Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Leakage-Safe K-Means-Assisted Multi-Objective Energy Management for a Solar and Wind–Battery Micro-grid

Pamela K. F. Barson

pamelabarson93@gmail.com

Kocaeli Üniversitesi

Nurettin Abut

Kocaeli Üniversitesi

Ismet Kandilli

Kocaeli Üniversitesi

Article

Keywords: Energy Management System, Solar–Wind Micro-grid, Battery Energy Storage, Random Forest, K-Means, Data Leakage, Multi-Objective Control, Chronological Validation

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11288771/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Leakage-Safe K-Means-Assisted Multi-Objective Energy Management for a Solar and Wind–Battery Micro-grid

Pamela K. F. Barson[1]*, Nurettin Abut[2] , Ismet Kandilli[3]

Abstract

This study evaluates a leakage-safe, K-Means-assisted multi-criteria energy management framework for a solar-wind-battery micro-grid. Random Forest provides one-hour-ahead load predictions, while design-only K-Means clustering identifies operating regimes for battery and generator dispatch. A chronological design-validation-blind-test protocol separates model development, controller feasibility screening, and final evaluation. Random Forest achieved a blind-test RMSE of 3.343 kW, MAE of 2.315 kW, and R^2 of 0.840, corresponding to 45.68% RMSE skill relative to persistence. On the blind test, the reported regime-adaptive candidate achieved zero unserved energy, reduced generator energy by 59.34% and renewable curtailment by 23.85% relative to prediction-only control, and reduced battery throughput by 37.49% relative to reactive control. These gains involved explicit trade-offs in generator use and battery utilization. However, the validation grid was non-discriminative: all 81 feasible candidates produced identical objective values and normalized scores. The study therefore reports the candidate as a feasibility-screened representative policy, not as a uniquely optimized solution. The results demonstrate the operational trade-offs obtainable within the stated simulator while identifying controller-parameter sensitivity as a necessary next methodological step.

Keywords: Energy Management System, Solar–Wind Micro-grid, Battery Energy Storage, Random Forest, K-Means, Data Leakage, Multi-Objective Control, Chronological Validation.

[1] Pamela K.F. Barson
pamelabarson93@gmail.com

Department of Electrical Engineering, Kocaeli University, Kocaeli, Türkiye

[2] Nurettin Abut
abut@kocaeli.edu.tr

Department of Electrical Engineering, Kocaeli University, Kocaeli, Türkiye

[3] Ismet Kandilli
kandilli@kocaeli.edu.tr

Department of Electronics and Automation, Kocaeli University, Kocaeli, Türkiye

*Corresponding Author: Pamela K. F. Barson

1. Introduction

Hybrid renewable-energy micro-grids combine variable renewable generation with energy storage and, where required, dispatchable generation. Their operation therefore requires continuous coordination of renewable availability, time-varying demand, battery state of charge (SoC), converter and generator constraints, and supply reliability. Energy management systems (EMSs) perform this coordination by converting measurements, forecasts, and operating constraints into dispatch decisions [1, 7–9]. Rule-based EMSs remain attractive because of their interpretability and low computational requirements. However, purely reactive strategies respond to current conditions and may therefore provide limited anticipation of near-term changes in demand or renewable availability.

Machine-learning methods provide an opportunity to introduce predictive information into EMS decision-making. Random Forest (RF), for example, can represent nonlinear relationships without imposing a predefined functional form [17–19]. Unsupervised clustering methods such as K-Means can additionally identify recurring operating conditions and provide compact representations of system regimes [10, 11, 13]. However, the presence of predictive or clustering components does not by itself guarantee improved control performance. In time-dependent energy data, inappropriate preprocessing, feature construction, clustering, or model selection can expose the model to information from future observations and produce overly optimistic performance estimates. Furthermore, a forecasting improvement does not necessarily translate into an operational improvement when battery utilization, renewable curtailment, generator dependence, and terminal SoC are considered simultaneously.

Previous studies have investigated clustering for operating-condition classification, scenario reduction, and energy-management optimization in renewable-energy and micro grid applications [26, 27]. Machine-learning-based forecasting has also been widely investigated as decision support for renewable-energy systems [2–6, 12, 23]. Nevertheless, the methodological separation between causal prediction, train-only regime identification, EMS parameter selection, and blind chronological evaluation remains insufficiently emphasized. In particular, it is important to establish whether clustering genuinely improves prediction before assigning it a role in the controller, rather than assuming that additional information necessarily improves forecasting performance.

Figure 1 is a representative of a hybrid micro grid system that includes variable solar and wind generation, time-varying demand, energy-storage constraints, and dispatchable generation which create a need for coordinated EMS decisions. Machine-learning-based forecasting provides predictive support for managing these operating conditions.

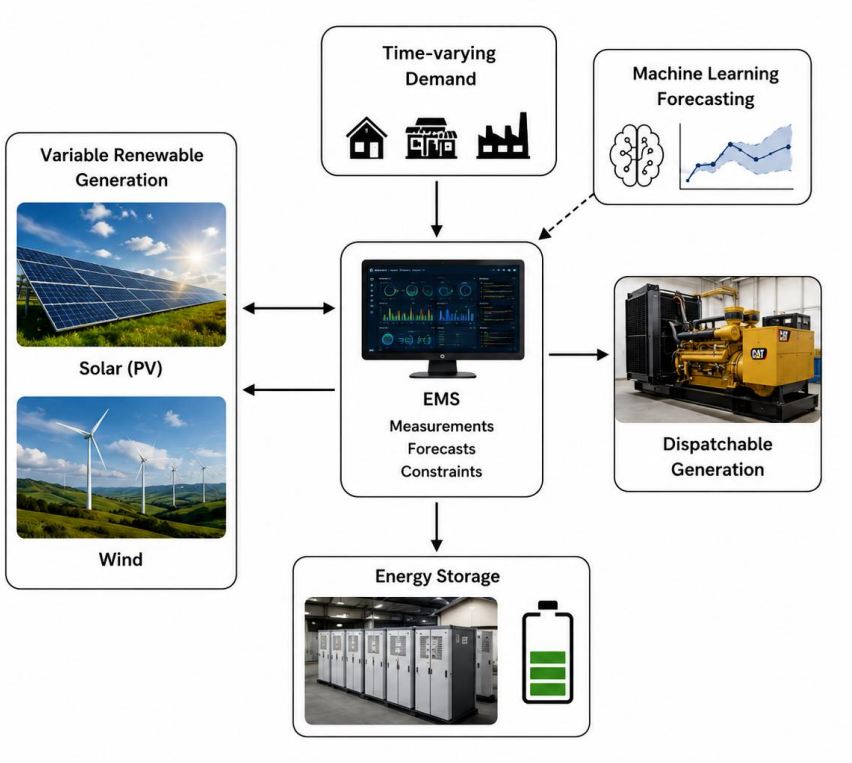


Figure 1. Conceptual motivation for predictive and regime-adaptive energy management in a hybrid renewable-energy micro-grid.

Accordingly, this study develops and evaluates a leakage-safe, regime-adaptive EMS for a solar–wind–battery microgrid. The first objective is to establish a causal and chronologically validated RF forecasting pipeline in which future information cannot influence model development. The second objective is to determine whether K-Means operating-regime information provides measurable predictive value and, if not, to assign clustering a justified operational role as an EMS regime selector. The third objective is to select EMS parameters using validation-only multi-objective analysis while maintaining reliability and terminal battery-state consistency. The final objective is to evaluate the frozen framework on an untouched blind-test period against reactive, prediction-only, and untuned regime-adaptive controllers.

