



AI-Based Energy Management for Renewable-Integrated Smart Grids using Load Forecasting and Battery Optimization

Abdullah Al Fathah

Sales Engineer, Forgent Power Solution, Waco, Texas, USA

Publication History: Received: 26.08.2026; Revised: 24.09.2026; Accepted: 29.09.2026; Published: 09.10.2026.

ABSTRACT: Renewable-integrated smart grids and commercial microgrids need operating strategies that keep photovoltaic (PV) surplus, time-of-use (TOU) prices and peak-demand charges under control. This paper develops and evaluates a forecast-driven energy management framework that combines (i) day-ahead load and PV forecasting, (ii) a linear-programming battery scheduler with incremental demand-charge accounting, hourly re-optimisation and calibrated uncertainty margins, and (iii) a real-time peak guard that protects execution from forecast errors. It is tested on one year of public hourly hospital-load data and NREL TMY3 irradiance processed through a physical PV model (1.95 MWp PV, 1 MW / 2 MWh battery) under an explicit TOU-plus-demand-charge tariff, using rolling-origin evaluation over 214 days. An ensemble of Ridge regression, random forest, LightGBM and a multilayer perceptron lowers day-ahead load RMSE by 19.8% relative to a seasonal-naïve benchmark (51.9 vs 64.7 kW; MAPE 2.61%), whereas the learned PV forecast improves on smart persistence by only 1.6%. The AI-driven scheduler reduces operating cost by 4.52% (\$26,243) and mean monthly peak import by 12.0% relative to PV-only operation, capturing 60.9% of a perfect-foresight bound (7.42%); a persistence-driven scheduler attains a comparable 4.98%, so forecast accuracy gains do not translate one-for-one into savings. Removing the real-time guard turns the saving into a 2.34% loss, and error-scaling experiments show that PV forecast error, rather than load-forecast error, governs the value of information.

KEYWORDS: smart grid; energy management system; load forecasting; battery energy storage; photovoltaic; peak shaving; model predictive control; machine learning

I. INTRODUCTION

Decarbonisation of power systems is increasing the share of variable photovoltaic generation, which complicates balancing and raises the cost of peak capacity (Borghini et al., 2021). Battery energy storage systems (BESS) can time-shift PV surplus, arbitrage time-varying energy prices and shave the demand peaks that dominate many commercial electricity bills (Almassalkhi et al., 2026), which is particularly valuable for facilities that must maintain service continuity, such as hospitals (Mustafa et al., 2021).

Operating a battery is a sequential decision problem under uncertainty: the controller must decide when to charge and discharge before load and PV output are known (Rosewater et al., 2019). Research on forecasting (tree ensembles, neural networks) (Beichter et al., 2022) and on scheduling (mathematical programming, model predictive control (MPC), reinforcement learning) has largely progressed in isolation, and forecasts are usually judged by statistical metrics although their economic value depends on tariff structure, battery size and the control law (Hong et al., 2020). Three gaps follow. First, the link between forecast accuracy and operating cost is rarely quantified in controlled experiments (Schmitt et al., 2021). Second, the execution layer that converts a day-ahead plan into real-time set-points is seldom studied, although forecast errors can turn an optimal plan into a costly one under demand charges (Langenmayr et al., 2020). Third, many studies rely on proprietary data and unspecified tariffs, which prevents replication (Mariano-Hernández et al., 2020).

This paper therefore contributes (i) an integrated, reproducible forecast–optimise–execute framework, evaluated end-to-end on open data (Pimm et al., 2019); (ii) an incremental demand-charge formulation with a real-time guard and calibrated, conformal-style uncertainty margins (Davarzani et al., 2019); and (iii) a controlled value-of-forecast study, based on rolling-origin evaluation, a perfect-foresight bound and error-scaling experiments, that also reports negative



results. Section 2 reviews related work, Sections 3–4 describe the data and methods, Section 5 reports results, and Sections 6–7 discuss and conclude.

II. RELATED WORK

Short-term load forecasting is dominated by calendar effects and lagged information, and carefully engineered features and ensembles are hard to beat (Vilmarest & Goude, 2022); for aggregated, regular loads such as institutional buildings, seasonal-naïve benchmarks are demanding and must be included when claiming improvements (Wood et al., 2023). Day-ahead PV forecasting is dominated by cloud uncertainty: skill relative to persistence must be demonstrated, and without weather or satellite inputs, forecasts based on the clear-sky index of preceding days offer only limited skill (Voyant et al., 2021). For storage scheduling, MPC re-solves the optimisation with updated forecasts and measurements, providing feedback against errors (Nair et al., 2020), whereas reinforcement learning avoids explicit forecasts but needs extensive training and careful safety handling. Few studies trace how forecast error propagates to realised cost under demand charges on a transparent open benchmark (Zhang et al., 2023), which is the focus here.

III. DATA AND SYSTEM DESCRIPTION

3.1 Load data

We use the open hourly load series of a hospital in San Francisco for 2015 (Neural Prophet repository): 8,760 consecutive hourly values without gaps, with a mean of 1,012 kW, a minimum of 716 kW and a maximum of 1,389 kW, treated as hourly mean power (Ghimire et al., 2024). Timestamps were relabelled to hour-beginning, so that hour h covers $h:00$ – $h+1:00$. The load shows a pronounced weekday–weekend pattern.

3.2 Solar resource and PV model

Irradiance and temperature come from the NREL TMY3 data set for Greensboro–Piedmont Triad International Airport, NC, shipped with the open-source library pvlib (Meiers & Frey, 2023). PV output is simulated with the pvlib Model Chain (PVWatts DC and AC models, fixed 30° tilt facing south, DC/AC ratio 1.2, temperature coefficient $-0.4\%/^\circ\text{C}$, default system losses $\approx 14\%$). The array is sized so that annual PV energy equals roughly 30% of annual load, giving 1.95 MWp, an annual yield of 2,639 MWh (1,353 kWh/kWp) and a peak AC output of 1,560 kW. Clear-sky PV output, used as a forecasting input, is simulated with the Ineichen–Perez model on the same geometry (Ledesma et al., 2020).

