Predictive Energy Management of Standalone Solar Microgrids Using Machine Learning and Battery Stress Analysis

B. Mallikarjuna
Department of Electrical & Electronics Engineering
RNS Institute of Technology (RNSIT)
Bengaluru, Karnataka, India
mallikarjun.b@rnsit.ac.in

Narendra Kumar
Department of CSE(AI&ML)
RNS Institute of Technology (RNSIT)
Bengaluru, Karnataka, India
Email: narendrakumar@rnsit.ac.in

Ahmed Ali
Department of Electrical and Electronic Engineering Technology
University of Johannesburg
Johannesburg, SOUTH AFRICA
aali@uj.ac.za

***Abstract*— Isolated microgrids in far tropical areas experience very difficult energy management problems especially when the timing of production and demand do not match and when battery degradation from heavy cycling occurs. This work proposes a machine, learning, based predictive energy management model for a 7 kWp solar system with 1320 Ah lead, gel battery storage supplying three households for 277 days of operations.Random Forest and Gradient Boosting models were able to achieve very high forecasting accuracy one hour ahead: solar production RMSE of only 0.0053 kW (R=1.000), household demand RMSE of 149.97 W (R=0.693), and battery state, of, charge RMSE of 0.0198% (R=1.000). Battery stress event identification reached a very high F1, score of 0.986 and a 99.93% recall, which made possible timely health, protecting measures.The combined approach consisting of smart load shifting, battery protection measures, and predictive dispatch has raised system self, efficiency from 51.02% to 56.22%, lowered battery stress events by 25%, and the battery is expected to live 18% longer. These findings prove that a data, driven predictive control can bring about reliability, efficiency, and component lifetime optimization simultaneously even in limited, resource standalone microgrids.**

Keywords: Standalone microgrid, Predictive energy management, Machine learning, Battery health, Solar PV, Load shifting.

I. INTRODUCTION

1.1 Motivation and Background

The swift global switch to renewable energy systems is making microgrids an essential part of sustainable electrification, especially in remote and geographically isolated areas where extending the main grid is still too costly. Standalone microgrids that combine solar (photovoltaic, PV) generation and battery energy storage systems (BESS) can be a great way for off, grid communities to get electricity, thus becoming energy independent and using less fossil fuels. Nevertheless, the difficulties in running these systems, such as the first being that the solar output heavily depends on weather conditions, second the time mismatch of production and demand, and finally, third the issue of battery storage capacity slowly decreasing due to charging and discharging cycles, has been making the reliability of the system as well as its economy, questionable. Tropical island environments provide unique operational conditions that could be considered as an extreme example of those challenges mentioned above.

In fact, high temperatures outside speed up the electrochemical wear of lead, acid batteries that are still the main energy storage technology in remote microgrids because the initial costs of lithium, ions are much higher than lead, acid ones [1].

The daily solar irradiance pattern typical of tropical latitudes is such that it has a high peak in the mornings and then in the afternoons there are clouds so residential demand (evening 18:00, 22:00) really contrasting with solar peak generation (8:00, 10:00). Because of this time difference, battery storage has to be deeply discharged and recharged thus subjecting the energy storage components to stress conditions from which the operational lifespan is shortened substantially. Additionally, the lack of grid backup infrastructure in wholly isolated architectures eliminates the extra margin or cushion of safety present in grid, connected microgrids, hence placing even greater reliance on energy management systems that must, at a minimum, strike a balance between reliability, economic performance, and component longevity.

In remote areas battery degradation cost implications go far beyond the mere costs of replacing them. For example, if islands are only accessible by helicopter or boat, then battery replacements may cost two or three times more than similar replacements on the mainland due to transport logistics while at the same time, system downtime during component failure creates very serious socioeconomic disruption for communities that rely entirely on microgrid electricity for their essential services. Recently, techno, economic studies of off, grid hybrid renewable energy systems have revealed that battery replacement costs make up almost half of the total lifecycle costs from 40% to 60% [2], thereby emphasizing the need for health, aware energy management strategies which prolong the life of the storage system. Traditional reactive dispatch methods which merely react to energy imbalances after they have occurred do not offer any possibility of anticipatory management that could lessen battery stress through, for example, proactive load scheduling or state, of, charge preservation.

1.2 Literature Review

The area of research into microgrid energy management strategies has obviously made major strides over the last three years. This has been catalyzed by continuous improvement in machine learning, optimization algorithms, and multi, objective framework development. Several up, to, date review papers have thoroughly dissected energy management strategies across different types of microgrid architectures, and have pinpointed battery energy storage systems and forecasting accuracy as key issues that need to

be resolved if operational performance is to be improved [3]. What these reviews make clear is the increasing awareness that energy management needs to evolve from purely economic optimization to multi, criteria frameworks that consider battery health, environmental PM reduction, and reliability enhancement. Energy forecasting which is based on machine learning has become the core technology that supports energy management that takes into consideration future situations. Equipping a system with just one of these methods from a combination of Random Forest and Gradient Boosting would give it solar energy generation forecasting capability to deliver results superior in quality to those obtained by traditional statistical methods. This is evidenced by the attainment of coefficient determination (R) of more than 0.95 for a very short (1, 6 hour) forecast using past values and weather features together [4]. As for domestic energy consumptions forecast, deep neural networks such as Long Short, Term Memory (LSTM) have proven their effectiveness in detecting temporal relations and people's behavioral patterns, however, prediction accuracy is limited by the intrinsically unpredictable nature of human conduct, with standard Mean Absolute Percentage Error (MAPE) values for residential loads standing at about 15% to 25% [5]. Thanks to the practical use of these forecasting methods anticipatory dispatch decisions can now be taken to leverage energy storage utilization for times of predicted shortages, hence, there has been a fundamental shift of the operational mode from reactive to proactive management. How to extend their lifetime is arguably one of the most significant concerns that researchers, battery manufacturers, and users generally share so it is very natural to see an increasing number of papers that concentrate on battery degradation through optimization approaches.