The main contributions are therefore:

The novelty claimed here is methodological integration rather than novelty of RF, K-Means, or rule-based control individually. The framework assigns each component a separate, auditable role: RF supplies causal one-hour-ahead load information; K-Means, fitted only on design data, identifies operating regimes; validation data select the multi-objective controller; and an untouched chronological block evaluates the frozen system. This separation enables the operational value of prediction and clustering to be tested without allowing the blind period to influence development, and a multi-metric interpretation that emphasizes operational trade-offs rather than claiming universal superiority of one controller.

2. Related Work and Research Gap

2.1 Prediction and decision support in renewable-energy EMS

Forecasting, optimization, and control are commonly identified as complementary EMS functions [1]. Statistical, machine-learning, and ensemble models have been applied to renewable generation and demand prediction [2–6, 12, 23]. LR provides a transparent baseline, whereas RF captures nonlinear interactions and is robust to heterogeneous predictors [17–19]. Nevertheless, a forecast should be assessed against a simple persistence model and through chronological rather than randomly shuffled evaluation. In this work, prediction is not presented as an end in itself; it is evaluated as causal decision support for the following dispatch interval.

2.2 Clustering and adaptive control

Clustering can separate operating points that would otherwise be governed by one fixed policy [10, 11, 13]. In power-system studies, clusters are often interpreted as resource, demand, or storage regimes. Such labels are useful only if fitted without access to the evaluation period and if their control role is stated explicitly. Here, K-Means is fitted using design-period observations only. Direct cluster-label augmentation is first tested in the prediction benchmark. Because it does not improve RF, the labels are not used to claim a hybrid forecasting gain; instead, they select interpretable dispatch policies.

2.3 Multi-objective EMS evaluation

Battery dispatch, generator operation, renewable curtailment, and reliability are coupled objectives. Predictive or optimization-based EMS studies therefore require multiple operational indicators and comparable boundary conditions [8, 9, 20–22, 24, 25]. The research gap addressed here is the absence of a unified, leakage-safe protocol that (i) keeps clustering causal, (ii) separates model development from EMS tuning, (iii) freezes the controller before blind testing, and (iv) reports terminal SoC and low-SoC exposure alongside conventional energy balances.

3. Materials and Methods

3.1 Dataset and scope

The analysis uses the public KU-HMG hybrid-microgrid dataset [28]. The file contains 8,760 hourly observations from 1 January to 31 December 2007. Variables include photovoltaic output, wind-turbine output, generator output, served load, excess production, total renewable output, inverter and rectifier power, battery charge and discharge power, and battery SoC. The source data were generated in HOMER Pro and must therefore be interpreted as a simulation benchmark rather than field measurements, as shown in Table 1. The storage technology described by the source is a CELLCUBE flow battery; the present paper consequently uses the general term BESS and does not claim lithium-ion behavior. The renewable-energy identity was internally consistent to

within 9×10^{-5} kW, and no timestamp showed simultaneous charging and discharging. Figure 2 presents representative profiles and the seasonal SoC behavior.

Table 1. Dataset integrity audit.

Check	Result
Rows	8760
Start	2007-01-01 00:00:00
End	2007-12-31 23:00:00
Duplicate timestamps	0
Missing values	0
Non-hourly gaps	0
Negative solar	0
Negative wind	0
Negative load	0
SoC outside 0-100	0
Simultaneous charge/discharge	0
Maximum |renewable-(solar+wind)| kW	9.0×10^{-5} kW

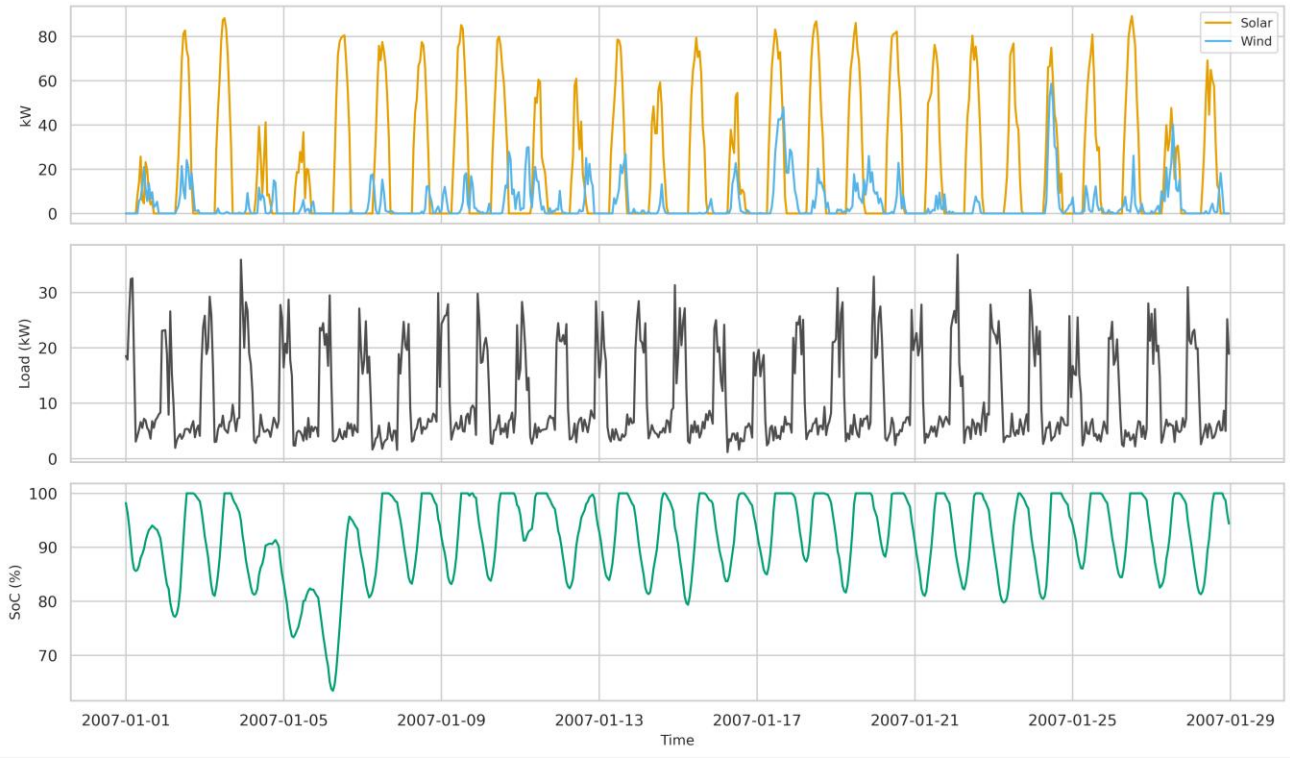


Figure 2. Representative dataset profiles for combined solar–wind generation, load, battery power, and SoC.

3.2 Leakage-safe chronological protocol

After causal lag and rolling-feature construction, the usable samples were divided in temporal order into three non-overlapping blocks. The design block supported prediction development, time-series cross-validation, and initial RF/K-Means fitting. The EMS-validation block was used exclusively to select controller parameters. The final block remained unopened until the controller was frozen and was then used once for blind-test reporting. No random temporal shuffling was applied. Figure 3 explains how causal telemetry is processed through independent prediction and clustering paths. The combined predicted load, operating regime, and SoC support dispatch; controller parameters are selected on validation data and frozen before blind testing (Table 2). The design period develops the predictor and regime model, the validation period selects EMS parameters, and the final period is retained as an untouched blind test, as indicated in figure 4.