3.3 Tariff and battery assumptions

Tariff and battery parameters are scenario assumptions representative of commercial TOU structures and lithium-ion systems, not taken from a specific utility or product; they are varied in Section 5.6 (Rosewater et al., 2019). Table 1 lists all assumptions. Federal holidays are treated as weekends; calendar ageing, self-discharge, thermal limits and inverter losses are not modelled.

Table.1. Scenario parameters (assumptions, not measured data)

Parameter	Value	Comment
PV array	1.95 MWp, 30° tilt, south	pvlib PVWatts chain
Battery power / energy	1,000 kW / 2,000 kWh	2-hour system
Charge / discharge efficiency	95% / 95% (RTE 90.25%)	Constant
SOC window; initial; end-of-day	10–90%; 50%; $\geq 50\%$	Usable 1,600 kWh
Degradation cost	\$0.04 per kWh discharged	Throughput proxy
Off-peak price	\$0.09/kWh	Nights; weekends, holidays
Mid-peak price	\$0.15/kWh	Weekdays 08–16 h, 21–23 h
Peak price	\$0.25/kWh	Weekdays 16–21 h
Export compensation	\$0.04/kWh	No export limit
Demand charge	\$12 per kW-month	Max hourly net import



3.4 Evaluation protocol

All forecasts follow a rolling-origin design: for each target month from March to December, models are retrained using only data preceding that month and then forecast every hour of the month at the day-ahead origin (00:00). March–May is a development period (used to estimate the error-persistence parameters and compare forecasters); June–December (214 days, 5,136 h) is the test period on which all results are computed. Economic outcomes are not annualised because the period excludes January–May.

IV. METHODOLOGY

4.1 Framework overview

Fig. 1 shows the three layers of the proposed energy management system. The data layer provides load history, PV-model inputs and hourly telemetry. The intelligence layer produces day-ahead load and PV forecasts and calibrates their errors. The decision layer solves a day-ahead LP (and, in one variant, re-solves it hourly), after which a real-time guard converts the plan into the executed battery set-point, which updates the next hour's measurements (Talluri et al., 2021).

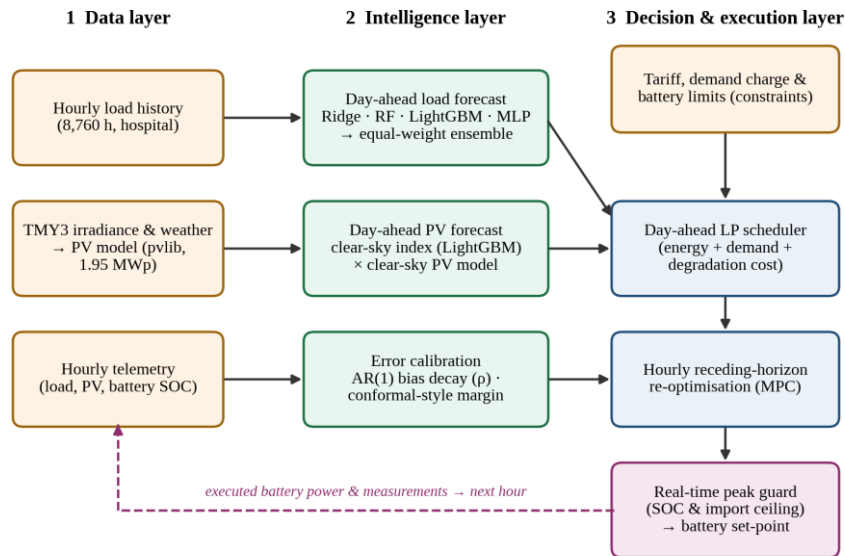


Fig.1. Proposed forecast–optimise–execute energy management framework

4.2 Day-ahead load forecasting

At 00:00 of each day a single direct model with hour of day as a feature predicts the 24 hours of that day, using only inputs available at the forecast origin: calendar variables (hour, day of week, weekend and holiday flags), loads 24, 48, 168 and 336 h earlier, same-hour mean and standard deviation over the previous 7 days, and recent level indicators. Four learners are compared: ridge regression, a random forest (300 trees), LightGBM (500 trees, learning rate 0.03, 15 leaves) and a multilayer perceptron (two hidden layers of 64 and 32 ReLU units) (Dalal et al., 2021). Their equal-weight mean forms the ensemble forecast (Eq. 1). Benchmarks are the 24-hour naïve forecast and the seasonal-naïve forecast that repeats the load of the same hour one week earlier. Hyperparameters were fixed a priori, not tuned on the test period, and chosen conservatively because training sets are small (about 2,000–7,000 samples) (Groß et al., 2021).

$$\hat{L}_t = \frac{1}{4} \sum_{m \in \{\text{Ridge, RF, LightGBM, MLP}\}} f_m(\mathbf{x}_t) \quad (1)$$

4

3 Day-ahead PV forecasting

PV power is forecast through the clear-sky index (Eq. 2). A LightGBM regressor (300 trees, 7 leaves) is trained on the index for daylight hours (clear-sky power > 20 kW) using hour, day of year, the daily mean index of the previous 1, 3



and 7 days, the index at the same hour on the previous day and the clear-sky power (Nejla et al., 2025). The forecast is the product of the predicted index and the clear-sky power (Eq. 3). Benchmarks are smart persistence (yesterday's daily mean index times today's clear-sky power), 24-hour persistence and a LightGBM regressing power directly. No numerical weather prediction is used, so forecast skill is inherently limited (Markovics & Mayer, 2022).

$$\kappa_t = P_t^{pv} / P_t^{cs}, \quad P_t^{cs} > 20 \text{ kW} \quad (2)$$

$$\hat{P}_t^{pv} = \hat{\kappa}_t P_t^{cs} \quad (3)$$

4.4 Error persistence and uncertainty margins

Forecast errors are positively autocorrelated, so at each re-optimisation the remaining horizon is corrected by a geometrically decaying fraction of the latest error (Eqs. 4–5), with decay rates equal to the lag-1 autocorrelations of development-period errors ($\rho_L = 0.84$, $\rho_P = 0.82$). For risk-aware scheduling, a margin m_h is added to the load in the demand-charge constraint (Eq. 6): it is the 90th percentile of past out-of-sample load errors for hour of day h , using only earlier months, in the spirit of split conformal prediction (mean margin 41.4 kW). The resulting upper bound covers 87.9% of test hours (nominal 90%), versus 80.3% for a LightGBM 0.9-quantile model, i.e. raw quantile regression was under-dispersed on this small sample (Jensen et al., 2022).