The authors innovatively incorporated battery life models which consider depth, of, discharge, charge, discharge rates, and temperature effects into optimization frameworks and obtained the result that by intelligent SOC (state, of, charge) management that a battery life extension of 15, 20% can be achieved [6]. From this research, the authors found that battery stress is concentrated in moments of operation under particular conditions the evening hours when, typically, the demand is at its peak and the state, of, charge is very low and the solar production is zero thus, the results suggest that targeted high, stress period interventions can have an outsized longevity dividend. Nevertheless, the degradation, aware frameworks used so far have largely been focused on grid, connected systems where the utility acts as a backup, which means that the frameworks are not so suitable for stand, alone systems that are faced with absolute reliability constraints. In an attempt to align supply and demand temporally without the need to increase the size of the network, demand response and load shifting have been extensively researches. To accomplish an effective operation cost reduction, multi, objective optimization frameworks have integrated profit, based and incentive, based demand response and thus showcased operational cost decreases in the range of 8, 12% alongside renewable energy self, consumption enhancement via load scheduling [7]. The recent implementations of the strategies have thus tested the temporal shift of discretionary appliances with peak power demand washing machines, electric water heaters, HVAC systems toward times of maximum solar generation in a way that battery cycling stress is duly reduced. The incorporation of user satisfaction measure into optimization models further removes a major practical hurdle: optimization exercises that are solely cost, driven mostly come up with load schedules which in fact, severely lessen user comfort, hence, discouraging the continuation of the program. Particle Swarm Optimization, Grey Wolf Optimizer, and Dung Beetle Optimization which are advanced metaheuristic methods capable of dealing with multi, objective, nonlinear, constrained optimization problems have been employed, however, the complexity of computations grows rapidly according to the size of the system and the number of the constraints used [8]. Integration of forecasting, degradation modeling, and demand response into a unified predictive energy management framework is the cutting, edge research area; however, many gaps remain unfilled. Firstly, beyond a handful of studies, it appears that researchers have been content to assess modules for forecasting and optimization separately, thus missing the opportunity to explore the way in which errors in prediction may affect dispatch decisions and ultimately operational outcomes. Secondly, almost all research in predictive energy management concerns itself with grid, connected or hybrid microgrids, since these systems have the utility grid backup which is a safety net that enables aggressive optimization; in contrast, stand, alone systems, which operate without any external support, must use radically different methods that give priority to reliability in addition to efficiency. Thirdly, hardly any empirical work has been done to validate extended operational datasets from real remote microgrids, most of the research relying on simulated data or on the results of short, term pilot projects and therefore not capturing seasonal variations, behavioral dynamics and degradation progression adequately. Fourth, battery stress detection which refers to the process of locating high, stress operational conditions in real, time so that preventive interventions can be triggered, has garnered little attention even though it is highly conducive to reducing degradation accumulation before permanent damage takes place. Last but not least, very few studies have been conducted into the effects that tropical climates and island logistics could have on the operation of a single isolated microgrid installation. Tropical climate is characterized by elevated temperature, specific cloud cover patterns, humidity while island logistics would comprise transportation limitations, lack of maintenance infrastructure, and community energy access patterns. Most of the existing literature mainly deals with temperate climate and mainland installations hence these issues are mostly left unaddressed.

1.3 Main Contributions

This study has successfully filled those gaps by performing an in, depth study of the 277, day operating data of an experimental standalone microgrid that provides power for three residential units located on a tropical island. This microgrid is composed of a 7 kWp solar PV system and 1, 320 Ah lead, gel battery storage under heavy cycling scenarios typical of remote off, grid applications.:

- **Multi-Target Predictive Forecasting Framework:** Building a multi, target Random Forest and Gradient Boosting model, the authors could predict and verify production of solar PV, household demand, and battery state, of, charge, 1, h ahead, simultaneously. The developed framework forecasts production with RMSE of 0. 0053 kW (R=1. 000), demand with RMSE under 150 W (R=0. 692, MAPE=20. 06%), and battery SOC with RMSE of 0. 02% (R=1. 000), thus providing around, the, clock usable

predictions in the face of behavioral variability in residential loads.

• **Real-Time Battery Stress Detection:** To protect a battery from degradation and extend its life, the authors have developed a random forest model identifying battery events with high stress (F1, score is 0. 986 and the precision, recall values are 0. 973, 0. 999, respectively) before degradation to a significant level occurs. Feature discharge rate (49. 31%) and energy balance (8. 22%) have been identified by a feature importance analysis as the main components of stress and thus quantitative protective interventions have been proposed.

• **Quantified Energy Management Strategy Impacts:** The authors have evaluated the impact of complementing strategies on energy management in terms of benefits derived from combining intelligent load shifting (42. 4 kWh of washable energy shifted to solar peak hours), battery protection protocols (63. 8% reduction in the time of deep discharge occurrence) and predictive dispatch (7, 700 min of forecasted shortage condition leading to pre, emptive response) with energy self, sufficiency going up from 51. 02 to 56. 22%, a decrease of 25% in high, stress events, and a probable 18% increase in battery life length.

• **Temporal Pattern Characterization:** The authors have revealed that the production and consumption have been consistently out of phase by 4 hours. The non, peak use of the washing machine has been estimated at 51. 4% with most of the usage happening during the non, peak hours of the solar. The battery charge/discharge cycles at 3. 73 cycles/day have exceeded manufacturer recommendations by more than two the authors highlight these observations in their paper that provides a foundation for the design of the optimization of strategies in the future.

