energies

MDPI

Article

Operational Value of Load Forecasting for Battery Dispatch and Energy Sharing in Renewable Energy Communities

Muhammad Akram [1], Amy Fahy [2], Chiara Martone [1], Fabiano Pallonetto [2], Ilenia Perugini [1], Maurizio Sasso [1,*] and Sania Scellato [3]

[1] Department of Engineering, University of Sannio, 82100 Benevento, Italy; m.akram2@studenti.unisannio.it (M.A.); cmartone@unisannio.it (C.M.); i.perugini@studenti.unisannio.it (I.P.)
[2] International Research on Energy System Integration, Education & Environment for Sustainability and Innovation Centre (IRESI) , School of Business, Maynooth University, W23 HW31 Maynooth, Ireland; amy.fahy@mu.ie (A.F.); fabiano.pallonetto@mu.ie (F.P.)
[3] KOALA SRL SB, 82030 Apollosa, Italy; sania.scellato@koala.it
* Correspondence: sasso@unisannio.it

Abstract

Residential batteries in energy communities usually run local rule-based controllers whose community-level effect is rarely checked against measured data. This study uses one year of smart-meter, PV and battery records from a real 20-household Irish community (24 kWp PV, 200 kWh storage) to test how much can be achieved with locally available data only. The existing controllers charged all batteries in the same night-tariff hours. With household flows netted hourly at the community level, this raised grid import by 14.5% and CO_2 emissions by 15.8% compared with no batteries, raised the maximum hourly import from 33 to 83 kW, and created a new morning peak. A lightweight framework, combining a drift-adaptive LightGBM load forecast and a mixed-integer dispatch with real-time correction, reduced grid import by 10.5–13.5%, emissions by 11.6–14.6%, and energy imported above a 14.1 kW reference threshold by 87% relative to the measured operation, with each scenario computed in seconds. A forecast-free controller acting on the same threshold reduced this exceedance at least as much, while the forecast-based optimisation lowered the import by a further 2.6%. LightGBM was 10.9% more accurate than a seven-day average, yet its dispatch did not differ significantly (Wilcoxon $p = 0.53$); correcting the systematic underestimation of daily peaks reduced the energy above the threshold by 14%. Sharing PV reduced imports by 9.7% and benefited mainly households without PV, whereas pooled storage mainly relieved the grid, with the effect increasing with better information.

Keywords: renewable energy community; battery dispatch; day-ahead load forecasting; mixed-integer linear programming; energy sharing; forecast value; concept drift; self-sufficiency; CO_2 emissions

Check for updates

Academic Editor: Mihaela Simionescu

Received: 15 October 2026
Revised: 2 November 2026
Accepted: 3 November 2026
Published: 5 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Decarbonising urban and community energy systems is a priority of the global energy transition [1]. Cities and residential communities consume a large and rising share of electricity, and their supply increasingly comes from decentralised renewables, mainly rooftop photovoltaics (PV) with residential battery storage. A system built on a few controllable central plants is being replaced by many small, loosely coordinated prosumers, making local operation harder to manage and increasing the need for data-driven tools to model,

forecast, and control community energy flows. Machine learning now provides most of these tools, since it can learn consumption and generation patterns from smart-meter data and inform operating decisions [2].

Renewable energy communities (RECs), defined in the recast Renewable Energy Directive (2018/2001) [3], let local members jointly produce, share, and use renewable electricity through collective self-consumption for environmental and economic benefit [4]. In most of these communities, each residential battery runs a default, rule-based self-consumption controller that uses only local, present information [5]. Whether this control still benefits the community when many batteries act at the same time has rarely been tested against measured data, although the combined behaviour of many batteries can change the community's aggregate exchange with the grid. Three gaps emerge from the literature. First, battery dispatch is usually evaluated with assumed or perfect forecasts rather than against measured operation, and few studies test whether a community's real, rule-based battery operation helps or harms at the community level [6–8]. Second, forecasting and storage control are usually studied separately: machine-learning forecasts are selected by statistical accuracy, although accuracy gains do not always translate into operational gains [9,10]. Third, the benefits of community sharing are mostly reported in aggregate, without distinguishing the value of sharing energy from that of pooling storage or indicating which households receive them [11,12].

This paper addresses all three gaps for a real REC, using measured PV, battery, and grid data from a 20-household Irish demonstration community [13]. It asks a practical question: with only the data already collected in the community, how much can a lightweight forecasting and control framework achieve, and where does it reach its limits? Because the battery operation is measured, the rule-based baseline reflects field behaviour, and forecast-driven alternatives can be compared against it on a scenario ladder that runs from no storage, through the measured control and a forecast-informed rule, to a forecast-driven mixed-integer optimisation and an ideal-information benchmark. Forecasts are produced with a recursive day-ahead scheme that uses only the information available when dispatch is planned, adapts to concept drift by retraining, and deliberately relies solely on local smart-meter data, without weather inputs. The dispatch is optimised for two grid-oriented objectives, the daily peak import and the import above a reference threshold, and is corrected in real time as the actual load is observed. Environmental impact is evaluated with the hourly carbon intensity of the Irish grid [14,15]. The main contributions of this paper are as follows:

1. An assessment of the measured operation of a real community, showing that its rule-based batteries increased grid import and CO_2 emissions;
2. A lightweight forecasting and control framework that uses only locally available data;
3. Evidence that better forecast accuracy does not necessarily give better dispatch, because the important peaks are still missed; and
4. A decomposition showing that energy sharing and pooled storage provide different kinds of value to the community and to different households.

The remainder of this paper is organised as follows. Section 2 reviews related work on AI-based forecasting, the operational value of forecasts, battery dispatch and energy sharing, adaptive learning, and positions the present contribution. Section 3 describes the case study, dataset, forecasting framework, dispatch models, and evaluation scenarios. Section 4 presents the results, Section 5 discusses their implications and limitations, and Section 6 concludes the paper.

2. Related Work

2.1. AI-Based Forecasting for Community Load and PV

Short-term forecasting of demand and generation is a prerequisite for operating community energy systems, and machine learning now dominates this task. Statistical models such as ARIMA capture linear temporal structure but handle the nonlinear, multi-scale behaviour of aggregated residential demand poorly. Tree-ensemble and gradient-boosting methods, including Random Forest, XGBoost, LightGBM, and CatBoost, address this at low computational cost, and several studies report that they match or exceed deep-learning architectures for community and residential load while training far faster [16–18]. Gradient boosting has been applied to REC forecasting and energy management [2], and to PV production forecasting from meteorological and historical inputs [19]. While these studies demonstrated accurate forecasting performance, they focused on prediction and did not link the forecasts to community-level operation. Day-ahead forecasts also differ in how they are produced: a model can predict all hours of the next day from information available the day before, or it can proceed recursively, feeding each prediction back as an input for the following hour. Both are legitimate for day-ahead operation only if every input is known at the moment the schedule is fixed, a condition that is easily violated when lagged observations are used in evaluation. Across the broader literature, weather inputs are commonly assumed to be available and beneficial; however, their actual contribution to day-ahead, hourly, aggregated community forecasting is case-dependent and often remains insufficiently evaluated.

2.2. Forecast Accuracy and Operational Value

Forecasting models are usually compared by statistical error [20], while the systems that use them are judged by operational outcomes [21]. In a study of local energy communities, Putz et al. [9] showed that improvements in accuracy metrics did not translate into proportional gains in cost or self-sufficiency and concluded that accuracy measures alone are insufficient to select a forecasting method. Task-based and value-oriented approaches formalise this view by training or evaluating forecasts through the downstream optimisation they feed [10]. For battery dispatch, the relevant question is therefore not only how large the forecast error is but also which part of it the controller is sensitive to, for example, the timing and magnitude of demand peaks when the objective is to limit grid import. Evidence of this kind, obtained with measured community data and compared against both the field operation and an ideal-information benchmark, remains scarce.

2.3. Battery Dispatch, Optimisation, and Energy Sharing in RECs

A separate body of work optimises storage, dispatch and energy sharing within RECs and microgrids. Battery scheduling is commonly posed as a linear or mixed-integer program that maximises self-consumption, lowers grid exchange, or shaves peaks under battery and network constraints [6–8]. Collective and peer-to-peer sharing schemes coordinate several prosumers, so local generation is consumed inside the community rather than exported [22]. Shared community storage has been shown to outperform private, per-household storage [23]. These studies demonstrate the technical potential of coordinated storage and sharing, but they usually evaluate control against an assumed or perfect forecast and rarely compare it against the battery's actual, measured operation. The benefit of sharing is also mostly reported for the community as a whole. How the benefit is divided among members has been addressed mainly from an economic perspective, through allocation rules based on members' contributions or on game-theoretic criteria [11,12], whereas the question of which part of the benefit comes from sharing energy and which from pooling storage, and which households receive each, has received less attention.

2.4. Concept Drift and Adaptive Forecasting

Electricity consumption is non-stationary: user behaviour, technology adoption, and external shocks change demand over time and degrade static models [24]. Adaptive methods, including online learning, sliding-window retraining, and drift-triggered updates, keep models current [25–27], and the COVID-19 period is a documented source of such drift in residential load [28]. Adaptive retraining is often assumed to help rather than be tested, even though the trigger and the retraining data decide whether it improves or worsens accuracy. With a short history, a model retrained only on recent weeks may also discard seasons it has already seen, which makes the choice between sliding and expanding retraining windows relevant for community-scale forecasting.

2.5. Research Gap and Positioning

Three gaps follow from this review. First, few studies use measured field data to test whether a community's real, rule-based battery operation helps or harms at the community level. Second, forecasting and dispatch optimisation are usually evaluated separately, often with external inputs such as weather data, so it is rarely tested whether a more accurate machine-learning forecast yields better dispatch than a simple benchmark or which forecast errors limit the gap to ideal information. Third, the benefits of community sharing are seldom decomposed into the value of sharing energy and the value of pooling storage, or traced to the households that receive them. The present work addresses all three using measured data from an Irish REC, evaluating each layer of the framework by its operational effect as well as by its accuracy.

3. Materials and Methods

3.1. Overall Framework

The study follows the five-stage workflow shown in Figure 1. Measured household data are first cleaned and aggregated to the community level. Day-ahead forecasts of community load and PV production are then produced for every day of a test period, using only information available at midnight. These forecasts drive battery dispatch, either via a simple rule or a mixed-integer linear program (MILP), whose decisions are fixed to the forecast and then applied to the measured load and PV for that day. Finally, the resulting grid exchange is compared across a ladder of scenarios, from no battery to ideal information, and across individual and community operation. All strategies are therefore judged on the same measured days and with the same metrics.

3.2. Case Study and Dataset

The case study is the StoreNet demonstration community in Dingle, County Kerry, Ireland, led by the International Energy Research Centre with Electric Ireland, Solo Energy, and ESB Networks [13]. Twenty households took part, each with a Sonnen battery rated at 3.3 kW and 10 kWh, and ten of them also have a 2.4 kWp rooftop PV system. The community therefore has 24 kWp of PV and 200 kWh/66 kW of storage in total (Table 1). The open dataset provides one-minute records of consumption, PV production, battery charge and discharge, grid imports and exports, and battery state of charge (SOC) for the entire year 2020. The batteries ran their default self-consumption control under a day/night tariff, and about 75% of their charging took place between 23:00 and 08:00. Over the year, the community consumed 100.0 MWh, of which PV covered 19%. The measured self-consumption ratio (SCR) was 81.1%, but the self-sufficiency ratio (SSR) was only 2.9% because most of the energy stored in the batteries was supplied by the grid. Figure 2 shows the daily community demand over 2020. Demand was mostly between about 200 and

400 kWh per day, higher in winter and lower in summer. Public-health restrictions changed this pattern three times during the year: a national stay-at-home order from 27 March until the first reopening phase on 18 May, a Level 5 lockdown from 22 October until 1 December, and the reintroduction of most Level 5 measures from 24 December [29]. Demand fell during the first lockdown and rose sharply in late December, when the new restrictions coincided with the holidays and the coldest weeks of the year. These shifts are the kind of concept drift that the forecasting framework must follow, and they motivate the lockdown flag and the drift-adaptive retraining described in Section 3.4. The isolated single-day drops to about 50–60 kWh are data gaps (Section 3.3), not real demand.

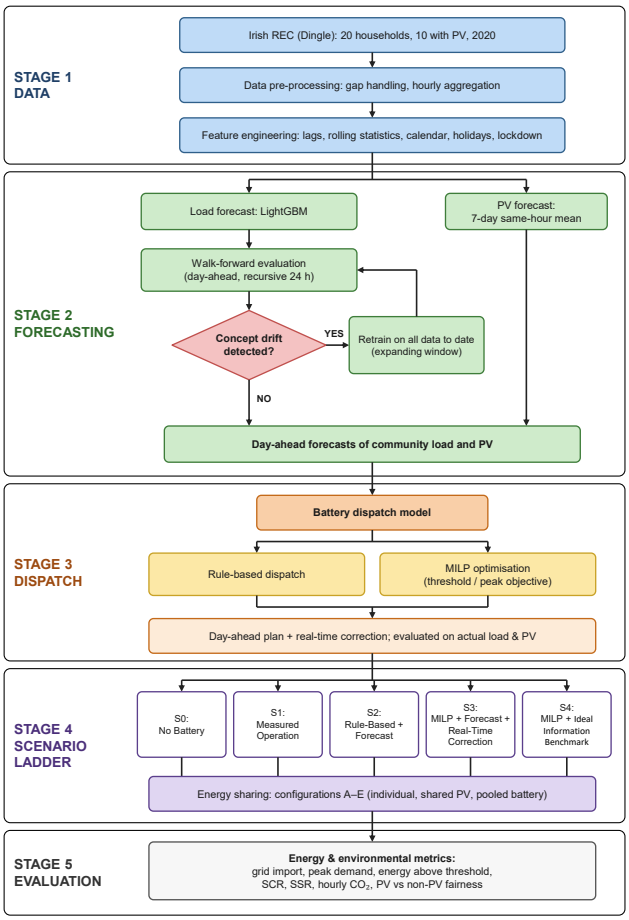


Figure 1. Overview of the framework: data processing, day-ahead forecasting with drift-adaptive retraining, battery dispatch, and evaluation of the scenario ladder and of the sharing configurations.

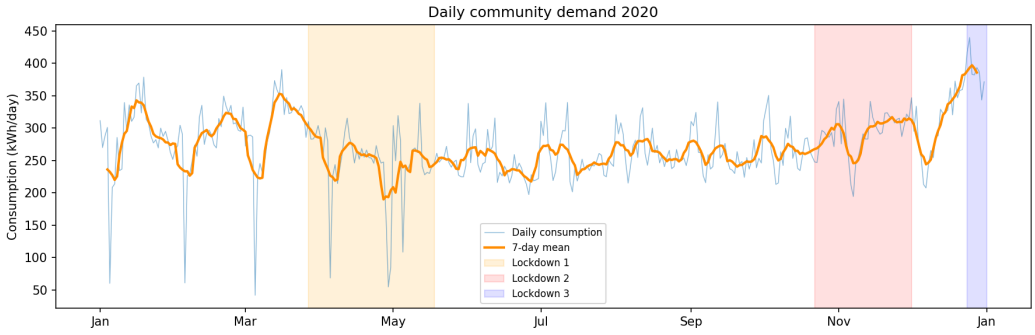


Figure 2. Daily community demand in 2020 with its seven-day mean. Shaded areas mark the stay-at-home order (27 March–17 May) and the Level 5 periods (22 October–1 December and from 24 December) [29].

Table 1. Main characteristics of the Dingle community and dataset.

Item	Value
Households	20 (10 with PV, 10 without)
PV capacity	10 × 2.4 kWp = 24 kWp
Battery (per household)	3.3 kW/10 kWh (Sonnen)
Battery (community)	66 kW/200 kWh
Estimated one-way efficiency $\eta_{ch} = \eta_{dis}$	0.885 (Equation (1))
Usable SOC window in dispatch	10–90%
Period and resolution	2020, 1 min (raw), 1 h (analysis)
Annual consumption/PV production	100.0 MWh/19.0 MWh
Measured SCR/SSR (whole year)	81.1%/2.9%

The capacity and efficiency of the batteries were checked against the data rather than taken from the datasheet. For each household, the hourly change in SOC was regressed on the charged and discharged energy,

$$\Delta SOC_{i,t} = a_1 E^{ch}_{el,i,t} - a_2 E^{dis}_{el,i,t} + \varepsilon_{i,t}, \qquad a_1 = \frac{\eta_{ch}}{C^{BAT}}, \quad a_2 = \frac{1}{\eta_{dis} C^{BAT}}, \tag{1}$$

using hours with SOC between 5% and 95%. Assuming equal charging and discharging efficiencies, $\eta_{ch} = \eta_{dis}$, the two coefficients give $C^{BAT} = (a_1 a_2)^{-1/2}$ and $\eta_{ch} = \eta_{dis} = (a_1/a_2)^{1/2}$. The median over the twenty households gives $C^{BAT} \approx 10.0$ kWh, which confirms the nominal rating, and $\eta_{ch} = \eta_{dis} \approx 0.885$, i.e., a battery round-trip efficiency of about 78%. The measured system round-trip efficiency over the year was lower, at 72%, which is consistent with standby consumption of the battery units. The difference corresponds to about 2.8 MWh per year, or a continuous standby consumption of about 0.31 kW for the community.