$$\tilde{L}_{t+j} = \hat{L}_{t+j} + \rho_L^{j+1} (L_{t-1} - \hat{L}_{t-1}) \quad (4)$$

$$\tilde{P}_{t+j}^{pv} = \max\{0, \hat{P}_{t+j}^{pv} + \rho_P^{j+1} (P_{t-1}^{pv} - \hat{P}_{t-1}^{pv})\} \quad (5)$$

$$m_h = \max\{0, Q_{0.9}(L_\tau - \hat{L}_\tau : \text{hour}(\tau) = h, \tau \in \mathcal{H}_m)\} \quad (6)$$

4.5 Battery scheduling model

Given forecasts over a horizon of H hours, the scheduler minimises energy cost, battery degradation cost and the incremental demand charge (Eq. 7) subject to power balance (Eq. 8), battery dynamics, SOC limits and an end-of-horizon SOC target (Eq. 9) and power limits (Eq. 10). Here π is the TOU price, π^x the export price, c_d the degradation cost, g^+ and g^- grid import and export, P^{ch} and P^{dis} battery charge and discharge power, S the stored energy, η the one-way efficiency and z a non-negative peak variable (Alhasnawi et al., 2024).

$$\min \sum_{k=0}^{H-1} [\pi_{t+k} g_k^+ - \pi^x g_k^- + c_d P_k^{\text{dis}}] + \lambda z \quad (7)$$

$$g_k^+ - g_k^- = \tilde{L}_{t+k} - \tilde{P}_{t+k}^{pv} + P_k^{\text{ch}} - P_k^{\text{dis}} \quad (8)$$

$$S_{k+1} = S_k + \eta P_k^{\text{ch}} - P_k^{\text{dis}} / \eta, \quad S_{\min} \leq S_k \leq S_{\max}, \quad S_H \geq S^{\text{end}} \quad (9)$$

$$0 \leq P_k^{\text{ch}}, P_k^{\text{dis}} \leq \bar{P}, \quad g_k^+, g_k^- \geq 0 \quad (10)$$

Demand charges are billed on the monthly maximum, which couples days within a month. We therefore use an incremental (ratchet) formulation: only import above the realised month-to-date peak P^{ref} is charged at the full rate λ , so the increments sum to the final peak (Eq. 11) (Chalendar et al., 2019). The margin m is zero for the standard schedulers. A simpler alternative that amortises the demand charge uniformly over the days of the month is evaluated as an ablation. The problem is a linear program with at most $4H+1$ variables solved with HiGHS through SciPy in about 10 ms per solve (Wang et al., 2019).

$$g_k^+ - g_k^- + m_{t+k} - z \leq P_m^{\text{ref}}, \quad z \geq 0 \quad (11)$$



4.6 Real-time peak guard

Plans are built on forecasts, so following them blindly can create costly import peaks, for example when the plan charges against a forecast PV surplus that does not materialise. The guard executes the planned set-points subject to the actual state: charging is limited to the room below the planned import ceiling C (planned peak plus P^{ref}) given the actual pre-battery net load $b = L - P^{pv}$, extra discharge is dispatched whenever b exceeds C , and actual SOC and power limits are always respected (Eqs. 12–13), with $[\cdot]^+ = \max(\cdot, 0)$ (Maheshwari et al., 2020).

$$P^{ch,x} = \min\{P^{ch}, [C - b]^+, \bar{P}, (S_{max} - S)/\eta\} \quad (12)$$

$$P^{dis,x} = \min\{\max(P^{dis}, [b - C]^+), \bar{P}, \eta(S - S_{min})\} \quad (13)$$

4.7 Compared strategies and metrics

Seven strategies are compared on identical actual data. **S0** is PV-only operation without a battery. **S1** is a forecast-free rule-based TOU controller (charge from PV surplus, discharge in peak-price hours or when net load exceeds 1,200 kW, charge from the grid at night only while import stays below 1,000 kW; thresholds not tuned). **S2** is the day-ahead LP driven by persistence forecasts (seasonal-naïve load, smart-persistence PV). **S3** is the same LP driven by the AI forecasts. **S4** re-solves the LP every hour with a shrinking horizon, bias-corrected with Eqs. 4–5. **S5** is S3 with the calibrated load margin of Eq. 6. **S6** is a perfect-foresight monthly LP that provides an upper bound. S2–S5 use the guard of Section 4.6 and the incremental demand formulation; the initial SOC is 50% and each day (month for S6) must end at $\geq 50\%$.

Forecasts are evaluated by MAE, RMSE, MAPE and RMSE normalised by mean load (PV: by the 1,560 kW peak AC output). Operating cost is energy cost plus degradation cost plus the demand charge on the monthly maximum of hourly net import, computed on actual data; battery use is expressed as equivalent full cycles (EFC). Forecast differences are tested with the Diebold–Mariano (DM) test (Newey–West, six lags) and paired cost differences with a 5,000-sample bootstrap over days. Oracle-gap capture is a strategy’s saving divided by that of S6.

V. RESULTS

5.1 Load forecasting accuracy

Table 2 and Fig. 2a summarise load forecasting on the test period. The seasonal-naïve benchmark is strong (MAPE 3.02%) because the load repeats weekly; the 24-hour naïve forecast is much worse (7.20%). The ensemble has the lowest MAE and RMSE (51.9 kW, nRMSE 5.15%, MAPE 2.61%), a 19.8% RMSE reduction relative to seasonal-naïve. The DM test indicates that this improvement is significant for squared-error loss ($p = 0.018$) and borderline for absolute-error loss ($p = 0.058$). The ensemble is significantly better than Ridge and the MLP, but not distinguishable from the random forest ($p = 0.21$) or LightGBM ($p = 0.10$).

Table.2. Day-ahead load forecast accuracy on the test period (Jun–Dec 2015, 5,136 h).