Leakage-control audit. Target construction shifts served load by one hour before partitioning. Every lag and rolling statistic is backward-looking and is computed from values available at time t or earlier; no centered window, backward fill, or future observation is used. Rows made incomplete by lag construction are removed before evaluation. The dataset contains no missing observations, so no imputation is applied. Scaling for K-Means is fitted on the design period only and then applied unchanged to validation and blind-test observations. Prediction-model hyperparameters and model choice are determined within forward-chaining folds of the design period. The number of clusters, cluster centroids, engineering label mapping, objective definitions, controller grid, feasibility thresholds, and objective weights are fixed before the blind test. The validation block is used only for controller selection, and the blind block is not used for preprocessing, feature selection, normalization, hyperparameter tuning, cluster fitting, policy selection, or compromise selection.

Information timing. Immediately after measurements at hour t become available, the controller forms the t+1 load forecast from causal features, uses the current renewable measurement as the t+1 persistence forecast, assigns the current operating vector to the nearest frozen design-period centroid, and issues the dispatch command for interval t+1. Actual load and renewable generation at t+1 are revealed only after that command and are used solely to settle the simulated power balance and update SoC.

Table 2. Chronological design–validation–blind-test protocol.

Partition	Rows	Start	End
Design	5497	2007-01-08 00:00:00	2007-08-25 00:00:00
EMS validation	1375	2007-08-25 01:00:00	2007-10-21 07:00:00
Blind test	1719	2007-10-21 08:00:00	2007-12-31 22:00:00

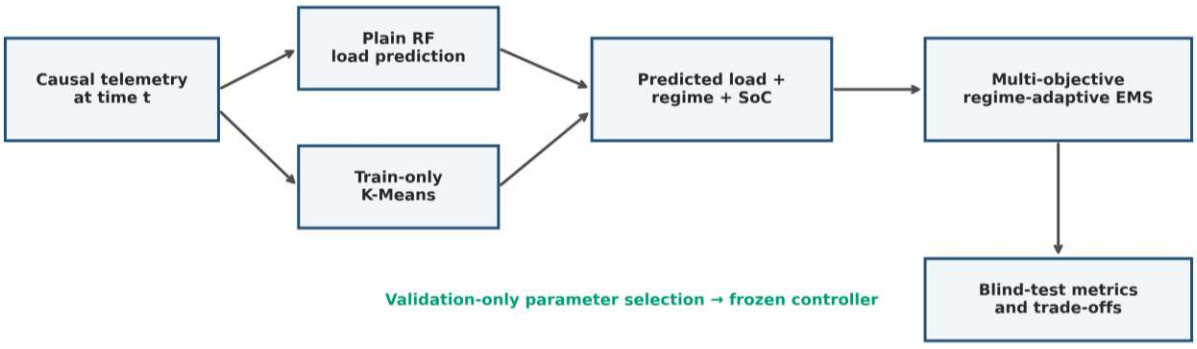


Figure 3. Proposed leakage-safe RF prediction and K-Means regime-adaptive multi-objective EMS.

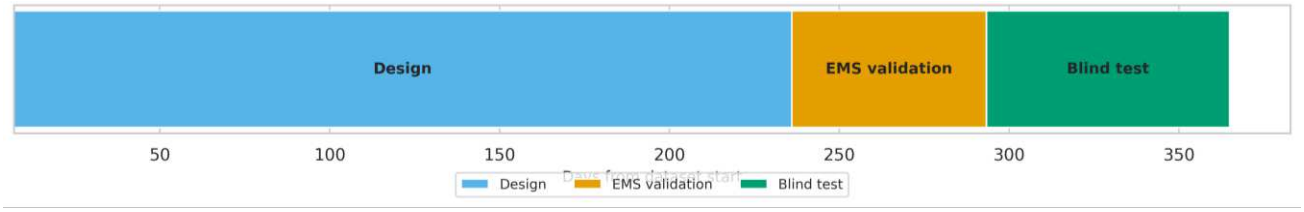


Figure 4. Chronological design–validation–blind-test protocol.

3.3 One-hour-ahead load prediction

Equation (1) defines the supervised learning task of predicting next-hour ($\hat{y}_{t+1}$) served load from information available at the present time or earlier.

$$\hat{y}_{t+1} = f(X_t, X_t^{-1}, \dots) \quad (1)$$

Predictors comprise current causal telemetry, cyclical hour/day-of-year terms, lags at 1, 2, 3, 24, and 168 h, and causal rolling summaries. Every fitted transformation is estimated on the corresponding training fold only. RF, raw LR, log-transformed LR, RF with K-Means augmentation (KM_RF), and log-LR with K-Means augmentation are compared with persistence. RF outputs are naturally nonnegative for the observed nonnegative target; negative predictions are audited explicitly for every candidate.

All candidate models use the same chronological folds and prediction target. The RF implementation uses a fixed random seed of 42; complete estimator settings, feature names, fold boundaries, preprocessing objects, and prediction outputs are retained in the accompanying executable notebook. Hyperparameter selection is restricted to the design-period folds. The persistence forecast is defined as the most recently observed load. Because the source table contains no missing values, no imputation model is fitted. Raw LR is allowed to expose negative outputs for diagnostic purposes, whereas nonnegative candidates are preferred for control integration.

Prediction error is quantified using root-mean-square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and RMSE skill relative to persistence, respectively represented in equations (2), (3), (4), and (5):

$$RMSE = \sqrt{\left(\frac{1}{n}\right)\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (2)$$

$$MAE = \left(\frac{1}{n}\right)\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (3)$$

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right] \quad (4)$$

$$Skill_p = 100\left[1 - \frac{RMSE_{model}}{RMSE_{persistence}}\right] \quad (5)$$

Where:

y_i: actual load

$\hat{y}$: predicted load

$\bar{y}$: mean of the actual load

n: number of evaluated observations

$Skill_p$: (positive skill) model outperforms persistence

3.4 Train-only K-Means regime identification

K-Means clustering partitions standardized operating vectors by minimizing the within-cluster sum of squared distances (calculated in equation (6)). Four clusters are fitted using the design period only, with random seed 42. The resulting centroids are interpreted using solar generation, wind generation, load, SoC, battery charging and discharging power, generator output, renewable generation, and net surplus, as illustrated in figure 5.

$$\min\sum_{k=1}^{k}\sum_{x_1 \in C_k}\Sigma_{-}\{x_i \in C_k\}\,||x_i - \mu_k||^2 \quad (6)$$

Where:

k: number of clusters

C_k: observations assigned to cluster k

μ_k: centroid of cluster k

$||x_i - \mu_k||^2$: squared distance between an observation and its centroid.

The resulting clusters represent deficit and renewable-surplus operating conditions with different magnitudes. Because K-Means augmentation does not improve RF prediction performance, the cluster labels are not used as

evidence of a forecasting improvement. Instead, they are used exclusively to select interpretable EMS policies during controller operation and are used only to select EMS rules, presented in Table 3. This distinction separates the predictive and operational roles of clustering and prevents the clustering component from being credited with a forecasting benefit that is not supported by the results.

Table 3. Design-period K-Means centroids and engineering policy labels.

