Kurdish Studies
September 2019
Volume: 07, No: 02, pp 189-212
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.67145/ks.v7i2.4154

An Algorithmic Benchmark of Tolerance-Bounded Minimum-Peak Battery Dispatch on a Minute-Resolution Household Load Series

Luis Eduardo Muñoz Guerrero[1*]

[1*]Universidad Tecnológica de Pereira, Facultad de Ingenierías, Pereira, Risaralda, Colombia, lemunozg@utp.edu.co

Abstract

We compute the minimum one-minute peak of a battery with power and energy limits by bisection with an exact greedy test, to within 10−4 kW, and compare an online threshold rule with it over 119 seven-day windows of one household. With the threshold at the training 99th percentile and a 4 kWh, 3 kW battery, the median relative gap is 34.5% (95% CI 20.4–43.7%), rejecting H1 (median below 10%). Whenever the threshold is achievable, the rule's peak equals it, so the gap mainly measures how the threshold is set. The bisection runs over ten times faster than a linear programme.

Keywords: peak shaving; energy storage; online algorithms; linear programming; threshold rules

1. Introduction

Some electricity tariffs bill customers for their highest demand as well as for the energy they use, and peak loads bear on investment in the network.1 A storage device that lowers the maximum metered load can reduce such costs without changing the energy delivered.2 Batteries are one storage technology proposed for this use,3 and Section 2 reviews how they are dispatched for peak shaving.

Dispatch rules differ in the information they use. An offline optimum knows the whole future load, so it gives the lowest peak that any feasible schedule can reach under the same constraints. An online rule acts on the present state and recent history. The gap be-tween the two tells a designer how much a practical rule leaves on the table, and so whether a more sophisticated controller could be worth its cost. Online algorithms are usually analysed by their worst-case ratio to the offline optimum,4 and peak minimization with a battery has been studied in that way. 5 A worst-case ratio bounds the gap for every input, but it does not say how large the gap is on a recorded load.

This paper asks two questions. For one household's load at one-minute resolution, how much higher is the peak reached by a simple online threshold rule than the offline minimum? And what does that minimum cost to compute? We answer the first with a hypothesis fixed in advance in a frozen plan, and with validity checks, and the second with a benchmark.

We make three contributions. We formulate minimum-peak dis-patch with power, energy and efficiency limits and give a bisection procedure whose output is optimal up to a stated tolerance, and prove that bound. We measure the online-to-offline gap over the 119 complete seven-day windows that follow the training period, in a public dataset, and show that wherever the threshold is achievable the gap is fixed by the threshold and the week's optimum. We benchmark the bisection against a linear programme and check the two against each other on every window. We do not claim priority for the problem, which Bar-Noy and colleagues studied before,6 for the offline procedure or for the threshold rule. The contribution is the measurement, its explanation and its computational cost.

2. Related Work

Strategies for shaving peak load have been reviewed.7 Most dis-patch methods optimize the schedule over a known or forecast load profile. Oudalov and colleagues size a battery and derive its operat-ing strategy by dynamic programming; Lu and colleagues combine rolling load forecasts with charge and discharge policies that mini-mize the peak–valley difference or the load variance; linear pro-grammes schedule storage for hourly peak shaving at home and on a distribution feeder; Joshi and Pindoriya dispatch the battery a day ahead from representative daily profiles; and Zheng and colleagues compare dispatch strategies for residential peak shaving.8 Nottrott and colleagues schedule a photovoltaic battery system by linear programming with forecasts, compare it with real-time rules that use no forecast, and calculate the financial value of forecasting; Park and Lappas find the battery utilization strategy critical to the demand reduction achieved; and Neubauer and Simpson bound the savings from demand charges with a perfect 48-hour forecast and leave the sensitivity to forecast error unexplored.9 Pimm and col-leagues estimate how much peak shaving storage could deliver in residential low-voltage networks, using synthetic high-resolution household profiles.10 For industrial consumers, Papadopoulos and colleagues observe that much of the literature assumes a known or accurately predicted load profile, and find that peak shaving based on predicted profiles becomes reliable for lower forecast errors; Bennett and colleagues pair a next-day forecast and schedule with an online control algorithm that corrects for forecast error.11

Other methods dispatch by rules or controllers. Shi and col-leagues propose a simple threshold algorithm that runs in real time while the battery also provides frequency regulation; Barzkar and Hosseini derive a simple real-time scheduling rule from the solution of an optimization problem and test it on measured loads; and Chua and colleagues reduce peak demand

with a fuzzy control algorithm for storage.12 For storage that responds to prices, the optimal control policy has a threshold structure.13 Model predictive control re-plans over a short horizon: it has been applied to a household with photo-voltaics and a battery, and to a large battery with a degradation cost.14 Leadbetter and Swan study battery storage for residential peak demand shaving, and a household can also flatten its demand by scheduling flexible loads.15

Online algorithms are usually compared with the offline optimum through worst-case guarantees. BarNoy and colleagues introduced peak minimization with a battery as an online problem. They give optimal offline algorithms, and online algorithms for the reduction of the peak whose competitive ratio matches a lower bound when the peak effective demand is revealed in advance; a companion paper extends the problem to batteries that lose part of each charge and evaluates online and heuristic algorithms against optimal ones on real and synthetic data.16 Online algorithms with worst-case guarantees also exist for peak-aware dispatch in microgrids, for storage that responds to prices and for peak-minimizing charging of electric vehicles, and data centres use their backup batteries to cut peak power or electricity cost.17

Table 1. The closest works as their abstracts describe them, and the present study. "Not stated" means that the abstract does not say.

Work and problem	Online policy	Offline benchmark	Data, as stated
Bar-Noy 2009, peak with a battery	Matching bounds	Optimal algorithms	Not stated
Bar-Noy 2008, lossy battery	Online and heuristic	Optimal algorithms	Real and synthetic
Zhang 2018, microgrid with peak pricing	Competitive algorithms	Offline optimum	Real traces
Zhao 2017, peak of vehicle charging	Robust online algorithms	Offline optimum	Traces
Chau 2016, storage under prices	Competitive algorithm	Offline optimum	Real data
Nottrott 2013, demand charges with photovoltaics	Real-time load response	Linear programme with forecasts	Not stated
Shi 2018, peak and frequency	Real-time threshold	Not stated	Real data
Pimm 2018, low-voltage networks	Not stated	Potential peak shaving	Synthetic households
This study, peak with a power limit	Training-percentile threshold	Optimum per window	One household, one minute

Table 1 summarizes the closest works as their abstracts describe them. Several compare online or forecast-free rules with optimal or forecast-based schedules, some on real traces.18 The present study adds one such measurement, for one household at one-minute resolution and a battery with a power limit, and traces the gap of a threshold rule to its threshold. We do not claim that it is the first.

3. Data

The data are the individual household electric power consumption records published in the UCI Machine Learning Repository, 19 measured in one house in Sceaux, near Paris; the repository does not describe its space heating or its tariff. The dataset has been used to benchmark streaming classifiers at several time scales and to forecast household consumption with neural networks. 20 The dataset was donated to the repository in August 2012. The file has 2,075,259 rows, one per minute, from 16 December 2006 at 17:24 to 26 November 2010 at 21:02. Each row carries a date, a time and seven electrical measurements. This study uses only the active power, Global_active_power, which is the mean power in kilowatts over the minute. The file was downloaded and its SHA-256 hash checked against a recorded value before analysis.

Table 2 summarizes the data. Time stamps are complete: every minute between the first and the last record is present, so missing-ness concerns only the power values. The median observed segment is 26,484.5 minutes (18.4 days), the median power 0.602 kW and the maximum 11.122 kW, and the file contains no zero readings.

Table 2. Size, coverage and missing values of the record, and the training and evaluation samples drawn from it.

Quantity	Value
Rows (one per minute)	2,075,259
Coverage	16 Dec 2006 17:24 to 26 Nov 2010 21:02
Missing active power	25,979 minutes (1.252%)
Contiguous observed segments	72
Longest observed segment	78,706 minutes
Complete seven-day windows, full record	134
Evaluation windows	119
99th percentile, full record	4.85 kW

We do not impute missing values. A battery state cannot be car-ried across an unknown interval, so the analysis uses only windows without gaps. Ignoring the missing minutes in this way is justified only under conditions on how they arise;21 the

plan assumed that the gaps are unrelated to the load, and Section 8 reports evidence against this assumption. A window is a block of 10,080 consecutive minutes, which is seven days, aligned to the first minute of the record, so that every window starts on a Saturday at 17:24, and containing no missing values. The full record contains 134 such windows. Figure 1 shows every block of the record and its role: 60 of the 179 blocks after the training period are excluded, 41 of them for at most two missing minutes.

