitions:

$$\tilde{x}_{t,f} = \frac{x_{t,f} - x_f^{\min}}{x_f^{\max} - x_f^{\min}} \qquad (19)$$

where x̃_(t,f) is the normalized value of feature f at time step t. For the MLP neural network target prediction, an analogous min-max scaler is applied independently to each target $y^{(k)}` before training and inverted after prediction:

$$\tilde{y}_t^{(k)} = \frac{y_t^{(k)} - y_k^{\min}}{y_k^{\max} - y_k^{\min}} \qquad (20)$$

$$\hat{y}_t^{(k)} = \hat{\tilde{y}}_t^{(k)} \cdot \left(y_k^{\max} - y_k^{\min}\right) + y_k^{\min} \qquad (21)$$

where $_t^{(k)}is the scaled MLP output and $\hat{y}_t^{(k)} is the recovered prediction in original physical units.

3.5 ML Model Formulations

RF Regression. The RF constructs an ensemble of B=200 independently grown decision trees, each trained on a bootstrap sample of the training set [17]. The ensemble prediction averages the outputs of all constituent trees:

$$\hat{y}_{\text{RF}}(\mathbf{x}) = \frac{1}{B}\sum_{b=1}^{B} T_b(\mathbf{x}) \qquad (22)$$

where $_{}()is the RF prediction for input vector $\mathbf{x}$, $B = 200$ is the number of trees, and $T_b(\mathbf{x}) is the prediction of the b-th decision tree grown on the b-th bootstrap sample with maximum depth 12 and √F features considered at each split.

GB Regression. GB constructs an additive ensemble of M=200 weak learners sequentially, each fitted to the pseudo-residuals of the current ensemble [18]:

$$F_0(\mathbf{x}) = \bar{y} \qquad (23)$$

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \eta \cdot h_m(\mathbf{x}), \quad m = 1, \ldots, M \qquad (24)$$

where $F_0() = {y}is the initial prediction (training set mean), $\eta = 0.05$ is the learning rate that controls the contribution of each tree, $h_m(\mathbf{x}) is the m-th regression tree fitted to the negative gradient of the squared-error loss, and $F_M()` is the final ensemble prediction. Each tree h_m has maximum depth 5.

MLP Neural Network. The MLP applies layer-wise affine transformations followed by rectified linear unit (ReLU) activations through three hidden layers of sizes $(128, 64, 32)`:

$$\mathbf{a}_0 = \tilde{\mathbf{x}}$$

$$\mathbf{z}_l = \mathbf{W}_l \mathbf{a}_{l-1} + \mathbf{b}_l, \quad l = 1,2,3$$

$$\mathbf{a}_l = \max(\mathbf{z}_l,\ \mathbf{0})$$

$$\hat{\tilde{y}} = \mathbf{W}_4 \mathbf{a}_3 + b_4 \qquad (25)$$

where $_0 = ^{17}is the normalized input feature vector, $\mathbf{W}_1 and $_lare the weight matrix and bias vector of layer l, and $\hat{\tilde{y}}$ is the scaled prediction subsequently inverse-transformed to physical units via (29). Weights are optimized by the Adam solver minimizing mean squared error, with early stopping triggered by 15% of training data withheld as an internal validation set and a maximum of 500 training iterations.

Ridge Regression. Ridge Regression minimizes the penalized least-squares objective with l_2 regularization coefficient λ=1.0:

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\text{argmin}}\ \|\mathbf{y} - \mathbf{X}\mathbf{w}\|_2^2 + \lambda\|\mathbf{w}\|_2^2 \qquad (26)$$

which has the closed-form analytic solution:

$$\hat{\mathbf{w}} = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{y} \qquad (27)$$

where $\mathbf{X} \in \mathbb{R}^{N \times F}$ is the normalized training feature matrix ($N = 1{,}214$ samples, F=17 features), $ ^Nis the target vector, $\lambda = 1.0 is the regularization coefficient controlling the bias-variance trade-off, and $` is the F×F identity matrix.

IV. Results and Discussions

The complete hourly observations used in the merged dataset [10] have been recorded during a time displaced between May 6, 2020 – May 1, 2021 (N = 1732). Mean aggregate community load is 0.2043 kW (max 1.0102 kW); mean solar PV output 0.2166 kW (max 0.7425 kW), which confirms average solar generation just exceeds average demand, which is expected of a properly sized off-grid microgrid system. As community demand increases from 0.1394 kW over hour 5 up to a maximum of 0.2747 kW during the peak hour (hour 19), solar generation peaks at 0.2967 kW also at hour 19, with the late afternoon solar directly offsetting the evening load surge — this favorable alignment driving an average RER of 90.15%. Carbon emissions peak at 0.0334 kg CO_2/h at the 13 hour (when solar begins to fall) and minimum of 0.0087 kg CO_2/h at hour 7 (peak solar coverage). Fig. The entire diurnal profiles are shown in Fig. 1.