Cluster	Policy	Solar (kW)	Wind (kW)	Load (kW)	SoC (%)	Net surplus (kW)
0	Deficit_Low_SoC	0.27	3.25	22.67	73.5	-19.14
1	Wind_Surplus	15.41	24.16	12.14	88.9	27.44
2	Solar_Surplus	50.44	18.72	11.19	73.8	57.98
3	Deficit_Low_SoC	0.62	1.21	29.84	33.5	-28.01

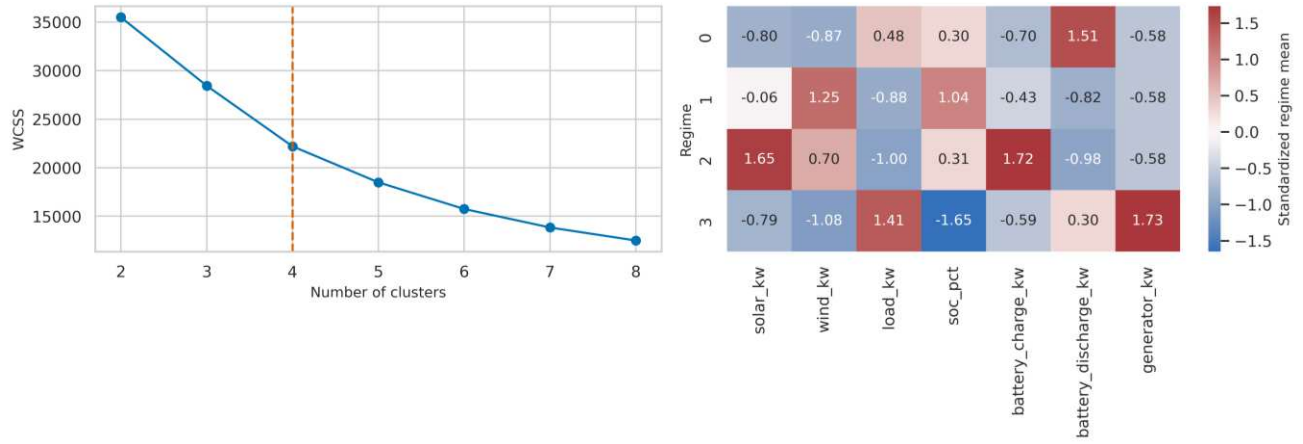


Figure 5. Training-only K-Means diagnostics and centroid heat map for the four operating clusters.

3.5 Regime-adaptive rule-based EMS

At each dispatch interval, the controller receives the predicted next-hour load, a persistence estimate of next-hour renewable power based on the current measurement, current SoC, and the train-only K-Means regime assignment. Actual next-hour load and renewable output are used only to settle the simulated energy balance after the decision. Positive battery power denotes discharge and negative power denotes charging in plots.

The battery update is represented by the simplified energy-balance in equation (7).

$$SoC_{t+1} = clip\{SoC_t + \frac{100\Delta t}{E_{cap}}[\eta_c P_{ch} - \frac{P_{dis},t}{\eta_d}], 0, 100\} \quad (7)$$

Where:

SoC_{t+1}, SoC_t: updated SoC after charging and discharging and battery SoC at the current time t, respectively.

P_{ch}, P_{dis}: battery charging and discharging, respectively.

$\eta_c = \eta_d = 0.92$ (charge and discharge efficiency, respectively).

E_{cap} = 1,200 kWh (effective battery capacity).

The capacity is an explicit scenario assumption inferred for consistent replay and is not a manufacturer-validated value for the number of CELLCUBE units. Charge, discharge, and generator limits are 119.545, 38.855, and 34.383 kW, respectively. A 72-h terminal recovery window targets 95% SoC, with a ±2-percentage-point tolerance. The initially considered ±1-point condition was infeasible on validation (closest error 1.559 points); the tolerance was therefore widened transparently before blind testing rather than forcing artificial battery dumping.

3.6 Validation-only multi-objective selection

Candidate controllers are evaluated on five lower-is-better objectives: generator energy, battery throughput, renewable curtailment, hours below the selected reserve SoC, and terminal-SoC error. Metrics are min–max normalized within the feasible validation candidates and combined as indicated in equation (8).

$$J = 0.30\bar{G} + 0.30\bar{B}_{thr} + 0.20\bar{C} + 0.15\bar{H}_{low} + 0.05\bar{E}_{terminal} \quad (8)$$

Where:

$\bar{G}$: normalized generator energy consumption

$\bar{B}_{thr}$: normalized battery energy throughput

$\bar{C}$: normalized curtailed renewable energy

$\bar{H}_{low}$: normalized number of hours below the reserve SoC

$\bar{E}_{terminal}$: normalized terminal-SoC error.

The objectives are min–max normalized among feasible validation candidates and combined using the predefined weighted score.

The controller feasibility screen requires LPSP ≤ 0.01% and terminal-SoC error ≤ 2 percentage points. The parameter grid contains 243 combinations, of which 81 satisfy both constraints. The executed notebook reveals that these 81 feasible candidates are operationally identical under the current controller logic: generator energy, battery throughput, curtailment, low-SoC exposure, and terminal-SoC error are constant; all normalized objective values and weighted scores equal zero; and every feasible candidate is classified as nondominated. Consequently, the weighted score cannot identify a unique compromise. The first feasible parameter set is retained only as a deterministic representative for blind-test replay. It must not be interpreted as evidence of successful parameter optimization.

The term multi-criteria evaluation refers to feasibility filtering and joint reporting of five operational objectives. Because the tested feasible parameter combinations do not change dispatch, the present notebook does not demonstrate a discriminative multi-objective optimization. A valid optimization claim requires decision variables and bounds that produce materially different dispatch trajectories before scalar weighting or NSGA-II comparison is applied.

For reproducible blind-test replay, the representative feasible parameter set is frozen before the blind-test period. Four controllers are compared: (i) reactive rule-based EMS; (ii) prediction-only rule-based EMS; (iii) untuned K-Means regime-adaptive EMS; and (iv) the feasibility-screened K-Means regime-adaptive candidate. The frozen values are listed in Table 4. Their presence does not imply that the validation grid discriminated among feasible policies.

Table 4. Validation-selected EMS parameters (frozen before blind testing).

Parameter	Selected value
Reserve SoC	10%
Low-SoC generator factor	0.5
High-SoC generator factor	0.0
Recovery fraction	0.1
Recovery horizon	12 h

Loss of Power Supply Probability (LPSP), System Supply Reliability (SSR), renewable energy utilization (REU), unserved energy, generator energy, battery throughput, curtailment, low-SoC hours, and terminal SoC are reported in equations (9), (10), and (11).

$$LPSP = 100\frac{E_{unserded}}{E_{load}} \tag{9}$$

$$SSR = 100 - LPSP \tag{10}$$

$$B_{thr} = \Sigma_t(P_{ch}, t + P_{dis}, t)\Delta t \tag{11}$$

4. Results

4.1 Prediction performance

Uncertainty and comparison are reported through the mean and standard deviation across five forward-chaining folds (Table 5), together with the held-out effect relative to persistence (Table 6). The RF-versus-KM_RF

difference is only 0.022 kW in blind-test RMSE and their cross-validation dispersions overlap; it is therefore interpreted as no meaningful clustering benefit, not as statistically significant RF superiority. RF shows a substantially larger practical improvement over persistence (45.68% blind-test RMSE skill), but the study does not attach a p-value to this single blind block. Confirmatory inference would require repeated rolling origins or block-resampling, which is identified as future work.

In figure 6, the prediction follows the principal demand variation, while residuals reveal the remaining interval-specific errors.

Table 5. Five-fold chronological cross-validation results.