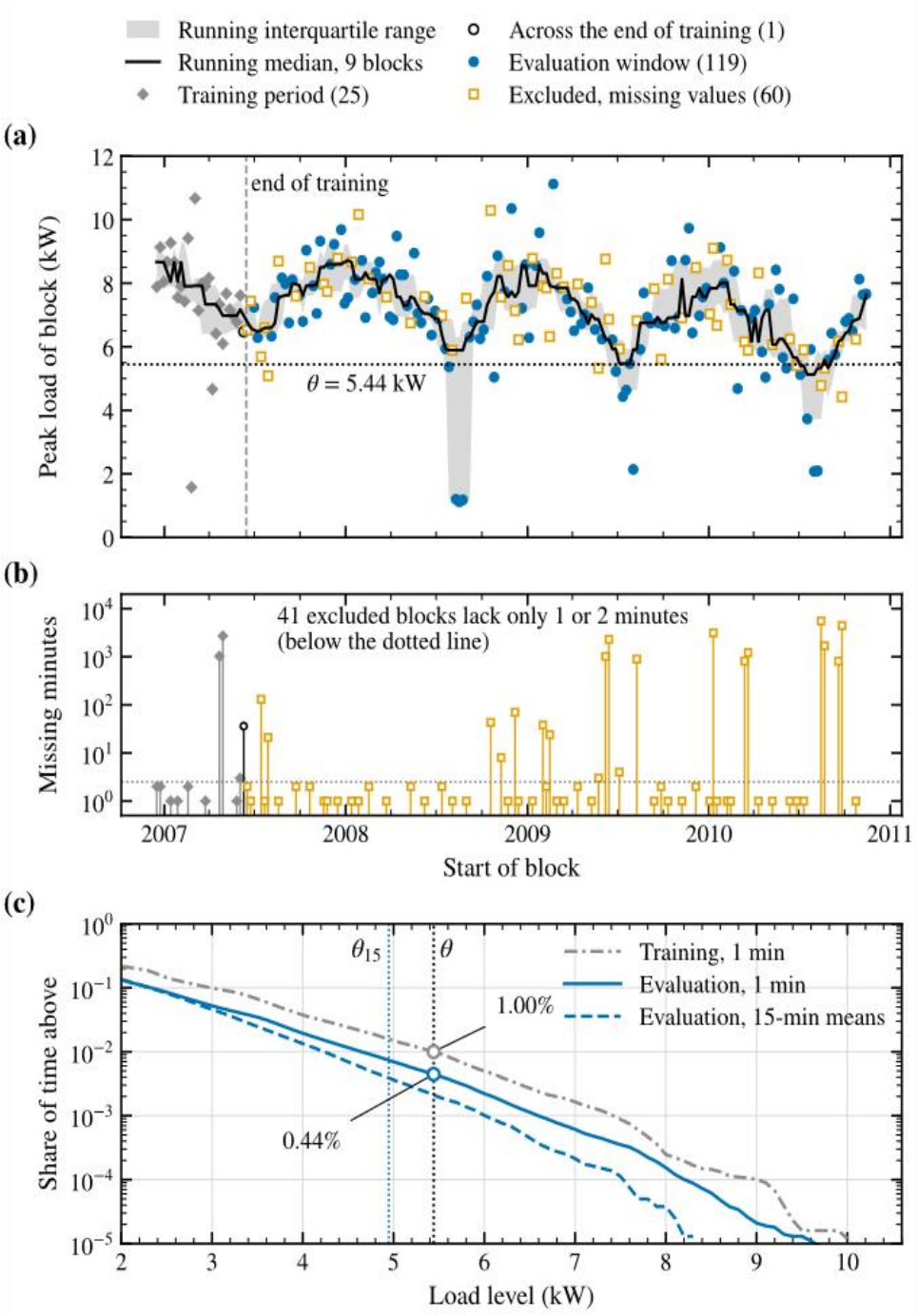


Figure 1. Coverage of the record in seven-day blocks aligned to its first minute.

Figure 1. (a) Peak load of each block, marked by its role, with the running median over nine blocks (black) and its interquartile range (shaded); the dotted line is the threshold θ and the dashed one the end of training. (b) Minutes without a power value in each block that has any, on a logarithmic axis; the dotted line separates the blocks that lack one or two minutes. (c) Share of time with the load above each level, in the training period and in the evaluation windows, at one minute and for 15-minute means, with the shares above θ marked; the blue dotted line is the threshold $\theta 15$ for 15-minute means.

The first 180 days, from 16 December 2006 at 17:24 to 14 June 2007 at 17:23, form the training period. It only fixes the threshold of the online rule. The evaluation sample is the 119 complete windows that start after the training period, the first of which starts on 30 June 2007. The two earlier windows aligned to the record, which start on 16 and 23 June 2007, contain missing values and are there-fore excluded.

Figure 2 shows the mean load in each hour of each week. The mean load is highest between 20:00 and 21:00 and in December, and lowest in August, and some summer weeks have very little load: the six evaluation windows with the smallest optimal peaks, dis-cussed in Section 7, have a mean load of 0.18 to 0.38 kW, against a median of 1.12 kW for the other windows.

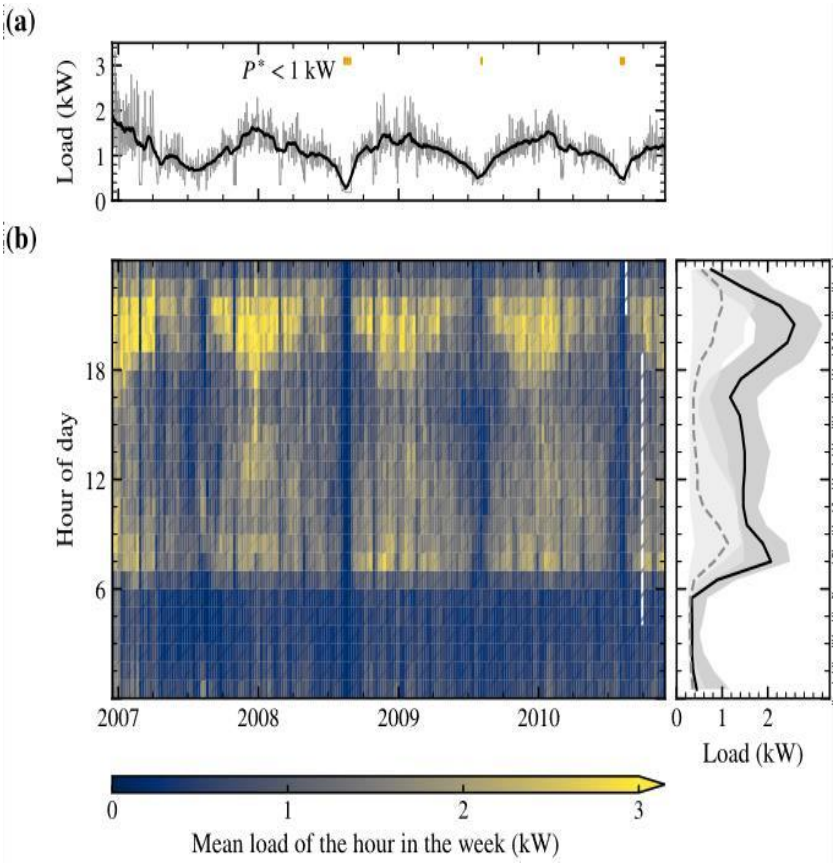


Figure 2. The load over the whole record.

4. Model And Algorithms

Formulation. Table 3 lists the notation. A household load is a se-quence $L1, \ldots, LT$ of non-negative powers in kilowatts, observed at steps of $\Delta t = 1/60$ h. A battery has usable energy capacity E in kil-owatt-hours, a power limit Pb in kilowatts for charging and for dis-charging, a one-way efficiency $\eta \in (0,1]$, and an initial state of charge $s0 \in [0, E]$. At step t the battery draws power $ct \geq 0$ from the grid or delivers power $dt \geq 0$ to the household, so the grid power is

$$g_t = L_t + c_t - d_t \qquad (1)$$

The state of charge evolves as

$$s_t = s_{t-1} + \eta\, c_t\, \Delta t - \frac{d_t\, \Delta t}{n}, \quad 0 \leq s_t \leq E \qquad (2)$$

and the power limits are

$$0 \leq c_t \leq P_b, \qquad 0 \leq d_t \leq \min(P_b, L_t) \qquad (3)$$

The bound on dt forbids export to the grid, so the grid power is never negative. For a target peak P, a schedule is feasible if the var-iables satisfy (2) and (3) and $gt \leq P$ at every step, and P is feasible if such a schedule exists. The minimum peak is

$$P^* = \min\{P: P \text{ is feasible}\} \qquad (4)$$

Algorithm 1. Minimum peak to within a tolerance by bisection with an exact greedy feasibility test.

Input: load $L1, \dots, LT$, step Δt, capacity E, power limit Pb, efficiency η, initial state $s0$ and tolerance τ

Output: $P^{\wedge}$, with $0 \le P^{\wedge} - P^{*} \le \tau$

```
hi ← maxtLt
lo ← max (0, hi − Pb)
while hi − lo > τ do
    P ← (lo + hi)/2; s ← s0; ok ← true
    for t = 1, … , T do
        if Lt > P then
            d ← Lt − P
            if d > Pb or d Δt/η > s then ok ← false; break
            s ← s − d Δt/η
        else
            c ← min (Pb, P − Lt, (E − s)/(η Δt))
            s ← s + η c Δt
    if ok then hi ← P else lo ← P
return P^ = hi
```

Problem (4) is equivalent to a linear programme: minimize P sub-ject to linear constraints in P, ct, dt and st, with $s0$ fixed. Its opti-mum is therefore computable in polynomial time,22 and no hardness argument is needed for the question studied here. The programme has $3T + 1$ variables and $2T$ constraints, the T inequalities $gt \le P$ with gt from (1) and the T equalities (2), plus the bounds on the variables in (2) and (3). The model allows charging and discharging in the same step, which Lemma 1 shows is never needed, and sets no condition on the final state of charge. The programme serves as an independent reference for the procedure below.

Table 3. Notation for the model, the algorithms and the measured gap.