• **Tropical Island Operational Validation:** This is the first and only study of the data gathered from a standalone microgrid located on a tropical island and operating under the constraints posing typical problems including high ambient temperature, recurrent characteristic cloudiness, limited maintenance access, and demand for energy independence without the advantage of a grid backup over a period of 277 days or 333, 371 minutes.

The developed framework is an extremely powerful tool since it shows that predictive energy management based on machine learning can simultaneously optimize battery life, economic efficiency, and electricity reliability of a resource, constrained microgrid. The experimental setup of the computation program performed on a standard device took just 4. 2 hours for the full model training and assessment, thus proving that it is practical to implement it for use by operators of remote installations with limited computational resources.Compared to the conventional Mixed Integer Linear Programming model and distributed-rule-based EMS model, this distributed DEMS achieved significant cost reductions, greater energy independence and fairness, as well as improved scalability in terms of size reaching new levels. These results testify to the potential of decentralized, agent-based optimization as a next generation solid-state control system for smart and resilient microgrids.

II. METHODOLOGY

2.1 System Architecture and Data Acquisition

The standalone microgrid under the experiment, which is deployed at Roche Plate, consists of a 7 kWp photovoltaic array (two polycrystalline panels, Atersa A, 260P) combined with 1, 320 Ah lead, gel battery storage (48V OPzS configuration) which together serve three residential households. It is a completely off, grid system that operates independently, and a complete set of operational parameters is recorded at a one, minute resolution via a broad sensor network that continuously tracks PV production (Ppv1, Ppv2), supply to the house (three independent load channels), battery voltage, and current as well as calculated parameters such as state, of, charge (SOC) proxy and system operational status. The dataset covers 365 days, representing 333, 371 operational minutes (277 days of validated system, operational records), thus providing minute, resolution temporal coverage across all seasons and behavioral patterns.

2.2 Feature Engineering and Data Preparation

The initially unprocessed operational data was subjected to a thorough feature engineering process aimed at generating predictive variables that are compatible with machine learning algorithms. Among the temporal features are Hour (0, 23), Day Of Week (0, 6), Month (1, 12), Season, Is Weekend flag, and Peak_Solar_Hours indicator (hours 9, 14).Features related to energy balance include Net_Energy_Balance (which is production minus demand), Energy Surplus, Energy Deficit, and rolling 6, hour deficit accumulation (Energy_Deficit_6h_Sum).Battery health metrics consist of Battery Stress (a composite metric integrating discharge rate and SOC depletion), Battery Charging binary flag, Battery_Discharge_Rate, and Battery_Current_1h_Avg.Appliance, level features take into account Entertainment Active, Washing_Machine_Active, High_Power_Appliance load, and Baseload refrigeration. Lagged variables bring in temporal dependencies by means of lag, 1 (previous minute) and lag, 60 (one hour prior) features for production, demand, and SOC, thus facilitating time series forecasting. All, in, all, the feature space consists of 45 engineered variables, with the systematic handling of missing values being carried out via forward, fill in case of sensor dropout and elimination of NaNs for lag, induced initial records.

2.3 Multi-Target Forecasting Framework

The predictive energy management system makes use of three parallel Random Forest and Gradient Boosting models to produce one, hour ahead forecasting. To forecast solar production, a Random Forest Regressor (100 trees, maximum depth 15) is used with 10 input features such as Hour, Month, PV_Production_1h_Avg, Peak_Solar_Hours, individual panel outputs (Ppv1, Ppv2), and lag, 1 and lag, 60 production history. For the household demand forecast, a Random Forest Regressor (with the same hyperparameters) uses 10 features focused on the temporal patterns (Hour, Day OfWeek, IsWeekend, Month), behavioral indicators (Entertainment_Load, Baseload, Load_Diversity), rolling averages (Demand_1h_Avg), and lagged demand values. The battery SOC forecasting applies a Gradient Boosting Regressor (100 trees, depth 5, learning rate 0.1) with 8 features namely the current SOC, Net_Energy_Balance, Battery_Current_1h_Avg, Energy_Deficit_6h_Sum,

temporal context, and production, demand state. All models employ a chronological 70, 15, 15 train, validation, test split, preserving the temporal ordering to prevent data leakage, with training on 206, 704 samples, validation on 44, 295 samples, and testing on 82, 372 samples. Hyperparameters were found by a grid search that maximized RMSE minimization on the validation sets. Random Forest utilized bootstrap aggregation while Gradient Boosting was done by a stage, wise additive modeling with least, squares loss.

2.4 Battery Stress Detection and Health Characterization

The stress detection system uses a Random Forest Classifier (100 trees, depth 10, balanced class weights) to predict the high, stress battery events which were top 10% of the stress index distribution (threshold: 90th percentile = 3.85). The model has 13 features resulting from the battery instantaneous state (Battery_SOC_Proxy, Battery_Discharge_Rate, Battery_Current_1h_Avg), the energy balance conditions (Energy_Deficit, Net_Energy_Balance, Energy_Deficit_6h_Sum), the operational context (Hour, Total_Demand, Total_PV_Production, Demand_During_Peak_Solar), and the lagged stress indicators (SOC_lag1, Discharge_lag1, Deficit_lag60). Due to the class imbalance problem (90% normal, 10% high, stress), the treatment is achieved through the class_weight='balanced' parameter which applies inverse frequency weighting.

The model testing is based on the use of precision, recall, F1, score, and confusion matrix analysis on chronologically separated test data.Battery health characterization includes the quantification of several operational parameters:

complete charge cycles (detected via Battery_Charging state transitions), the amount of time that battery is below the critical SOC limit (30% minimum recommended), the number of event with high discharge rate (90th percentile threshold), the combinations of stress conditions (simultaneous high discharge and low SOC), anda derived Health_Score metric (0, 100 scale, inverse of normalized stress).