3.3. Data Pre-Processing

Duplicate timestamps were removed, SOC values outside 0–100% were set as missing, and negative energy readings were set to zero. The one-minute data were aggregated to a 30-min resolution by summing energy quantities and averaging the SOC. Missing intervals of up to two hours were forward-filled; longer gaps were set to zero for energy quantities, and the SOC was linearly interpolated over gaps of up to four hours. The timestamps follow a fixed offset without daylight-saving changes. For forecasting and dispatch, the data were aggregated to hourly resolution. Hours with community load below 0.5 kWh correspond to filled gaps rather than real demand. They were excluded from model training and all error metrics, and in the model inputs they were replaced with the value from the same hour on the most recent previous day, so only past information is used. At the household level, the zero-filled intervals accounted for 0–2.7% of the year and 0–0.9% of the test period, and only one household had no record. No test day contained a community-level gap, so the dispatch is evaluated on all 239 test days, and the zero-filled household intervals are entered as recorded. In addition, the records of one household without PV (H14) are missing for all variables from 12 to 31 December 2020 (20 days). In the community aggregates, these hours enter with zero load, PV, and battery flows for that household; the omitted load is estimated at about 279 kWh, 2.7% of the community consumption in December and 0.4% of the test-period consumption. Since dispatch decisions, import peaks and the allocation of shared energy depend nonlinearly on the load, the effect of the missing records cannot be assumed to cancel across scenarios. All dispatch and sharing comparisons were therefore repeated over a reduced test period ending on 11 December 2020 (219 days), in which the records of all twenty households and the availability of all twenty batteries are complete (Appendix A, Table A1). Over this period, the ranking of the dispatch scenarios and of the

sharing configurations, as well as the distribution of benefits between households with and without PV, remained unchanged.

3.4. Forecasting Framework

The dispatch requires two day-ahead inputs for each hour of the following day: the community load and the PV production. Both are forecast at midnight using only the smart-meter, PV and calendar data that the community already records, without weather forecasts or other external inputs. The load is forecast with a LightGBM model applied recursively over the 24 h of the next day and retrained when concept drift is detected. The PV production is forecast with a seven-day same-hour mean, since PV covers only a small share of the demand and deviations are handled by the real-time correction. The following subsections describe the load model and its features, the recursive scheme, the walk-forward evaluation with drift-adaptive retraining, the benchmark forecasts and the PV forecast.

3.4.1. Model and Features

Community load is forecast with LightGBM, a gradient-boosted tree ensemble. The model uses 400 trees, a learning rate of 0.05, 64 leaves per tree, feature subsampling of 0.9, and an absolute-error objective, with a fixed random seed. These settings were fixed before the test period; a validation-based tuning on the training period did not improve them (Section 4.2). The features in Table 2 were deliberately restricted to data already available within the community: historical load and calendar information. No weather or other external data is used. This tests how far a lightweight framework can go with local data alone, keeping it deployable without external data services.

Table 2. Input features of the load model. $E_{\mathrm{el},t}$ is the community electric load in hour t.

Group	Features
Lags	$E_{\mathrm{el},t-1}$, $E_{\mathrm{el},t-24}$, $E_{\mathrm{el},t-48}$, $E_{\mathrm{el},t-168}$
Rolling statistics	mean of $E_{\mathrm{el},t-3\ldots t-1}$; mean and standard deviation of $E_{\mathrm{el},t-24\ldots t-1}$
Same-hour history	mean of $E_{\mathrm{el},t-24\tau}$ for $\tau = 1, \ldots, 7$
Calendar	hour, day of week, month, weekend flag; sine and cosine of hour and day of week
Calendar events	Irish public holiday flag; lockdown flag (27 Mar–17 May, 21 Oct–30 Nov, from 24 Dec 2020) [29]

3.4.2. Recursive Day-Ahead Scheme

The dispatch for a day is planned at midnight, when the load is known only up to the last hour of the previous day, denoted t_0. The next $N_{\mathrm{h}} = 24$ hourly values are therefore predicted recursively, as shown in Figure 3: each forecast is added to the history and used as an input for the following hour,

$$\hat{E}_{\mathrm{el},t_0+m} = f\Big(\mathrm{cal}_{t_0+m};\ \hat{E}_{\mathrm{el},t_0+m-1}, \ldots, \hat{E}_{\mathrm{el},t_0+1};\ E_{\mathrm{el},t_0}, \ldots, E_{\mathrm{el},t_0-(N_{\mathrm{w}}-m)}\Big), \qquad m = 1, \ldots, N_{\mathrm{h}}, \tag{2}$$

where m is the forecast step, f is the LightGBM model, cal_{t_0+m} are the calendar variables of hour $t_0 + m$, which are known in advance, and $N_{\mathrm{w}} = 168$ hourly values is the longest history used by the features of Table 2. For $m = 1$, the input contains only measured loads. No measured value from the forecast day enters its own forecast. The model is trained on measured data, as is standard for recursive forecasting.

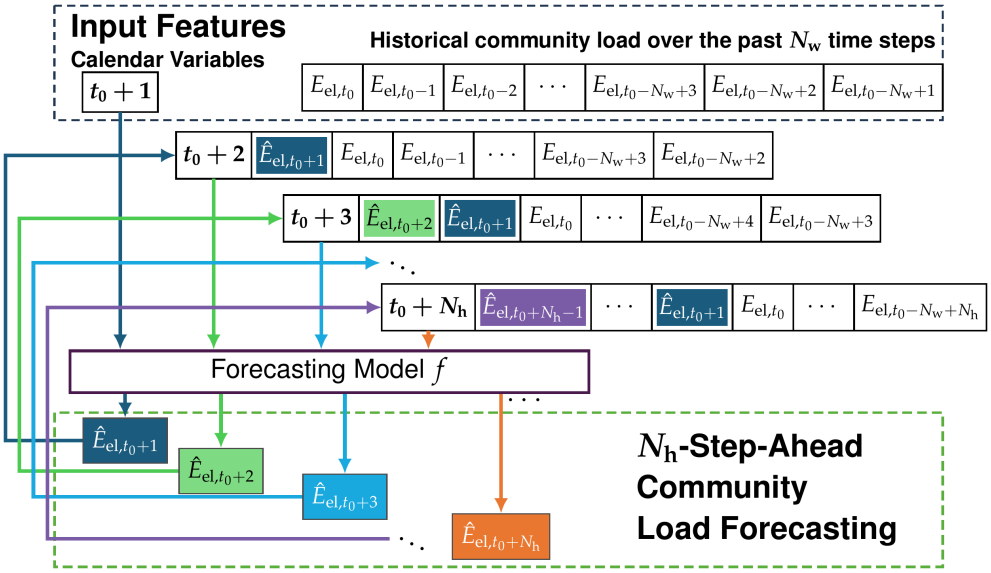


Figure 3. Recursive day-ahead forecasting. At midnight (t_0 = last hour of the previous day) the model predicts $\hat{E}_{\mathrm{el},t_0+1}$ from calendar variables and from lagged and rolling-window features computed over the last $N_w = 168$ hourly loads; each prediction is then appended to this history window and used for the next hour, up to $N_h = 24$ steps. Each colour marks the forecast of one step, which is reused as an input in the following steps; white cells are measured loads.

3.4.3. Walk-Forward Evaluation and Drift-Adaptive Retraining

The model was first trained for 120 days—from 8 January to 6 May 2020—after the first week was used to build the lags. It then forecasts each of the 239 days from 7 May to 31 December, in chronological order. After each day, the mean absolute error e_g is compared with the mean μ_g and the sample standard deviation σ_g of the daily errors of the seven preceding days, $g-7$ to $g-1$. Day g is not part of this reference, so the current error cannot affect the level against which it is judged. Detection starts once seven previous daily errors are available, so no retraining is triggered during the first seven test days. Concept drift is declared when

$$e_g > \kappa_1 \mu_g \quad \text{or} \quad e_g > \mu_g + \kappa_2 \sigma_g, \qquad \kappa_1 = 1.6,\ \kappa_2 = 2.5, \tag{3}$$

with

$$\mu_g = \frac{1}{7}\sum_{k=1}^{7} e_{g-k}, \qquad \sigma_g = \sqrt{\frac{1}{6}\sum_{k=1}^{7}\left(e_{g-k}-\mu_g\right)^2}. \tag{4}$$

The model is then retrained on all data available up to the end of day g and used from day $g+1$. An expanding window was chosen because a single year of data does not capture every season before it is forecast; retraining only on recent weeks would discard seasonal patterns the model has already seen. The two thresholds were fixed before the test period and not tuned: κ_1 reacts to a sustained rise of the daily error of about 60% above its recent level, and κ_2 to isolated outliers. Their effect is examined in Section 4.2.

3.4.4. Benchmarks

Two simple forecasts serve as benchmarks: persistence, which repeats the load of the same hour on the previous day, and a seven-day same-hour mean,

$$\hat{E}^{7\mathrm{d}}_{\mathrm{el},g,h} = \frac{1}{7}\sum_{\tau=1}^{7} E_{\mathrm{el},g-\tau,h}. \tag{5}$$

The seven-day mean is also used as an alternative load forecast in the dispatch scenario S3n (Section 3.6), to directly test the operational value of the machine-learning forecast. The

hyperparameters of LightGBM (Section 3.4) are common settings for hourly load data and were not selected for the test period. To check them and provide a second machine-learning benchmark, LightGBM and XGBoost were tuned on the training period only: the models were fitted on the first 90 days and evaluated on the last 30 days (7 April to 6 May 2020) using the same recursive day-ahead scheme. For LightGBM, 27 combinations of the number of leaves (15, 31, 64), the learning rate (0.03, 0.05, 0.1) and the number of trees (200, 400, 800) were tested; for XGBoost, 54 combinations of the maximum depth (3, 4, 6), the same learning rates and numbers of trees, and a squared-error or absolute-error objective were tested. The best setting on the validation days was then run once on the test period, with the same features, drift rule and retraining. The selected XGBoost model uses 200 trees of depth 3, a learning rate of 0.1, and an absolute-error objective.

3.4.5. Photovoltaic Forecasting

PV production is forecast with the same seven-day same-hour mean that serves as the load benchmark. The forecast for hour h of day g is

$$\hat{E}^{\mathrm{PV}}_{\mathrm{el},g,h} = \frac{1}{7}\sum_{\tau=1}^{7} E^{\mathrm{PV}}_{\mathrm{el},g-\tau,h}, \tag{6}$$

where $E^{\mathrm{PV}}_{\mathrm{el},g-\tau,h}$ is the measured PV production in the same hour τ days earlier. It is applied to the community total in the dispatch analysis and to each PV household in the sharing analysis. The forecast uses only past measured production and needs no weather data; it follows the daily solar cycle and the seasonal change in day length. Deviations caused by day-to-day cloud variability are handled by the real-time correction. Because PV covers only 19% of demand, the dispatch is driven mainly by the load forecast. The effect of PV forecast errors is isolated in scenario S3b, which uses the realised PV profile.

3.5. Battery Model and Dispatch Strategies

The community battery is modelled as a single aggregated unit, a virtual pool of the twenty household batteries whose feasibility is discussed in Section 3.7, with capacity $C^{\mathrm{BAT}} = 200\,\mathrm{kWh}$, power limit $P^{\mathrm{BAT}} = 66\,\mathrm{kW}$, and one-way efficiencies $\eta_{\mathrm{ch}} = \eta_{\mathrm{dis}} = 0.885$. The time step is $\Delta t = 1\,\mathrm{h}$, so power in kW equals energy in kWh per step. For each hour t, $E_{\mathrm{el},t}$ and $E^{\mathrm{PV}}_{\mathrm{el},t}$ are the community load and PV production; $E^{\mathrm{ch}}_{\mathrm{el},t}$ and $E^{\mathrm{dis}}_{\mathrm{el},t}$ are the charged and discharged energy; $E^{\mathrm{PG}}_{\mathrm{el},t}$ and $E^{\mathrm{TG}}_{\mathrm{el},t}$ are the grid import and export; E^{BAT}_{t} is the stored energy; and $u_t \in \{0,1\}$ indicates charging. The model is

$$E_{\mathrm{el},t} + E^{\mathrm{ch}}_{\mathrm{el},t} + E^{\mathrm{TG}}_{\mathrm{el},t} = E^{\mathrm{PV}}_{\mathrm{el},t} + E^{\mathrm{dis}}_{\mathrm{el},t} + E^{\mathrm{PG}}_{\mathrm{el},t}, \tag{7}$$

$$E^{\mathrm{BAT}}_{t} = E^{\mathrm{BAT}}_{t-1} + \eta_{\mathrm{ch}}\,E^{\mathrm{ch}}_{\mathrm{el},t} - E^{\mathrm{dis}}_{\mathrm{el},t}/\eta_{\mathrm{dis}}, \tag{8}$$

$$0 \le E^{\mathrm{ch}}_{\mathrm{el},t} \le P^{\mathrm{BAT}}\Delta t\,u_t, \qquad 0 \le E^{\mathrm{dis}}_{\mathrm{el},t} \le P^{\mathrm{BAT}}\Delta t\,(1-u_t), \tag{9}$$

$$0.1\,C^{\mathrm{BAT}} \le E^{\mathrm{BAT}}_{t} \le 0.9\,C^{\mathrm{BAT}}, \tag{10}$$

with all flows non-negative. The binary variable prevents simultaneous charging and discharging, and the SOC window of 10–90% limits the battery degradation. The initial stored energy is taken from the measured SOC at the start of the test period.

3.5.1. Rule-Based Dispatch with Forecast (S2)

Between 00:00 and 08:00, the battery does not discharge. It charges from the grid only as much as the forecast requires for the rest of the day, namely the target

$$E^{\mathrm{BAT},*}_{g} = \min\left(0.9\,C^{\mathrm{BAT}},\ 0.1\,C^{\mathrm{BAT}} + \frac{1}{\eta_{\mathrm{dis}}}\sum_{h=8}^{23}\max\left(\hat{E}_{\mathrm{el},g,h} - \hat{E}^{\mathrm{PV}}_{\mathrm{el},g,h},\ 0\right)\right), \tag{11}$$

spread evenly over the remaining hours of the night. From 08:00 to 24:00, the battery stores PV surplus and discharges to cover any deficit. The rule represents a forecast-informed controller without optimisation.

3.5.2. Mixed-Integer Linear Programming Dispatch (S3, S4)

Each day at midnight, a MILP is solved over the next 24 h, $t \in \mathcal{T} = \{1, \ldots, 24\}$, with the forecasts $\hat{E}_{\mathrm{el},t}$ and $\hat{E}^{\mathrm{PV}}_{\mathrm{el},t}$ in place of $E_{\mathrm{el},t}$ and $E^{\mathrm{PV}}_{\mathrm{el},t}$. The decision variables are $E^{\mathrm{ch}}_{\mathrm{el},t}, E^{\mathrm{dis}}_{\mathrm{el},t}, E^{\mathrm{PG}}_{\mathrm{el},t}, E^{\mathrm{TG}}_{\mathrm{el},t}, E^{\mathrm{BAT}}_{t} \geq 0$ and $u_t \in \{0,1\}$; E^{BAT}_{0} is the stored energy at midnight, carried over from the previous day. The constraints are Equations (7)–(10) for all $t \in \mathcal{T}$, together with

$$E^{\mathrm{BAT}}_{24} \geq E^{\mathrm{BAT}}_{0}, \tag{12}$$

which prevents the optimiser from emptying the battery at the end of each horizon. Two grid-oriented objectives are considered. The peak objective minimises the largest hourly import of the day,

$$\min \; P^{\mathrm{pk}} \Delta t + w \sum_{t \in \mathcal{T}} E^{\mathrm{PG}}_{\mathrm{el},t} \quad \text{s.t.} \quad E^{\mathrm{PG}}_{\mathrm{el},t} \leq P^{\mathrm{pk}} \Delta t \;\; \forall t \in \mathcal{T}, \tag{13}$$

while the threshold objective penalises only the import above a reference threshold P^{thr},

$$\min \sum_{t \in \mathcal{T}} E^{\mathrm{exc}}_{\mathrm{el},t} + w \sum_{t \in \mathcal{T}} E^{\mathrm{PG}}_{\mathrm{el},t} \quad \text{s.t.} \quad E^{\mathrm{exc}}_{\mathrm{el},t} \geq E^{\mathrm{PG}}_{\mathrm{el},t} - P^{\mathrm{thr}} \Delta t, \;\; E^{\mathrm{exc}}_{\mathrm{el},t} \geq 0 \;\; \forall t \in \mathcal{T}, \tag{14}$$

with $w = 0.1$ in both cases, so the grid-oriented term dominates and the import term provides a secondary penalty on total import. The effect of this choice is tested in Section 4.3.5 (Table A3 in Appendix A) with $w = 0.01, 0.05, 0.1, 0.2$ and 0.5. The threshold is a statistical reference level, not a network capacity verified by a power-flow analysis; exceeding it identifies hours of high community import, not a violation of a physical limit. It was fixed before the test period, as the 80th percentile of the hourly net load $\max(E_{\mathrm{el},t} - E^{\mathrm{PV}}_{\mathrm{el},t}, 0)$ over the first 120 days of 2020, which gives $P^{\mathrm{thr}} = 14.1$ kW. The 80th percentile means the battery operates during about a fifth of the hours with the highest net load, a range that the 160 kWh of usable storage can cover on most days, while the base load below it is left to the grid. The threshold acts as a soft limit: import above it is penalised but not forbidden, so the problem remains feasible on days when the load exceeds what the battery can cover. The threshold objective is the main case; the peak objective is reported for comparison. No lower bound is imposed on the import. Summing Equation (7) over the day and using Equations (8) and (12) gives

$$\sum_{t \in \mathcal{T}} E^{\mathrm{PG}}_{\mathrm{el},t} \;\geq\; \sum_{t \in \mathcal{T}} \left(E_{\mathrm{el},t} - E^{\mathrm{PV}}_{\mathrm{el},t} \right) + \sum_{t \in \mathcal{T}} E^{\mathrm{TG}}_{\mathrm{el},t} + (1 - \eta_{\mathrm{ch}}\, \eta_{\mathrm{dis}}) \sum_{t \in \mathcal{T}} E^{\mathrm{ch}}_{\mathrm{el},t}, \tag{15}$$

so the daily import cannot fall below the daily net demand. This bound applies when the terminal stored energy is at least as high as the initial stored energy. For the realised dispatch after correction, the energy balance must also account for the change in stored energy. The battery can change when energy is drawn from the grid, but it can reduce the amount drawn only by storing PV surplus and only after conversion losses. The problems were solved with CBC through PuLP, with a time limit of 10 s per day; a daily problem is typically solved in well under one second.