Symbol	Meaning
L_t, T	household load at step t (kW); number of steps in a window
$\Delta t, \tau$	step length (h); bisection tolerance (kW)
E, P_b, η	usable capacity (kWh), power limit (kW), one-way efficiency
s_t, s_0	state of charge after step t, and at the start (kWh)
c_t, d_t, g_t	charging, discharging and grid power (kW)
P, P^*	target peak, and minimum feasible peak (kW)
$\widehat{P}, \theta$	output of Algorithm 1; threshold of the online rule (kW)
M, P_w^c	raw peak and online peak of window w (kW)
w, γ_w	evaluation window, and its relative gap
ρ_w, t_{rec}	share of the achievable reduction captured; time to recharge from s_0 to E (h)
$T_{\text{tr}}, \theta_{\text{H1}}$	length of the training period; threshold at which the median gap crosses 10% (kW)

Greedy feasibility test. For a fixed target P, the greedy schedule works as follows. When $Lt > P$, the battery must deliver at least
$Lt - P$, and the greedy schedule delivers exactly that. When $Lt \le$
P, the greedy schedule charges as much as the power limit and the headroom under P allow, up to the capacity. The test fails when a required discharge exceeds Pb or the stored energy cannot cover it. Algorithm 1 uses this test inside a bisection on P.

Lemma 1 (dominance). Fix P. Let $\sigma = (c, d)$ be any schedule that satisfies (2) and (3) and has $gt \le P$ for $t = 1, \dots, T$, with state trajectory $s\sigma$. Then the greedy schedule satisfies (2) and (3) with
$gt \le P$, and its state trajectory sG satisfies $sG \ge_G s\sigma$ for σevery t.

Proof. Both trajectories start at $s0$. Assume $st-1 \geq st-1$.
Case $Lt > P$. The greedy schedule delivers $nt = Lt - P$, which requires $s_{t-1}^{G} \geq n_t \Delta t/\eta$. The schedule σ satisfies $L_t + c_t - d_t \leq P$, so $d_t \geq L_t + c_t - P \geq n_t$. Its state update satisfies

$$s_t^{\sigma} - s_{t-1}^{\sigma} = \eta c_t \Delta t - \frac{d_t \Delta t}{\eta} \leq \eta c_t \Delta t - \frac{(L_t + c_t - P)\Delta t}{\eta} = -\frac{n_t \Delta t}{\eta} + c_t \Delta t \left(\eta - \frac{1}{\eta}\right) \leq -\frac{n_t \Delta t}{\eta},$$

because η≤1"." Since s_t^σ≥0"," we get s_(t-1)^σ≥n_t Δt/η"," and therefore s_(t-1)^G≥n_t Δt/η"." Since n_t≤d_t≤P_b "," the power limit holds, and the greedy discharge is feasible, and s_t^G=s_(t-1)^G-n_t Δt/η≥s_(t-1)^σ-n_t Δt/η≥s_t^σ "."

Case L_t≤P"." Let c ¯=min(P_b,P-L_t)"." The greedy schedule charges c^G=min(c ¯,(E-s_(t-1)^G)/(ηΔt))"," so that (2) holds and s_t^G=min(E,s_(t-1)^G+ηc ¯Δt)"." The schedule σ has c_t≤P_b and c_t≤P-L_t+d_t "." If d_t=0"," its net energy gain is at most ηc ¯Δt"." If d_t>0"," its net gain is at most η(P-L_t+d_t)Δt-d_t Δt/η=η(P-L_t)Δt+d_t Δt(η-1/η)≤η(P-L_t)Δt"," and also at most ηP_b Δt"." In both sub-cases the net gain is at most ηc ¯Δt"." Hence s_t^σ≤min(E,s_(t-1)^σ+ηc ¯Δt)≤s_t^G "."

Corollary. The greedy test succeeds if and only if a feasible schedule exists for P"."
Proof. If σ is feasible, the greedy schedule never fails, by the case analysis above. If the greedy test succeeds, its schedule, with the charge capped by the free capacity, satisfies (2) and (3) by construction.
Theorem 1 (minimum to within a tolerance). In exact arithmetic, for any tolerance τ>0"," Algorithm 1 returns P̂ such that P̂ is feasible and 0≤P̂-P^*≤τ"," after at most max(0,⌈log_2 (P_b/τ)⌉) feasibility tests.
Proof. Feasibility is monotone in P":" a schedule with g_t≤P also satisfies g_t≤P^' for every P^'≥P"." The initial upper bound hi=max_t L_t is feasible, because with P=hi no discharge is needed and the greedy charging keeps every state within capacity. The initial lower bound lo=max(0,max_t L_t-P_b) satisfies lo≤P^* ":" for the step t at which the maximum load occurs, any feasible schedule has P≥L_t-d_t≥L_t-P_b "," and P≥0 because grid power is never negative. By the Corollary, each test decides feasibility exactly, so the invariant lo≤P^*≤hi with hi feasible holds throughout. Each iteration halves the interval. Its initial width is min(max_t L_t,P_b)≤P_b "," so after n iterations the width is at most P_b/2^n "." The loop stops once the width is at most τ"," which takes at most max(0,⌈log_2 (P_b/τ)⌉) iterations. At that point P^*∈[hi-τ,hi] and hi is feasible, so P̂=hi has the claimed properties.

Algorithm 1 has one greedy pass per test, so its time is O(Tlog(P_b/τ))"." The extra memory is constant if the series can be read again on each pass. Our implementation converts the series to a list of floating-point numbers for speed, which makes its memory grow linearly with the series length, and compares stored energy with a slack of 10^(-12) kWh. When P^* equals the lower bound, every test succeeds and P̂ exceeds it by the final bracket width, P_b/2^15≈9.2×10^(-5) kW here; a single test at the lower bound would certify such optima exactly.
Figure 3 runs Algorithm 1 on the evaluation window in which stored energy limits the optimum most, the week from 17 July 2010, whose optimum lies 0.73 kW above the lower end of the initial bracket, max(0,M-P_b)"." The lowest stored energy reached by the greedy pass, computed as if the battery could go below empty, never falls as the target rises and reaches zero at P^* "," so each test reads the sign of one number; 6 of the 15 tests fail. The greedy state of charge at P̂ never falls below that of the optimal schedule returned by the linear programme, as Lemma 1 requires, and the third test, at 1.10 kW, runs out of stored energy at minute 6,208.

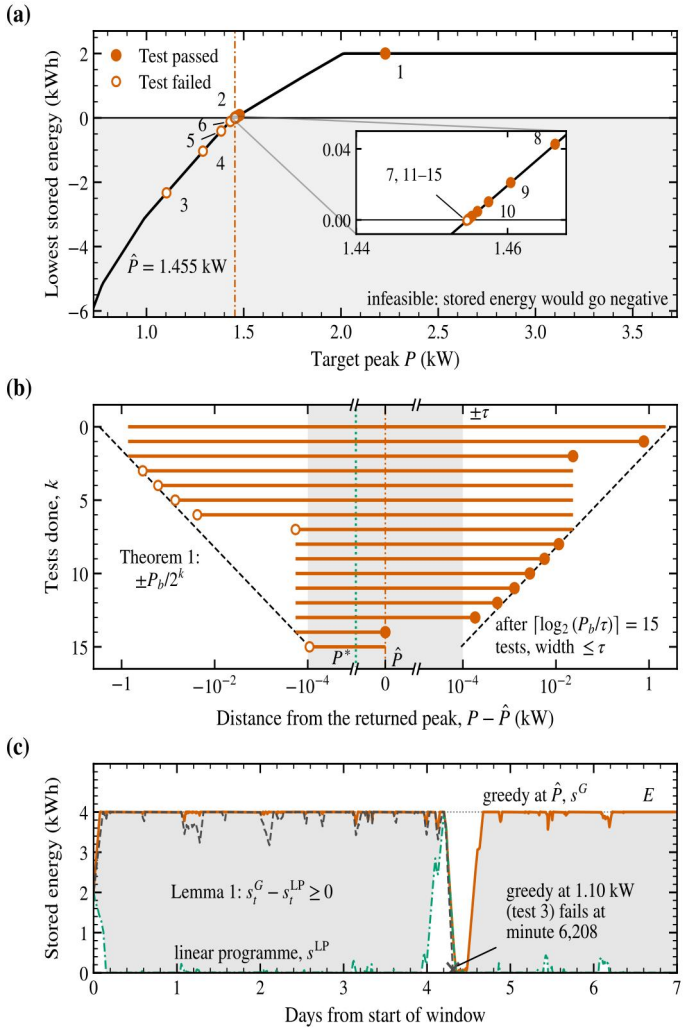


Figure 3. Algorithm 1 at work on one evaluation window. The window is the one in which stored energy limits the optimum most, from 17 July 2010, with E=4 kWh. (a) Lowest stored energy of the greedy pass for each target P in the initial bracket [M-P_b,M]"," computed as if the battery could go below empty (shaded: infeasible), with the 15 tests, filled when passed; the inset enlarges tests 7 to 15. (b) The bisection as a funnel: row k is the bracket after k tests, as distance from the returned peak P^"," on an axis linear within ±10^(-5) kW (double bars) and logarithmic outside. The end moved by test k shows its result; the dashed lines are the bound ±P_b/2^k of Theorem 1, the grey band is ±τ"," and the green dotted line is the linear programme's optimum P^* "," inside every bracket. (c) State of charge of the greedy schedule at P^"," of the linear programme's optimal schedule, never above it (Lemma 1; grey margin), and of the greedy schedule at test 3, which runs out of stored energy at the cross.