Model	MAE (kW)	RMSE (kW)	MAPE (%)	nRMSE (%)	Bias (kW)	RMSE Mar–May	DM p vs. ens.
Naïve (24 h)	75.1	134.2	7.20	13.32	+0.9	128.3	<0.001
Seasonal naïve (168 h)	30.3	64.7	3.02	6.42	-0.5	46.2	0.018
Ridge	29.7	56.2	2.94	5.58	+0.6	47.0	0.039
Random forest	26.6	57.1	2.59	5.67	-3.7	54.3	0.210
LightGBM	29.4	57.4	2.90	5.70	-4.9	63.7	0.103
MLP (64-32)	32.1	58.2	3.19	5.77	-0.2	69.4	0.009
Ensemble (proposed)	26.4	51.9	2.61	5.15	-2.0	50.9	–

DM p: Diebold–Mariano test of equal squared-error loss against the ensemble; small p means the model is significantly worse. The advantage is regime-dependent: in the development period (March–May) the seasonal-naïve forecast was actually more accurate (RMSE 46.2 vs 50.9 kW), and in December the neural and boosted models degrade. With a



single year of data, models cannot fully learn annual seasonality or holiday effects. LightGBM's feature importance is dominated by the weekly lags (168 h and 336 h), consistent with the periodic load structure.

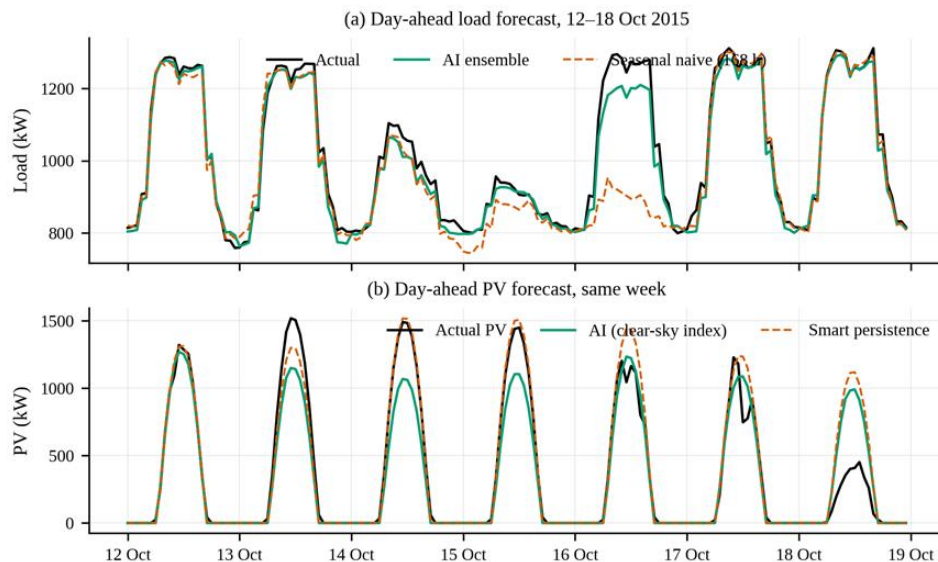


Fig.2. Day-ahead forecasts versus actual values for 12–18 October 2015: (a) load, (b) PV generation.

5.2 PV forecasting accuracy

PV forecasts (Table 3) are far less accurate than load forecasts because cloud cover cannot be anticipated from previous days' clear-sky indices alone. Regressing power directly produced a biased forecast (+42.5 kW), whereas predicting the clear-sky index removed the bias and gave the lowest RMSE (217.4 kW, nRMSE 13.9%), only 1.6% better than smart persistence (220.8 kW); its MAE (106.8 kW) is actually higher than that of smart persistence (100.7 kW). The learned PV forecaster therefore offers little skill beyond a strong persistence reference, in line with the literature's emphasis on exogenous weather information (Voyant & Notton, 2018).

Table.3. Day-ahead PV forecast accuracy on the test period (all hours; normalised by the 1,560 kW peak AC output).

Model	MAE (kW)	RMSE (kW)	nMAE (%)	nRMSE (%)	Bias (kW)
Persistence (24 h)	105.8	237.9	6.78	15.25	+1.4
Smart persistence (kt)	100.7	220.8	6.45	14.16	+1.4
LightGBM on power	160.3	235.1	10.27	15.07	+42.5
LightGBM on clear-sky index	106.8	217.4	6.85	13.94	-8.5

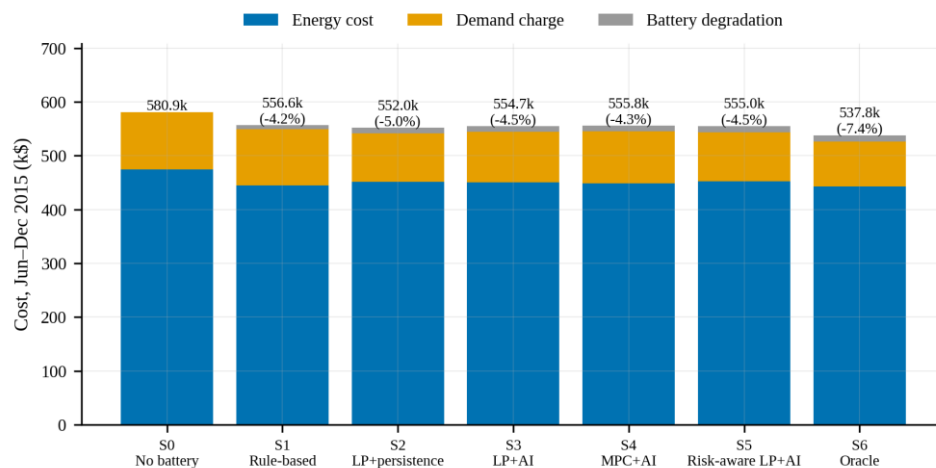
5.3 Economic performance of the battery strategies

Table 4 and Fig. 3 give the cost of the strategies over the 214 test days. PV-only operation costs \$580,919. The AI-driven day-ahead scheduler S3 reduces the cost by 4.52% (\$26,243), lowers the mean monthly peak import from 1,263 kW to 1,112 kW (12.0%) and captures 60.9% of the perfect-foresight saving (7.42%, peak reduced by 21.4%). The rule-based controller S1 saves 4.19%, almost entirely through energy cost: its monthly peak falls only 2.2% because it has no peak-aware logic (Fig. 4). The persistence-driven scheduler S2 achieves 4.98% (67.2% of the bound), slightly more than S3.