3.5.3. Execution and Real-Time Correction

The plan is applied hour by hour to the measured load and PV, with net load $n_t = E_{\mathrm{el},t} - E^{\mathrm{PV}}_{\mathrm{el},t}$ and $[x]^+ = \max(x, 0)$. Because the controller observes the actual load each hour, the plan is corrected against a target import Π_g, equal to the planned daily

peak (peak objective) or to the larger of P^{thr} and the planned maximum import (threshold objective). With $E^{\mathrm{ch}}_{\mathrm{el},t}$ and $E^{\mathrm{dis}}_{\mathrm{el},t}$ the planned values, the applied charging and discharging are

$$\tilde{E}^{\mathrm{ch}}_{\mathrm{el},t} = \min\left\{E^{\mathrm{ch}}_{\mathrm{el},t},\ \frac{0.9\,C^{\mathrm{BAT}} - E^{\mathrm{BAT}}_{t-1}}{\eta_{\mathrm{ch}}},\ [\Pi_g \Delta t - n_t]^+\right\}, \tag{16}$$

$$\tilde{E}^{\mathrm{dis}}_{\mathrm{el},t} = \min\Big\{\max\left(E^{\mathrm{dis}}_{\mathrm{el},t},\ \mathbf{1}[\tilde{E}^{\mathrm{ch}}_{\mathrm{el},t} = 0]\,[n_t - \Pi_g \Delta t]^+\right),\ P^{\mathrm{BAT}}\Delta t,$$
$$\eta_{\mathrm{dis}}\left(E^{\mathrm{BAT}}_{t-1} + \eta_{\mathrm{ch}}\,\tilde{E}^{\mathrm{ch}}_{\mathrm{el},t} - 0.1\,C^{\mathrm{BAT}}\right),\ [n_t]^+\Big\}. \tag{17}$$

Equation (16) reduces planned charging that would raise the import above the target and Equation (17) adds discharge when the actual load pushes the import above it. Discharge is limited to the actual net demand, so the battery never exports to the grid. The grid exchange then follows from Equation (7); any remaining PV surplus is stored if there is room, and only the rest is exported. The corrections use the net load of the current hour, assuming the controller responds to the measured load within the hour, as battery inverters do with instantaneous meter readings. Because the energy of an hour is known only at its end, a conservative variant is also evaluated in which every correction uses the net load of the previous hour (Section 4.3.8).

3.6. Scenario Ladder

Table 3 lists the scenarios. Moving up the ladder isolates one effect at a time: the effect of the batteries as they were operated (S0 to S1), the value of a forecast with a simple rule (S1 to S2), the value of optimisation (S2 to S3), and the cost of forecast errors (S3 to S4). S4 is a non-deployable reference in which the optimiser is given the realised load and PV profiles instead of forecasts; it is referred to as the ideal-information benchmark. Scenarios S3n and S3b change only one forecast and show which forecast limits the dispatch. All scenarios are evaluated on the same 239 test days. For S1, the measured household imports and exports are netted hourly at the community level, so S0–S4 share the same accounting boundary; the sum of the individual household imports is reported as configuration M in Section 4.4.

Table 3. Dispatch scenarios.

Scenario	Dispatch	Load/PV Information	Purpose
S0	No battery	–	Reference
S1	Measured Sonnen control	None (as operated)	Field baseline
S2	Rule, Equation (11)	LightGBM/7-day mean	Forecast without optimisation
S3	MILP + correction	LightGBM/7-day mean	Main strategy
S3n	MILP + correction	7-day mean/7-day mean	Value of the ML forecast
S3b	MILP + correction	LightGBM/realised PV	Cost of PV forecast error
S4	MILP + correction	Realised/realised	Ideal-information benchmark

3.7. Energy Sharing and Allocation of Benefits

Energy sharing is represented as hourly virtual netting of household imports and exports within the community; internal network constraints and losses are not modelled. The pooled battery in configuration E is likewise a virtual pool of household batteries operated as a single unit, not a physical central battery. Its schedule is feasible for every household battery when the aggregate charging and discharging are divided equally among the twenty identical units (10 kWh, 3.3 kW): each unit then follows one twentieth of the aggregate schedule, whose energy, power and SOC limits are twenty times those of a unit. This requires the units to share the same state of charge; therefore, the pool starts from the total measured stored energy, which corresponds to balancing the units at the start of the test period. The energy that the pool moves between households is accounted for by the same hourly virtual netting. To separate the

value of sharing energy from that of pooling storage, the same controller is applied to the five configurations in Table 4. In C and D, each household i runs its own 10 kWh/3.3 kW battery with the MILP of Section 3.5, using its own threshold P_i^{thr}, defined as the 80th percentile of its own net load over the first 120 days. In D, the resulting household imports and exports are netted within the community in each hour. In E, the aggregated PV and a single pooled battery of 200 kWh/66 kW serve the whole community. To keep the comparison fair, households and the community use the same forecast method: the seven-day same-hour mean for load and PV. An ideal-information version of each configuration is also computed. The measured operation (M) is included for reference.

Table 4. Sharing configurations.

Config.	PV	Storage
A	Own use only	None
B	Shared within the community	None
M	Own use only	Own battery, measured operation
C	Own use only	Own battery, MILP
D	Shared within the community	Own battery, MILP
E	Shared within the community	One pooled battery, MILP

In the netted configurations (B, D, and E), the energy exported by one household can cover the import of another household in the same hour, so only the remaining deficit is drawn from the grid. The community grid import in hour t is therefore

$$E_{\text{el},t}^{\text{PG}} = \max\left(\sum_i \left(E_{\text{el},i,t}^{\text{PG}} - E_{\text{el},i,t}^{\text{TG}}\right), 0\right), \tag{18}$$

where $E_{\text{el},i,t}^{\text{PG}}$ and $E_{\text{el},i,t}^{\text{TG}}$ are the import and export of household i before sharing. Without sharing (A, C, and M), the community import is simply the sum of the household imports, $\sum_i E_{\text{el},i,t}^{\text{PG}}$. To see which households benefit from sharing, the net import is allocated back to the households. In B and D, the energy shared in hour t,

$$E_{\text{el},t}^{\text{sh}} = \min\left(\sum_i E_{\text{el},i,t}^{\text{PG}}, \sum_i E_{\text{el},i,t}^{\text{TG}}\right), \tag{19}$$

is given to the importing households in proportion to their import,

$$E_{\text{el},i,t}^{\text{PG,REC}} = E_{\text{el},i,t}^{\text{PG}}\left(1 - \frac{E_{\text{el},t}^{\text{sh}}}{\sum_j E_{\text{el},j,t}^{\text{PG}}}\right). \tag{20}$$

In E, where no household-level battery flows exist, the community import is split in proportion to each household's net deficit $\delta_{i,t} = \max(E_{\text{el},i,t}^{\text{Us}} - E_{\text{el},i,t}^{\text{PV}}, 0)$,

$$E_{\text{el},i,t}^{\text{PG,REC}} = E_{\text{el},t}^{\text{PG}} \frac{\delta_{i,t}}{\sum_j \delta_{j,t}}. \tag{21}$$

The self-sufficiency of PV and non-PV households is then computed from the allocated imports. The results of the allocation depend on these rules, which are energy-accounting rules under hourly virtual netting and not a tariff, settlement, or compensation scheme.

3.8. Evaluation Metrics

Forecast accuracy is measured over all valid test hours with the mean absolute error (MAE), root mean square error (RMSE), normalised MAE, mean absolute percentage error (MAPE), and coefficient of determination R^2:

$$\text{nMAE} = \frac{\sum_t |E_{\text{el},t} - \hat{E}_{\text{el},t}|}{\sum_t E_{\text{el},t}}, \qquad \text{Skill} = 1 - \frac{\text{MAE}}{\text{MAE}_{\text{persistence}}}. \tag{22}$$

The error of the forecast daily maximum is also reported, together with four metrics for the high-load hours: the error of the peak hour, the error of the daily net energy above the threshold, the error of the number of hours above it, and the error of the evening energy (16:00–21:00). To isolate the contribution of the peak error to the dispatch, a peak-bias correction is applied to the forecast of day g,

$$\hat{E}^{\mathrm{c}}_{g,t} = \hat{E}_{g,t} + b_g\,\omega_{g,t}, \qquad \omega_{g,t} = \min\left\{1,\ \max\left[0,\ \frac{\hat{E}_{g,t} - q_g}{\max_\tau \hat{E}_{g,\tau} - q_g}\right]\right\}, \tag{23}$$

where q_g is the 75th percentile of the hourly forecast of day g. The forecast maximum thus rises by b_g, while the hours below q_g and the timing of the peak stay unchanged. In the causal case, b_g is the mean underestimation of the daily maximum over the previous 28 days (at least seven), bounded below by zero. In the controlled case, b_g is the error of the daily maximum of day g itself. This case uses information that is not available at the time of planning and is used only for diagnosis. The dispatch is evaluated over the test period by the total grid import and export, the maximum hourly import, the mean daily peak import, the energy imported above the community threshold, and the number of hours above it,

$$E^{\mathrm{exc}}_{\mathrm{el}} = \sum_t \max\left(E^{\mathrm{PG}}_{\mathrm{el},t} - P^{\mathrm{thr}}\,\Delta t,\ 0\right), \tag{24}$$

and by the self-consumption and self-sufficiency ratios,

$$\mathrm{SCR} = 1 - \frac{\sum_t E^{\mathrm{TG}}_{\mathrm{el},t}}{\sum_t E^{\mathrm{PV}}_{\mathrm{el},t}}, \qquad \mathrm{SSR} = 1 - \frac{\sum_t E^{\mathrm{PG}}_{\mathrm{el},t}}{\sum_t E_{\mathrm{el},t}}. \tag{25}$$

Because the controllers end the test period with different stored energies, the import is also reported for a common terminal state. A deficit relative to the initial stored energy E_0 is assumed to be recharged from the grid, and a surplus is credited as avoided import:

$$E^{\mathrm{PG,adj}}_{\mathrm{el}} = \sum_t E^{\mathrm{PG}}_{\mathrm{el},t} + \begin{cases} (E_0 - E_T)/\eta & \text{if } E_T < E_0, \\ -(E_T - E_0)\,\eta & \text{otherwise,} \end{cases} \tag{26}$$

where E_T is the stored energy at the end of the test period and η is the one-way efficiency. The recharge is assumed to take place in hours below the threshold, so $E^{\mathrm{exc}}_{\mathrm{el}}$ is not adjusted. The controllers are then compared on the realised value of the threshold objective,

$$J = E^{\mathrm{exc}}_{\mathrm{el}} + w\,E^{\mathrm{PG,adj}}_{\mathrm{el}}, \qquad w = 0.1, \tag{27}$$

with the same weight as in the dispatch optimisation. Operational CO_2 emissions are computed with the hourly carbon emission factor of the Irish grid (Republic of Ireland, 2020) published by the transmission system operator [14,15],

$$M_{\mathrm{CO_2}} = \sum_t E^{\mathrm{PG}}_{\mathrm{el},t}\,\alpha_{\mathrm{CO_2},t}, \tag{28}$$

where $\alpha_{\mathrm{CO_2},t}$ is the 15-min intensity averaged to hourly values (0.254 kg/kWh on average over the test period). The average daily profile of $\alpha_{\mathrm{CO_2},t}$ is almost flat (0.249–0.260 kg/kWh), so a one-hour offset between the time conventions of the two datasets has a negligible effect. The dispatch obtained from the LightGBM forecast (S3) and the seven-day mean (S3n) is compared using paired daily values over the 239 test days. Two daily quantities are tested: the peak import and the energy imported above the threshold. The stored energy is carried from one day to the next, and the forecasts use overlapping histories, so the daily differences d_g (S3 minus S3n) may be serially dependent. This is assessed using the lag-1 autocorrelation r_1 and the Ljung–Box test at lag 7, with the effective sample size approximated as $n(1-r_1)/(1+r_1)$. The two-sided Wilcoxon signed-rank test is applied to days with a nonzero

difference, and the number of such days is reported. The median interpretation of this test assumes a symmetric distribution of the differences, so the skewness of the nonzero differences is tested, and a sign test is reported as a distribution-free alternative. The mean difference is reported with a 95% confidence interval from a circular moving-block bootstrap with a block length of 7 days (10,000 resamples); block lengths of 3 and 14 days and an independent bootstrap serve as checks. Practical similarity is assessed with two one-sided tests at the 5% level: the two forecasts are considered equivalent for a quantity when the 90% block-bootstrap interval of the mean difference lies within ±0.5 kW for the daily peak and ±1.0 kWh for the daily energy above the threshold. The margins come from the battery hardware, not from the results: 0.5 kW is about 15% of the 3.3 kW power of one household battery, and 1.0 kWh is 10% of its 10 kWh capacity.

3.9. Implementation

All computations were carried out in Python 3.13 with LightGBM 4.6.0, XGBoost 3.4.1, scikit-learn 1.6.1, pandas 2.2.3, NumPy 2.1.3, SciPy 1.16.3, and PuLP 3.3.2 with the CBC solver. Random seeds were fixed for reproducibility. All runs were performed on a cloud notebook with two CPU cores (x86_64) and 13.6 GB of memory. Each reported time covers the complete test period of 239 days and includes model training and drift-triggered retraining but not data loading and pre-processing: the walk-forward load forecast takes about 15–25 s, the community dispatch about 4–10 s per scenario (below 0.1 s per daily MILP on average), and the twenty household-level dispatches about 1.5 min per configuration.

4. Results

4.1. Measured Operation of the Community

The measured data show that the batteries did not match the community's load. Figure 4a gives the average hourly energy flows over 2020. Consumption peaks in the evening, at about 17 kWh around 18:00, while PV production peaks at about 7 kWh around midday. Grid import follows neither of them. It is highest at 06:00–07:00, at about 28 kWh per hour, and high again between 01:00 and 03:00, at about 18–22 kWh per hour. In these hours, consumption is only 7–12 kWh: most of the import goes into the batteries, which charge at up to 19 kWh per hour. During the evening peak, the batteries discharge, and the import falls to about 11 kWh. The batteries therefore shifted the community's largest grid demand from the evening to the early morning, thereby increasing it. The monthly balance (Figure 4b) shows the energy cost of this operation. PV covered part of the demand from March to September. In January, October, and November, however, the grid import exceeded the total consumption, and in February and December, the two were almost equal: without PV surplus to store, the batteries only added conversion losses.

Figure 5 and Table 5 show that this pattern holds across the households. Nine of the ten non-PV households charged mostly at 06:00–07:00, and eight of the ten PV households charged mostly between 01:00 and 03:00. Only two PV households (H1 and H4) charged mainly around midday from their own PV. As a result, the non-PV households imported 14% more electricity than they consumed, whereas the PV households imported 79% of their consumption. The measured round-trip efficiency of the batteries was 69–73%, and the mean SOC was below 50% in both groups.