Model	RMSE mean	RMSE SD	MAE mean	R^2 mean	Negative outputs
RF	4.391	0.390	3.282	0.642	0
KM_RF	4.400	0.377	3.293	0.640	0
LR_raw	6.157	1.306	4.651	0.301	82
Persistence	6.743	0.254	4.831	0.133	0
LR_log	7.122	1.395	5.149	-0.002	0
KM_LR_log	10.112	6.031	5.497	-1.530	0

Table 6. Blind-test next-hour load-prediction performance.

Model	RMSE (kW)	MAE (kW)	R^2	Negative outputs	Skill vs persistence (%)
RF	3.351	2.320	0.839	0	45.55
KM_RF	3.374	2.333	0.837	0	45.19
LR_raw	4.808	3.540	0.669	0	21.89
KM_LR_log	5.421	3.614	0.579	0	11.93
LR_log	5.637	3.843	0.544	0	8.42
Persistence	6.155	3.765	0.457	0	0.00

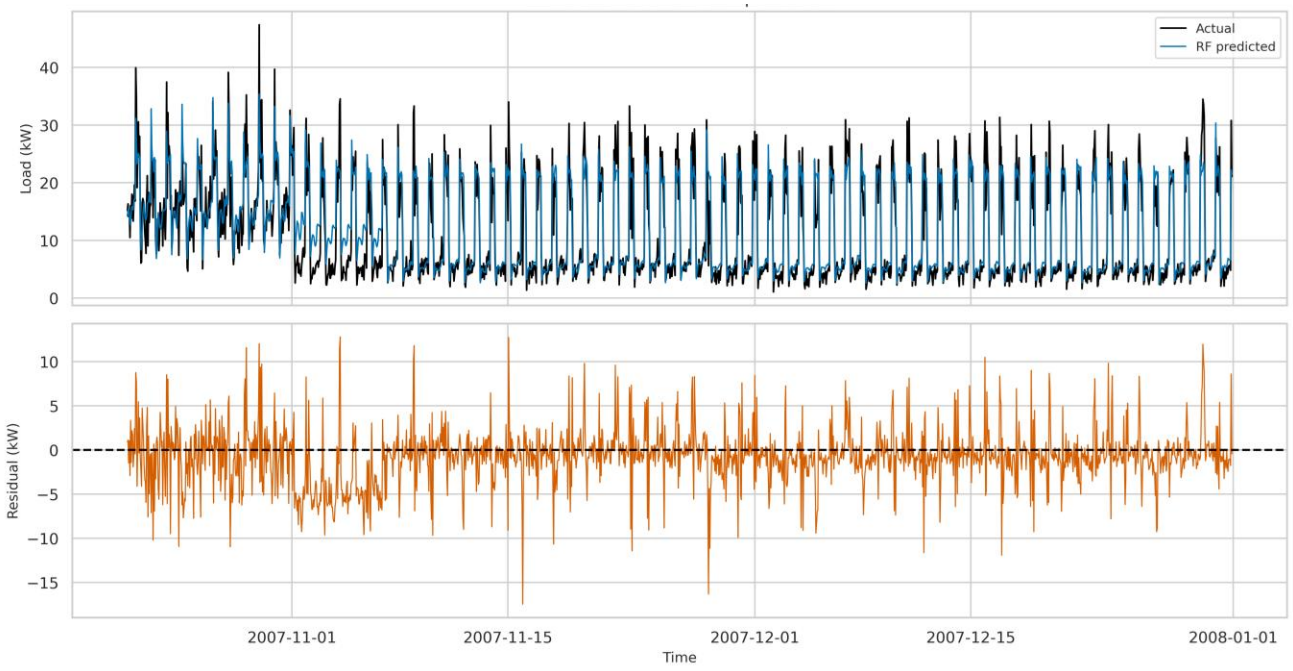


Figure 6. Blind-test RF load predictions and residual diagnostics.

4.2 Validation search and controller selection

The validation search retains 81 feasible parameter combinations. All satisfy the reliability and terminal-energy constraints, but no single candidate strictly dominates the others over all five objectives. Figure 7 visualizes the trade-off surface and the validation-selected operating point. The selected controller uses a 10% reserve, activates half of the available generator contribution in low-SoC deficit operation, applies no high-SoC generator contribution, and uses a 0.1 recovery fraction over 12 h.

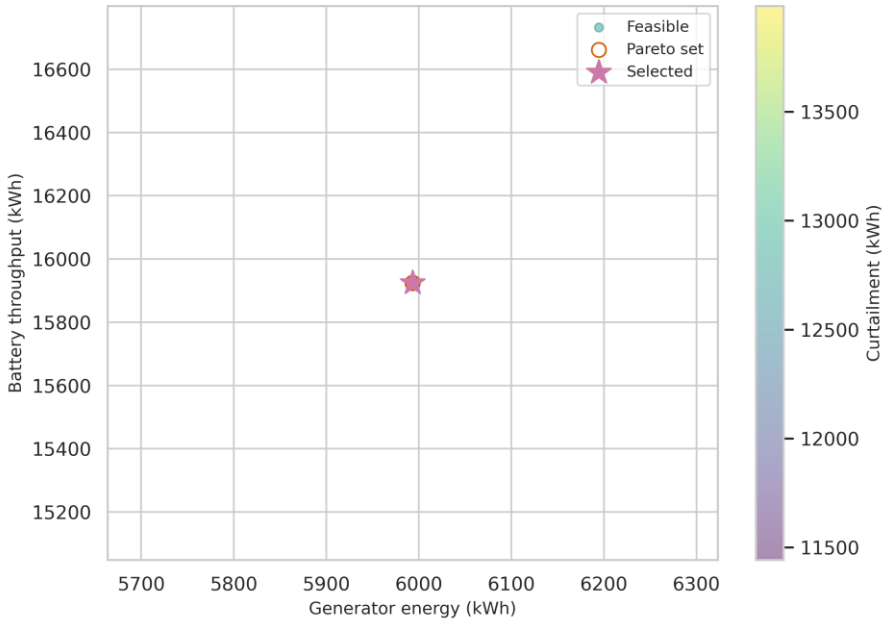


Figure 7. Validation-only multi-objective trade-offs. The highlighted point is selected using the prespecified weighted score after feasibility filtering.

4.3 Blind-test EMS comparison

The results in Table 7 show all controllers achieve zero unserved energy, LPSP = 0%, and SSR = 100%. REU is also identical (15.458%) because the controllers face the same large renewable surplus and system limits. Their internal resource allocation nevertheless differs markedly.

Relative to prediction-only control, the proposed multi-objective controller reduces generator energy from 12,466.72 to 5,068.11 kWh (59.35%), curtailment from 32,583.18 to 24,815.62 kWh (23.83%), and low-SoC exposure from 15 to 2 h. These gains require greater battery use: throughput increases from 5,181.21 to 17,669.36 kWh. Relative to the reactive controller, the proposed approach reduces battery throughput from 28,286.82 to 17,669.36 kWh (37.54%) and low-SoC exposure from 35 to 2 h, but generator energy rises by 4,941.64 kWh and curtailment rises by 5,813.26 kWh. Its terminal error is the smallest of all controllers at 0.431 percentage points.

Figure 8 and 9 consequently show that the proposed controller is not universally superior. Reactive control is generator-efficient but battery-intensive; prediction-only control is battery-conservative but generator- and curtailment-intensive; and the proposed controller occupies an intermediate region while providing the lowest low-SoC exposure and terminal error.