Online threshold rule. The online rule compares the load with a threshold θ fixed before the window. It charges when the load is below the threshold and discharges when the load is above it. When L_t<θ"," it sets d_t=0 and charges

$$c_t = \min\left(P_b,\ \theta - L_t,\ \frac{E - s_{t-1}}{\eta\,\Delta t}\right), \qquad (5a)$$

and when L_t>θ"," it sets c_t=0 and discharges

$$d_t = \min\left(P_b,\ L_t - \theta,\ \frac{\eta\, s_{t-1}}{\Delta t}\right). \qquad (5b)$$

When L_t=θ both are zero. The state of charge follows (2). The rule uses only the thresholxd and the current state, so it can run in real time. Its peak on a window is $\max_t g_t$ ";" Figure 4 shows the rule and the offline optimum on one window. Because the rule produces a feasible schedule for its own peak, that peak can never be below P^*.

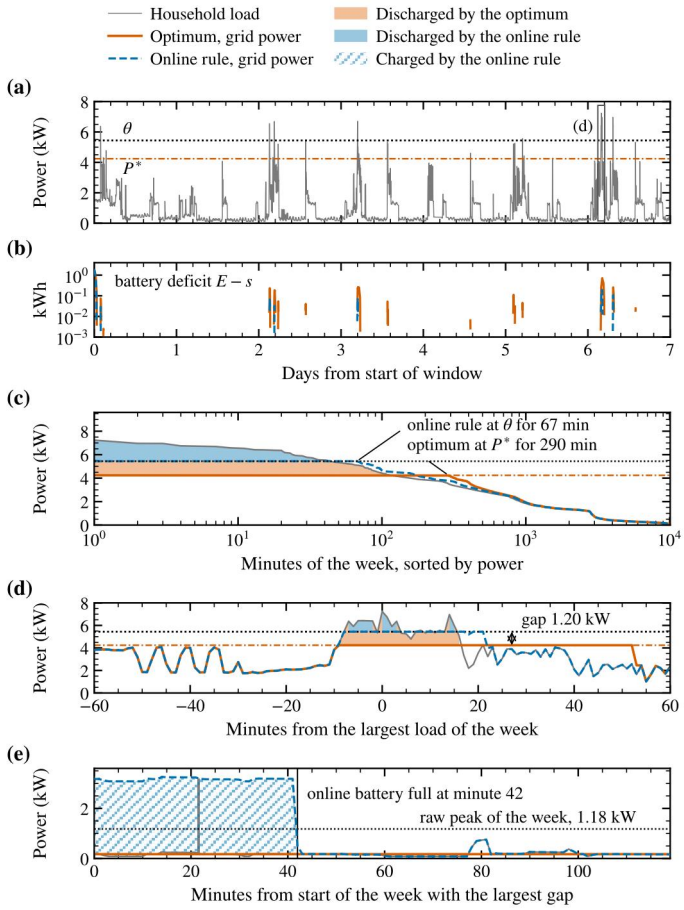


Figure 4. The online rule and the offline optimum over one week. The week is the first evaluation window, from minute 282,240, with E=4 kWh and P_b=3 kW. (a) Household load (grey line) with the energy each policy discharges shaded: in vermillion the optimum, the greedy schedule at the optimal peak, 4.24 kW (dash-dot), and in blue the online rule, which discharges only above θ"," 5.44 kW (dotted); the box marks the two hours of (d). (b) Shortfall of stored energy below capacity, E-s"," on a logarithmic axis; minutes with a full battery are blank. (c) Load-duration curves of the week, each series sorted from highest to lowest on a logarithmic axis: the flat tops show how long each policy holds the grid at its peak. (d) Grid power of the online rule (dashed) and of the optimum (solid) in the two hours around the largest load. (e) The first two hours of the week with the largest gap: the online rule recharges from half capacity (hatched) and lifts the grid above the raw peak of the week (dotted).

Relative gap. For a window w"," let P_w^* be the offline optimum, which Algorithm 1 returns as P^_w to within τ"," and let P_w^c be the peak of the online rule. The relative gap is

$$\gamma_w = \frac{P_w^c - \hat{P}_w}{\hat{P}_w}. \qquad (6)$$

Using P^_w for P_w^* lowers γ_w by at most (1+γ_w)τ/P^_w "," which is below 10^(-3) wherever the optimum exceeds 1 kW and at most 0.011 at 4 kWh. We use relative gaps because they are scale-free. Section 7 shows that they become unstable when the optimal peak is small, and we therefore also report the absolute gap of the primary case.

Structure of the online peak. Let M=max_t L_t be the raw peak of a window. The online rule is the greedy schedule of Algorithm 1 run at the target θ"," which gives its peak in closed form.

Proposition 1. Run the online rule from the same initial state as Algorithm 1. If P^*≤θ"," the rule never runs short of power or stored energy, and its peak is θ when M≥θ and lies between M and min(θ,M+P_b) when M<θ"." In particular, if P^*≤θ≤M"," the relative gap is

$$\gamma = \frac{\theta}{P^*} - 1. \qquad (7)$$

If instead θ<P^*=M-P_b and the rule never runs out of stored energy, its peak is P^* and the gap is zero.

Proof. With the target θ"," (5a) is the greedy charge with its cap, and (5b) delivers the greedy discharge L_t-θ whenever that is within P_b and the stored energy. If P^*≤θ"," the target θ is feasible, so by the Corollary the greedy test never fails, every discharge is the greedy one, and the rule is the greedy schedule. Its grid power is then θ when L_t≥θ"," and L_t+c_t "," between L_t and min(θ,L_t+P_b)"," otherwise, which gives the peak in both cases. If θ<P^*=M-P_b and energy never

runs out, (5b) delivers min(P_b,L_t-θ)"," so the grid power never exceeds max(θ,M-P_b)=P^* "," and no schedule has a lower peak.
An optimum sits at the power bound when it equals the lower end of the bounds in the proof of Theorem 1,

$$\max(0, M - P_b) \le P^* \le M. \qquad (8)$$

At 4 kWh, 84 of the 119 evaluation windows have P^≤θ≤M"," and in every one of them the online peak equals θ exactly, so (7) with P^ in place of P^* gives their gaps. In 19 windows the threshold lies below the optimum, which sits at the power bound; the rule never ran out of energy there, and their gaps are zero to within 2×10^(-5) "." The other 16 windows have M<θ"," and their online peaks lie between M and min(θ,M+P_b)"," the upper bound being reached in four. Figure 5a draws both policies as maps from load to grid power in the minutes in which they use the battery, and Figure 5b plots the observed peak minus the predicted one, or minus its upper bound.

Figure 5. The online rule against Proposition 1. (a) Grid power against household load in the minutes of the first evaluation week, with E=4 kWh, in which each policy uses the battery; in all other minutes the battery is idle and the point lies on the diagonal g=L"." The shaded band is the grid power any schedule can reach, from max(0,L-P_b) to L+P_b "." The rule caps the grid at θ and the optimum at P^";" above the diagonal the battery charges, below it discharges. (b) Observed online peak minus its prediction, or minus its upper bound, in the 119 evaluation windows, by regime of Proposition 1. In the 103 windows with M>θ the difference is rounding error, whether the peak is θ (84 windows, one circle) or P^*=M-P_b (19 windows); for the 16 windows with M<θ"," min(θ,M+P_b) is only an upper bound, reached in four (the square around the circle).

5. Design
Battery and policy parameters. Table 4 lists the design. The battery parameters are assumptions chosen to span a range of ca-pacities, not measurements of any device, and a zero-capacity battery serves as a negative control. The threshold is the 99th per-centile of the observed load of the training period,

$$\theta = Q_{0.99}\{L_t : t \le T_{\mathrm{tr}},\ L_t \text{ observed}\} = 5.442 \text{ kW}, \qquad (9)$$

with linear interpolation between order statistics. The training period precedes every evaluation window, so the rule is online.

Table 4. Design of the evaluation, as fixed in the plan unless marked.

Setting	Value
Capacity E	1, 2, 4 (primary), 8 and 16 kWh; 0 as control
Power limit P_b; efficiency η	3 kW both ways; 0.95 one way
Initial and final state	$s_0 = E/2$ in every window; no final condition
Step; window	1 min; 10,080 min from a Saturday at 17:24
Training period	first 180 days; threshold (9)
Bisection tolerance τ	10^{-4} kW, at most 15 tests
H1	median relative gap below 10% at 4 kWh
Interval	percentile bootstrap, blocks of 4 windows, 10,000 resamples, seed 2018
Validity sample (V1)	30 complete days, seed 2018
Timing repetitions	5 (bisection) and 3 (linear programme); 21 interleaved, after a review

Hypothesis and decision rule fixed in advance
The primary hypothesis, H1, states that the median relative gap over the evaluation windows is below 10% when the battery has E=4 kWh. We chose 10% as the margin of practical relevance, and we do not present it as an external standard. The decision rule uses a 95% percentile bootstrap interval for the median, for weeks generated like the observed ones, drawing blocks of four consecutive windows that wrap around the end of the series. If the upper limit of the interval is below 0.10, H1 is supported; if the lower limit is above 0.10, it is rejected; otherwise the result is inconclusive. The plan with this hypothesis and rule was frozen in a version-controlled tag before the primary analysis was run, following the logic of preregistration for data that already exist; it is not publicly registered, so readers cannot verify its date. Before the freeze only the audit of the record had been run (coverage, missing values, complete windows and percentiles of the whole record); no window peak or gap had been computed.