4.2. Day-Ahead Load Forecasting

Table 6 compares the recursive LightGBM forecast with the two benchmarks over the 239 test days. The final model, with all features and drift-adaptive retraining, reached an nMAE of 18.2% and an R^2 of 0.605. Its MAE was 16.9% lower than that of persistence and 10.9% lower than that of the seven-day same-hour mean.

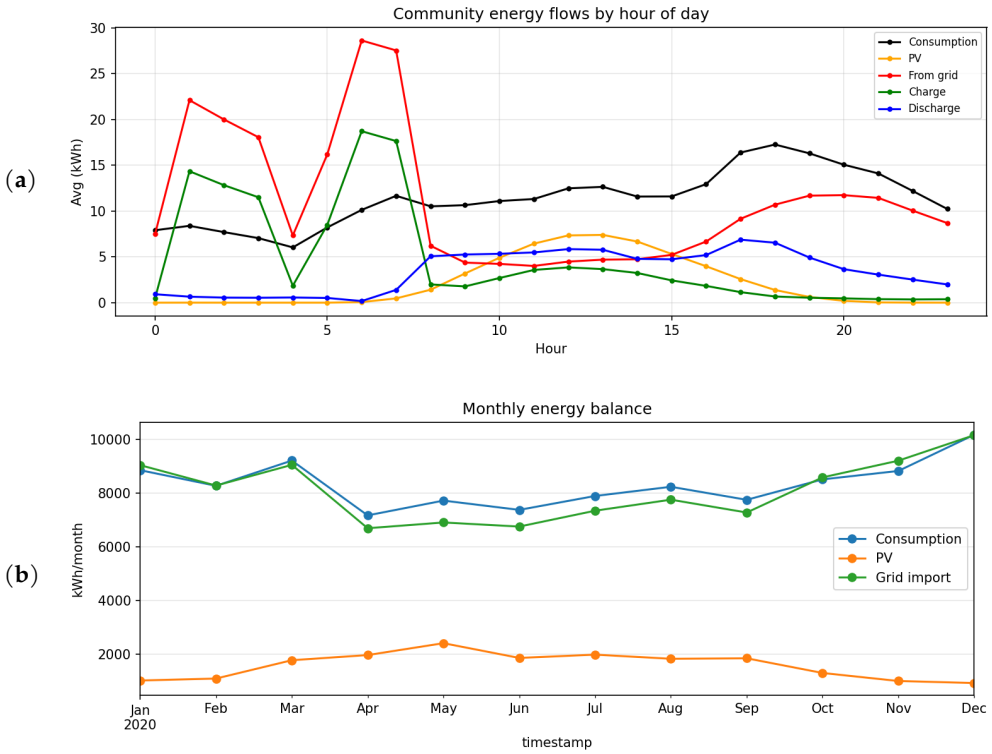


Figure 4. Measured community operation in 2020: (**a**) average energy flows by hour of the day; (**b**) monthly consumption, PV production, and grid import.

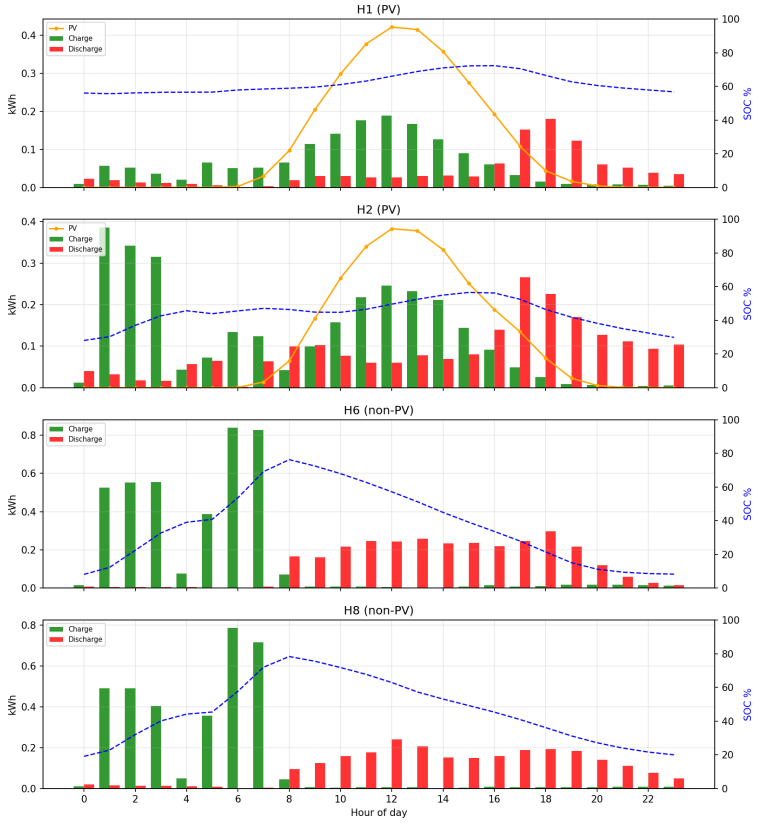


Figure 5. Average hourly battery charging (green) and discharging (red) in 2020 for four representative households: H1 and H2 (with PV, PV production in orange) and H6 and H8 (without PV). The dashed blue line is the mean state of charge (right axis).

Table 5. Measured operation in 2020 by household group.

	PV Households (10)	Non-PV Households (10)
Consumption (MWh)	48.9	51.1
PV production (MWh)	19.0	0.0
Grid import (MWh)	38.8	58.3
Import/consumption (%)	79	114
Battery charge (MWh)	16.9	25.1
Battery round-trip (%)	69	73
Hour of maximum charging	01–03 (8), 12–14 (2)	06–07 (9), 02 (1)
Mean SOC (%)	45	38

Table 6. Day-ahead load forecast accuracy over the test period (hourly values). Skill is the reduction in MAE relative to persistence. Bold indicates the final model used in the dispatch, which has the lowest MAE, RMSE and nMAE.

Forecast	MAE (kWh)	RMSE (kWh)	nMAE (%)	MAPE (%)	R^2	Retrains	Skill (%)
Persistence (same hour, previous day)	2.484	3.433	21.92	22.51	0.411	–	0.0
Seven-day same-hour mean	2.320	3.110	20.47	21.71	0.517	–	6.6
LightGBM, base features	2.088	2.860	18.42	19.05	0.591	24	15.9
LightGBM, + calendar events	2.090	2.860	18.44	19.04	0.591	25	15.9
LightGBM, + seven-day mean	2.079	2.841	18.34	18.95	0.597	24	16.3
LightGBM, all features, no drift	2.218	2.945	19.57	21.22	0.567	0	10.7
LightGBM, all features, tuned	2.093	2.856	18.47	19.07	0.592	20	15.7
XGBoost, all features, tuned	2.174	2.892	19.18	20.64	0.582	18	12.5
LightGBM, all features	**2.066**	**2.810**	**18.22**	**18.99**	**0.605**	**20**	**16.9**

The additional features made little difference: adding the calendar events and the seven-day mean lowered the nMAE from 18.42% to 18.22%. Drift-adaptive retraining had a larger effect. Without it, the nMAE rose to 19.57% and the MAPE to 21.22%. Figure 6 shows the daily error and the 20 retraining events. Each event followed a sharp rise in the daily error, for example, at the start of the Level 5 period in late October. Six of the events fell on the first day of a month (June to October and December), when the month feature took a value not present in the training data; after retraining on the expanded history, the error returned to its usual level. Of the 20 events, 14 met both conditions of Equation (3), three only the relative condition, and three only the deviation condition. The three events found only by the deviation condition had error ratios e_g/μ_g of 1.46–1.52, below κ_1, and standardised deviations of 3.4–4.3. Table A2 (Appendix A) shows the effect of the drift thresholds. Lower thresholds triggered more retraining events, up to 37, but the MAE changed only between 2.066 and 2.075 kWh, within 0.5% of the default setting, and every setting remained more accurate than the seven-day mean (2.320 kWh). The deviation condition added no events for $\kappa_1 = 1.4$, three for $\kappa_1 = 1.6$ and eight for $\kappa_1 = 1.8$. It therefore acts as a safeguard against isolated outliers when the relative condition is less sensitive, and it did not change the accuracy: with $\kappa_1 = 1.6$, the relative condition alone gave 17 events and an MAE of 2.064 kWh. The settings $\kappa_2 = 2.5$ and $\kappa_2 = 3.0$ gave identical results, because every day found only by the deviation condition had a standardised deviation above 3.0. The thresholds were fixed before the test period, and the conclusions on forecast accuracy do not depend on this choice.

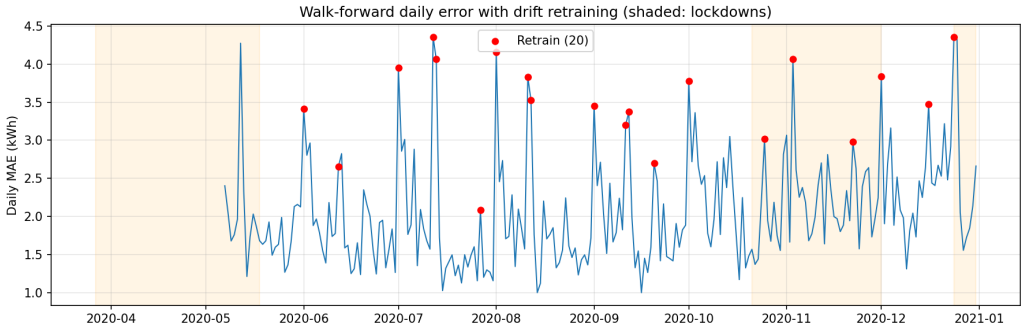


Figure 6. Daily MAE of the load forecast over the test period. Red dots mark drift-triggered retraining; shaded areas mark the lockdown periods.

Figure 7a shows a typical summer week. The forecast follows the daily shape well, but it misses unusually high evening peaks, such as the 32 kWh peak on 12 July, which it predicted at about 16 kWh. The hourly error profile in Figure 7b confirms this pattern. The error is lowest around midnight (about 1.2 kWh), rises at the morning ramp (2.8 kWh at 07:00), and is highest at the evening peak (2.9 kWh at 17:00). LightGBM was more accurate than the seven-day mean in 23 of the 24 h, and the two were equal within 1% at 00:00. The gain was uneven: 19–28% at 01:00–03:00 and 05:00, where the recursive scheme still has recently measured values, and only 8–14% between 17:00 and 19:00. These missed evening peaks mark the main limitation of a forecast based solely on local data.

For peak-oriented dispatch, errors during high-load hours matter more than the average hourly error. Here, the ranking of the forecasts is reversed. The MAE of the forecast daily maximum was 3.43 kW for LightGBM and 3.26 kW for the seven-day mean. Both forecasts underestimated the daily maximum on average, LightGBM by 2.86 kW and the seven-day mean by 2.00 kW. Most of the LightGBM error on the daily maximum is therefore a systematic bias rather than random scatter. The timing of the peak was equally uncertain for both forecasts: the forecast peak hour differed from the measured one by 2.3 h on average and matched it on only 28% of the days. The validation-tuned models confirm these results. The LightGBM setting selected on the validation days (15 leaves, learning rate 0.03, 200 trees) was less accurate on the test period than the default setting (MAE 2.093 against 2.066 kWh), and the default was therefore retained. The validation days fall entirely within the first lockdown, and such a short and atypical window does not allow for reliable hyperparameter selection. The tuned XGBoost model reached an MAE of 2.174 kWh, between LightGBM and the seven-day mean. For the daily maximum, defined as the largest hourly load of each test day, the seven-day mean had the smallest mean absolute error (3.26 kW), followed by XGBoost (3.36 kW), LightGBM (3.43 kW), and the tuned LightGBM (3.69 kW); with hourly resolution, an hourly energy of 1 kWh corresponds to an average power of 1 kW. The average underestimation of the daily maximum by XGBoost (2.58 kW) lay between those of the seven-day mean (2.00 kW) and LightGBM (2.86 kW). The ranking of the forecasts therefore depends on whether hourly or peak accuracy is measured, and the metrics for high-load hours are more relevant for peak-oriented dispatch (Sections 4.3.7 and 5.1). Deep-learning forecasters, such as LSTM networks or Temporal Fusion Transformers, were not included. Gradient-boosting models have been reported to match them for aggregated residential load [16,17], and the dispatch results in Section 4.3 show that differences in hourly accuracy have little effect on the dispatch; for such models, too, the relevant test would be their accuracy on the daily peak.

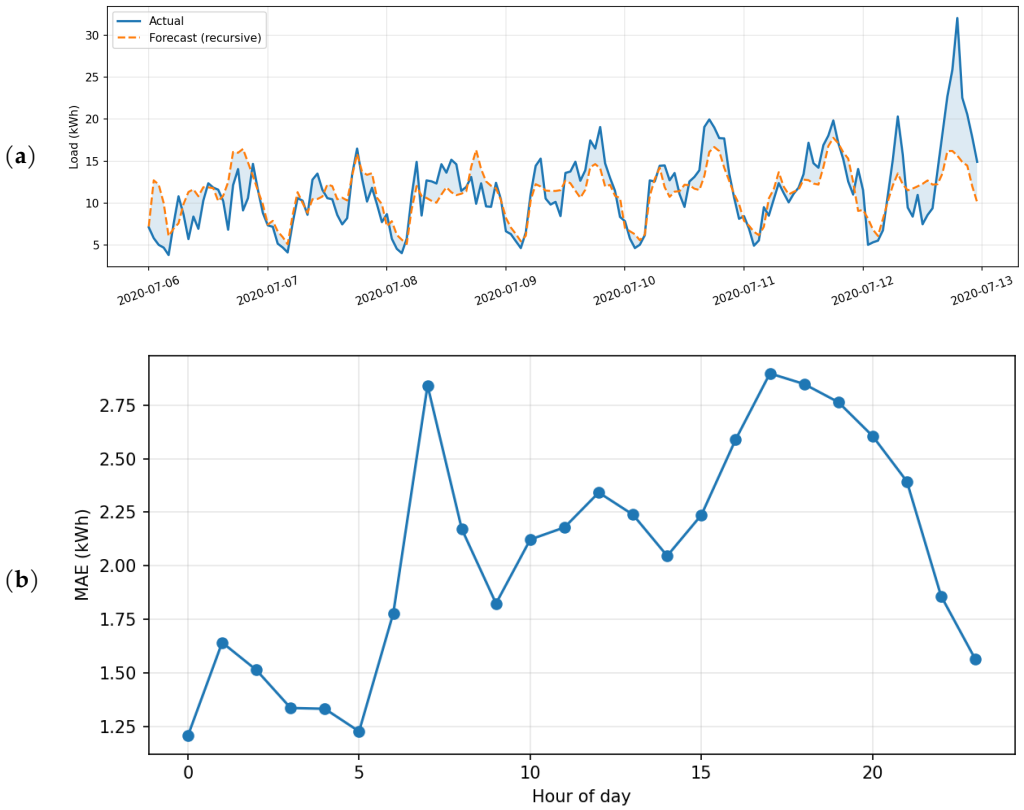


Figure 7. (**a**) Measured and forecast community load for 6–12 July 2020; (**b**) MAE of the LightGBM forecast by hour of the day over the test period.

4.3. Dispatch Scenario Ladder

Table 7 and Figure 8 summarise the dispatch results. The threshold objective is the main case; the peak objective is listed for comparison.

4.3.1. Measured Operation (S1)

Compared with no battery, the measured operation increased grid imports by 14.5% and CO_2 emissions by 15.8%. The maximum import rose from 33 to 83 kW, the mean daily peak from 18.5 to 44.6 kW, and the energy above the threshold almost fivefold. The SSR fell from 18.4% to 6.6%.

4.3.2. Rule with Forecast (S2)

Adding a forecast to a simple rule reduced the import of the measured operation by 1.4%, but the result remained worse than without a battery: 12.9% higher import, a mean daily peak of 31.8 kW, and the largest energy above the threshold among all scenarios. On 178 of the 239 days, the daily peak occurred at 06:00 or 07:00, when the rule completes its night charging and the morning load is rising.

4.3.3. Optimised Dispatch (S3)

With the same forecasts, the MILP yielded a different outcome. Compared with the rule, it reduced grid imports and CO_2 emissions by about 12–13%, the mean daily peak by 48%, the energy above the threshold by 91%, and the hours above it by 75%. Compared with the measured operation, imports fell by 13.5%, emissions by 14.6%, and energy above the threshold by 87%, while the SSR rose from 6.6% to 19.2%. Compared with no battery, S3 reduced the energy above the threshold by 39% and the hours above it by 43%. The import of S3, however, was only 0.9% below S0, which is close to the physical limit given by

Equation (15). Over the test period, the net demand was 52,089 kWh, and the PV surplus of 949 kWh could return at most about 743 kWh after conversion losses, so no strategy that ends the test period with its initial stored energy could import less than about 52,300 kWh. Because the terminal condition, Equation (12), is imposed on the plan but not enforced after the real-time correction, S3 ended the test period with 20.0 kWh stored, against 69.4 kWh at its start (76 of the 239 days ended below their starting level); this lowers its limit by about 44 kWh after discharge losses, and S3 was 0.6% above the adjusted limit. The simulated scenarios in Table 7 do not include the standby consumption of the battery systems (about 0.31 kW; Section 3.2), which is accounted for in the measured operation. When it is simulated explicitly (column "Import incl. standby" in Table 7 and Section 4.3.6), the import of every battery scenario rises by about 1.8 MWh, and S3 imports 10.5% less than the measured operation and 2.4% more than S0.