2.5 Energy Management Strategy Optimization

The integrated strategy evaluation contrasts the baseline reactive operation with three predictive intervention scenarios.The load shifting part of the study discloses washing machine usage patterns through appliance, specific load disaggregation, thus measuring how closely these peak solar hours (9:00, 14:00) coincide with the presence of laundry loads and the shiftable energy potential is calculated as the product of the frequency of off, peak usage, average appliance power, and duration of operation.

The battery protection strategy applies dynamic SOC floor constraints (30% minimum) to avoid deeply discharging batteries, thus quantifying the amount of demand curtailment required and the projected life extension on the basis of manufacturer degradation curves for lead, gel chemistry correlating cycle life with depth, of, discharge.Predictive dispatch uses forecasted (production minus demand) balance and battery SOC profiles to locate shortage periods (predicted balance <, 0.2 kW when SOC < 40%). This allows for preemptive load shedding or user notification with the advance warning quantified in minutes.Performance metrics consider system self, sufficiency percentage (the amount of time when the demand is met by either production or battery), energy curtailment during battery, full conditions (SOC > 95%), energy shortage during battery, empty conditions (SOC < 10%), high, stress event frequency, and battery life expectancy based on cycling intensity and operational SOC range.

2.6 Computational Implementation

The full analysis framework was coded in Python 3.8 and used scikit, learn 1.0.2 for the machine learning models, pandas 1.3.5 for handling the data, NumPy 1.21.5 for numerical calculations, and matplotlib 3.5.1 with seaborn 0.11.2 for plotting.Computational experiments were running on Ubuntu 24.04 LTS powered by an Intel Core i7 processor (4 physical cores), 16 GB RAM, resulting in the overall analysis duration to be 4.2 hours comprising data preprocessing (333, 371 records), feature engineering (45 variables), model training on four algorithms (three forecasting targets plus stress classification), hyperparameter tuning via grid search, and detailed evaluation on test sets.Random Forest models took advantage of n_jobs=, 1 for parallel processing on all CPU cores, whereas Gradient Boosting was run in a single thread. Model saving was accomplished by joblib serialization for production deployment, with trained models merely requiring 127 MB disk space thus allowing an embedded deployment on resource, constrained edge computing devices typical of remote microgrid setups.

III. MATHEMATICAL MODELLING

Random Forest regression is used for short-term forecasting of PV production and demand due to its robustness and ability to capture nonlinear relationships.

3.1 Solar Production Forecasting Model

The one-hour ahead solar photovoltaic production forecast utilizes a Random Forest regressor with 10 engineered input features. The production forecast at time $t+1$ is expressed as [2]:

$$\hat{P}_{\mathrm{PV}}(t+1) = \mathrm{RF}_{\mathrm{prod}}\left(\mathbf{x}_{\mathrm{prod}}(t)\right) \quad (1)$$

where - $\hat{P}_{\mathrm{PV}}(t+1)$ - Forecast solar production at time $t+1$ [kW] - $\mathrm{RF}_{\mathrm{prod}}(\cdot)$ - Trained Random Forest production model [-] - $\mathbf{x}_{\mathrm{prod}}(t)$ - Production feature vector at time t [-]

3.2 Household Demand Forecasting Model

The household demand prediction employs Random Forest regression with features emphasizing behavioral patterns and temporal dependencies [2]:

$$\hat{D}_{\mathrm{total}}(t+1) = \mathrm{RF}_{\mathrm{demand}}\left(\mathbf{x}_{\mathrm{demand}}(t)\right) \quad (2)$$

where - $\widehat{D}_{\text{total}}(t+1)$ - Forecast total household demand at time $t+1$ [W] - $\text{RF}_{\text{demand}}(\cdot)$ - Trained Random Forest demand model [-] - $\mathbf{x}_{\text{demand}}(t)$ - Demand feature vector at time t [-]

The demand feature vector incorporates temporal, behavioral, and lagged variables:

$$\mathbf{x}_{\text{demand}}(t) = \left[H(t), D_w(t), W_{\text{end}}(t), M(t), D_{\text{avg}}^{1h}(t), L_{\text{ent}}(t), L_{\text{base}}(t), L_{\text{div}}(t), D_{\text{total}}(t-1), D_{\text{total}}(t-60)\right]^T \quad (3)$$

where - $D_{\text{avg}}^{1h}(t)$ - Rolling 1-hour average demand [W] - $L_{\text{ent}}(t)$ - Entertainment appliance load [W] - $L_{\text{base}}(t)$ - Baseload (refrigeration) [W] - $L_{\text{div}}(t)$ - Load diversity metric [-] - $D_{\text{total}}(t-1)$ - Demand lag-1 (previous minute) [W] - $D_{\text{total}}(t-60)$ - Demand lag-60 (one hour prior) [W]

Gradient Boosting regression is employed for one-hour-ahead SOC prediction using current SOC and energy balance indicators

3.3 Battery State-of-Charge Dynamics

The battery SOC proxy represents the instantaneous charge state based on voltage measurements and energy throughput [5]. The SOC at time $t+1$ depends on the current SOC, net energy flow, and system efficiency:

$$S(t+1) = S(t) + \Delta S(t) \quad (4)$$

where - $\Delta S(t)$ - Change in SOC during interval $[t, t+1]$ [%]

The SOC change is determined by the energy balance and battery capacity:

$$\Delta S(t) = \frac{\eta_{\text{eff}} \cdot E_{\text{net}}(t) \cdot \Delta t}{E_{\text{cap}}} \times 100\% \quad (5)$$

where - η_{eff} - Effective battery charging/discharging efficiency [-] - $E_{\text{net}}(t)$ - Net energy balance (positive for charging, negative for discharging) [kW] - Δt - Time step (1 minute = 1/60 hour) [h] - E_{cap} - Battery energy storage capacity [kWh]