Validity checks. Three checks must pass for any result to count as valid, and Table 5 gives their criteria and results. V1 compares the bisection with the linear programme on 30 complete 1,440-minute days with E=4 kWh, V2 checks that no online peak lies below the computed optimum minus τ"," and V3, a weak negative control that only exercises the zero-capacity path, requires a zero gap with zero capacity; V1 is itself checked with a deliberately defective variant that never charges. Before any results were computed, we amended V2 from a tolerance of 10^(-6) kW to τ"," because the bisection

can legitimately place the optimum up to τ above the true minimum; a count made after a later review found that 93 of the 595 window–capacity pairs with a battery would have failed the original criterion.

Table 5. Validity checks, their criteria and their results. Rows marked post hoc were added after a review.

Check	Criterion	Result
V1, optimum	within 10^{-3} kW of the linear programme	largest difference 9.2×10^{-5} kW
V1, defective variant	must fail	fails on 17 of 30 days, by up to 0.743 kW
V2, dominance	online peak at least the optimum minus τ	no violation in 595 pairs; smallest -9.2×10^{-5} kW
V3, zero capacity	gap exactly zero	zero in every window
All pairs, post hoc	bisection at most τ above the linear programme	1.3×10^{-6} to 9.2×10^{-5} kW above, in 595 pairs
Second implementation, post hoc	same decision on H1	same, median within 10^{-4}; not independent
Clean clone, two runs	identical result files	identical apart from timings and memory

Exploratory analyses. After the primary analysis, we evaluated a double-threshold variant that discharges above a threshold θ_d and charges only below a lower threshold θ_c "," set to the training median. We varied θ_d over the 90th, 95th and 99th training percentiles. These exploratory analyses do not alter H1. A sensitivity analysis on the power limit, the efficiency, the threshold percentile and the training length, one at a time and jointly (Section 6), was added in the same way, with the same windows and bootstrap seed, as were Proposition 1 and its numerical check, the shares of windows at the power bound, and the checks of the window protocol, of missing data and of timing reported below.

Benchmark protocol. Run-time non-determinism makes execution times differ from run to run, so every timing is repeated. The bisection was timed over all 119 evaluation windows with five repetitions. Both procedures were timed on one, two, four and seven days of the first evaluation window; each timing of the linear programme includes building its constraint matrices in the interpreter, and the linear programme was not run on the full archive. The full set of 2,049,280 observed minutes was timed three times as a scale check, with the segment boundaries ignored. Timing runs had memory tracing off, because tracing inflates the time. Memory was measured in separate passes with the Python allocation tracer, which does not see memory allocated inside the compiled solver, together with the peak working set of the process. The bisection is a single-threaded loop, and the linear programme, solved by the dual revised simplex method, used about as much processor time as wall-clock time on seven days, so neither gained from the other hardware threads. After a review we added diagnostics that time the two procedures interleaved, 21 times each, also with the bisection run to tolerances from 10^(-1) to 10^(-8) kW, the solver's default feasibility tolerance being 10^(-7) "," and on every evaluation window. Timings follow the reporting rules of Hoefler and Belli: every ratio comes with its absolute times, medians with their interquartile ranges, and the interleaved diagnostic gives the confidence intervals that their fifth rule asks for.

The hardware is described by its characteristics: an eight-core processor with sixteen hardware threads and 16 GiB of installed memory, running 64-bit Microsoft Windows. The software is Py-thon with NumPy and SciPy, whose versions are recorded with the code, and the figures are drawn with Matplotlib.
Reproducibility. Figure 6 shows the stages of the pipeline, the data at each one and what is checked there. All random elements are seeded: the validity sample and the bootstrap use seed 2018. Every number in the text comes from a result file written by the pipeline, which is available from the corresponding author on request, and a check compares the decimal values in the text with those files. The bisection is tested against the linear programme on 200 random instances, and the tests check monotonicity and the state-of-charge invariants. A second implementation, written after the primary analysis from the frozen plan and its amendments, without the analysis code, and a clean clone of the repository, with the data downloaded and verified again, reproduce the results (Table 5). The second implementation is not independent: the amendments it read report the primary result, and it ran with the same software and the same copy of the data. Reproducibility of this kind is a minimum standard when independent replication is not possible, and the repository follows common recommendations for research computing.

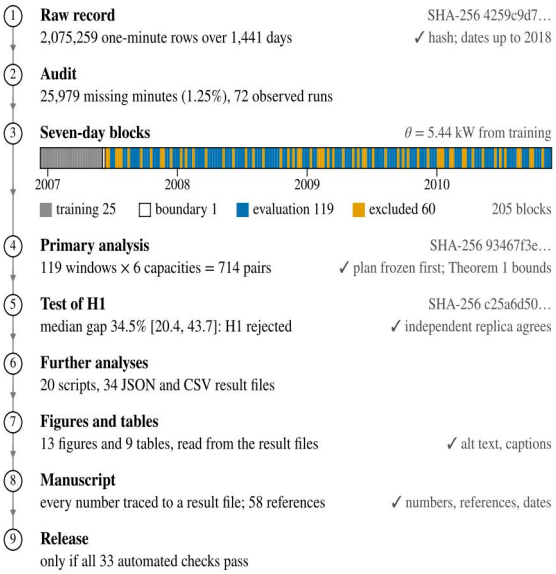


Figure 6. Architecture of the analysis pipeline. The stages run in order down the left axis, each with the size of its data; the strip shows the 205 seven-day blocks of the record in time order, col-oured by role as in Figure 1. On the right are the first eight hexa-decimal digits of the SHA-256 fingerprints of the raw data and of the primary results and, after a check mark, what an automated check verifies at that stage before the document can be released. Counts and fingerprints are read from the files when the figure is drawn.

6. Results

Validity. All checks passed (Table 5). The defective variant fails, so V1 detects a faulty procedure, and over all window–capacity pairs the bisection stays within the bound of Theorem 1.

Primary result. For E=4 kWh, the median relative gap over the 119 evaluation windows is 34.5%, with a 95% bootstrap interval of 20.4% to 43.7%. The interquartile range is 7.6% to 64.0%, and only 21.0% of windows have a gap below 5% and 26.9% below 10%. The lower limit of the interval lies above 10%, so H1 is rejected (Table 6). Figure 7a shows the distribution of the gap for every capacity.

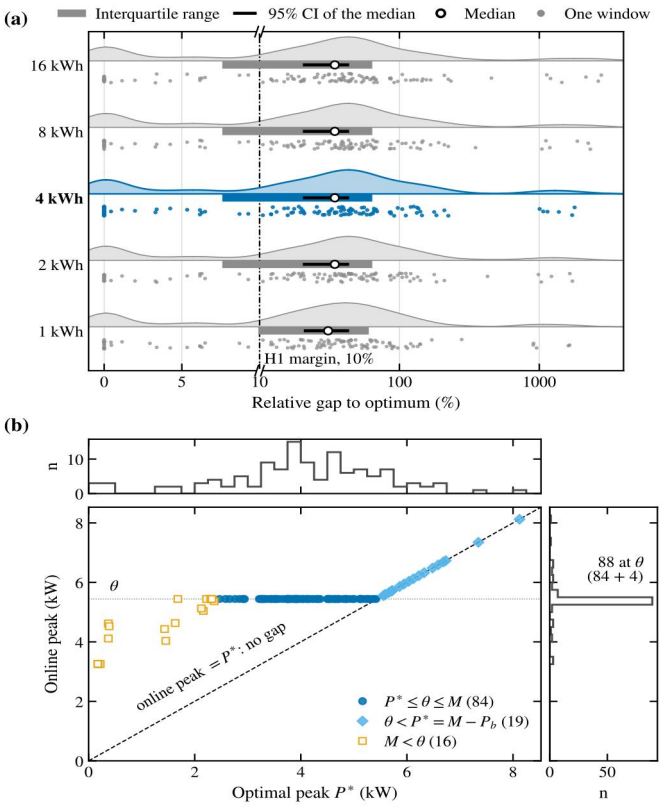


Figure 7. Distribution of the relative gap and of the online peak. (a) The relative gap in the 119 evaluation windows for each capacity, on an axis linear up to 10% and logarithmic above (double bars): a kernel density (top), the interquartile range (bar) with the median (circle) and its 95% bootstrap interval (black), and the windows themselves (dots). The primary capacity is in blue, and the dash-dot line is the 10% margin of H1. (b) Online peak against the offline optimum for E=4

kWh, by regime of Proposition 1, with marginal distributions: 88 online peaks sit at θ"," 84 in the first regime and four with M<θ"." Every point lies on or above the identity (dashed) up to the bisection tolerance, as V2 requires.

Table 6. Results for the evaluation windows. Relative gaps as defined in equation (6). Rows marked exploratory were added after the primary analysis, and the intervals for capacities other than 4 kWh after a review.

Setting	Median gap, % [95% CI][b]	IQR, %	Under 5%, %
Battery 1 kWh	30.9 [20.7, 43.7]	9.9–60.4	19.3
Battery 4 kWh (primary)	34.5 [20.4, 43.7]	7.6–64.0	21.0
Battery 2, 8 or 16 kWh	34.5 [20.4, 43.7]	7.6–64.0	21.0
Zero capacity (control)	0.0	0.0	100
Exploratory,[a] θ_d = 2.99 kW	2.0 [0.0, 12.5]	–	53.8
Exploratory,[a] θ_d = 3.75 kW	0.0 [0.0, 5.9]	–	58.8

The median offline optimum is 4.05 kW and the median raw peak 7.05 kW. Per window, the median reduction of the peak is 42.6% for the optimum and 22.8% for the online rule, whose me-dian peak is 5.44 kW; the rule achieves a median 53.5% of the achievable reduction, measured per window as

$$\rho_w = \frac{M_w - P_w^c}{M_w - P_w^*}, \qquad (10)$$

with an interquartile range of 29.2% to 87.2%. The online peak equals the threshold in 73.9% of windows. The median absolute gap between the online peak and the optimum is 1.40 kW, with an interquartile range of 0.38 to 2.12 kW.