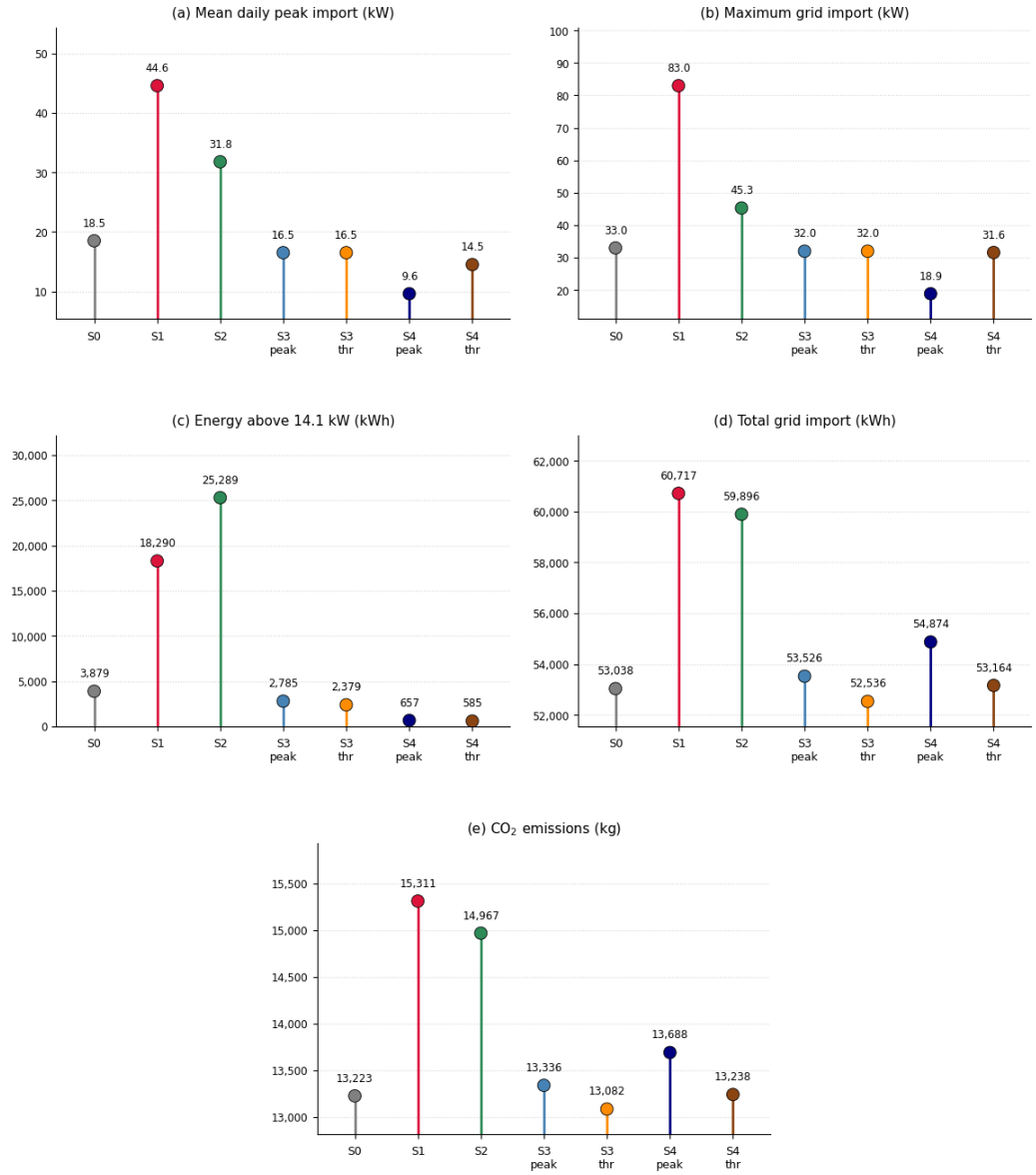


Figure 8. Dispatch scenario ladder: (**a**) mean daily peak import; (**b**) maximum grid import; (**c**) energy imported above the 14.1 kW threshold; (**d**) total grid import; (**e**) CO_2 emissions. S0: no battery; S1: measured operation; S2: rule + forecast; S3: MILP with forecasts; S4: ideal-information benchmark; peak/thr: peak or threshold objective.

Table 7. Dispatch results over the test period (239 days). E_{el}^{exc}: energy imported above the threshold of 14.1 kW; h^{exc}: hours above the threshold. Import incl. standby: import with the standby consumption of the battery systems (0.31 kW) simulated, computed for the threshold objective (Section 4.3.6); S1 includes it by measurement.

Scenario	Import (kWh)	Import Incl. Standby (kWh)	Max Peak (kW)	Daily Peak (kW)	E_{el}^{exc} (kWh)	h^{exc} (h)	CO_2 (kg)	SCR (%)	SSR (%)
S0: No Battery	53,038	53,038	33.0	18.5	3879	967	13,223	92.7	18.4
S1: Measured	60,717	60,717	83.0	44.6	18,290	1169	15,311	95.6	6.6
S2: Rule + Forecast	59,896	61,747	45.3	31.8	25,289	2168	14,967	99.4	7.9
MILP, threshold objective									
S3: LightGBM Forecast	52,536	54,317	32.0	16.5	2379	551	13,082	100	19.2
S3n: Seven-day Mean Forecast	52,658	54,446	32.0	16.5	2362	570	13,107	100	19.0
S3b: LightGBM + Realised PV	52,619	–	32.0	16.9	2471	604	13,108	100	19.1
S4: Ideal-information Benchmark	53,164	54,968	31.6	14.5	585	95	13,238	100	18.2
MILP, peak objective									
S3: LightGBM Forecast	53,526	–	32.0	16.5	2785	712	13,336	100	17.7
S3n: Seven-day Mean Forecast	53,607	–	32.0	16.3	2741	756	13,354	100	17.6
S3b: LightGBM + Realised PV	53,704	–	32.0	16.1	2415	613	13,389	100	17.4
S4: Ideal-information Benchmark	54,874	–	18.9	9.6	657	432	13,688	100	15.6

Figure 9 shows a sample day. The measured batteries drew 36 kWh from the grid at 06:00, three times the load. The optimised plans spread the charging over hours with spare capacity and kept the import at or below the threshold for most of the day; the remaining exceedance occurred at the evening peak. The duration curve in Figure 10 shows the same picture over the whole period. S1 and S2 import more than 30 kW for hundreds of hours, S3 lies below S0 over the highest-import hours, and S4 holds the import at the threshold after its first 95 h.

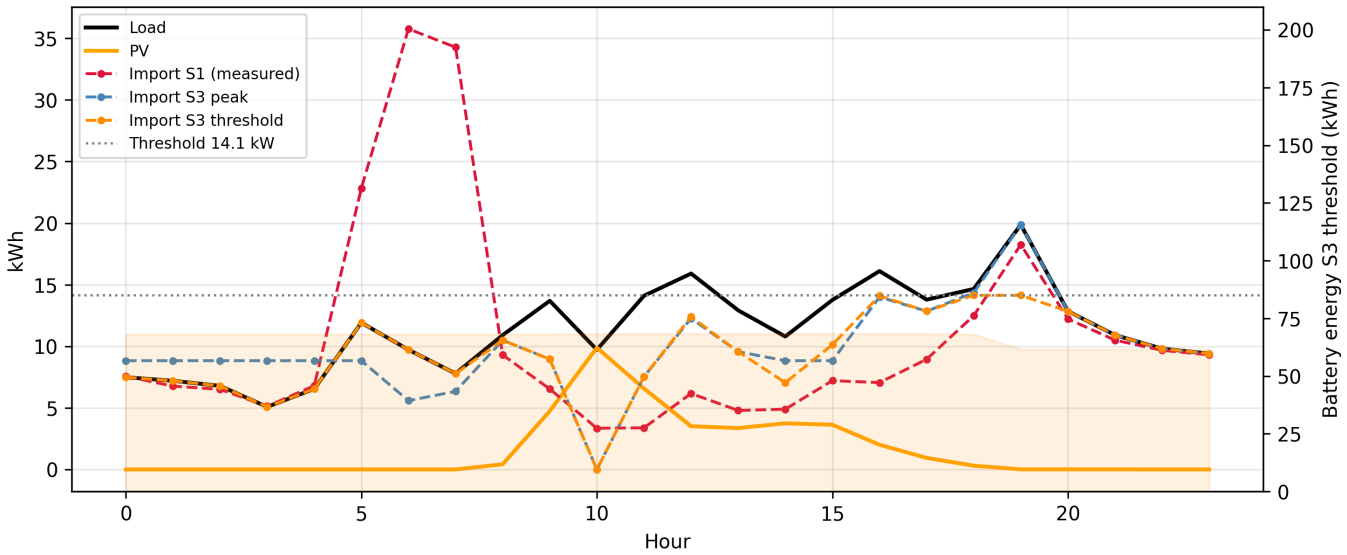


Figure 9. Load, PV, and grid import on 10 November 2020 for the measured operation (S1) and the optimised dispatch with both objectives (S3). The shaded area represents the stored energy below the threshold objective; the dotted line represents the threshold.

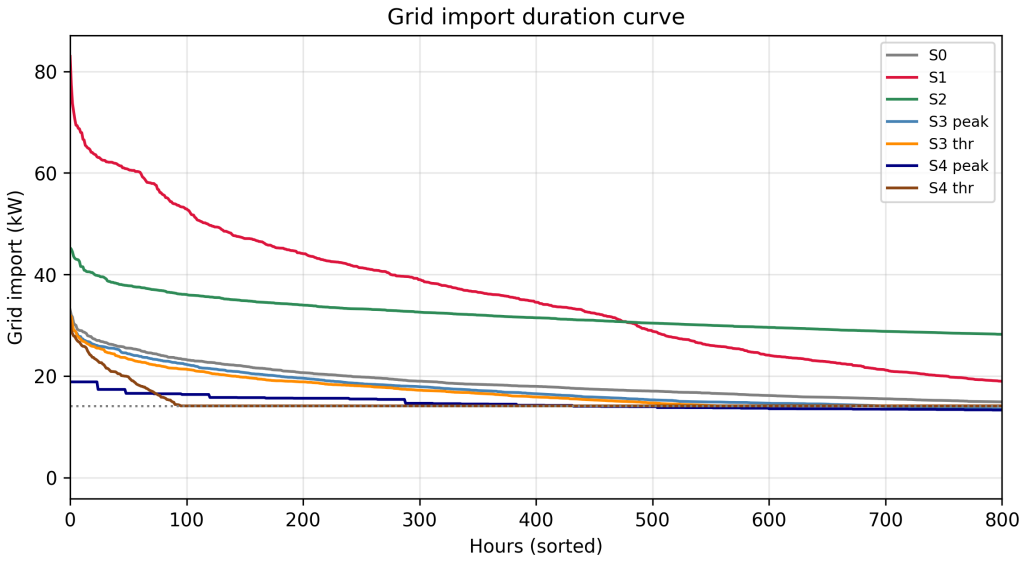


Figure 10. Duration curve of hourly grid import (highest 800 h of the test period). The dotted line marks the 14.1 kW threshold. S0: no battery; S1: measured operation; S2: rule + forecast; S3: MILP with forecasts; S4: ideal-information benchmark; peak/thr: peak or threshold objective.

4.3.4. Choice of Objective

The threshold objective performed better than the peak objective on every energy metric: 1.8% less import, 15% less energy above the threshold, and a higher SSR (19.2% against 17.7%), with a same mean daily peak (16.5 kW). The peak objective charges and discharges more to flatten each day, and the extra losses increase the import. This trade-off is most visible with ideal information: the peak objective reached the lowest daily peak (9.6 kW) but also the highest import of all optimised cases (54,874 kWh).

4.3.5. Sensitivity to the Import Weight

The weight w in Equations (13) and (14) was varied between 0.01 and 0.5 for S3 and S4 (Table A3 in Appendix A). With the threshold objective, the results were practically insensitive to w. The import of S3 changed by less than 0.02%, and the energy above the threshold changed by about 1.5% at most. S4 gave the same import and exceedance for every value; only its hours above the threshold varied, between 70 and 95. With the peak objective, the results did not change up to $w = 0.2$. At $w = 0.5$, the import term was no longer a secondary penalty: the import fell by 1.1% for S3 and 2.3% for S4, but the energy above the threshold rose by 7% and 141%, and the mean daily peak of S4 rose from 9.6 to 11.5 kW. The ranking was the same for every w. The threshold objective placed less import and less energy above the threshold than the peak objective, and S4 always yielded the lowest exceedance. The charged energy supports the explanation given above. With $w = 0.1$, S3 charged 2264 kWh under the threshold objective and 6833 kWh under the peak objective and S4 charged 4962 and 12,850 kWh. The default value $w = 0.1$ therefore lies in the range where the import term remains secondary, and the conclusions do not depend on it.

4.3.6. Sensitivity to Battery Losses

The battery model assumes a one-way efficiency of 0.885 and omits standby consumption for battery systems (Section 3.5). Table A4 (Appendix A) repeats the threshold-objective scenarios with the estimated standby consumption of 0.31 kW added to the load of every battery scenario, with a standby consumption 20% higher (0.37 kW), and with the standby consumption combined with a lower one-way efficiency of 0.85. With the standby consumption, S3 imported 54,317 kWh, 10.5% less than the measured operation and 2.4%

more than S0. With the higher standby consumption and the lower efficiency, the import of S3 was at most 3.1% above S0 and still about 10% below the measured operation, and its energy above the threshold remained 28–30% lower than with no battery and about 85% lower than in the measured operation. The ranking of the scenarios did not change in any case. The rule-based dispatch S2 remained the worst and was the most sensitive to losses, because it cycles the most energy; S3 and S3n remained practically equal; and S4 kept the lowest exceedance. The conclusions drawn from the scenario ladder are therefore robust to these assumptions.

4.3.7. Value of the Forecasts

Replacing the LightGBM load forecast with the seven-day mean (S3n) did not measurably affect dispatch (Table A5). With the threshold objective, the mean daily peak was 16.55 kW with LightGBM and 16.52 kW with the benchmark (Wilcoxon $p = 0.53$; mean paired difference 0.03 kW, 95% block-bootstrap interval −0.15 to 0.22 kW), and the daily energy above the threshold was 9.96 and 9.88 kWh ($p = 0.50$; 0.07 kWh, −0.47 to 0.63 kWh). With the peak objective, the benchmark was slightly better: 16.35 against 16.52 kW ($p = 0.13$) and 11.47 against 11.65 kWh above the threshold ($p = 0.49$). Only 82–117 of the 239 days had a nonzero difference. The differences were serially correlated (lag-1 autocorrelation 0.12–0.41; Ljung–Box $p < 0.001$ in three of the four cases), which reduces the effective sample size to about 100–190 days. For the daily peak under the peak objective, the independent bootstrap interval of the mean difference (0.02 to 0.34 kW) excluded zero, but the dependence-aware interval (−0.01 to 0.38 kW) did not. The differences were nearly symmetric, except for the energy above the threshold under the peak objective (skewness 1.17, $p < 0.001$); for this quantity, the sign test gave $p = 0.70$. For all four quantities, the 90% block-bootstrap interval lay within the equivalence margins, so the two forecasts led to a practically equivalent dispatch.

A realised PV forecast (S3b) did not help either. The ideal-information benchmark (S4), by contrast, reduced the energy above the threshold from 2379 to 585 kWh and the hours above it from 551 to 95. The optimised dispatch with realistic forecasts thus achieved 46% of the reduction in exceedance that was possible relative to S0, and the remaining gap to S4 is due to the load forecast; Section 4.3.8 shows that this gap persists when the real-time correction uses delayed information. S4 imported 628 kWh (1.2%) more than S3, and two effects account for this difference. With the evening peaks known in advance, S4 charged the battery more to keep the import below the threshold: 4962 kWh against 2264 kWh for S3 (Table A3 in Appendix A). At a round-trip efficiency of about 78%, the additional 2698 kWh of charging caused about 584 kWh of extra conversion losses. In addition, S3 ended the test period with 20.0 kWh in storage, 49.4 kWh less than at the start, and this energy covered about 44 kWh of its load after discharge losses, whereas S4 returned to its initial 69.4 kWh. Together, the two effects amount to about 628 kWh. This result should not be read as a loss of performance caused by better information. On the quantity that the threshold objective targets, S4 reduced the energy above the threshold from 2379 to 585 kWh (75%) and the hours above it from 551 to 95. The additional import of 1.2% is the price of this reduction, because the objective prioritises energy above the threshold and treats the total import only as a secondary term.