Table.4. Operating cost, peak demand and battery use, Jun–Dec 2015 (214 days). Costs in US\$.

Strategy	Energy	Demand	Degr.	Total	Saving (%)	Mean peak (kW)	EFC	Gap capt. (%)
S0 No battery	474,799	106,120	0	580,919	—	1,263	0	—
S1 Rule-based TOU	444,939	103,803	7,844	556,585	4.19	1,236	123	56.5
S2 LP + persistence	450,943	90,206	10,815	551,963	4.98	1,074	169	67.2
S3 LP + AI forecasts	450,748	93,431	10,497	554,676	4.52	1,112	164	60.9
S4 MPC + AI forecasts	449,013	96,408	10,374	555,796	4.32	1,148	162	58.3
S5 Risk-aware LP + AI	452,410	91,438	11,150	554,999	4.46	1,089	174	60.2
S6 Oracle (bound)	442,340	83,395	12,102	537,837	7.42	993	189	100.0

EFC: equivalent full cycles. Mean peak: average of the seven monthly peak net imports. Gap captured: saving relative to that of the oracle S6.


Fig.3. Cost composition of the strategies over the test period (percentages relative to S0).

Paired bootstrap results (daily energy plus degradation cost) clarify these differences. S3 reduces energy-related cost relative to PV-only operation by \$13,554 (95% CI \$11,660 to \$15,491). Relative to S2 the difference is \$513 (95% CI –449 to 1,471), i.e. indistinguishable from zero. S3 is more expensive than the rule-based S1 on energy-related cost but cheaper on demand charges (\$93,431 vs \$103,803). The S2–S3 ranking is therefore decided by demand charges, which depend on only seven monthly peaks and are a high-variance outcome; we do not read it as evidence that persistence forecasts are better in general.

The hourly MPC variant S4 (4.32%) did not outperform the day-ahead plan S3, indicating that, with the real-time guard in place, re-optimisation with error-persistence correction adds little on this system. The risk-aware variant S5 (4.46%) cut the demand charge by about \$2.0k relative to S3 but spent about \$2.3k more on energy and degradation, leaving the total essentially unchanged.

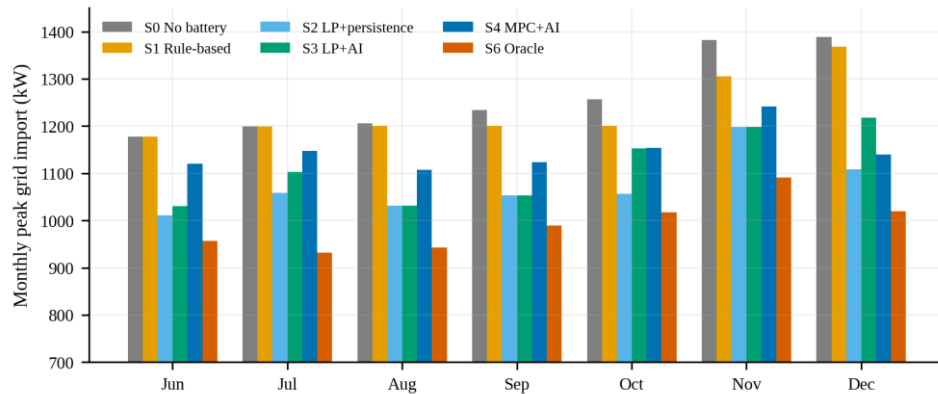


Fig.4 Monthly peak grid import by strategy

5.4 Dispatch behaviour

Fig. 5 illustrates two representative days. On the high-PV August weekday (Fig. 5a) the AI scheduler charges at night and from the midday PV surplus, then discharges almost fully during the 16:00–21:00 peak-price window, flattening the evening import. On the high-load November weekday (Fig. 5b) PV is negligible and the battery trims the afternoon and evening net-load peak while recharging off-peak. The perfect-foresight schedule uses a different SOC trajectory because it can pre-position energy for the exact peak hours of the month.

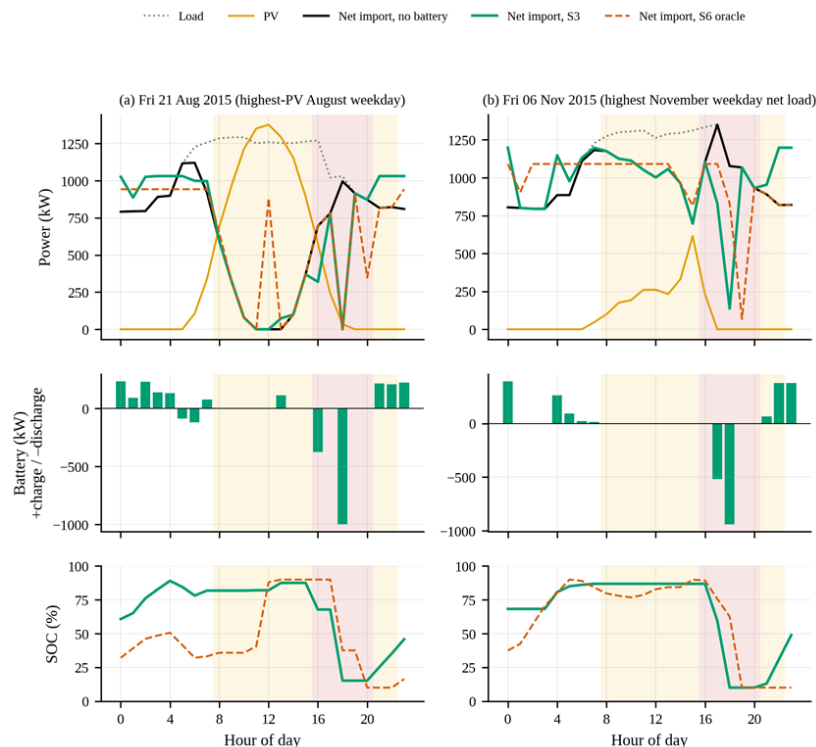


Fig.5. Example dispatch under S3 (green) compared with no battery (black) and the oracle S6 (dashed orange). Shaded bands denote mid-peak (yellow) and peak (red) price periods. (a) Highest-PV August weekday; (b) highest-net-load November weekday



5.5 Ablation of design elements

Table 5 shows that the real-time guard is the most consequential element. Without it, the S3 plan produces a monthly peak of up to 1,977 kW (the load alone never exceeds 1,389 kW) and the 4.52% saving becomes a 2.34% loss; the persistence-driven plan behaves similarly (−1.33%). The failure mechanism is charging against forecast PV surplus or forecast-flat net load that does not materialise, so that charging power adds directly to a high real net load. Replacing the incremental demand formulation by the daily-amortised variant lowers the saving to 3.14%, because plans that treat the demand charge as a small daily cost allow higher peaks. Extending the margin to include PV error is also counter-productive (2.93%) because the large midday PV uncertainty (margins of about 600 kW) distorts the plan.