Figure 10 shows operating profiles of the hybrid renewable-energy microgrid showing load demand and renewable generation, generator and BESS charging/discharging power, and battery state of charge. The profiles illustrate the temporal mismatch between renewable generation and demand and the resulting need for coordinated storage and dispatchable generation.

Table 7. Untouched blind-test controller comparison.

Metric	Reactive	Prediction-only	Untuned regime	Multi-objective regime
LPSP (%)	0.000	0.000	0.000	0.000
SSR (%)	100.000	100.000	100.000	100.000
REU (%)	15.458	15.458	15.458	15.458
Generator energy (kWh)	126.472	12,466.721	10,041.850	5,068.111
Battery throughput (kWh)	28,286.821	5,181.208	9,296.312	17,669.364
Curtailment (kWh)	19,002.364	32,583.180	30,425.133	24,815.620
Low-SoC exposure (h)	35.000	15.000	7.000	2.000
Final SoC (%)	93.596	96.243	95.520	94.569
Terminal error (percentage points)	1.404	1.243	0.520	0.431

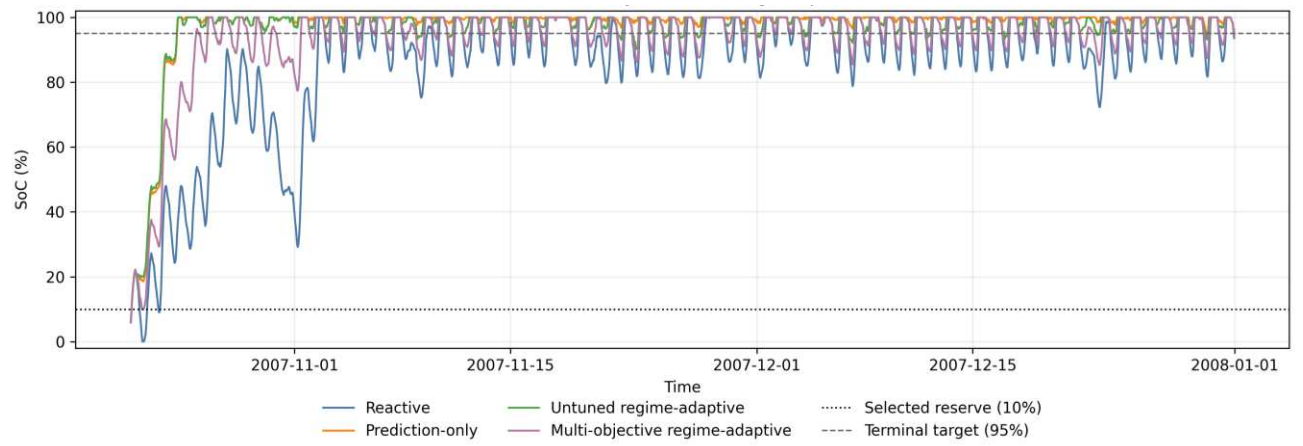


Figure 8. Blind-test SoC trajectories for all controllers. Similar terminal values do not imply similar battery histories; the multi-objective controller limits low-SoC exposure while meeting the terminal tolerance.

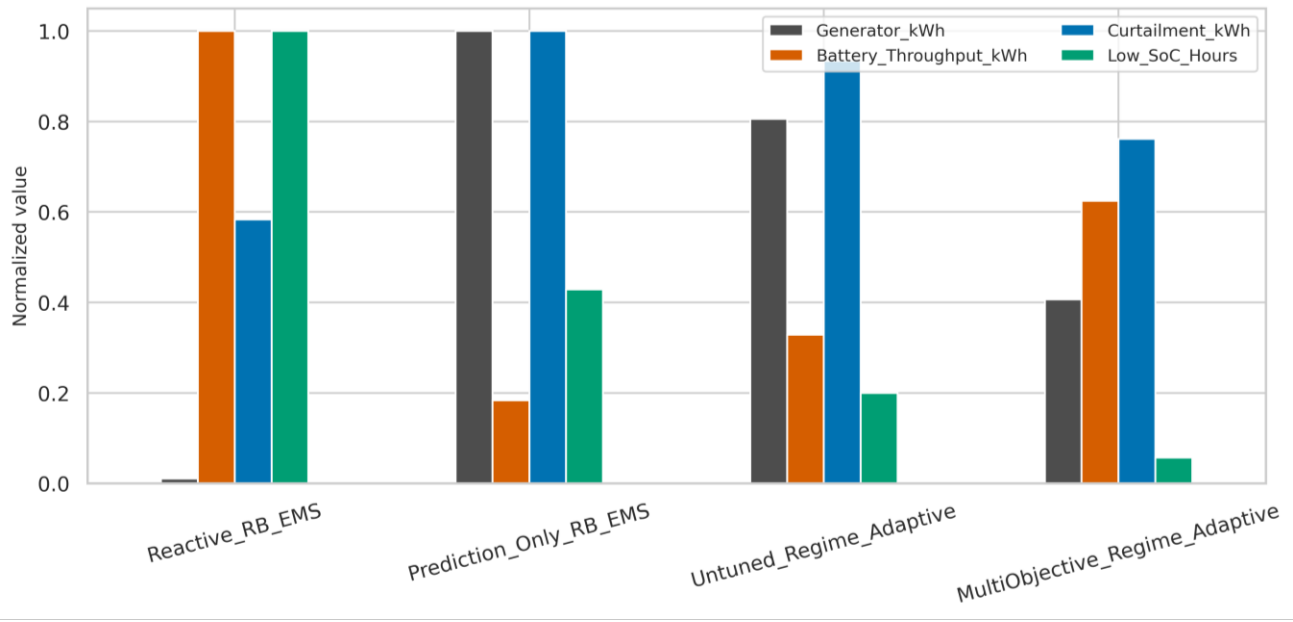


Figure 9. Blind-test operational trade-offs. Each metric is independently normalized by its maximum across controllers; bar heights must be compared between controllers within the same metric. Lower is preferable for all displayed metrics.