To isolate the contribution of the peak error, the LightGBM forecast was modified with Equation (23) and the dispatch was repeated with the same controller (Table 8). Setting the forecast daily maximum to the measured one, with the shape and timing of the forecast unchanged, reduced the energy above the threshold from 2379 to 1892 kWh (20%) and the hours above it from 551 to 479. This closes 27% of the gap to S4. The remaining 73% comes from errors in the hourly profile that the correction leaves unchanged, such as the

duration of the high-load periods and the distribution of energy across the evening. The causal correction, which uses only the errors of previous days, reduced the energy above the threshold by 14%, to 2051 kWh, at a 0.5% higher import. With it, the dispatch based on LightGBM was 13% better than the one based on the seven-day mean (2362 kWh); applied to the seven-day mean, the same correction gave 2237 kWh (5% lower). Over the test days, the daily exceedance of S3 correlated more strongly with the forecast error of the net energy above the threshold (Spearman $\rho = 0.61$) and of the evening energy ($\rho = 0.52$) than with the error of the daily maximum ($\rho = 0.50$), and it did not correlate with the timing error of the peak ($\rho = -0.06$). The underestimation of the daily peaks therefore explains part of the limited value of the more accurate forecast but not all of it.

Table 8. Contribution of the daily peak error (threshold objective, same-hour correction, 239 days). Causal correction: Equation (23) with the mean underestimation of the previous 28 days; controlled: forecast daily maximum set to the measured one (diagnostic only).

Forecast Used by the MILP	Import (kWh)	E_{el}^{exc} (kWh)	h^{exc} (h)	Daily Peak (kW)
S3: LightGBM	52,536	2379	551	16.5
S3: LightGBM, causal peak-bias correction	52,791	2051	497	16.3
S3: LightGBM, measured daily maximum (controlled)	52,812	1892	479	16.1
S3n: Seven-day mean	52,658	2362	570	16.5
S3n: Seven-day mean, causal peak-bias correction	52,846	2237	549	16.3
S4: Ideal-information benchmark	53,164	585	95	14.5

4.3.8. Execution Timing and a Forecast-Free Controller

Two further cases separate the contributions of the forecast and of the real-time correction (Table 9). The first is a forecast-free real-time threshold controller (S_{RT}), which charges whenever the net load is below the threshold, up to the threshold, and discharges whenever it is above it, without any forecast or optimisation. The second repeats each controller with a one-hour delay in the information used by the correction, so every decision uses the net load of the previous hour.

With same-hour information, S_{RT} gave the lowest energy above the threshold of all scenarios, 418 kWh, lower even than S4. Most of this result, however, came from the same-hour information. With a one-hour delay, its exceedance rose sixfold to 2532 kWh, while S3 rose by 28% to 3044 kWh and S4 to 1333 kWh. The forecast-free controller depends most on same-hour information, because it reacts to the measured load alone, whereas the optimised dispatch prepares the battery in advance. With delayed information, S_{RT} and S3 exceeded the threshold in a similar number of hours (915 and 893): S_{RT} with 17% less energy above the threshold and S3 with 2.6% less grid import and a higher SSR (19.2% against 17.1%).

To compare the controllers on a common basis, Table 9 also reports the stored energy at the end of the test period, the import adjusted to the initial stored energy, Equation (26), and the realised objective, Equation (27), with and without the standby consumption. All controllers except S4 with same-hour information ended the test period at the lower state-of-charge limit, 20.0 kWh. The adjustment therefore adds the same 56 kWh to each of them and does not change their ranking. On the realised objective, S_{RT} did better than S3 and S3n in all four cases, for example, 7931 against 8303 kWh with delayed information. The benefit of the batteries thus has two parts: coordinating them against a capacity signal reduces the energy above the threshold, and the forecast-driven optimisation adds a lower import and lower conversion losses. The ideal-information benchmark remained the best case with delayed information, with 56% less energy above the threshold than S3, so better forecasts of the high-load hours still have substantial value. Because S_{RT} outperformed S4 with the

same-hour information, S4 is not an upper bound for every controller; it is the best result of the daily MILP with a terminal-energy constraint. The advantage over the measured operation remained large in all cases: with delayed information, S3 still reduced the energy above the threshold by 83% relative to S1.

Table 9. Effect of the execution timing and of the standby consumption (threshold objective, 239 days). Same hour: the real-time correction uses the net load of the current hour (main case); one-hour delay: it uses the net load of the previous hour. E_T: stored energy at the end of the test period (E_0 = 69.4 kWh); $E_{el}^{PG,adj}$: import adjusted to E_0, Equation (26); J: realised objective, Equation (27). S_{RT}: real-time threshold controller without forecast or optimisation.

Scenario	Import (kWh)	$E_{el}^{PG,adj}$ (kWh)	E_{el}^{exc} (kWh)	h^{exc} (h)	E_T (kWh)	J (kWh)
Reference						
S0: No battery	53,038	–	3879	967	–	9183
S1: Measured [1]	60,717	60,773	18,290	1169	20.0	24,367
Same hour						
S_{RT}: Real-time rule, no forecast	53,938	53,994	418	62	20.0	5817
S3: LightGBM forecast	52,536	52,592	2379	551	20.0	7638
S3n: Seven-day mean forecast	52,658	52,713	2362	570	20.0	7633
S4: Ideal-information benchmark	53,164	53,164	585	95	69.4	5901
Same hour, standby 0.31 kW						
S_{RT}: Real-time rule, no forecast	55,670	55,726	521	72	20.0	6094
S3: LightGBM forecast	54,317	54,373	2703	608	20.0	8140
S3n: Seven-day mean forecast	54,446	54,502	2715	660	20.0	8165
S4: Ideal-information benchmark	54,968	54,968	717	110	69.4	6214
One-hour delay						
S_{RT}: Real-time rule, no forecast	53,932	53,988	2532	915	20.0	7931
S3: LightGBM forecast	52,533	52,589	3044	893	20.0	8303
S3n: Seven-day mean forecast	52,633	52,688	3002	897	20.0	8271
S4: Ideal-information benchmark	52,989	53,044	1333	384	20.0	6637
One-hour delay, standby 0.31 kW						
S_{RT}: Real-time rule, no forecast	55,665	55,721	2773	984	20.0	8345
S3: LightGBM forecast	54,304	54,360	3377	963	20.0	8813
S3n: Seven-day mean forecast	54,424	54,480	3343	980	20.0	8791
S4: Ideal-information benchmark	54,778	54,834	1499	422	20.0	6982

[1] Measured stored energy and measured losses; the stored energy of H14 is missing from 12 December.

4.4. Energy Sharing and Storage Pooling

Table 10 and Figure 11 compare the sharing configurations under the threshold objective. The main columns use the seven-day forecast; the last two columns repeat the key metrics with ideal information.

Table 10. Sharing configurations under the threshold objective. E_{el}^{exc} and h^{exc} refer to the community threshold of 14.1 kW.

Config.	Seven-Day Forecast						Ideal Information	
	Import (kWh)	Daily Peak (kW)	E_{el}^{exc} (kWh)	h^{exc} (h)	CO_2 (kg)	SSR (%)	Daily Peak (kW)	E_{el}^{exc} (kWh)
A: no battery, individual	58,702	18.7	3938	995	14,694	9.7	18.7	3938
B: no battery, shared PV	53,038	18.5	3879	967	13,223	18.4	18.5	3879
M: measured, individual	62,623	44.6	18,339	1179	15,810	3.7	44.6	18,339
C: own battery	56,572	16.3	2558	777	14,122	13.0	15.6	2669
D: own battery + sharing	54,324	16.1	2542	768	13,535	16.5	15.6	2661
E: pooled battery	52,656	16.5	2361	570	13,106	19.0	14.5	585

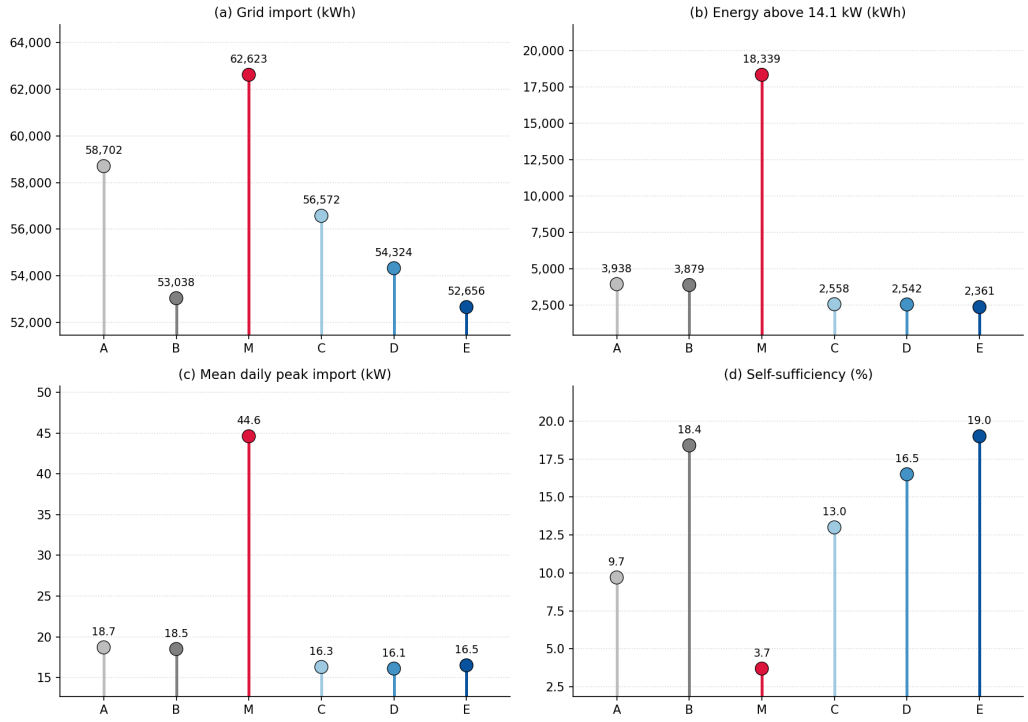


Figure 11. Sharing configurations under the threshold objective with the seven-day forecast: (**a**) grid import; (**b**) energy above the 14.1 kW threshold; (**c**) mean daily peak import; (**d**) self-sufficiency. A: no battery; B: shared PV; M: measured operation; C: own battery; D: own battery + sharing; E: pooled battery.

Hourly virtual sharing of PV without any battery (A to B) reduced the calculated residual community import by 9.7% relative to the sum of individual household imports and almost doubled the SSR, from 9.7% to 18.4%, because the export fell from 6614 to 949 kWh. It did not relieve the grid at critical times: the energy above the threshold fell by only 1.5%, since the PV surplus occurs at midday and the excess occurs in the evening. Individual batteries (C) had the opposite effect. They cut energy use above the threshold by 35% but imports by only 3.6%, because households without PV had little to store. Sharing the surplus of these batteries (D) lowered imports by a further 4.0% but left the excess unchanged. Pooling PV and storage (E) combined both effects. It gave the lowest import (10.3% below A), the highest SSR (19.0%), and the lowest energy above the threshold. Its advantage over individual batteries in grid terms was, however, small with realistic forecasts: 8% less energy above the threshold and 27% fewer hours above it. With ideal information, the advantage became large: the pooled battery reduced the energy above the threshold to 585 kWh, 78% less than individual batteries.

The value of pooling storage for the grid therefore depends strongly on the information available to the controller (Section 4.3.8). With the peak objective, the pooled battery did not outperform individual batteries at all with the seven-day forecast, 2741 against 2439 kWh above the threshold. Configuration E with the seven-day forecast is the same as S3n in Table 7; the 2 kWh difference in imports comes from the initial SOC, which was clipped to the 10–90% household-level window.

4.5. Distribution of Benefits

Table 11 and Figure 12 show how the benefits were distributed between the two household groups under the allocation rules in Section 3.7.

Table 11. Self-sufficiency and change in grid import relative to configuration A, by household group (threshold objective, seven-day forecast). Negative SSR values indicate that grid import exceeded consumption because of battery conversion losses.

Config.	PV Households		Non-PV Households	
	SSR (%)	Import Change (%)	SSR (%)	Import Change (%)
A: no battery, individual	20.4	0.0	0.0	0.0
B: no battery, shared PV	23.2	−3.5	14.1	−14.1
M: measured, individual	23.1	−3.3	−13.9	+13.9
C: own battery	29.6	−11.5	−2.1	+2.1
D: own battery + sharing	30.6	−12.8	3.6	−3.6
E: pooled battery	24.0	−4.6	14.5	−14.5

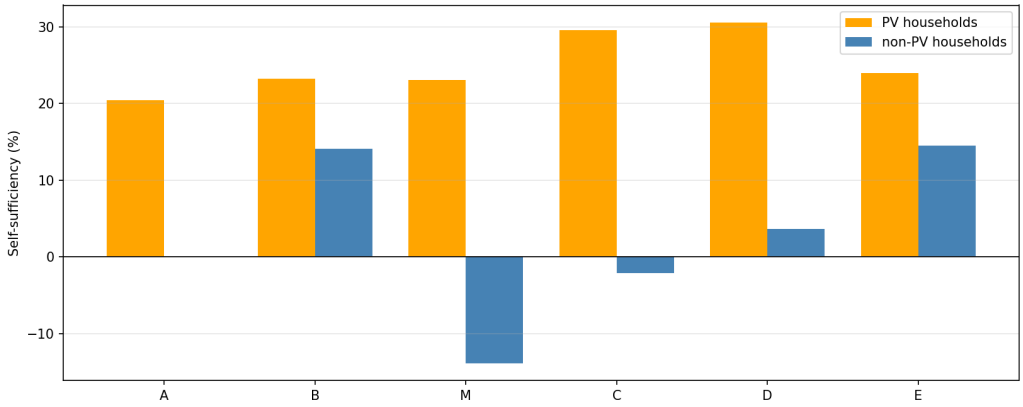


Figure 12. Self-sufficiency of PV and non-PV households in each configuration (threshold objective, seven-day forecast).

Under the measured operation, the non-PV households were worse off than with no battery at all: their batteries charged only from the grid, and conversion losses increased their imports by 13.9%. An optimised battery of their own (C) still increased their imports slightly (2.1%), whereas the PV households gained most from their own batteries, reaching an SSR of 29.6%. Energy sharing reversed this picture. Sharing PV alone (B) reduced imports for non-PV households by 14.1%, and the pooled configuration (E) by 14.5%. Almost all of the benefits that households without PV received, therefore, came from sharing energy, not from pooling storage. The PV households, in turn, reached an SSR of 24.0% under pooling, compared with 29.6% with their own batteries.

4.6. CO_2 Emissions

Because the carbon intensity of the Irish grid varied little over the day (Section 3.8), the CO_2 results follow the grid import closely (Tables 7 and 10). With all scenarios netted hourly at the community level, the measured operation emitted 15.8% more than the no-battery case, and the optimised dispatch emitted 14.6% less than the measured operation. Under the adopted accounting convention, the virtual pool of PV and storage (E) showed 10.8% lower emissions associated with imported electricity than individual operation without batteries (A); since E is netted hourly at the community level and A is not, this difference combines allocated accounting benefits with operational changes and should not be read as a physical reduction of system-wide emissions. The timing of the import made only a small difference. The effective intensity of the imported electricity was highest for the measured operation (0.2522 kg/kWh), which charged the batteries in the early morning, when the grid intensity was highest, and lowest for the optimised dispatch (0.2489 kg/kWh). The 1.3% difference is small compared with the differences in imports between the scenarios.

5. Discussion

5.1. Forecast Accuracy Is Not Operational Value

The most surprising result of this study concerns the forecast. The LightGBM model was clearly more accurate than the simple benchmark, by 10.9% in hourly MAE and in 23 of the 24 h of the day, yet the dispatch it produced showed no statistically significant difference (mean daily peak 16.55 against 16.52 kW, Wilcoxon $p = 0.53$), and the paired daily differences lay within the equivalence margins. This agrees with the finding of Putz et al. [9] that improved accuracy metrics do not translate into proportional operational gains in energy communities, and the present results explain why this occurs in a peak-oriented setting. The optimiser uses the hourly forecast in its energy balance and state-of-charge constraints, but the outcome depends mostly on the high-load hours: the forecast decides how much energy is kept for the hours in which the import would exceed the target and when the battery is charged. On the daily maximum, the model was worse than the benchmark and underestimated it by 2.86 kW on average. This bias explains part of the result, not all of it. Removing it in a controlled test reduced the energy above the threshold by 20% and closed about a quarter of the gap to ideal information (Section 4.3.7); the rest is linked to the duration and shape of the high-load periods.