Sensitivity to battery capacity. From 2 to 16 kWh the median and the quartiles of the gap are identical (Table 6), as Proposition 1 implies where the optimum sits at the power bound (8) and the threshold is achievable. Table 7 shows why: from 2 kWh on, more than 90% of windows sit at the bound, and only the minimum optimum keeps falling. The energy limit matters mainly at 1 kWh and in a few windows with a low raw peak: at 4 kWh, the nine windows whose optimum is off the bound all have a raw peak of about 5.04 kW or less, and 110 of the 113 windows whose raw peak exceeds 3 kW sit at the bound. A window counts as at the bound when its optimum is within 10^(-3) kW of M-P_b.

Table 7. The offline optimum and the online rule by battery capacity. Over the 119 evaluation windows: share of windows at the power bound (8), median and minimum optimum, share of online peaks at the threshold, and windows where the rule does worse than no battery.

E, kWh	At bound, %	Optimum, kW	At θ, %	Worse
1	74.8	4.21 / 0.18	71.4	8
2	90.8	4.05 / 0.18	72.3	9
4	92.4	4.05 / 0.17	73.9	11
8	94.1	4.05 / 0.16	73.9	12
16	95.0	4.05 / 0.14	75.6	13

Exploratory variants. The double-threshold variants in Table 6 show that the gap depends strongly on the discharge threshold. With the 99th percentile, 5.44 kW, the median gap equals the pri-mary value, but the two rules differ in 11 of the 119 windows, where the variant charges only below the training median. These variants were chosen after seeing the primary result, so they show where the gap comes from, not how a confirmed rule would per-form.

Computational cost. Figure 8 reports the timings. The bisection takes a median of 6.72 ms for a seven-day window and 0.93 ms for a one-day window, and the linear programme 144.28 ms for the same window, about 21.5 times longer. Over one to seven days, the log–log slope of the bisection is 1.02 and that of the linear programme is 1.03, and at every length the interquartile ranges of the two procedures do not overlap. In a separate run, building the matrices of the programme took 15.2 ms and solving it 128.7 ms for the seven-day window, so construction accounts for 10.6% of its time and the solver for the rest. Absolute times varied by up to about a factor of two between runs, and the ratio at seven days ranged from 17 to 26 over five separate runs; in the interleaved diagnostic it was 22.2 (95% interval 21.3–24.3). Tightening the tolerance from 10^(-1) to 10^(-8) kW raises the number of feasibility tests from 5 to 29 and the median time from 2.31 to 12.67 ms, still less than a tenth of the 145.49 ms of the linear programme (Figure 8b). Over the 119 windows the median ratio is 21.1 (interquartile range 20.1–23.6), but in the low-load week with the largest gap the linear programme needs 4.17 s and 24,048 simplex iterations (Figure 8c). The bisection over all 119 windows takes a median of 0.776 s, and the full set of observed minutes takes 1.39 s, consistent with the per-window rate, because 2,049,280/10,080×6.72 ms is about 1.37 s.

Memory differs by more than an order of magnitude. The bisec-tion allocates a Python peak of 0.31 MiB for one seven-day win-dow, mostly the list of 10,080 floats, and the linear programme at least 13.2 MiB, since the tracer misses the solver's own memory.

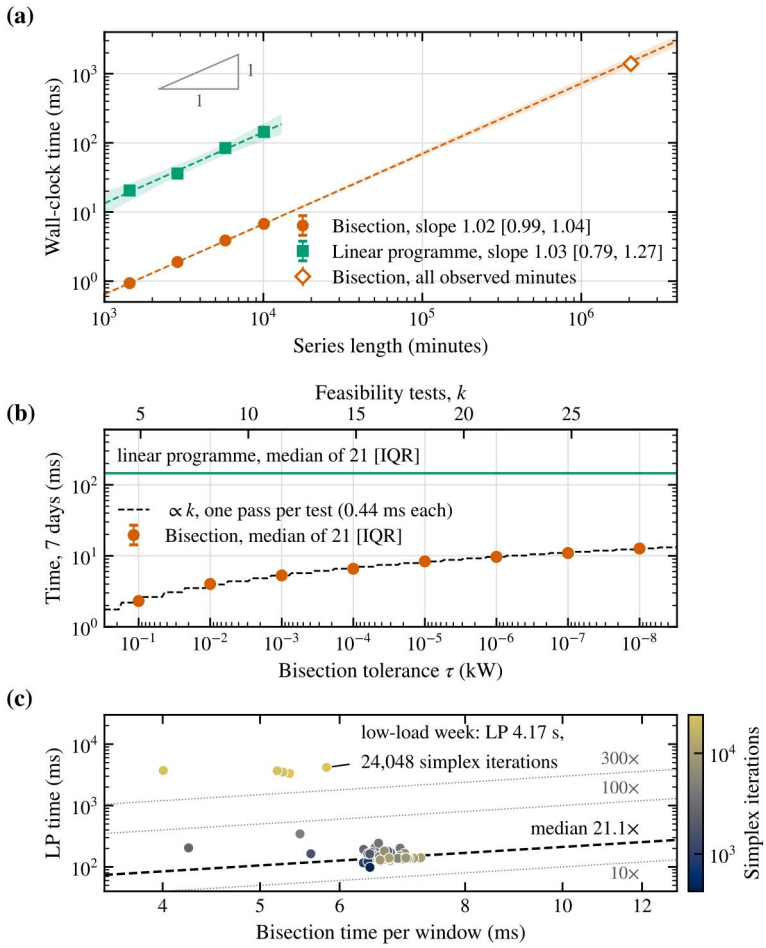


Figure 8. Computational cost. (a) Wall-clock time against series length for the bisection (circles) and the linear programme (squares) on the first evaluation window, one to seven days, with interquartile bars smaller than the markers, and for the bisection on all observed minutes (diamond). Dashed lines: least-squares log–log fits with 95% bands; the legend gives each slope with its interval (t"," two degrees of freedom), and the triangle has slope 1. (b) Bisection time on the seven-day window against its tolerance, with the number of tests on the top axis and a line proportional to that number fitted through the origin (dashed), and the linear programme with its interquartile band; 21 interleaved repetitions. (c) Linear-programme against bisection time in each of the 119 windows, on logarithmic axes, coloured by the number of simplex iterations; dotted lines mark constant ratios and the dashed line the median ratio.

For the full set of observed minutes the bisection reaches a Python peak of 62.5 MiB, almost all of it from converting the series to a list of Python floats. The whole process reached a peak working set of 393.8 MiB, which includes the interpreter and the loaded data. Table 8 gives the measured time and memory of every stage of the pipeline: run on a clean clone, the whole pipeline takes about 9 minutes, most of them in the validation against the linear pro-gramme and in the timing diagnostics.

Table 8. Measured cost of each stage of the pipeline. Run on a clean clone with 3 repetitions of each step: wall-clock time as median [interquartile range], and the median of the peak working set of the largest process of the stage; the word processor that exports the PDF is not counted.

Stage	Time, s	Memory, MiB
Download check and data audit	2.6 [2.6–2.6]	217
Primary analysis and benchmark	27.2 [27.1–27.3]	394
Linear-programme validation	199.2 [199.1–202.2]	260
Threshold sweep	38.9 [38.8–39.0]	261
Multiverse	25.6 [25.6–25.8]	261
Timing diagnostics	146.8 [146.8–146.8]	262
Other secondary analyses	64.2 [64.2–64.3]	262
Figures	13.6 [13.5–13.6]	260
Tests	6.2 [6.1–6.3]	101
Document, DOCX and PDF	7.6 [7.6–8.3]	173
Gates	1.8 [1.8–2.7]	73
Total, sum of medians	533.7	–

Sensitivity to design choices. These analyses are exploratory. Figure 9 crosses the design choices in a multiverse of 54 specifica-tions, for a 4 kWh battery: power limits of 1.5, 3 and 6 kW, thresholds at the 90th, 95th and 99th training percentiles, 90 or 180 training days, and efficiencies of 0.90, 0.95 and 1.00. All use the same windows and bootstrap draws, and the primary specification reproduces Table 6. The decision on H1 turns on two choices. With a 1.5 kW limit all 18 specifications have an upper interval limit below 10%, and with a 6 kW limit all 18 have a lower limit above it; at 3 kW the threshold decides, and the 90th and 95th percentiles bring the gap below 10%. Efficiency and training length never change the decision. Overall, 30 specifications would support H1 and 24 reject it. Figure 10 sweeps the threshold percentile and crosses it with the power limit: at 3 kW the median gap crosses 10% at a threshold of 4.45 kW, the 97.5th training percentile, and H1 would be supported up to the 96.5th percentile (4.07 kW) and is rejected from the 98.5th (5.00 kW).