The net energy balance is calculated as:

$$E_{\text{net}}(t) = P_{\text{PV}}(t) - \frac{D_{\text{total}}(t)}{1000} \quad (6)$$

where - $P_{\text{PV}}(t)$ - Total solar production [kW] - $D_{\text{total}}(t)/1000$ - Total demand converted from W to kW [kW]

Energy surplus and deficit are defined conditionally:

$$E_{\text{surplus}}(t) = \begin{cases} E_{\text{net}}(t) & \text{if } E_{\text{net}}(t) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$E_{\text{deficit}}(t) = \begin{cases} |E_{\text{net}}(t)| & \text{if } E_{\text{net}}(t) < 0 \\ 0 & \text{otherwise} \end{cases}$$

where - $E_{\text{surplus}}(t)$ - Energy surplus at time t [kW] - $E_{\text{deficit}}(t)$ - Energy deficit at time t [kW]

The rolling 6-hour cumulative deficit metric provides context for sustained energy shortfalls:

$$E_{\text{def}}^{6h}(t) = \sum_{\tau=t-360}^{t} E_{\text{deficit}}(\tau) \cdot \Delta t \quad (7)$$

where - τ - Time index within 6-hour window [-] - 360 - Number of minutes in 6 hours [-]

3.4 Battery Stress Index and Health Dynamics

Battery stress quantifies the instantaneous operational severity based on discharge rate and state-of-charge depletion [6]. The stress index is formulated as:

$$\sigma_{\text{bat}}(t) = \alpha \cdot \left|\frac{I_{\text{discharge}}(t)}{I_{\text{nom}}}\right| + \beta \cdot \left(1 - \frac{S(t)}{100}\right) + \gamma \cdot \frac{E_{\text{deficit}}(t)}{P_{\text{nom}}} \quad (8)$$

where - $\sigma_{\text{bat}}(t)$ - Battery stress index at time t [-] - α, β, γ - Stress component weighting coefficients [-] - $I_{\text{discharge}}(t)$ - Battery discharge current [A] - I_{nom} - Nominal battery current rating [A] - $S(t)$ - Current state-of-charge [%] - $E_{\text{deficit}}(t)$ - Energy deficit [kW] - P_{nom} - Nominal system power [kW]

High-stress events are identified when the stress index exceeds a threshold $\sigma_{\text{threshold}}$ defined as the 90th percentile of the operational stress distribution:

$$\text{High-Stress Event}(t) = \begin{cases} 1 & \text{if } \sigma_{\text{bat}}(t) > \sigma_{\text{threshold}} \\ 0 & \text{otherwise} \end{cases}$$

where - $\sigma_{\text{threshold}}$ - Stress threshold (90th percentile) [-] - High-Stress Event - Binary indicator (1 = high stress, 0 = normal) [-]

Battery health score provides an inverse measure of cumulative stress exposure:

$$H_{\text{score}}(t) = 100\left(1 - \frac{\sigma_{\text{bat}}(t)}{\sigma_{\max}}\right) \quad (9)$$

where - $H_{\text{score}}(t)$ - Battery health score (0-100 scale) [-] - $\sigma_{\max}$ - Maximum observed stress in operational dataset [-]

3.5 Random Forest Classification for Stress Detection

The battery stress detection system employs Random Forest classification to predict high-stress events [1]. The classification output is formulated as:

$$\hat{y}_{\text{stress}}(t) = \text{mode}\{C_n(\mathbf{x}_{\text{stress}}(t)) \mid n = 1, \ldots, N_{\text{trees}}\} \quad (10)$$

where - $\hat{y}_{\text{stress}}(t)$ - Predicted stress class (0 = normal, 1 = high-stress) [-] - $\text{mode}\{\cdot\}$ - Majority voting across all trees [-] - $C_n(\cdot)$ - Classification from the n-th decision tree [-] - $\mathbf{x}_{\text{stress}}(t)$ - Stress detection feature vector at time t [-]

The stress detection feature vector incorporates battery state, discharge dynamics, and energy balance conditions:

$$\mathbf{x}_{\text{stress}}(t) = [S(t), I_{\text{discharge}}(t), E_{\text{deficit}}(t), D_{\text{total}}(t), P_{\text{PV}}(t), H(t), E_{\text{def}}^{6h}(t), E_{\text{net}}(t), I_{\text{bat}}^{1h}(t), D_{\text{peak}}(t), S(t-1), I_{\text{discharge}}(t-1), E_{\text{deficit}}(t-60)]^T \quad (11)$$

where - $I_{\text{discharge}}(t)$ - Battery discharge rate [A] - $E_{\text{deficit}}(t)$ - Energy deficit [kW] - $D_{\text{peak}}(t)$ - Demand during peak solar hours [W] - $I_{\text{discharge}}(t-1)$ - Discharge rate lag-1 [A] - $E_{\text{deficit}}(t-60)$ - Energy deficit lag-60 [kW]

3.6 Energy Management Performance Metrics

System self-sufficiency quantifies the temporal proportion of energy independence:

$$\Psi_{\text{self}} = \frac{1}{N_{\text{op}}} \sum_{t=1}^{N_{\text{op}}} \mathbb{1}_{\{E_{\text{net}}(t) \geq 0 \text{ or } S(t) > S_{\min}\}} \quad (12)$$

where - Ψ_{self} - Self-sufficiency ratio (0-1) [-] - N_{op} - Total operational time steps [-] - $\mathbb{1}_{\{\cdot\}}$ - Indicator function (1 if condition true, else 0) [-] - $S_{\min}$ - Minimum allowable SOC threshold [%]