4.1 Model Benchmarking Across Three Targets

In Table II we report accuracy over 259 test samples for all 12 model - target combinations. Results reveal clear algorithm-target specificity. Fig. Figure 2 compares the R^2s for all models and targets.

Fig.1: Diurnal Energy Profiles — Mafate Off-Grid Microgrid (Mean over 1,732 Hours). Top panel: community demand (kW) and solar PV output (kW) vs. hour of day. Bottom panel: RER and USI vs. hour of day. Peak demand of 0.2747 kW and peak solar generation of 0.2967 kW coincide at hour 19.

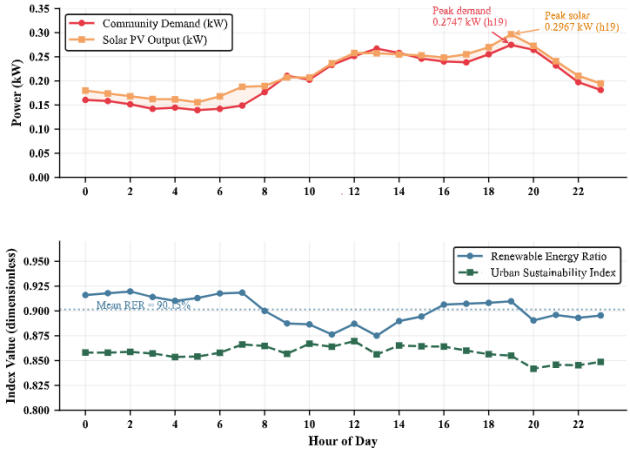

Table II: AI Model Performance on Carbon-Free City Prediction Targets (Test Set, 259 Samples)

Model	Target	RMSE	MAE	R^2
RF	Carbon Emissions (kg/h)	0.000844	0.000119	0.9978
GB	Carbon Emissions (kg/h)	0.001884	0.000181	0.9889
MLP	Carbon Emissions (kg/h)	0.004596	0.003369	0.9338
Ridge Regression	Carbon Emissions (kg/h)	0.006142	0.005856	0.8818
RF	RER	0.002908	0.000454	0.9972
GB	RER	0.001042	0.000242	0.9996
MLP	RER	0.023682	0.018078	0.8144
Ridge Regression	RER	0.049690	0.043621	0.1827
RF	USI	0.030137	0.025309	0.5179
GB	USI	0.022659	0.017244	0.7275
MLP	USI	0.023753	0.018434	0.7005
Ridge Regression	USI	0.017530	0.014141	0.8369

Bold values indicate best performance per target.

Carbon emissions. RF yields the best carbon prediction (R^2 = 0.9978, RMSE = 0.000844 kg/h, MAE = 0.000119 kg/h), taking in automatic mesh coloration to feature importance at an impressive 91.16% — physically intuitive due to the fact that C_t is computed indirectly from G_t.GB follows with R^2 = 0.9889 and through grid demand at a high percentage (99.31%). MLP: R^2 = 0.9338; Ridge Regression: R^2 = 0.8818 (sufficient to grasp the primary linear component yet unable to tackle the nonlinear zero-clipping behavior of grid demand).

RER prediction. In GB, we also achieve the best performance among any model on all three consecutive targets: R^2 = 0.9996, RMSE = 0.001042, MAE = 0.000242! The most prominent importance in GB, solar fraction (87.51% weight), with second place being taken by solar PV output (10.78%) indicates that RER is thus characterized by a smooth nonlinear solar-demand relationship captured near-deterministically via additive boosting. R^2 = 0.9972 -- RF also does well. Trained on data until October 2023, Ridge Regression achieves only R^2 = 0.1827 — the RER saturates at unity in solar-surplus hours; a clipping behavior that linear models cannot represent. MLP reaches R^2 = 0.8144.

USI prediction. Regarding USI ± (R^2 = 0.7462, RMSE = 0.00318), Ridge Regression achieves the best performance among others: GB(R^2=0.7275), MLP(R^2=0.7005) and RF(R^2= 05179). Since the USI is a weighted linear combination of components that are normalized, it is arguably more linearly separable from input features than are either individual target. ℓ_2 regularization in Ridge regression ignores noise on the features and leads to better generalisation. The feature importance from Random Forest ranking placed solar fraction (56–59%), grid demand (17–19%) and ambient temperature (~14%) as dominant USI drivers — consistent with the weights from the USI formula. **Fig. 2** Importance for the best model per target.