Three mechanisms contribute to the limited value of the more accurate model. First, the model is trained to minimise the absolute error, so it predicts something close to the conditional median of the load. This smooths out the rare, high evening peaks that matter most for the battery. The seven-day mean also smooths, but less, because it averages past days that contained such peaks. Second, the forecast is recursive: each hour of the 24-h horizon uses the forecasts of the earlier hours as lagged inputs, so errors accumulate along the horizon. Because the plan is made at midnight, the evening peak is forecast 17 to 19 h ahead, with the largest accumulated error. A model that already smooths high values passes these smoothed values on as inputs for the following hours. This is consistent with the error profile in Figure 7b: the gain over the benchmark was largest after midnight (19–28% at 01:00–03:00 and 05:00) and smallest at the evening peak (8–14% between 17:00 and 19:00), which is where the dispatch needs it. The evening load is also more variable, so the present design cannot fully separate the two effects; a comparison with a direct multi-horizon forecast would do so. Third, the model uses only historical load and calendar features. Evening peaks in a residential community depend on weather and occupancy, and without exogenous inputs the model has no information about the conditions that produce an unusually high peak. This is the main bottleneck of a framework based solely on local data and the price of its independence from external data services. The real-time correction then absorbed much of the remaining difference between the two forecasts.

Two practical points follow. First, a forecast for battery dispatch should be judged using the metrics the controller depends on, such as daily peak error and resulting exceedance, rather than only hourly MAE or MAPE. The seven-day mean illustrates this point: it was the least accurate forecast hour by hour, yet it had the smallest daily peak error, and a causal correction of the peak bias improved the dispatch more than the gain in hourly accuracy did. Second, forecasting still offers considerable value. Ideal information reduced the energy above the threshold by a further 56–75%, depending on the timing of the real-time correction, and almost all of this gap came from the load forecast rather than PV. The missed peaks show where the local-data-only approach reaches its limit. Weather variables are the first candidates for improving peak prediction, such as temperature and irradiance, from official meteorological services or from sensors installed in future smart-meter infrastructure. They were excluded here on purpose to test a framework based solely on local data, and their value to the dispatch remains to be quantified. The gain will come from forecasts that capture the size and timing of peaks, for example, models trained on a peak-weighted or quantile loss, or trained directly

through the dispatch objective as in task-based learning [10]. For the present community, a seven-day average is enough to run the optimiser as well as a machine-learning model does, which lowers the entry barrier for operators.

5.2. Control Changes the Timing of Grid Use, Sizing Changes the Amount

The measured batteries worsened the community's performance in grid import, peak import, energy above the threshold, emissions, and self-sufficiency. They were not faulty; they followed a reasonable rule under a day/night tariff and charged from the grid at night. Because all twenty units followed the same rule, their charging coincided, and the community import at 06:00–07:00 reached almost three times the load. Each battery was sensible on its own, but together they created a new morning peak. Similar default self-consumption controllers are common in residential storage [5], so this effect is likely to appear in other communities where many batteries respond to the same price signal. With the Irish grid's nearly flat carbon intensity in 2020, night charging also provided no CO_2 benefit; the measured operation had the highest effective intensity of all scenarios.

Optimisation reversed this, but the size of its gain depends on the metric. For grid exchange at critical hours, the gain was large: 39% less energy above the threshold than with no battery, or 23% when the real-time correction uses delayed information. For the total import, there was practically no gain: the optimised dispatch imported 0.9% less than S0 without battery standby consumption and about 2.4% more with it. The bound in Equation (15) explains this. A battery cannot create energy; it can only move it over time, and it can reduce imports only by storing PV surplus that would otherwise be exported. In this community, PV covered 19% of demand, and the surplus over 239 days was under 1 MWh, and the optimised dispatch came within 0.6% of the resulting limit. Better control cannot go much further. Higher self-sufficiency requires more PV, while well-operated batteries are a tool for grid relief. The choice between the two objectives reflects the same trade-off: flattening every daily peak incurs additional cycling and losses, while a reference threshold asks the battery to act only when it matters and yields better results across all energy metrics. The rule-based controller S2 is deliberately simple; a more refined heuristic could narrow its gap to the MILP, but the optimisation serves different objectives without hand-tuned rules. The forecast-free controller S_{RT} (Section 4.3.8) shows how much of this effect comes from the capacity signal itself: acting only on the measured load against the threshold, it reduced the energy above the threshold at least as much as the optimised dispatch, with same-hour and with delayed information, and had a lower realised objective than S3 in every case, at the cost of a 2.6% higher import. For an operator without forecasting tools, coordinating the batteries against a capacity signal is therefore the essential step; forecasting and optimisation add a lower import and, with bias-corrected peak forecasts, further grid relief (Section 4.3.7).

5.3. Energy Sharing and Storage Pooling Do Different Things

The sharing analysis separates two effects that are often reported together. Hourly virtual sharing of PV without any battery reduced calculated residual community import by approximately 9.7% under the adopted accounting convention, and it was the only measure that allocated to households without PV a share of local generation. Pooling storage added little energy on top of this, but it gave the lowest exceedance and the fewest hours above the threshold. Shared storage is often found to outperform private storage [23]. The present results confirm this in principle, since with ideal information, the pooled battery reduced the energy above the threshold by 78% compared with individual batteries. With realistic forecasts, however, the advantage shrank to 8%. The benefit of pooling storage is therefore closely tied to the quality of the forecast that drives it, and studies evaluating shared storage

with ideal information are likely to overstate its benefits. For a community manager, this suggests a priority order. Energy sharing brings most of the energy benefit and requires only metering and settlement. Pooled or coordinated storage adds grid value, and that value grows as the forecast of demand peaks improves.

5.4. Who Benefits

The distribution of benefits raises a question that aggregate results hide. Under the measured operation, the batteries of households without PV increased their grid import by 13.9%, because they charged only from the grid. With hourly virtual sharing, the same households reached an allocated self-sufficiency of about 14%, and almost all of this gain came from the allocated surplus of their PV neighbours. PV owners, on the other hand, were best off with their own optimised battery (SSR 29.6% in C and 30.6% in D) and lost part of this advantage when storage was pooled (24.0%). A virtual pool of batteries is thus better in terms of threshold-exceedance indicators and the allocated benefits to households without PV but not automatically for households that own generation. Households without rooftop PV include those that cannot install it, such as tenants or residents of shaded or shared buildings, although the data do not identify which households are vulnerable. Energy communities can extend the benefits of local generation to their members, but PV owners need a reason to participate. The allocation rules used here apportion energy only.

They are accounting rules applied under hourly virtual netting: they show how the shared energy is assigned to each member, not what anyone should pay or receive, and they are not an economic compensation or settlement scheme. Such a scheme would have to reward the members whose assets create the value. Under pooling, PV owners give up part of their own benefit, and without compensation they have little reason to join a community with shared storage. Cooperative game theory provides suitable tools for this purpose. The Shapley value, for example, assigns each member the average marginal contribution of its PV and storage to the community, and related methods have been proposed for energy communities [11,12]. The energy allocation reported here is a starting point for such a scheme because it shows that the value created by PV owners is primarily energy, while the value created by pooled storage is primarily a reduction in the energy imported above the threshold.

5.5. Regulatory Context

Collective self-consumption and renewable energy communities are enabled by the recast Renewable Energy Directive [3], which Ireland has transposed into national law [30]. The detailed rules for sharing energy among members, such as data access and settlement, determine which of the configurations studied here can be implemented in practice: B and D correspond to virtual sharing of energy through settlement, while E assumes that the batteries of the community can be coordinated as one fleet, which requires an aggregator or community operator with access to the household devices. The results also suggest that tariff design matters. A simple day/night tariff, applied to many identical controllers, produced the harmful morning peak observed in the field. Signals that reflect local grid capacity, such as the reference threshold used in this study, would lead the same batteries to relieve the grid instead.

5.6. Limitations and Future Work

The study has several limitations. It covers one community and one year, and 2020 was shaped by COVID-19 restrictions. The 239-day test period covers late spring through winter but not a full year. With only one year of data, the model could not learn each season before forecasting it; six of the twenty retraining events occurred on the first day of a month,

when the month feature took a new value. Multi-year data would remove this effect and allow the full year to be tested. The findings differ in how far they can be transferred. The synchronised night-time charging of the measured batteries follows from many identical controllers responding to the same tariff signal and is therefore expected wherever default self-consumption controllers operate under a common time-of-use tariff. The gap between hourly accuracy and dispatch value and the distinction between the energy value of sharing and the grid value of pooled storage follow from the structure of the problem, namely an objective driven by daily peaks and a PV surplus that is small relative to demand, and are likely to hold in other residential communities with similar characteristics. The magnitudes, by contrast, are specific to this case: the low PV share (19% of demand), the Irish climate and day/night tariff, the nearly flat carbon intensity of the Irish grid in 2020, and the COVID-19 restrictions all affect the absolute reductions in import, exceedance and emissions. In communities with a larger PV share, storage could reduce imports more, and in grids with a more variable carbon intensity, the timing of charging would have a stronger effect on emissions. The dispatch model treats the community battery as a single unit at hourly resolution and excludes the low-voltage network, so it cannot capture feeder constraints or sub-hourly peaks. The standby consumption of the batteries was estimated from the annual energy balance, and the battery efficiency was estimated from data, assuming equal charging and discharging efficiencies; both were treated as constant, although in practice they vary with state of charge, temperature, and power. Their effect on the results was tested in Section 4.3.6 and did not change the conclusions. Battery degradation was not modelled, because the ageing parameters of the installed units are not available and the study does not include costs. The cycling implied by the dispatch nevertheless indicates its direction. Scaled to one year, the threshold objective corresponds to about 18 equivalent full cycles of the 200 kWh pool, the peak objective to about 52, and the measured operation, which charged 42 MWh in 2020, to about 210. Including degradation would therefore penalise the measured operation and the peak objective most and would strengthen rather than weaken the preference for the threshold objective. The analysis covers energy and CO_2 only; costs and tariffs were not evaluated, and the carbon intensity used is average rather than marginal. The threshold was set at the 80th percentile of past net load as a statistical reference level rather than a verified network capacity, and it was not optimised. Finally, the allocation of benefits relies on simple accounting rules under hourly virtual netting and does not include an economic settlement.

Future work should test weather and other exogenous inputs, peak-aware and probabilistic forecasts, direct multi-horizon forecasts, and deep-learning forecasters such as LSTM networks and Temporal Fusion Transformers, evaluating them on daily peak accuracy and training them on the dispatch objective. The analysis should be extended to several years and communities, with network constraints and an economic evaluation, and compensation mechanisms, for example, based on the Shapley value, should be designed to make pooled storage attractive for PV owners.

6. Conclusions

This study used measured data from a 20-household Irish energy community to test how forecasting, battery control, and energy sharing affect the community's exchange with the grid, using only locally available data and a framework that runs in seconds per scenario. Four conclusions can be drawn.

First, the batteries as they were operated harmed the community. Twenty identical rule-based controllers responded to the same night tariff, charged at the same hours, and shifted the largest grid demand to the early morning. Compared with no batteries, grid import rose by 14.5%, CO_2 emissions by 15.8%, and the mean daily peak more than doubled,

while self-sufficiency fell from 18.4% to 6.6%. Adding a forecast to a simple rule did not solve the problem. Second, optimisation with real-time correction reversed this outcome with the same batteries. Relative to the measured operation, grid import fell by 10.5–13.5% and emissions by 11.6–14.6%, with the lower values including battery standby consumption, and energy imported above the reference threshold fell by 87%, at a computational cost of below one second per day. A forecast-free controller acting on the same threshold reduced the exceedance at least as much, so coordinating the batteries against a capacity signal provides most of the grid relief, and the optimisation adds a lower import. The total import, however, stayed close to that of a community without batteries and within 0.6% of the physical limit set by the small PV surplus. Control decides when the community draws energy from the grid; the amount is set by PV capacity and sharing.

Third, a more accurate forecast did not give a better dispatch. The LightGBM model reduced the hourly error by 10.9% relative to a seven-day average, but the resulting dispatch showed no significant difference, and the paired daily differences lay within the equivalence margins. The underestimation of the daily peaks explains part of this result: removing the peak-size error alone closed about a quarter of the gap to ideal information, and a causal peak-bias correction reduced the energy above the threshold by 14%. The remainder comes from the duration and shape of the high-load periods. Local smart-meter data are therefore sufficient for a useful forecasting framework but not for capturing all critical peaks. Ideal information would still reduce the energy above the threshold by a further 56–75%, so better forecasts of the high-load hours, rather than better average accuracy, are where future gains lie. Fourth, energy sharing and storage pooling have different roles. Hourly virtual sharing reduced the calculated residual community import by approximately 9.7% under the adopted accounting convention and delivered most of the allocated benefit to households without PV. Pooling storage added little energy but reduced grid exceedance, and its advantage over individual batteries grew from 8% with realistic forecasts to 78% with ideal information. PV owners did better with their own batteries than in a pooled community, underscoring the need for compensation schemes if pooling is to be adopted. For practitioners, the results suggest an order of priority. Energy sharing should come first, since it requires only metering and settlement and delivers most of the energy benefit. Storage should then be coordinated against a grid-capacity signal rather than a flat time-of-use tariff, so many batteries do not charge in the same hours. Forecasts should be judged by the operation they produce, in particular by the error of the daily peak, and not by hourly accuracy alone.

These conclusions rest on one community and one year shaped by COVID-19 restrictions, a 239-day test period, an aggregated battery model without network constraints, and an evaluation of energy and CO_2 emissions without costs or degradation. Future work should add weather and other exogenous inputs; develop peak-aware and probabilistic forecasts trained on the dispatch objective; extend the analysis to several years; consider communities with network constraints and economic evaluation; and design compensation mechanisms, for example, based on the Shapley value, that make pooled storage attractive for PV owners.

Author Contributions: Conceptualisation, M.A.; methodology, M.A., C.M., F.P., I.P. and M.S.; software, M.A., C.M. and S.S.; validation, M.A., A.F., C.M., F.P., I.P. and M.S.; formal analysis, M.A., A.F., I.P., M.S. and S.S.; investigation, M.A., A.F., C.M., F.P., I.P., M.S. and S.S.; resources, M.A.; data curation, M.A., C.M., I.P. and S.S.; writing—original draft preparation, M.A.; writing—review and editing, M.A., A.F., C.M., F.P., I.P., M.S. and S.S.; visualisation, M.A., A.F., C.M., F.P., I.P., M.S. and S.S.; supervision, A.F., F.P. and M.S.; project administration, M.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Campania Region, Project "Innovazione industriale per lo sviluppo sostenibile della Campania" (PR FSE+ 2021–2027), CUP F85E25000100006—SURF 23097AP000000008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors thank for the Campania Region, Project "Innovazione industriale per lo sviluppo sostenibile della Campania" (PR FSE + 2021-2027), CUP F85E25000100006—SURF 23097AP000000008.

Conflicts of Interest: Author S.S. is employed by KOALA SRL SB. The remaining authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Nomenclature

Symbol	Description
Acronyms	
AI	Artificial intelligence
ARIMA	Autoregressive integrated moving average
CBC	COIN-OR branch and cut (solver)
COVID-19	Coronavirus disease 2019
LightGBM	Light gradient-boosting machine
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MILP	Mixed-integer linear programming
ML	Machine learning
nMAE	Normalised mean absolute error
PV	Photovoltaic
REC	Renewable energy community
RMSE	Root mean square error
SCR	Self-consumption ratio
SOC	State of charge
SSR	Self-sufficiency ratio
XGBoost	Extreme gradient boosting
Latin symbols	
C^{BAT}	Usable battery capacity (kWh)
e_g	Mean absolute forecast error of day g (kWh)
$E_{\mathrm{el},t}$, $\hat{E}_{\mathrm{el},t}$	Measured and forecast community load (kWh)
$\hat{E}^{\mathrm{7d}}_{\mathrm{el},g,h}$	Seven-day same-hour mean load forecast (kWh)
$E^{\mathrm{Us}}_{\mathrm{el},i,t}$	Load of user i (kWh)
$E^{\mathrm{PV}}_{\mathrm{el},t}$, $\hat{E}^{\mathrm{PV}}_{\mathrm{el},t}$	Measured and forecast PV production (kWh)
$E^{\mathrm{PG}}_{\mathrm{el},t}$, $E^{\mathrm{TG}}_{\mathrm{el},t}$	Energy drawn from and fed to the grid (kWh)
$E^{\mathrm{PG,REC}}_{\mathrm{el},i,t}$	Grid import of user i after sharing (kWh)
$E^{\mathrm{sh}}_{\mathrm{el},t}$	Energy shared within the community (kWh)
$E^{\mathrm{ch}}_{\mathrm{el},t}$, $E^{\mathrm{dis}}_{\mathrm{el},t}$	Scheduled battery charge and discharge; a tilde denotes the values applied after real-time correction (kWh)
E^{BAT}_t, $E^{\mathrm{BAT},*}_g$	Stored energy at the end of hour t; night-time charging target of day g (kWh)
$E^{\mathrm{exc}}_{\mathrm{el},t}$, $E^{\mathrm{exc}}_{\mathrm{el}}$	Import above the threshold, hourly and annual total (kWh)
h^{exc}	Number of hours with import above the threshold (h)
$M_{\mathrm{CO_2}}$	Operational CO_2 emissions (kg)
n_t	Net load, $E_{\mathrm{el},t} - E^{\mathrm{PV}}_{\mathrm{el},t}$ (kWh)

N_w, N_h	Load-history window (168 h) and forecast horizon (24 h)
P^{BAT}	Battery power limit (kW)
P^{pk}	Daily peak import (kW)
P^{thr}, P_i^{thr}	Import threshold of the community and of user i (kW)
t_0	Forecast origin (last hour of the previous day)
u_t	Binary charging indicator (1 = charging)
w	Weight of the total-import term in the objectives (0.1)
$\Delta SOC_{i,t}$	Hourly change in state of charge of user i (%)
Greek symbols	
$\alpha_{CO_2,t}$	Hourly CO_2 emission factor of the grid (kg CO_2/kWh$_{el}$)
$\delta_{i,t}$	Net deficit of user i (kWh)
η_{ch}, η_{dis}	Battery charging and discharging efficiency (–)
κ_1, κ_2	Drift-detection thresholds (1.6 and 2.5)
μ_g, σ_g	Mean and sample standard deviation of the daily errors of the seven days preceding day g (kWh)
Π_g	Target import of the real-time correction on day g (kW)
Superscripts and subscripts	
BAT; el	Battery; electric energy
ch, dis	Charging, discharging
exc; thr	Above the threshold; threshold
PG, TG	Drawn from, fed to the grid
pk; sh	Peak; shared
REC; Us	With energy sharing in the REC; user

Appendix A. Supplementary Results

This appendix reports the sensitivity analyses and supporting results referred to in Sections 3.3, 4.2, 4.3.5, 4.3.6 and 4.3.7.