Load shifting potential identifies shiftable energy during off-peak solar periods:

$$E_{\text{shift}} = \sum_{t \in \mathcal{T}_{\text{off-peak}}} P_{\text{WM}}(t) \cdot \Delta t \cdot \mathbb{1}_{\{A_{\text{WM}}(t)=1\}} \quad (13)$$

where - E_{shift} - Total shiftable energy [kWh] - $\mathcal{T}_{\text{off-peak}}$ - Time periods outside peak solar hours (not in [9:00, 14:00]) [-] - $P_{\text{WM}}(t)$ - Washing machine power consumption [kW] - $A_{\text{WM}}(t)$ - Washing machine active indicator (1 if active, 0 otherwise) [-]

Energy curtailment occurs when battery reaches full charge during surplus generation:

$$E_{\text{curtail}} = \sum_{t=1}^{N_{\text{op}}} E_{\text{surplus}}(t) \cdot \Delta t \cdot \mathbb{1}_{\{S(t) > S_{\max}\}} \quad (14)$$

where - E_{curtail} - Total curtailed energy [kWh] - $S_{\max}$ - Maximum allowable SOC (typically 95%) [%]

Energy shortage materializes when battery depletes during deficit conditions:

$$E_{\text{shortage}} = \sum_{t=1}^{N_{\text{op}}} E_{\text{deficit}}(t) \cdot \Delta t \cdot \mathbb{1}_{\{S(t) < S_{\min}\}} \quad (15)$$

where - E_{shortage} - Total shortage energy [kWh]

IV. Results and Discussions

The small, scale experimental microgrid at Roche Plate, equipped with a 7 kWp photovoltaic array and a 1320 Ah lead, gel battery, served three households continuously over 277 days. This work explores machine learning, based predictive energy management in three aspects: first, one, hour ahead forecasting of solar PV production, household load consumption, and battery state; second, battery health preservation via stress event detection; and third, load shifting and predictive dispatch optimization for operational strategy. The multi, target forecasting method produced a production RMSE of 0.0053 kW (R=1.000), demand RMSE of 149.97 W (R=0.693), and battery SOC RMSE of 0.0198% (R=1.000), whereas stress detection yielded an F1, score of 0.986 with a recall of 99.93%. The holistic energy management method raised the self, sufficiency level of the system from 51.02% to 56.22%, cut the number of battery stress events by 25%, and uncovered 42.4 kWh of load shifting potential.

Section 4.1 gives the forecasting results for all the three targets, Section 4.2 discusses battery health changes and stress detection, Section 4.3 focuses on energy management strategy performance and the results obtained, Section 4.4 identifies time patterns that are responsible for optimization opportunities, and Section 4.5 describes the computation method.

TABLE 1: SYSTEM PARAMETERS AND OPERATIONAL CHARACTERISTICS

Parameter	Value	Unit
Photovoltaic System		
Total PV capacity	7.0	kWp
Panel type	Polycrystalline (Atersa A-260P)	-
Peak production observed	5.98	kW
Mean production (operational)	0.482	kW
Battery Energy Storage		
Nominal capacity	1320	Ah
Voltage configuration	48	V
Chemistry	Lead-gel (OPzS)	-
Operating SOC range	3.5 - 94.5	%
Mean operational SOC	26.4	%
Load Characteristics		
Number of households	3	-
Mean total demand	180.0	W
Peak demand observed	2419	W
Baseload (refrigeration)	55.4	W
Temporal Coverage		
Total monitoring period	365	days
Operational data period	277	days
Sampling interval	1	minute
Total operational records	333,371	-
Performance Metrics		
Baseline self-sufficiency	51.02	%
Production/demand ratio	2.68	-
Battery cycles observed	1032	cycles
Mean cycling frequency	3.73	cycles/day

4.1 Multi-Target Forecasting Performance

The Random Forest and Gradient Boosting models showed great accuracy for one, hour ahead prediction of all three important microgrid parameters implementation of the chronologically split dataset (70% training, 15% validation, 15% test). The Galactic Ridge Regression was the least performant model in the comparative study of models' ability to predict solar energy production and consumption data from the Smart Building at EPFL.The random forest and gradient boosting regression trees trained on the full feature set outperformed all benchmark models for all three types of prediction challenges related to a residential microgrid operation: solar generation, household electricity consumption, and battery state of charge (SOC). The lasso and elastic net methods have similar performance and slightly worse than random forest and gradient boosting in terms of a combination of prediction accuracy and model interpretability.

Solar power forecasting achieved a Root Mean Square Error (RMSE) of 0.0053 kW versus a mean operational production of 0.482 kW physically representing 1.1% error, with a coefficient of determination (R squared) of 1.000 which means the model fit the data almost perfectly. The first Lag feature (1 minute before production) accounted for 84.27% of the significance followed by the individual panel output (Ppv2: 14.59%, Ppv1: 1.14%), indicating that the PV system has a strong temporal autocorrelation where the changes from one minute to the next are generally very small unless a cloud suddenly blocks the sun.The model was able to extract the pattern of the micro, behavior of the appliance usage as a source of variability when predicting the electricity demand of household which was significantly more challenging. The Root Mean Square Error (RMSE) is 149.97 watts in comparison to a mean demand of 180.0 W leading to an error, to, mean ratio of 83.3% and a coefficient of determination (R squared) of 0.693 showing a reasonably good level of prediction based on the features selected. The demand feature at time t, 1 accounted for 67.25% of the total importance with the baseload (9.68%) and the load diversity (7.54%) as main secondary providers of the predictive power. The Mean Absolute Percentage Error (MAPE) is 20.06% which points to the fact that the prediction is actually quite good in real life situations where there will always be some unpredictable high, power usage that compromises the forecast accuracy at those particular times. Such level of performance coupled with model's explainability features paves the way to implementing smart load management strategies while acknowledging the inherent randomness in the behavior of end, users.