Table.5. Ablation results (test period).

Variant	Total cost (\$)	Saving (%)	Demand cost (\$)	Max peak (kW)
S3 (reference)	554,676	+4.52	93,431	1,217
S3 without real-time guard	594,495	-2.34	139,924	1,977
S3 with daily-amortised demand charge	562,689	+3.14	99,659	1,382
S2 without real-time guard	588,649	-1.33	133,714	1,866
S5 margin on net load (load+PV error)	563,896	+2.93	99,479	1,429

5.6 Sensitivity analysis

Battery capacity. Table 6 varies the battery energy from 0.5 to 6 MWh at a fixed 2-hour duration. Savings of S3 rise from 1.97% to 6.72% with diminishing returns: the saving per installed kWh falls from \$22.8 to \$6.5, and cycling falls from 203 to 99 equivalent full cycles. The gap to the perfect-foresight bound widens with capacity (1.6 percentage points at 0.5 MWh versus 5.3 points at 6 MWh), so larger batteries make forecast and control quality more important. S2 and S3 stay within about 0.5 percentage points of each other at all sizes.

Table.6. Sensitivity to battery size (power = energy / 2): cost saving relative to S0 (%), cycling and saving per kWh.

Energy (MWh) / power (MW)	S1	S2	S3	S6	EFC (S3)	S3 saving (\$/kWh)
0.5 / 0.25	1.64	2.09	1.97	3.61	203	22.8
1.0 / 0.50	3.06	3.06	3.20	5.35	189	18.6
2.0 / 1.00	4.19	4.98	4.52	7.42	164	13.1
3.0 / 1.50	5.26	5.59	5.65	8.88	137	10.9
4.0 / 2.00	5.92	6.04	6.36	10.17	120	9.2
6.0 / 3.00	6.24	7.14	6.72	12.01	99	6.5

Demand charge and peak price. The value of optimisation depends on the tariff (Fig. 6b–c). As the demand charge rises from \$0 to \$24 per kW-month, the saving of the forecast-free rule S1 falls from 4.64% to 3.88%, whereas LP-based schedulers gain (S3: 4.93% → 5.67%; S2: 4.95% → 6.53%) and the oracle bound rises from 5.94% to 9.85%. At a low demand charge of \$6 the rule-based controller (4.39%) beats S2 (3.98%) and S3 (3.77%); when peaks matter little, a simple TOU rule is hard to beat, and forecast errors can make an optimiser worse than the rule. All strategies profit from a higher peak price.

5.7 Value of forecast accuracy

To isolate the effect of forecast quality, forecasts were constructed as $\text{actual} + \alpha \times (\text{AI forecast error})$, with α scaling the observed errors of the load and PV forecasts ($\alpha = 1$ reproduces S3, $\alpha = 0$ is perfect foresight within the day-ahead scheduler). Table 7 and Fig. 6a show that a perfect day-ahead forecast would lower the cost from \$554,676 to \$548,634 (5.56% saving), while doubling both errors raises it to \$568,373 and tripling to \$581,732, almost back to PV-only operation. Load and PV errors play very different roles: scaling only the load error from $\alpha = 1$ to 2 leaves the cost virtually unchanged (\$554,676 → \$554,161), whereas scaling only the PV error raises it to \$564,507. In this system the



value of forecast information is therefore determined mainly by PV forecast quality, and the sophistication of the load forecaster is of secondary economic importance once it reaches the level of a good seasonal benchmark.

Table.7. Total cost (US\$, test period) when the observed forecast errors are scaled by α . RMSE columns refer to the “both” experiment.

α	Load (kW)	RMSE	PV (kW)	RMSE	Cost: both (\$)	Cost: load only (\$)	Cost: PV only (\$)
0	0.0		0.0		548,634	552,394	551,371
0.5	26.0		108.7		552,015	552,125	553,544
1	51.9		217.4		554,676	554,676	554,676
1.5	77.9		322.9		559,151	552,386	558,364
2	103.8		417.6		568,373	554,161	564,507
3	155.7		580.5		581,732	564,797	569,089

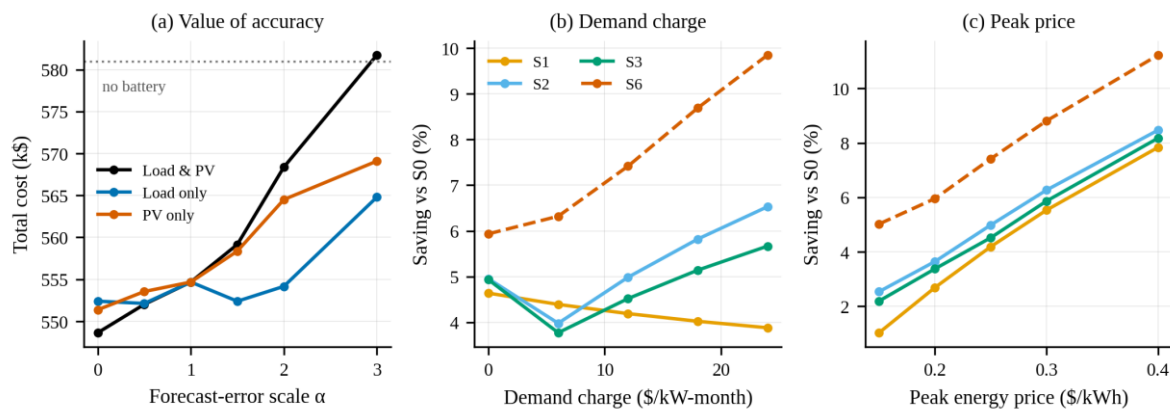


Fig.6. Sensitivity analyses: (a) total cost versus forecast-error scale α for load and PV errors scaled jointly or separately; (b) saving versus demand charge; (c) saving versus peak energy price. S6 is the perfect-foresight bound.