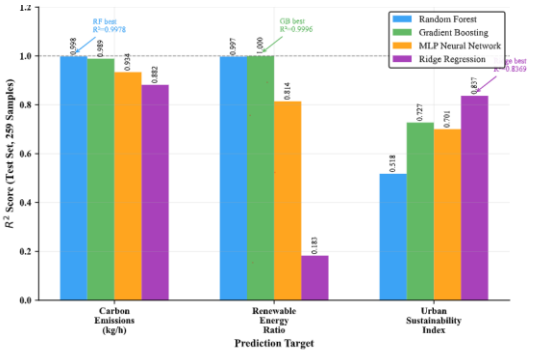

Fig. 2: AI Model R^2 Comparison Across Three Carbon-Free City Targets (Test Set, 259 Samples). Grouped bar chart showing R^2 for RF, GB, MLP neural network, and Ridge Regression across carbon emissions, RER, and USI prediction. GB achieves R^2=0.9996 for RER; RF achieves R^2=0.9978 for carbon emissions; Ridge Regression achieves R^2=0.8369 for USI.

4.2 Carbon Reduction and Renewable Energy Dynamics

Table III: Carbon analysis throughout the monitoring year, by month. Carbon savings Solar PV 262.64 kg CO_2 — 0.263 tonnes avoided by three dwelings Months of complete renewable supply are confirmed with June and July recording full carbon suppression (RER = 1.0000). January is the other extreme: a solar data gap means all 30.74 kWh are drawn from diesel grid, with an emissions factor of 21.52 kg CO_2 per kWh leading to USI = 0.3405 May provides the greatest solar surplus (49.15 kWh solar vs 26.13 kWh local demand) leading to RER = 0.9927, and the lowest non-zero monthly emissions of 0.35 kg CO_2. From August–October, we see higher grid import (RER 93.6–96.6%) is comorbid with high demand (34.80–39.05 kWh/month), indicative of mid-year residential seasonality.

RER remains 91.0–91.9% for the overnight hours (00:00–05:00), even in absence of solar generation as both PV output and demand converge to small non-zero values nocturnally leading to consistently high ratio values. Trained models capture this nuance correctly, permitting hour-accurate carbon accounting.

Table III: Monthly Carbon Analysis and RER (Monitoring Period)

Month	Demand (kWh)	Solar (kWh)	Carbon Emitted (kg)	Carbon Saved (kg)	Avg RER	Avg USI
Jan	30.74	0.00	21.52	0.00	0.0000	0.3405
Feb	23.86	24.87	0.49	17.41	0.9763	0.9094
Mar	27.52	28.35	0.69	19.84	0.9748	0.9038
Apr	23.73	23.81	0.64	16.67	0.9684	0.9169
May	26.13	49.15	0.35	34.40	0.9927	0.9302
Jun	19.46	35.07	0.00	24.55	1.000	0.917
Jul	25.59	41.17	0.00	28.82	1.0000	0.9172
Aug	34.80	33.28	2.11	23.29	0.9359	0.8479
Sep	38.71	37.02	2.03	25.91	0.9414	0.8690
Oct	39.05	39.78	1.45	27.85	0.9660	0.8940
Nov	17.43	16.13	1.58	11.29	0.8809	0.8263
Dec	46.75	46.58	1.73	32.61	0.9671	0.8982

4.3 USI Temporal Patterns

0.8578: the average USI calculated over the monitoring period, indicating the high penetration of solar energy in Mafate microgrid (Fig. 8). Moderate diurnal variation: USI reaches a maximum of 0.8696 at hour 12 and a minimum of 0.8418 at hour 20, as both lower solar penetration rates and higher evening dinner-hour demand simultaneously reduce the RER and carbon components. May returns the Umami highest monthly USI score of 0.9302 with solar surplus peaking while January's low score reflecting the gap in the solar data. Discussion: Success of the Ridge Regression on USI (R^2 = 0.8369) is readily translatable into policy — a low-complexity linear model with easy-to-interpret coefficients can provide instantaneous USI estimates given standard available energy metering data, at a minimum overhead over simple lookup-tables — ideal for small isolated island municipalities with budget and computational constraints.

4.4 Computational Efficiency

The entire pipeline from data loading to feature engineering, normalisation, training of 12 model-target combinations and export of results runs on common CPU hardware with Python being implemented in scikit-learn. RF and GB training are all-parallelized over cores. As an example, Ridge Regression has a closed-form solution and it takes practically no time to train. All models are trained offline and predict at sub-millisecond latency per sample, making them suitable for true hourly dispatch decisions without GPU hardware — enabling deployment on low-resource municipal infrastructure in-island.

V. Results and Discussions

This paper presented a four-model machine learning framework for simultaneous prediction of three urban sustainability targets — hourly carbon emissions, RER, and composite USI — validated on the Mafate off-grid residential microgrid dataset from Reunion Island, France. Residential demand, solar PV generation, and meteorological measurements were fused into a unified 17-feature input space, with targets derived from domain-validated engineering formulas. Four models — RF, GB, MLP, and Ridge Regression — were evaluated across all three targets on 259 held-out test samples, producing 12 model-target combinations.