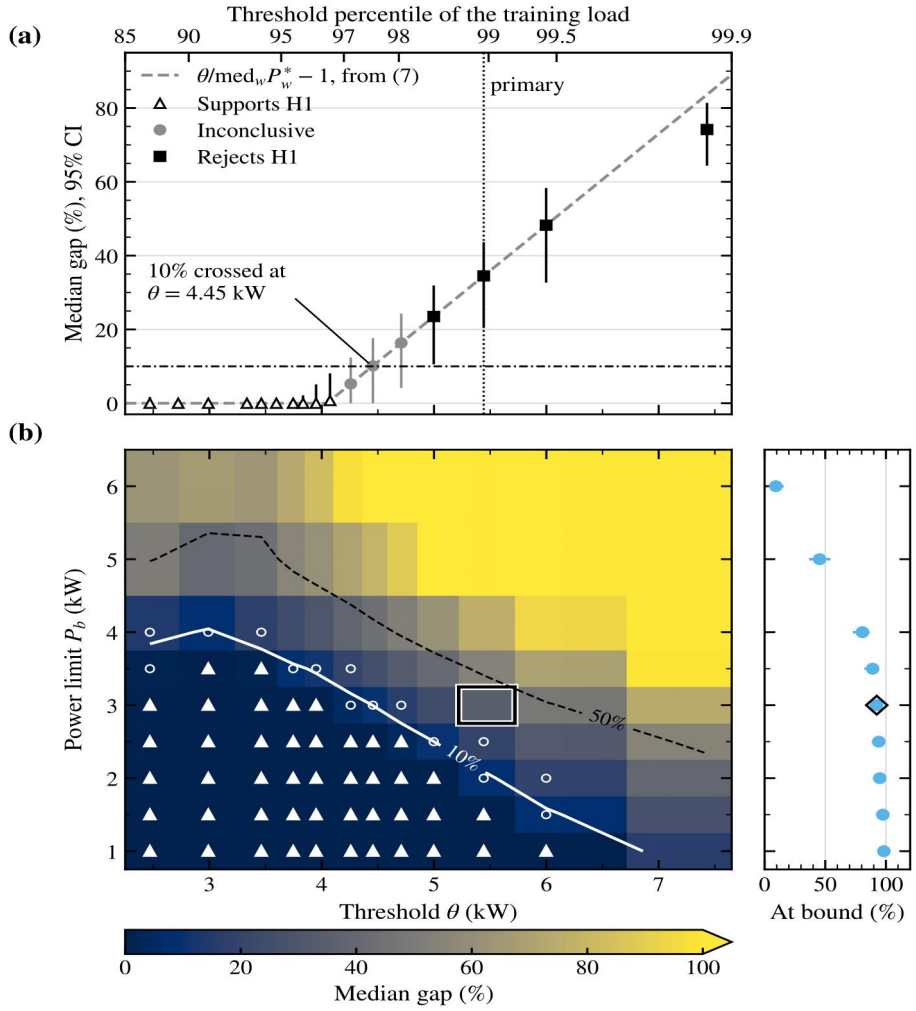


Figure 10. Threshold sweep for a 4 kWh battery (exploratory). (a) Median relative gap with its 95% bootstrap interval against the threshold, at P_b=3 kW, with the training percentile that sets it on the top axis; open triangles support H1, grey circles are inconclusive and black squares reject it. The dashed line is θ/med_w P_w^*-1"," the median gap that (7) gives if every window has P^*≤θ≤M"." (b) Median gap over thresholds and power limits, each cell as wide and tall as the step between settings; white lines are the 10% (solid) and 50% (dashed) contours, triangles and open circles mark settings that support H1 or are inconclusive, and the black box marks the primary specification. The right margin of (b) gives the share of windows whose optimum sits at the power bound (8) for each power limit, with 95% Wilson intervals (n=119); the diamond marks the primary 3 kW.

Rolling threshold. This analysis is exploratory. Recomputing the threshold before each window as the 99th percentile of the observed load in the preceding 180 days, as in (9), keeps the rule online. The rolling threshold ranges from 3.78 to 5.51 kW, with a median of 4.75 kW; it is highest in spring, after a winter of training data, and lowest in autumn (Figure 11a). It lowers the median gap to 13.5%, with a 95% interval of 7.0% to 26.8%, which contains the 10% margin, so H1 would be neither supported nor rejected, and 47.1% of windows have a gap below 10%, against 26.9% with the fixed threshold. The gap falls in 89 windows and rises in 6, and the rule does worse than no battery in 8 windows instead of 11 (Figure 11b, c).

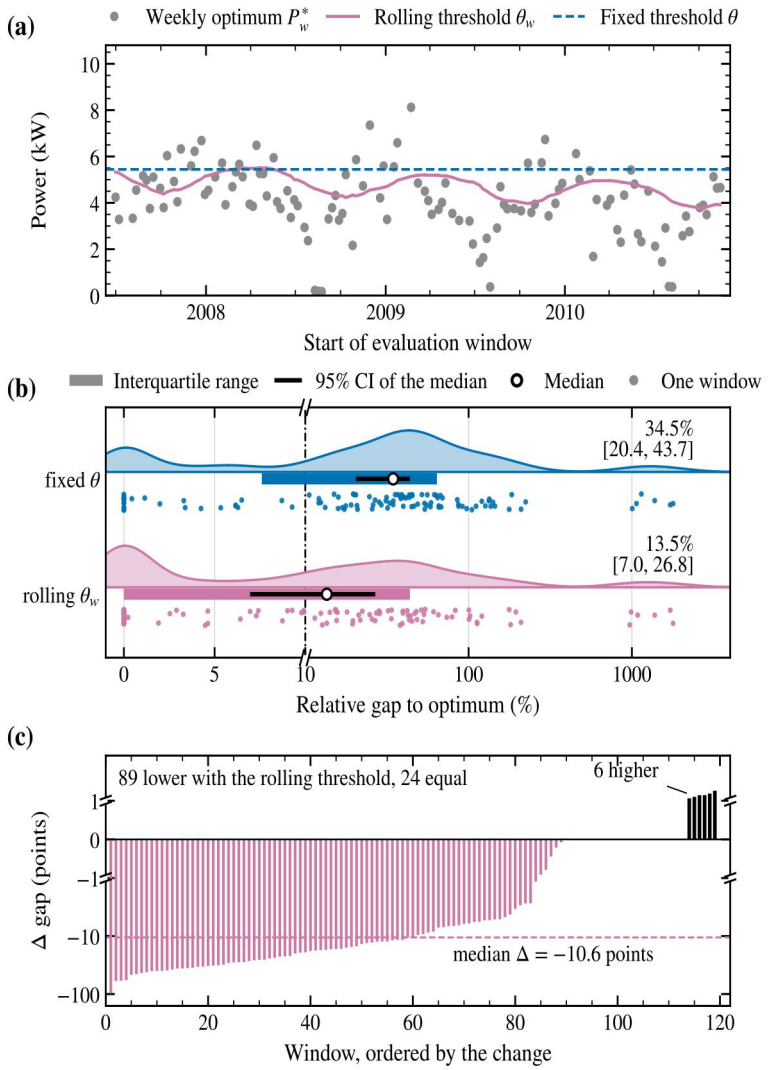


Figure 11. Fixed and rolling thresholds for a 4 kWh battery (exploratory). (a) The rolling threshold θ_w "," the 99th percentile of the observed load in the 180 days before each window, the fixed threshold θ and the optimal peak P_w^* of each window. (b) Distribution of the relative gap with each threshold, drawn as in Figure 7a, with the median and its 95% bootstrap interval written beside each. (c) Change in the gap of each window from the fixed to the rolling threshold, in percentage points, in increasing order, on an axis linear within ±1 point (double bars) and logarithmic outside; the dashed line is the median change.

Sensitivity to the window protocol. These checks are exploratory (Table 9). Neither a full battery at the start nor an end condition on the optimum moves the median gap; the end condition raises the optimum in five windows, by at most 0.012 kW. Shifting the start of every window by one to six days leaves 119 to 124 complete windows, with median gaps of 31.3% to 34.6% and lower interval limits of at least 18.4%, so H1 is rejected at every shift. At 15-minute resolution, with the threshold at the 99th percentile of the 15-minute training means, 4.95 kW, the gap grows, because aver-aging lowers the weekly peak to a median 76.7% of its one-minute value but the threshold only to 90.9% (Figure 1c). Only 3 of the 44 four-week blocks after training are complete, too few to evaluate monthly windows.

Table 9. Robustness of the primary result to the window protocol and to missing data, for a 4 kWh battery (exploratory). Shifted windows start one to six days later than the primary ones, and their row gives the range over the six shifts. Near-complete blocks lack at most two minutes, interpolated linearly.

Variant	Windows	Median gap, % [95% CI]	H1
Primary	119	34.5 [20.4, 43.7]	rejected
Full battery at the start	119	34.5 [20.4, 43.7]	rejected
Optimum ends at least as charged	119	34.5 [20.4, 43.7]	rejected
Windows shifted by 1 to 6 days	119–124	31.3–34.6, lower limits ≥ 18.4	rejected
Near-complete blocks added	160	28.8 [18.9, 39.8]	rejected
15-minute means	119	91.9 [72.1, 136.0]	rejected
Four-week windows	3	not estimable	–

7. Discussion

Where the gap comes from. Proposition 1 locates the gap. Wherever the threshold is achievable the online peak equals it, so the gap is θ/P^*-1":" it measures how far the training percentile sits above the week's optimum, not how the battery is operated within the week. The median gap therefore falls below the 10% margin when the threshold is below

$$\theta_{\mathrm{H1}} = 1.1\,\mathrm{med}_w\, P_w^* = 4.45\ \mathrm{kW}, \qquad (11)$$

approximately, because windows change regime as θ moves; the sweep of Figure 10 crosses 10% at the same value. At 4 kWh the optimum is the raw peak minus 3 kW in 110 of the 119 windows, so H1 amounts to asking whether the median weekly raw peak exceeds about 7.95 kW; it is 7.05 kW. Figure 7b shows the consequence: online peaks cluster at 5.44 kW, while optimal peaks spread from 0.17 kW to 8.12 kW, with a median of 4.05 kW and an interquartile range of 3.32 to 5.06 kW. The training period runs from December to June, so its 99th percentile, 5.442 kW, exceeds that of the full record, 4.85 kW: the load exceeds θ in 1.00% of training minutes but in only 0.44% of evaluation minutes (Figure 1c).