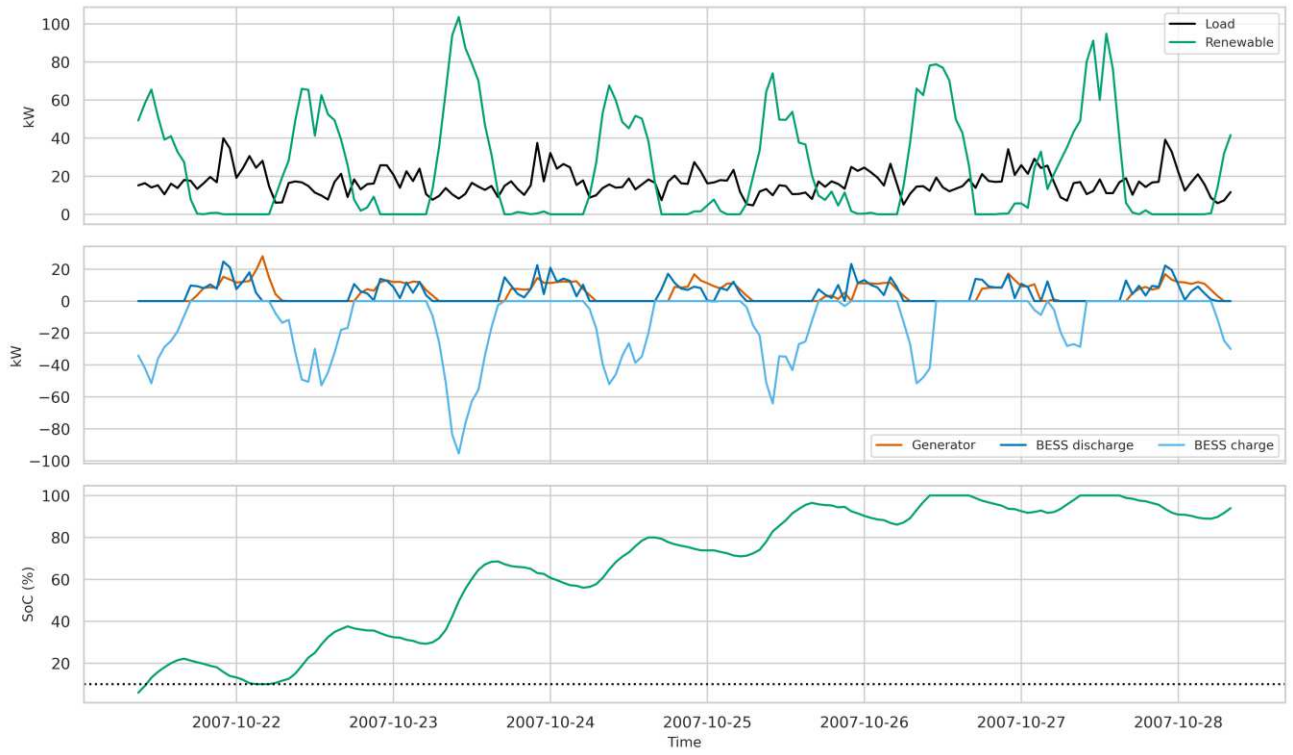


Figure 10. Representative one-week dispatch of the validation-selected controller during the blind test. The upper panel shows load and combined solar–wind generation; the middle panel shows generator output and BESS discharge/charge (charging plotted negative); and the lower panel shows SoC and the selected 10% reserve.

5. Discussion

The blind-test results demonstrate that EMS performance cannot be adequately characterized by a single operational metric. The reactive controller minimizes generator energy but does so through intensive battery utilization. Conversely, prediction-only control reduces battery throughput while increasing generator dependence and renewable curtailment. The proposed multi-objective regime-adaptive controller distributes the operational burden differently, reducing generator energy relative to prediction-only control while requiring less battery throughput than reactive control and maintaining the lowest low-SoC exposure.

Because all controllers achieve zero LPSP, reliability does not distinguish their blind-test performance. The principal differences therefore concern generator dependence, battery utilization, renewable curtailment, low-SoC exposure, and terminal-SoC consistency. The proposed controller should consequently be interpreted as a balanced operational compromise rather than a universally superior strategy.

K-Means contributes to this operational strategy by providing a compact representation of recurring operating conditions. Solar-surplus and wind-surplus regimes can support battery charging and reduce unnecessary generator intervention, whereas deficit regimes can justify protective generator support. However, these regimes are statistical representations rather than physical states derived from first-principles models. Their interpretation and stability may therefore change under different demand profiles, climates, component characteristics, or system configurations.

The methodological credibility of the results is strengthened by the explicit separation of model development, controller selection, and final evaluation. Prediction and clustering are developed without access to the blind-test period, EMS parameters are selected exclusively from chronological validation data, and the selected controller is frozen before blind testing. Nevertheless, the findings remain conditional on the dataset and simulation assumptions. Independent datasets, sensitivity analysis, multi-seed evaluation, and real-time or hardware-based validation are required before the proposed strategy can be considered for field deployment.

The selected chronology also limits the strength of the inference. The explicit dates make the experiment reproducible, but one design-validation-test boundary cannot establish that the ranking is invariant to season or split location. Accordingly, all percentage improvements in this paper refer only to the stated blind period. A submission-grade extension should repeat the complete development-and-freeze procedure over multiple forward rolling origins; merely moving the test boundary after observing results would compromise the blind-test claim.

6. Limitations and Future Work

The following limitations bound the interpretation:

- The study uses one HOMER Pro-simulated site and one historical year (2007); it does not establish generalization across sites, climates, years, or measured sensor noise.
- The source storage system is a CELLCUBE flow battery. Results should not be transferred directly to lithium-ion systems, supercapacitor HESSs, or other chemistries.
- The 1,200-kWh effective capacity and simplified efficiency/SoC equations are scenario assumptions. Temperature, voltage, internal resistance, converter dynamics, calendar aging, rain-flow cycle depth, and validated capacity fade are not modeled.
- Renewable power is represented by persistence, and the EMS uses point forecasts without uncertainty intervals or robust/stochastic decision rules.
- K = 4 and the cluster mapping are static. Cluster stability, alternate clustering methods, online adaptation, and distribution drift are not evaluated.
- The objective weights are judgment-based. Because all 81 feasible validation solutions are nondominated, weight and constraint sensitivity is necessary before deployment.
- Generator fuel cost, emissions, ramping, start-up cost, minimum up/down time, and maintenance are omitted; REU is also limited by the implemented accounting definition.
- Only one final chronological block and one random seed are reported. Multi-seed, rolling-origin, seasonal, and independent-dataset evaluation are required for stronger statistical evidence.
- The work is offline replay, not controller-in-the-loop or hardware-in-the-loop validation; computation latency, communications, sensor faults, and safety interlocks remain untested.

Priority future work should therefore examine battery-capacity and objective-weight sensitivity; predict solar, wind, and load separately with calibrated uncertainty; repeat optimization over multiple chronological origins and seeds; and validate the frozen controller on a measured microgrid or real-time digital-twin platform.

7. Conclusion

This study aimed to develop a leakage-safe, regime-adaptive multi-objective EMS for a solar–wind–battery microgrid and to determine whether combining causal next-hour load prediction with K-Means-based operating regimes could provide a more balanced operational strategy. The stated objectives were substantially achieved. The chronological design–validation–blind-test protocol prevented future information from entering prediction development, clustering, controller tuning, or final evaluation. Random Forest provided the strongest forecasting performance, achieving a blind-test RMSE of 3.351 kW, MAE of 2.320 kW, and R^2 of 0.839, corresponding to

45.55% RMSE skill over persistence. The explicit comparison also showed that K-Means augmentation did not improve forecasting accuracy; consequently, clustering was assigned the more defensible role of a train-only operating-regime selector within the EMS rather than being presented as a source of forecasting improvement.

The second objective, namely the development of a reliable and operationally balanced controller, was also achieved. EMS parameters were selected exclusively from the chronological validation block and frozen before blind testing. During the untouched blind test, the proposed multi-objective regime-adaptive controller supplied the complete load with LPSP = 0% and SSR = 100%, while achieving the lowest low-SoC exposure (2 h) and the smallest terminal-SoC error (0.431 percentage points) among the evaluated controllers. Relative to prediction-only control, it reduced generator energy by 59.35%, renewable curtailment by 23.83%, and low-SoC exposure from 15 to 2 h, although battery throughput increased. Relative to reactive control, it reduced battery throughput by 37.54% and low-SoC exposure from 35 to 2 h, at the cost of increased generator use and curtailment. These results confirm that the proposed framework achieves its intended goal of providing a transparent operational compromise rather than optimizing a single metric at the expense of other system requirements.

Overall, the findings support the central premise of the study: leakage-safe prediction, train-only operating-regime identification, and validation-based multi-objective controller selection can be integrated into a reproducible EMS framework while maintaining a clear separation between model development, controller tuning, and final evaluation. The results should nevertheless be interpreted within the limits of the single-year HOMER Pro simulation and the stated battery and control assumptions; independent datasets, sensitivity analysis, multi-seed chronological evaluation, and real-time or hardware-based validation are required before generalization to field operation.