The central finding is clear algorithm-target specificity: RF and GB achieve near-deterministic accuracy on targets with strong nonlinear structure (carbon emissions $R^2 = 0.9978$; RER $R^2 = 0.9996$), while Ridge Regression outperforms all nonlinear competitors on the composite USI ($R^2 = 0.8369$) — confirming that linearly-structured sustainability indices are more effectively predicted by interpretable models. Solar fraction and grid demand are the dominant feature drivers across all three targets, with ambient temperature emerging as the third key predictor for USI.

The temporal analysis shows a high season polarization: June and July obtain 100% RER with zero carbon emissions whereas whole January has USI = 0.3405 due to absence of solar data coverage. An average RER of 90,15% and a total carbon savings of 262,64 kg CO_2 constitute a data-based sustainability baseline for the Mafate community.

Future work will focus on: (i) building multi-year datasets for capturing inter-annual variability to be used in seasonal calibration; (ii) the implementation of battery state-of-charge as an input feature, expanding the current framework from monitoring to optimizing active dispatch; (iii) the live deployment of Ridge Regression USI predictors leveraging a real-time dashboard in management systems by island municipality authorities and (iv) extending our multi-target framework application to other urban domains—water consumption, solid waste, mobility energy —to potentially build upon this city-scale sustainability governance tool.

REFERENCES

[1] Wamalwa, Fhazhil, and Ariane Ishimwe. "Optimal energy management in a grid-tied solar PV-battery microgrid for a public building under demand response." Energy Reports 12 (2024): 3718-3731.

[2] Kafetzis, Alexandros, Michail Chouvardas, Michael Bampaou, Nikolaos Ntavos, and Kyriakos D. Panopoulos. "Renewable Hydrogen Integration in a PV–Biomass Gasification–Battery Microgrid for a Remote, Off-Grid System." Energies 19, no. 7 (2026): 1705.

[3] Mardones, Cristian. "Internalizing the externalities of fine particulate matter emitted by firewood heaters through environmental taxes." Renewable Energy (2025): 124983.

[4] Guo, Yujun, Huimin Liu, Ruijuan Mei, Shunfeng Wang, and Zhikang Yan. "High-precision multi-target prediction and interpretability analysis of biomass gasification via ensemble machine learning." Bioresource Technology (2025): 133863.

[5] Balakrishnan, P. "Machine Learning in Energy Storage Optimization for Carbon Neutrality." Renewable Energy (2026): 125772.

[6] Lei, Lei, Suola Shao, and Lixia Liang. "An evolutionary deep learning model based on EWKM, random forest algorithm, SSA and BiLSTM for building energy consumption prediction." Energy 288 (2024): 129795.

[7] Pardo, M. A., and Francisco José Navarro-Gonzalez. "Sizing and scheduling optimisation method for off-grid battery photovoltaic irrigation networks." Renewable Energy 221 (2024): 119822.

[8] Li, Yu, Zhenghuang Shi, Wenhui Chen, Yibai Wang, and Yiwen Ye. "The Low-Carbon Policy and Urban Industrial Transformation: Evidence from China's Low-Carbon City Pilot Using Double Machine Learning." Sustainability 18, no. 8 (2026): 4088.

[9] Azam, Md Esmay, Md Feroz Ali, and Md Abdur Razzak. "Smart hybrid microgrid for island electrification: integrated techno-economic optimization, dynamic stability, and demand response." Energy Conversion and Management: X (2026): 101656.

[10] Calogine, Didier, Johann Francou, Cedric Abbezzot, and Tovondahiniriko Fanjirindratovo. "Data in experimental stand-alone microgrid: Solar production, domestic loads, battery storage and meteorological series." Data in Brief 51 (2023): 109643.

[11] Biswal, Biswajit, Subhasish Deb, Subir Datta, Taha Selim Ustun, and Umit Cali. "Review on smart grid load forecasting for smart energy management using machine learning and deep learning techniques." Energy Reports 12 (2024): 3654-3670.

[12] Sahay, Kishan Bhushan, and Mukhtiar Singh. "Quantum-Inspired Deep Learning Model for Short-Term Load Forecasting in Smart Grid Energy Management." IEEE Access (2026).

[13] Tan, Ling, Ruixing Kang, Jingming Xia, and Yue Wang. "Application of multi-source data fusion on intelligent prediction of photovoltaic power." Solar Energy 277 (2024): 112706.

[14] Bertolini, Marina, Pierdomenico Duttilo, and Francesco Lisi. "Accounting carbon emissions from electricity generation: A review and comparison of emission factor-based methods." Applied Energy 392 (2025): 125992.