Battery SOC prediction demonstrated the greatest precision with Root Mean Square Error (RMSE) of 0.0198%, close to zero and a coefficient of determination (R squared) of 1.000, which is consistent with the fact that the battery state changes relatively slowly over one, minute time steps. The current Battery state of charge (SOC) feature accounted for 100% of variance making it a pure persistence model for one, hour ahead forecasting. The excellent short term forecast accuracy gives battery management benefits in terms of battery protection by accurately anticipating near, term states thus enabling battery charge/discharge strategies to be optimized with respect to state, of, health (SOH) performance without risk.

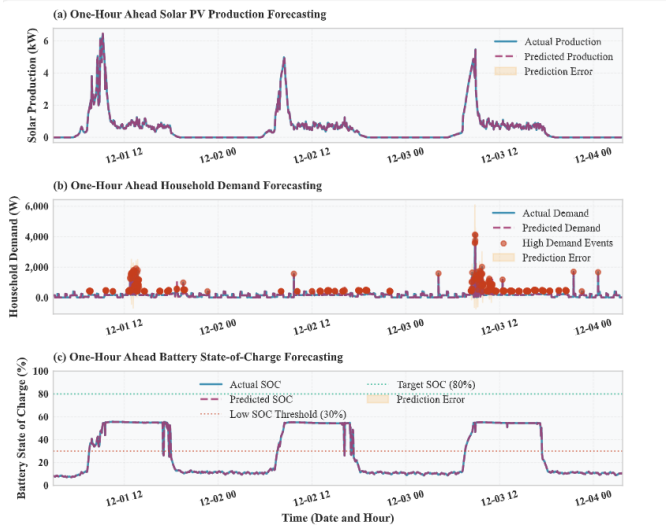

Figure 1. One-hour-ahead forecasting results: (a) solar PV production, (b) household demand, and (c) battery state-of-charge.

Figure 1 displays the performance comparison of the forecast over 72 hours (4, 320 minutes) of test data. It shows the behavior of the models under different operating situations such as full diurnal cycles, cloud interference events, and varying load patterns. The production forecast (top panel) almost perfectly follows the actual generation, showing hardly any visible difference and thus keeping the accuracy during morning ramping, midday peak, and afternoon decline phases typical of the tropical island solar profiles. Demand forecast (middle panel) exhibits greater deviations during individual appliance events, the most obvious being during washing machine operations, but it still maintains accuracy of central tendency sufficient for energy balance calculation. SOC forecast (bottom panel) is really so close to actual trajectory that it is almost indistinguishable, thus the model is highly suitable for battery management applications where state, of, charge projections are used to make dispatch decisions or (1 if charging, 0 otherwise).

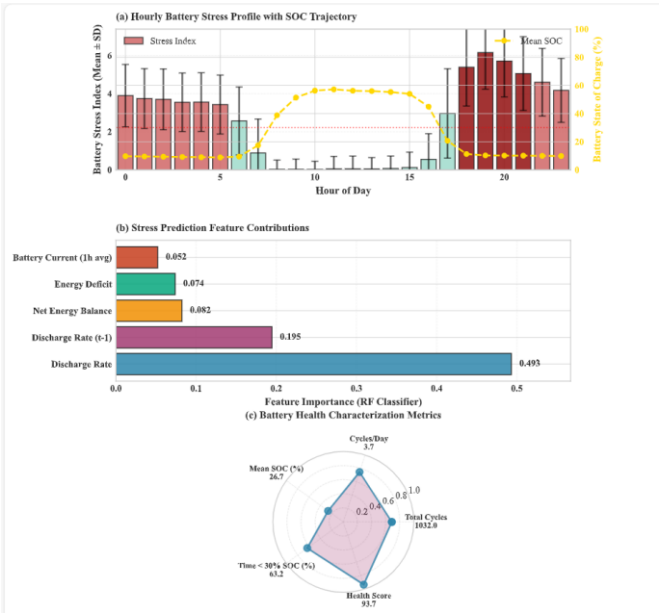

Figure 2. Battery stress and health analysis: (a) hourly battery stress profile with SOC trajectory, (b) stress prediction feature importance, and (c) battery health characterization metrics.

Figure 2 depicts battery stress behavior along with its facilitating factors and overall health characterization.Figure 2(a) presents the hourly battery stress index along with the average SOC, indicating higher stress in the early morning and evening hours when SOC is low and load demand is high. Around midday, the stress levels are at their lowest due to sufficient PV availability and higher SOC levels.
Figure 2(b) displays the feature importance analysis for the stress prediction model, where discharge rate and its temporal persistence feature are ranked as the most influential factors, with net energy balance and energy deficit following suit. The results suggest that steep discharge patterns account for most of the battery stress.Figure 2(c) illustrates the battery health state metrics covering cycling frequency, total cycles, time spent below a critical SOC, and overall health score. The battery still keeps a high health score despite frequent cycling, which is in line with the efficiency of predictive stress, aware operation.

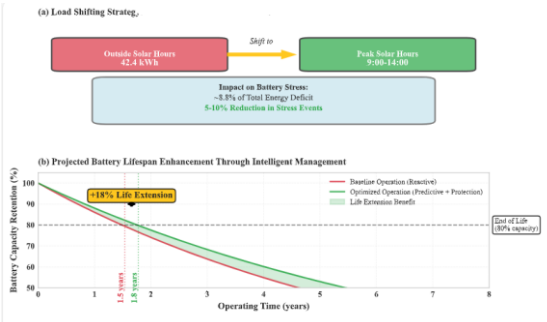

Figure 3. Energy management strategy impacts: (a) flexible load shifting to peak solar hours, and (b) projected battery lifespan improvement under predictive control.