VI. DISCUSSION

6.1 Principal findings

- **Forecast accuracy is necessary but not sufficient.** The load ensemble improves RMSE over the seasonal benchmark by 19.8%, but this does not yield a lower cost than a persistence-driven scheduler (4.52% vs 4.98%). The system is limited more by information no day-ahead model has (cloud cover, the timing of the monthly peak) than by load-forecast error.
- **Execution robustness dominates.** A plan that is optimal for the forecast can be harmful in reality (−2.34% without the guard). Safeguards that cap charging at the planned import ceiling are cheap and should be a baseline requirement of any forecast-driven EMS.
- **The tariff determines the value of optimisation.** Where the demand charge is small a simple TOU rule is competitive; where it is large, peak-aware optimisation pays off. Comparisons of optimisers against rules should therefore be tariff-specific.
- **Bigger batteries raise the stakes.** The gap to the perfect-foresight bound grows with capacity, so forecasting and control quality matter more for larger installations.

6.2 Practical implications and limitations

Facility operators should allocate effort in this order: (i) a robust real-time execution layer; (ii) an incremental, ratchet-aware demand-charge treatment; (iii) improved solar forecasting through numerical weather prediction, satellite or sky-imager information; and only then (iv) more sophisticated load models. Limitations are the single year of data, the non-



co-located load and irradiance sources, the assumed tariff and battery parameters, and the simple throughput-based degradation model. Future work should use co-located multi-year measurements with weather covariates, compare stochastic and reinforcement-learning controllers on the same benchmark, and extend the framework to feeder-level coordination.

VII. CONCLUSIONS

This paper presented an integrated forecasting-and-optimisation framework for battery energy management in a renewable-integrated facility and evaluated it transparently on open data. An ensemble of four learners reduced day-ahead load RMSE by 19.8% against a strong seasonal benchmark, while a learned PV forecaster brought only marginal gains over smart persistence. The AI-driven scheduler reduced operating cost by 4.52% and the monthly peak by 12.0%, capturing 60.9% of the perfect-foresight potential, yet it did not outperform a persistence-driven scheduler. Ablation and sensitivity analyses showed that the real-time guard, the demand-charge formulation, the tariff structure and above all the PV forecast error determine realised value. The findings caution against equating statistical forecast improvements with economic ones and provide a reproducible benchmark for future work.

REFERENCES

1. Alhasnawi, B. N., Jasim, B. H., Alhasnawi, A. N., Hussain, F. F. K., Homod, R. Z., Hasan, H. A., Khalaf, O. I., Abbassi, R., Bazooyar, B., Zanker, M., Bureš, V., & Sedhom, B. E. (2024). A novel efficient energy optimization in smart urban buildings based on optimal demand side management. *Energy Strategy Reviews*, 54, 101461. <https://doi.org/10.1016/j.esr.2024.101461>
2. Almassalkhi, M., Madsen, H., Dvorkin, Y., Oral, H. G., Gayme, D. F., Stoustrup, J., Junker, R. G., Ritschel, T., Tohidi, S., Ebrahimi, R., Hodge, B.-M., Hamilton, D., & Mathiesen, B. V. (2026). Cross-Atlantic research agenda for scalable grid architectures and distributed flexibility. *Smart Energy*, 22, 100236. <https://doi.org/10.1016/j.segy.2026.100236>
3. Beichter, M., Phipps, K., Frysztański, M., Mikut, R., Hagenmeyer, V., & Ludwig, N. (2022). Net load forecasting using different aggregation levels. *Energy Informatics*, 5. <https://doi.org/10.1186/s42162-022-00213-8>
4. Borghini, E., Giannetti, C., Flynn, J., & Todeschini, G. (2021). Data-Driven Energy Storage Scheduling to Minimise Peak Demand on Distribution Systems with PV Generation. *Energies*, 14(12), 3453. <https://doi.org/10.3390/en14123453>
5. Chalendar, J. A. de, Glynn, P. W., & Benson, S. M. (2019). City-scale decarbonization experiments with integrated energy systems. *Energy & Environmental Science*, 12(5), 1695–1707. <https://doi.org/10.1039/c8ee03706j>
6. Dalal, N., Mølnå, M., Herrem, M., Røen, M., & Gundersen, O. E. (2021). Day-ahead forecasting of losses in the distribution network. *AI Magazine*, 42(2), 38–49. <https://doi.org/10.1609/aimag.v42i2.15097>
7. Davarzani, S., Granell, R., Taylor, G., & Pisica, I. (2019). Implementation of a novel multi-agent system for demand response management in low-voltage distribution networks. *Applied Energy*, 253, 113516. <https://doi.org/10.1016/j.apenergy.2019.113516>
8. Ghimire, S., Deo, R. C., Pourmousavi, S. A., Casillas-Pérez, D., & Salcedo-Sanz, S. (2024). Point-based and probabilistic electricity demand prediction with a Neural Facebook Prophet and Kernel Density Estimation model. *Engineering Applications of Artificial Intelligence*, 135, 108702. <https://doi.org/10.1016/j.engappai.2024.108702>
9. Groß, A., Lenders, A., Schwenker, F., Braun, D., & Fischer, D. S. (2021). Comparison of short-term electrical load forecasting methods for different building types. *Energy Informatics*, 4. <https://doi.org/10.1186/s42162-021-00172-6>
10. Hong, T., Pinson, P., Wang, Y., Weron, R., Yang, D., & Zareipour, H. (2020). Energy Forecasting: A Review and Outlook. *IEEE Open Access Journal of Power and Energy*, 7, 376–388. <https://doi.org/10.1109/oajpe.2020.3029979>
11. Jensen, V., Bianchi, F. M., & Anfinson, S. N. (2022). Ensemble Conformalized Quantile Regression for Probabilistic Time Series Forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, 35(7), 9014–9025. <https://doi.org/10.1109/tnnls.2022.3217694>
12. Langenmayr, U., Wang, W., & Jochem, P. (2020). Unit commitment of photovoltaic-battery systems: An advanced approach considering uncertainties from load, electric vehicles, and photovoltaic. *Applied Energy*, 280, 115972. <https://doi.org/10.1016/j.apenergy.2020.115972>