Table A1 compares the main dispatch and sharing results for the full test period with those for a reduced period ending on 11 December 2020 (219 days), before the records of H14 are missing. The ranking of the dispatch scenarios by the energy imported above the threshold, the ranking of the sharing configurations by import and by exceedance, and the direction and approximate magnitude of the benefits for households with and without PV are identical in both periods. Only two orderings separated by small margins are reversed: the import of S4 is 0.2% below that of S0 instead of 0.2% above it, and, with same-hour information, the realised objective of S4, Equation (27), is slightly lower than that of S_{RT}. The lower absolute values of the reduced period result mainly from the exclusion of the last 20 days of December, which contain the highest winter loads and account for 44% of the energy above the threshold of S3, 93% of that of S4 and all of that of S_{RT}. The conclusions of Sections 4.3 and 4.4 are therefore independent of the treatment of the missing records.

Table A1. Robustness to the missing records of H14: full test period (239 days, 7 May–31 December) and test period ending before the gap (219 days, 7 May–11 December), threshold objective. (**a**) Dispatch: grid import and energy imported above 14.1 kW (E_{el}^{exc}) with same-hour and one-hour-delayed information. (**b**) Sharing: change in grid import relative to configuration A. Wilcoxon p for S3 vs. S3n: 0.53 (239 days) and 0.71 (219 days) for the daily peak; 0.50 and 0.16 for the daily E_{el}^{exc}.

(a) Dispatch						
	Import (kWh)		**E_{el}^{exc}, Same Hour (kWh)**		**E_{el}^{exc}, One-Hour Delay (kWh)**	
Scenario	**239 d**	**219 d**	**239 d**	**219 d**	**239 d**	**219 d**
S0	53,038	46,112	3879	2503	3879	2503
S1	60,717	53,415	18,290	16,466	18,290	16,466
S_{RT}	53,938	46,921	418	0	2532	1639
S3	52,536	45,496	2379	1324	3044	1916
S3n	52,658	45,587	2362	1303	3002	1893
S4	53,164	46,009	585	39	1333	679
(b) Sharing: Change in Grid Import Relative to A (%)						
	All Households		**Without PV**		**With PV**	
Configuration	**239 d**	**219 d**	**239 d**	**219 d**	**239 d**	**219 d**
B. Shared PV	−9.7	−10.8	−14.1	−15.4	−3.5	−4.0
C. Individual battery	−3.6	−4.4	+2.1	+2.0	−11.5	−13.8
D. Individual + sharing	−7.5	−8.7	−3.6	−4.2	−12.8	−15.3
E. Fully pooled	−10.3	−11.8	−14.5	−15.9	−4.6	−5.8

Table A2. Sensitivity of the LightGBM forecast to the drift thresholds (test period, 239 days). Reference window: the seven days preceding day g, sample standard deviation, Equation (4). The default setting is shown in bold.

κ_1	κ_2	Retrains	MAE (kWh)	nMAE (%)	R^2
1.4	2.0	37	2.075	18.30	0.598
1.4	2.5	31	2.075	18.31	0.599
1.4	3.0	31	2.075	18.31	0.599
1.6	2.0	25	2.069	18.25	0.602
1.6	**2.5**	**20**	**2.066**	**18.22**	**0.605**
1.6	3.0	20	2.066	18.22	0.605
1.8	2.0	26	2.068	18.25	0.603
1.8	2.5	18	2.066	18.22	0.605
1.8	3.0	18	2.066	18.22	0.605

Table A3. Sensitivity of the optimised dispatch to the import weight w (test period, 239 days). E_{el}^{exc}: energy imported above the threshold of 14.1 kW; h^{exc}: hours above the threshold; E_{el}^{ch}: energy charged into the battery. The values for $w = 0.1$ correspond to Table 7.

Scenario	**w (–)**	**Import (kWh)**	**Daily Peak (kW)**	**E_{el}^{exc} (kWh)**	**h^{exc} (h)**	**SSR (%)**	**E_{el}^{ch} (kWh)**
MILP, threshold objective							
S3: LightGBM Forecast	0.01	52,532	16.6	2395	552	19.2	2246
	0.05	52,536	16.5	2383	547	19.2	2264
	0.1	52,536	16.5	2379	551	19.2	2264
	0.2	52,536	16.5	2383	547	19.2	2264
	0.5	52,529	16.6	2413	564	19.2	2234

Table A3. *Cont.*

Scenario	w (–)	Import (kWh)	Daily Peak (kW)	E_{el}^{exc} (kWh)	h^{exc} (h)	SSR (%)	E_{el}^{ch} (kWh)
S4: Ideal-information Benchmark	0.01	53,164	14.4	585	95	18.2	4962
	0.05	53,164	14.6	585	70	18.2	4962
	0.1	53,164	14.5	585	95	18.2	4962
	0.2	53,164	14.6	585	70	18.2	4962
	0.5	53,164	14.5	585	95	18.2	4962
MILP, peak objective							
S3: LightGBM Forecast	0.01	53,528	16.5	2787	710	17.7	6843
	0.05	53,526	16.5	2785	712	17.7	6833
	0.1	53,526	16.5	2785	712	17.7	6833
	0.2	53,529	16.5	2785	712	17.7	6845
	0.5	52,957	16.7	2975	782	18.6	4209
S4: Ideal-information Benchmark	0.01	54,874	9.6	657	432	15.6	12,850
	0.05	54,874	9.6	657	432	15.6	12,850
	0.1	54,874	9.6	657	432	15.6	12,850
	0.2	54,874	9.6	657	432	15.6	12,850
	0.5	53,639	11.5	1582	812	17.5	7154

Table A4. Sensitivity of the dispatch to battery losses (threshold objective, 239 days). Standby: constant consumption of the battery systems added to the load; η: one-way efficiency. S0 and S1 are unaffected (S1 includes the real losses).

Case	Scenario	Import (kWh)	E_{el}^{exc} (kWh)	h^{exc} (h)
Reference	S0	53,038	3879	967
	S1	60,717	18,290	1169
Main case ($\eta = 0.885$, no standby)	S2	59,896	25,289	2168
	S3	52,536	2379	551
	S3n	52,658	2362	570
	S4	53,164	585	95
Standby 0.31 kW	S2	61,747	26,554	2216
	S3	54,317	2703	608
	S3n	54,446	2715	660
	S4	54,968	717	110
Standby 0.37 kW (+20%)	S2	62,105	26,796	2219
	S3	54,663	2764	619
	S3n	54,794	2777	667
	S4	55,317	746	114
Standby 0.31 kW, $\eta = 0.85$	S2	64,630	29,086	2247
	S3	54,475	2802	653
	S3n	54,653	2814	681
	S4	55,348	792	125

Table A5. Paired daily differences S3 minus S3n (239 days). $n_{\neq 0}$: days with a nonzero difference; r_1: lag-1 autocorrelation; LB: Ljung–Box test at lag 7; $n_{eff} = n(1 - r_1)/(1 + r_1)$; CI: circular moving-block bootstrap with a block length of 7 days. Equivalence margins: ±0.5 kW for the daily peak and ±1.0 kWh for the daily energy above the threshold; all four cases are equivalent.

Objective	Quantity	$n_{\neq 0}$	Mean	r_1	LB p	n_{eff}	Skewness (p)	Wilcoxon p	Sign p	95% CI	90% CI
Threshold	Daily peak (kW)	82	0.025	0.20	<0.001	161	−0.28 (0.28)	0.53	0.22	[−0.15, 0.22]	[−0.12, 0.18]
Threshold	E_{el}^{exc} (kWh/day)	93	0.074	0.30	<0.001	130	0.25 (0.30)	0.50	0.15	[−0.47, 0.63]	[−0.39, 0.54]
Peak	Daily peak (kW)	117	0.176	0.12	0.056	186	0.38 (0.09)	0.13	0.58	[−0.01, 0.38]	[0.02, 0.34]
Peak	E_{el}^{exc} (kWh/day)	105	0.184	0.41	<0.001	101	1.17 (<0.001)	0.49	0.70	[−0.29, 0.74]	[−0.22, 0.64]

References

1. International Energy Agency. World Energy Outlook 2025. Available online: https://www.iea.org/reports/world-energy-outlook-2025 (accessed on 22 September 2026).
2. Matos, M.; Almeida, J.; Gonçalves, P.; Baldo, F.; Braz, F.J.; Bartolomeu, P.C. A Machine Learning-Based Electricity Consumption Forecast and Management System for Renewable Energy Communities. *Energies* **2024**, *17*, 630. [CrossRef]
3. European Union. Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 On the Promotion of the Use of Energy from Renewable Sources (Recast). 2018. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018L2001 (accessed on 22 September 2026).
4. Esposito, P.; Marrasso, E.; Martone, C.; Pallotta, G.; Roselli, C.; Sasso, M.; Tufo, M. A roadmap for the implementation of a renewable energy community. *Heliyon* **2024**, *10*, e28269. [CrossRef] [PubMed]
5. Coccato, S.; Barhmi, K.; Lampropoulos, I.; Golroodbari, S.; van Sark, W. A Review of Battery Energy Storage Optimization in the Built Environment. *Batteries* **2025**, *11*, 179. [CrossRef]
6. Pasqui, M.; Gerini, F.; Jacobs, M.; Carcasci, C.; Paolone, M. Self-dispatching a renewable energy community by means of battery energy storage systems. *J. Energy Storage* **2025**, *114*, 115837. [CrossRef]
7. Pocola, T.O.; Robu, V.; Rietveld, J.; Norbu, S.; Couraud, B.; Andoni, M.; Flynn, D.; Poor, H. Optimal sizing and control of a grid-connected battery in a stacked revenue model including an energy community. *Appl. Energy* **2025**, *397*, 126122. [CrossRef]
8. Bahloul, M.; Daoud, M.; Khadem, S.K. Optimal dispatch of battery energy storage for multi-service provision in a collocated PV power plant considering battery ageing. *Energy* **2024**, *293*, 130744. [CrossRef]
9. Putz, D.; Gumhalter, M.; Auer, H. The true value of a forecast: Assessing the impact of accuracy on local energy communities. *Sustain. Energy Grids Netw.* **2023**, *33*, 100983. [CrossRef]
10. Donti, P.L.; Amos, B.; Kolter, J.Z. Task-based end-to-end model learning in stochastic optimization. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 5484–5494.
11. Casalicchio, V.; Manzolini, G.; Prina, M.G.; Moser, D. From investment optimization to fair benefit distribution in renewable energy community modelling. *Appl. Energy* **2022**, *310*, 118447. [CrossRef]
12. Volpato, G.; Carraro, G.; Dal Cin, E.; Rech, S. On the Different Fair Allocations of Economic Benefits for Energy Communities. *Energies* **2024**, *17*, 4788. [CrossRef]
13. Trivedi, R.; Bahloul, M.; Saif, A.; Patra, S.; Khadem, S. Comprehensive Dataset on Electrical Load Profiles for Energy Community in Ireland. *Sci. Data* **2024**, *11*, 621. [CrossRef] [PubMed]
14. EirGrid. Smart Grid Dashboard: CO_2 Intensity, Ireland (ROI), 2020. Available online: https://www.smartgriddashboard.com/all/co2/ (accessed on 25 September 2026).
15. Marrasso, E.; Martone, C.; Pallotta, G.; Perugini, I.; Roselli, C.; Sasso, M. Environmental impact reduction of residential end-users: An innovative methodology based on the power grid CO_2 emission factor. *Energy Build.* **2025**, *348*, 116419. [CrossRef]
16. Shiblee, M.F.H.; Koukaras, P. Short-Term Load Forecasting in the Greek Power Distribution System: A Comparative Study of Gradient Boosting and Deep Learning Models. *Energies* **2025**, *18*, 5060. [CrossRef]
17. Koukaras, P.; Mustapha, A.; Mystakidis, A.; Tjortjis, C. Optimizing Building Short-Term Load Forecasting: A Comparative Analysis of Machine Learning Models. *Energies* **2024**, *17*, 1450. [CrossRef]
18. Muqtadir, A.; Li, B.; Zhou, Y.; Chen, S.; Kazmi, S.N. Nowcasting the next hour of residential load using boosting ensemble machines. *Sci. Rep.* **2025**, *15*, 7157. [CrossRef] [PubMed]
19. Yang, Y.; Lee, S.H.; Choi, Y.S.; Lee, K.M. LightGBM Medium-Term Photovoltaic Power Prediction Integrating Meteorological Features and Historical Data. *Energies* **2025**, *18*, 5526. [CrossRef]
20. Waseem, M.; Lin, Z.; Liu, S.; Jinai, Z.; Rizwan, M.; Sajjad, I.A. Optimal BRA based electric demand prediction strategy considering instance-based learning of the forecast factors. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12967. [CrossRef]
21. Waseem, M.; Lin, Z.; Liu, S.; Sajjad, I.A.; Aziz, T. Optimal GWCSO-based home appliances scheduling for demand response considering end-users comfort. *Electr. Power Syst. Res.* **2020**, *187*, 106477. [CrossRef]
22. Guo, F.; Gomes, L.; Ma, L.; Tian, Z.; Vale, Z.; Pang, S. Optimizing battery storage for sustainable energy communities: A multi-scenario analysis. *Sustain. Cities Soc.* **2025**, *118*, 106030. [CrossRef]
23. Chang, H.C.; Ghaddar, B.; Nathwani, J. Shared community energy storage allocation and optimization. *Appl. Energy* **2022**, *318*, 119160. [CrossRef]
24. Bayram, F.; Aupke, P.; Ahmed, B.S.; Kassler, A.; Theocharis, A.; Forsman, J. DA-LSTM: A dynamic drift-adaptive learning framework for interval load forecasting with LSTM networks. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106480. [CrossRef]
25. Azeem, A.; Ismail, I.; Mohani, S.S.; Danyaro, K.U.; Hussain, U.; Shabbir, S.; Bin Jusoh, R.Z. Mitigating concept drift challenges in evolving smart grids: An adaptive ensemble LSTM for enhanced load forecasting. *Energy Rep.* **2025**, *13*, 1369–1383. [CrossRef]
26. Oliveira, M.B.; Silva, L.A. Two-phase drift monitoring and automatic retraining (TPDM-AR) for electric load forecasting. *MethodsX* **2026**, *16*, 103966. [CrossRef] [PubMed]

27. Zhou, S.; Ni, S.; Han, Y.; Dong, Z.; Lai, C. Adaptive electricity consumption forecasting approach for universal environments. *Sci. Rep.* **2025**, *15*, 24400. [CrossRef] [PubMed]
28. Cao, C.; He, Y.; Yang, X. Online decoupling feature framework for optimal probabilistic load forecasting in concept drift environments. *Appl. Energy* **2025**, *392*, 125952. [CrossRef]
29. Central Statistics Office. COVID-19 in Ireland: Timeline 2020–2023. 2023. Available online: https://www.cso.ie/en/releasesandpublications/ep/p-covid19/covid-19informationhub/covid-19inirelandtimeline2020-2023 (accessed on 27 September 2026).
30. Government of Ireland. S.I. No. 76/2022—European Union (Renewable Energy) Regulations 2022. Available online: https://www.irishstatutebook.ie/eli/2022/si/76/made/en/print (accessed on 27 September 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.