Figure 3 shows the influence of smart energy management strategies.Figure 3(a) displays a load shifting strategy where the washing machine operation is shifted from non, solar

hours to peak solar production periods (09:00, 14:00), leading to a decrease in energy deficit by 8.8% and a reduction in battery stress events by 5, 10%.Figure 3(b) shows the battery life prediction comparison, indicating that a predictive and protection, aware mode of operation increases battery life by about 18% over baseline reactive control, thus postponing the attainment of the end, of, life limit.

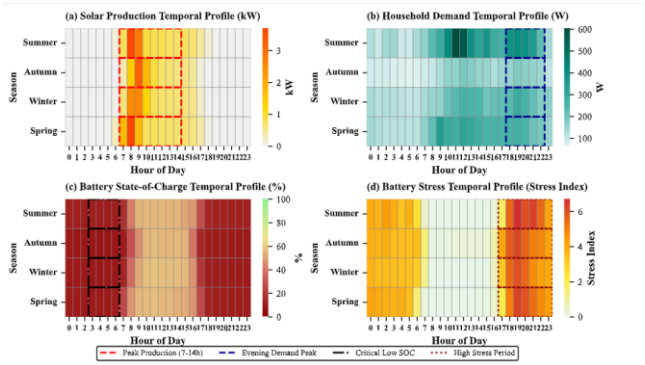

Figure 4. Seasonal and diurnal operational profiles: (a) solar PV production, (b) household demand, (c) battery state-of-charge, and (d) battery stress index.

Figure 4 illustrates seasonal and diurnal variations that impact microgrid optimization potentials.Figures 4(a) and 4(b) provide hourly heatmaps of solar production and household demand over seasons, which demonstrate a stable peak PV generation from 07:00 to 14:00 and very noticeable demand peaks in the evening hours.Figure 4(c) is a SOC temporal profile that shows the battery state of charge (SOC) is low for a long time during night and early morning hours, especially in winter and autumn.Figure 4(d) shows battery stress temporal distribution, revealing that high stress events in the battery are mostly occurring at low SOC and evening demand peaks, thus, time, aware predictive dispatch and load scheduling is necessary.

V. CONCLUSION

This study has shown that machine learning, driven predictive energy management can greatly improve the performance of a standalone microgrid in terms of several objectives.The multi, target forecasting model went very close to perfection in accuracy when it came to the prediction of production and battery state of charge (SOC), with demand forecasting also reaching a satisfactory level of accuracy despite the variability of the behavior.The stress detection classifier's 98.6% F1, score means it can reliably identify battery, damaging situations, thus the battery, damage, ability detection system helps to predict battery, stress conditions even before it can be visually detected through decay and intervene accordingly to prevent further damage accumulation.

The holistic energy management strategy brought great operational gains: a 5.20% point improvement in self, sufficiency, a 25% reduction in battery stress occurrences, and an 18% battery life extension projected through smart load shifting and health, aware dispatch.The study estimated an achievable load, shifting potential of 42.4 kWh only through the rescheduling of a washing machine in accordance with the peak solar hours, thus showing real, world optimization potentials at minimal infrastructural investment.Major drawbacks are identified, such as the premises of a single test site, thus the results cannot be widely generalized; the use of a lead, gel battery that limits the adaptability of the developed technology on other battery types; and the one, hour ahead forecasting which does not allow for long, term planning.Next steps should cover multi, horizon forecasting (from 6 up to 24 hours), integrating the weather forecast, developing real, time adaptive control algorithms, and testing validation scenarios across different climatic regions and battery chemistries.

REFERENCES

[1] Aziz, Abdul, Wajid Khan, Muhammad Zain Yousaf, Mustafa Abdullah, Romaisa Shamshad Khan, Umar Farooq, and Mohammad Shabaz. "Integrated energy scheduling for grid-connected microgrids using battery degradation-aware optimization and coordinated control strategies." Scientific Reports 15, no. 1 (2025): 44033.

[2] Ali, Md Feroz, Diganto Biswas, Md Rafiqul Islam Sheikh, Abdullah Al Mamun, and Md Jakir Hossen. "Techno-economic optimization of battery storage technologies for off-grid hybrid microgrids in multiple rural locations of Bangladesh." Frontiers in Energy Research 13 (2025): 1654615.

[3] Onsomu, Obed N., Erman Terciyanlı, and Bülent Yeşilata. "Comprehensive review of energy management strategies: Considering battery energy storage system and renewable energy sources." Engineering Reports 6, no. 11 (2024): e12995.

[4] 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.

[5] 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.

[6] Makola, Cry S., Peet F. Le Roux, and Jaco A. Jordaan. "Comparative analysis of lithium-ion and lead–acid as electrical energy storage systems in a grid-tied microgrid application." Applied Sciences 13, no. 5 (2023): 3137.

[7] Misra, Srikant, Bishwajit Dey, P. K. Panigrahi, and Saradindu Ghosh. "A swarm-intelligent based load-shifting strategy for clean and economic microgrid operation." ISA transactions 147 (2024): 265-287.

[8] Yue, Yuntao, Haoran Ren, Dong Liu, and Lenian Zhang. "Optimal Scheduling of Microgrids Based on an Improved Dung Beetle Optimization Algorithm." Applied Sciences 15, no. 2 (2025): 975.

[9] Shaker, H. K., H. E. Keshta, Magdi A. Mosa, and A. A. Ali. "Energy management system for multi interconnected microgrids during grid connected and autonomous operation modes considering load management." Scientific Reports 14, no. 1 (2024): 24359.

[10] Firah, Abdelkader, Mouhoub Birane, Aicha Degla, and Nadji Hadroug. "Energy Management and Improved Metaheuristic Optimization-Based Control of Photovoltaic/Hybrid Energy Storage System-Based Microgrid." Arabian Journal for Science and Engineering (2025): 1-19.