Data availability: The source dataset is available from Mendeley Data at DOI 10.17632/x8v796pjsx.2

Processed chronological splits and result tables are generated by the accompanying notebook.

Code availability: The executable Jupyter notebook, exported tables, dispatch histories, metadata, and figures accompany this manuscript.

Reproducibility note: the archived release should include the dependency versions, RF parameter dictionary, random seeds, chronological row indices, K-Means scaler and centroids, complete 243-candidate validation table, frozen controller configuration, and blind-test dispatch history. These artifacts are necessary to reproduce the reported tables without reopening controller selection on the blind period.

Author contributions:

All authors contributed equally on this research.

Department of Electrical Engineering, Kocaeli University, Kocaeli, Türkiye
Pamela K. F. Barson
Department of Electrical Engineering, Kocaeli University, Kocaeli, Türkiye
Nurettin Abut
Department of Electronics and Automation, Kocaeli University, Kocaeli, Türkiye
İsmet Kandilli

Corresponding author: Pamela K. F. Barson

All authors reviewed and approved the final manuscript.

Funding: This research received no external funding.

Conflicts of interest: The authors declare no conflict of interest.

8. References

[1] Mischos, S., Dalagdi, E., Vrakas, D. Intelligent energy management systems: a review. Artificial Intelligence Review 56 (2023) 11635–11674. https://doi.org/10.1007/s10462-023-10441-3

[2] Tascikaraoglu, A., Uzunoglu, M. A review of combined approaches for prediction of short-term wind speed and power. Renewable and Sustainable Energy Reviews 34 (2014) 243–254. https://doi.org/10.1016/j.rser.2014.03.033

[3] Li, G., Shi, J. On comparing three artificial neural networks for wind speed forecasting. Applied Energy 87 (2010) 2313–2320. https://doi.org/10.1016/j.apenergy.2009.12.013

[4] Clifton, A., Kilcher, L., Lundquist, J.K., Fleming, P. Using machine learning to predict wind turbine power output. Environmental Research Letters 8 (2013) 024009. https://doi.org/10.1088/1748-9326/8/2/024009

[5] Bošnjaković, M., Martinović, M., Đokić, K. Application of artificial intelligence in wind power systems. Applied Sciences 15 (2025) 2443. https://doi.org/10.3390/app15052443

[6] Chander, B., Gopalakrishnan, K. Integration of artificial intelligence techniques for energy management. In: Sustainable Management of Electronic Waste. Wiley (2024). https://doi.org/10.1002/9781394166923.ch1

[7] Kanchev, H., Lu, D., Colas, F., Lazarov, V., Francois, B. Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications. IEEE Transactions on Industrial Electronics 58 (2011) 4583–4592. https://doi.org/10.1109/TIE.2011.2119451

[8] Sassi, A., Zaidi, N., Nasri, O., Ben Hadj Slama, J. Energy management of PV/wind/battery hybrid energy system based on batteries utilization optimization. GECS (2017) 1–7. https://doi.org/10.1109/GECS.2017.8066133

[9] Aghmadi, A., Hussein, H., Mohammed, O.A. Enhancing energy management system for a hybrid wind solar battery based standalone microgrid. EEEIC/I&CPS Europe (2023) 1–6. https://doi.org/10.1109/EEEIC/ICPSEurope57605.2023.10194769

[10] Miraftabzadeh, S.M., Colombo, C.G., Longo, M., Foiadelli, F. K-Means and alternative clustering methods in modern power systems. IEEE Access 11 (2023) 119596–119633. https://doi.org/10.1109/ACCESS.2023.3327640

[11] Arthur, D., Vassilvitskii, S. K-means++: The advantages of careful seeding. Proceedings of SODA (2007) 1027–1035.

[12] Forootan, M.M., Larki, I., Zahedi, R., Ahmadi, A. Machine learning and deep learning in energy systems: a review. Sustainability 14 (2022) 4832. https://doi.org/10.3390/su14084832

[13] Jain, A.K. Data clustering: 50 years beyond K-means. Pattern Recognition Letters 31 (2010) 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

[14] Riahy, G.H., Abedi, M. Short term wind speed forecasting for wind turbine applications using linear prediction method. Renewable Energy 33 (2008) 35–41. https://doi.org/10.1016/j.renene.2007.01.014

[15] Montgomery, D.C., Peck, E.A., Vining, G.G. Introduction to Linear Regression Analysis, 5th ed. Wiley, 2012.

[16] Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning, 2nd ed. Springer, 2009. https://doi.org/10.1007/978-0-387-84858-7

[17] Breiman, L. Random forests. Machine Learning 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324

[18] Iannace, G., Ciaburro, G., Trematerra, A. Wind turbine noise prediction using random forest regression. Machines 7 (2019) 69. https://doi.org/10.3390/machines7040069

[19] Shi, K., Qiao, Y., Zhao, W., Wang, Q., Liu, M., Lu, Z. An improved random forest model of short-term wind-power forecasting. Wind Energy 21 (2018) 1383–1394. https://doi.org/10.1002/we.2261

[20] Faria, J., Pombo, J., Calado, M.R., Mariano, S. Power management control strategy based on artificial neural networks for standalone PV applications with a hybrid energy storage system. Energies 12 (2019) 902. https://doi.org/10.3390/en12050902

[21] Aguilera-Gonzalez, A., Vechiu, I., Rodriguez, R.H.L., Bacha, S. MPC energy management system for a grid-connected renewable energy/battery hybrid power plant. ICRERA (2018) 738–743. https://doi.org/10.1109/ICRERA.2018.8566710

[22] Zhou, S., Chen, Z., Huang, D., Lin, T. Model prediction and rule-based energy management strategy for a plug-in hybrid electric vehicle with hybrid energy storage system. IEEE Transactions on Power Electronics 36 (2021) 5926–5940. https://doi.org/10.1109/TPEL.2020.3028154

[23] Lateko, A.A.H., Yang, H.-T., Huang, C.-M. Short-term PV power forecasting using a regression-based ensemble method. Energies 15 (2022) 4171. https://doi.org/10.3390/en15114171

[24] Lund, P.D., Lindgren, J., Mikkola, J., Salpakari, J. Review of energy system flexibility measures to enable high levels of variable renewable electricity. Renewable and Sustainable Energy Reviews 45 (2015) 785–807. https://doi.org/10.1016/j.rser.2015.01.057

[25] Luo, X., Wang, J., Dooner, M., Clarke, J. Overview of current development in electrical energy storage technologies and application potential in power system operation. Applied Energy 137 (2015) 511–536. https://doi.org/10.1016/j.apenergy.2014.09.081

[26] Peter Anuoluwapo Gbadega, Yanxia Sun, Olufunke Abolaji Balogun, Optimized energy management in Grid-Connected microgrids leveraging K-means clustering algorithm and Artificial Neural network models, Energy Conversion and Management, Volume 336, 2025, 119868, ISSN 0196-8904, https://doi.org/10.1016/j.enconman.2025.119868

[27] Mathur, A., Kumari, R., Meena, V.P. et al. Data-driven optimization for microgrid control under distributed energy resource variability. Sci Rep 14, 10806 (2024). https://doi.org/10.1038/s41598-024-58767-4

[28] KU-HMG hybrid microgrid dataset, Version 2. Mendeley Data. https://doi.org/10.17632/x8v796pjsx.2