13. Ledesma, J. R., Almeida, R. H., Martínez-Moreno, F., Rossa, C. H., Martín-Rueda, J., Narvarte, L., & Lorenzo, E. (2020). A simulation model of the irradiation and energy yield of large bifacial photovoltaic plants. *Solar Energy*, 206, 522–538. <https://doi.org/10.1016/j.solener.2020.05.108>
14. Maheshwari, A., Heleno, M., & Ludkovski, M. (2020). The effect of rate design on power distribution reliability considering adoption of distributed energy resources. *Applied Energy*, 268, 114964. <https://doi.org/10.1016/j.apenergy.2020.114964>
15. Mariano-Hernández, D., Hernández-Callejo, L., Zorita-Lamadrid, Á., Duque-Pérez, Ó., & García, F. S. (2020). A review of strategies for building energy management system: Model predictive control, demand side management, optimization, and fault detect & diagnosis. *University of Valladolid Documentary Repository (University of Valladolid)*, 33, 101692. <https://doi.org/10.1016/j.job.2020.101692>
16. Markovics, D., & Mayer, M. J. (2022). Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. *Renewable and Sustainable Energy Reviews*, 161, 112364. <https://doi.org/10.1016/j.rser.2022.112364>
17. Meiers, J., & Frey, G. (2023). A Case Study of the Use of Smart EV Charging for Peak Shaving in Local Area Grids. *Energies*, 17(1), 47. <https://doi.org/10.3390/en17010047>
18. Mustafa, M. B., Keatley, P., Huang, Y., Agbonaye, O., Ademulegun, O. O., & Hewitt, N. (2021). Evaluation of a battery energy storage system in hospitals for arbitrage and ancillary services. *Journal of Energy Storage*, 43, 103183. <https://doi.org/10.1016/j.est.2021.103183>
19. Nair, U. R., Sandelić, M., Sangwongwanich, A., Dragičević, T., Costa-Castelló, R., & Blaabjerg, F. (2020). Grid Congestion Mitigation and Battery Degradation Minimisation Using Model Predictive Control in PV-Based Microgrid. *IEEE Transactions on Energy Conversion*, 36(2), 1500–1509. <https://doi.org/10.1109/tec.2020.3032534>
20. Nejla, E., Saidani, K., & Besbes, M. (2025). Overcoming the integration bottleneck: a global review of renewable energy and grid adaptation strategies. *International Journal of Sustainable Energy*, 44(1). <https://doi.org/10.1080/14786451.2025.2569922>
21. Pimm, A., Palczewski, J., Morris, R., Cockerill, T., & Taylor, P. (2019). Community energy storage: A case study in the UK using a linear programming method. *Energy Conversion and Management*, 205, 112388. <https://doi.org/10.1016/j.enconman.2019.112388>
22. Rosewater, D., Baldick, R., & Santoso, S. (2019). Risk-Averse Model Predictive Control Design for Battery Energy Storage Systems. *IEEE Transactions on Smart Grid*, 11(3), 2014–2022. <https://doi.org/10.1109/tsg.2019.2946130>
23. Rosewater, D., Copp, D. A., Nguyen, T. A., Byrne, R. H., & Santoso, S. (2019). Battery Energy Storage Models for Optimal Control. *IEEE Access*, 7, 178357–178391. <https://doi.org/10.1109/access.2019.2957698>
24. Schmitt, T., Rodemann, T., & Adamy, J. (2021). The Cost of Photovoltaic Forecasting Errors in Microgrid Control with Peak Pricing. *Energies*, 14(9), 2569. <https://doi.org/10.3390/en14092569>
25. Talluri, G., Lozito, G. M., Grasso, F., Garcia, C. I., & Luchetta, A. (2021). Optimal Battery Energy Storage System Scheduling within Renewable Energy Communities. *Energies*, 14(24), 8480. <https://doi.org/10.3390/en14248480>
26. Vilmarest, J. de, & Goude, Y. (2022). State-Space Models for Online Post-Covid Electricity Load Forecasting Competition. *IEEE Open Access Journal of Power and Energy*, 9, 192–201. <https://doi.org/10.1109/oajpe.2022.3141883>
27. Voyant, C., Lauret, P., Notton, G., Duchaud, J.-L., Foulloy, A., David, M., Yaseen, Z. M., & Soubdhan, T. (2021). A Monte Carlo based solar radiation forecastability estimation. *Journal of Renewable and Sustainable Energy*, 13(2). <https://doi.org/10.1063/5.0042710>
28. Voyant, C., & Notton, G. (2018). Solar irradiation nowcasting by stochastic persistence: A new parsimonious, simple and efficient forecasting tool. *Renewable and Sustainable Energy Reviews*, 92, 343–352. <https://doi.org/10.1016/j.rser.2018.04.116>
29. Wang, Z., Jochem, P., & Fichtner, W. (2019). A scenario-based stochastic optimization model for charging scheduling of electric vehicles under uncertainties of vehicle availability and charging demand. *Journal of Cleaner Production*, 254, 119886. <https://doi.org/10.1016/j.jclepro.2019.119886>
30. Wood, M. J., Ogliari, E., Nespoli, A., Simpkins, T., & Leva, S. (2023). Day Ahead Electric Load Forecast: A Comprehensive LSTM-EMD Methodology and Several Diverse Case Studies. *Forecasting*, 5(1), 297–314. <https://doi.org/10.3390/forecast5010016>
31. Zhang, J., Wang, Y., Jiang, B., He, H., Huang, S., Wang, C., Zhang, Y., Han, X., Guo, D., He, G., & Ouyang, M. (2023). Realistic fault detection of li-ion battery via dynamical deep learning. *Nature Communications*, 14(1), 5940. <https://doi.org/10.1038/s41467-023-41226-5>