The charging rule also matters. Charging adds up to 3 kW to the load of the minute, so the grid peak can rise to that load plus 3 kW, capped at θ"," even in windows that need no peak reduction at all. Of the 16 windows whose raw peak is below θ"," the online peak sits at θ in 4, at the load of its minute plus 3 kW in 7, and at another value in 5. In 11 of them it exceeds the raw peak, so the rule does worse than no battery. In all 11 the peak falls, at minutes 0 to 36, within the time the battery needs to recharge from half capacity at full power,

$$t_{\mathrm{rec}} = \frac{E - s_0}{\eta P_b} = 42\ \mathrm{min}. \qquad (12)$$

Figure 4e shows the week with the largest gap, whose raw peak is 1.18 kW. The relative gap is therefore very large in windows with a small optimum. The largest relative gap at 4 kWh is 1,788%, in a window whose optimum is 0.17 kW and whose online peak is 3.26 kW. Only 5.0% of windows have an optimal peak below 1 kW, the summer weeks of very low load in Figure 2, yet these six windows all have gaps above 100%, with a median gap of 1,261%, against a median of 31.4% for the other windows. For this reason the absolute gap of 1.40 kW, or the bounded share (10), summarizes the cost of the online rule better than the relative gap.

The exploratory variants indicate that a lower discharge thresh-old brings the online peak close to the optimum in the median window, with or without a lower charging threshold. This charging threshold did not help in these data: at 3.75 kW the share of win-dows within 5% is 58.8% for the double-threshold variant and 66.4% for the single threshold, and with the discharge threshold at the 90th and 95th percentiles 46% and 41% of windows keep a gap of at least 5%. The repository records that they came after the primary result only in the amendment that adds them, because both entered the same commit. A confirmatory test would have to fix such a design in advance and evaluate it on data not used to choose it.

Practical reading. For a designer, the offline optimum is a bound on what any rule can achieve for a given battery and load, computed in milliseconds per week, against which a rule and its threshold can be checked before a design is adopted. The power bound holds in most windows only because the limit is 3 kW: with a 6 kW limit and a 4 kWh battery, only 11 of the 95 windows whose raw peak exceeds 6 kW sit at the bound (9.2% of all windows), and the share of windows at the bound falls from 92.4% at 3 kW to 80.7% at 4 kW and 45.4% at 5 kW (Figure 10b, right margin), so this conclusion depends on the assumed power limit. It is not a competitive analysis: the worst-case ratio of this rule is unbounded, because on a constant load ϵ with ϵ+P_b<θ and a battery that is not full it draws ϵ+P_b while the optimum is at most ϵ"," and we did not test the algorithms with worst-case guarantees.

Battery use. Reaching the optimum asks more of the battery. Every minute above P^needs a discharge of at least L_t-P^"," so no schedule reaches P^with less energy than the greedy schedule, which delivers exactly that. At 4 kWh this least energy has a median of 2.00 kWh per week, against 0.31 kWh for the online rule, or 0.53 against 0.08 equivalent full cycles of the battery (Figure 12). The largest deliveries, up to 18.9 kWh in a week, come in summer weeks in which the optimum holds the grid below 0.4 kW, and in 9 weeks even the greedy schedule, which keeps the most energy in store (Lemma 1), must draw the battery to within 0.01 kWh of empty. The model does not cost this cycling (Section 10).

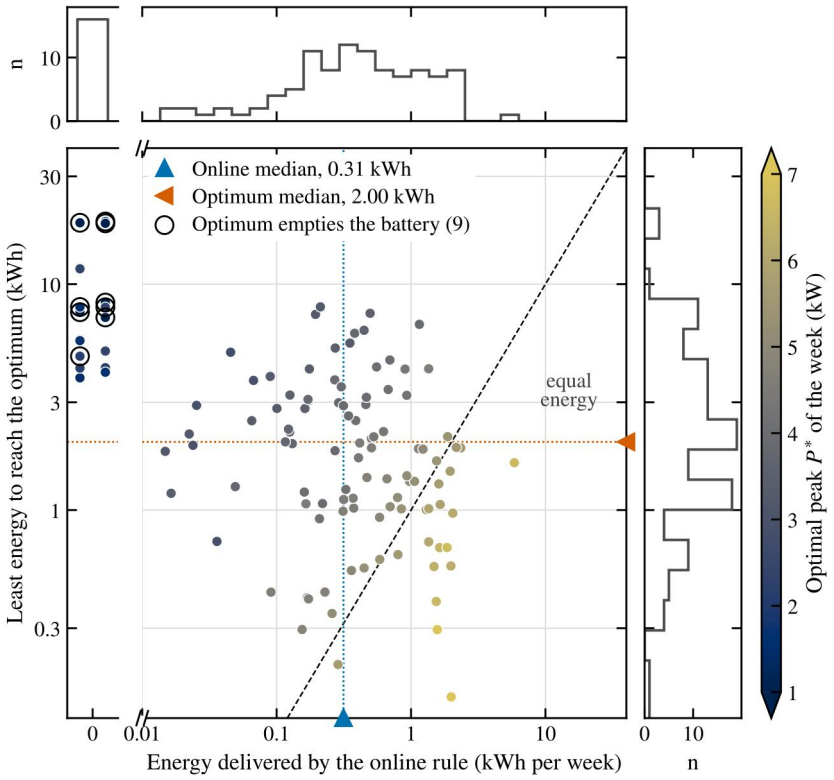


Figure 12. Energy delivered per week by the 4 kWh battery. Vertical axis: the energy delivered by the greedy schedule at P̂, the least with which any schedule reaches the optimum; horizontal axis: that of the online rule, on a logarithmic axis, with the 16 windows in which it delivers nothing in the strip at zero. Colour gives the optimal peak of the week, from 1 to 7 kW; the six weeks below 1 kW take the darkest shade. Triangles on the axes and dotted lines: the two medians; dashed: equal energy. Circles: the 9 windows where the greedy schedule brings the battery to within 0.01 kWh of empty. The margins give the distribution of each energy in logarithmic bins, with the count at zero.

Computation. The bisection is fast because each test is one pass, linear in the window length, and the number of tests is logarithmic in P_b/τ"." The linear programme is more general and returns the same optimum, to within τ"," on every window, but it is about 17 to 26 times slower at the analysis tolerance, more than ten times slower at 10^(-8) kW, and far slower in low-load weeks, where its iteration count grows (Figure 8). Both times grow about linearly over one to seven days, but four lengths of one window cannot establish growth rates, and we did not time the linear programme on longer series. For a single offline bound both procedures are fast enough; the difference matters when many windows, batteries or thresholds are evaluated.

Choices that could change the answer. Four choices shape the headline number, and Figure 9 crosses them. Training on 90 days instead of 180 raises the threshold to 5.75 kW and the median gap to 42.1%, as (7) predicts from the median optimum, 5.75/4.05-1≈0.42"." With a 6 kW limit the threshold sits far above the achievable peak.

8. Threats to validity

Construct validity. The offline optimum is a perfect-foresight benchmark for each window, but the gap of one fixed rule does not measure the cost of not knowing the future. Because the rule's peak equals its threshold wherever the threshold is achievable, the gap mostly measures how the threshold is calibrated to the week: with the 90th or 95th training percentile, the same rule with the same information has a median gap near zero (Figure 9). The gap is therefore an upper bound on the value of foresight for this load, not an estimate of it, and we did not seek the best online rule. The relative gap is unstable when the optimum is small (Section 7), and the primary metric was the relative gap although the absolute gap proved more informative.

Internal validity. Each window starts with the battery half charged and has no condition on its final state, so the optimum, which knows when the window ends, may finish emptier, while the online rule recharges at the start for the time (12); Table 9 shows that neither choice changes the median gap. Excluding windows with missing values is not neutral: the 41 blocks that lack at most two minutes have a higher median raw peak than the evaluation windows (7.58 against 7.05 kW) and more weeks that start in winter (31.7% against 18.5%), and adding them lowers the median gap (Table 9), so the primary estimate is probably too high. The bisection tolerance of 10^(-4) kW bounds the error of every optimum; V2 uses the same tolerance, V1 uses 10^(-3) kW and V3 an exact zero.

External validity. The findings concern one household over almost four years at one-minute resolution. They do not establish how the gap behaves for other households, climates, tariffs or battery designs, or at the coarser resolutions, such as 15 minutes, on which demand charges can be assessed, where Section 6 finds a larger gap.

Conclusion validity. H1 is a single test fixed in advance, and the secondary capacities carry intervals but no tests; the plan called for a correction only if they were tested as a family. The block bootstrap treats windows as dependent within blocks of four: the autocorrelation of the gap falls from 0.58 at lag one to 0.02 at lag three, and with every block length from 1 to 26 windows the lower limit stays between 17.7% and 20.9% (Figure 13). The interval still assumes stationarity, which the seasonal pattern only approximates. The exploratory analyses are not corrected for multiplicity and carry no inferential weight. The timings come from one machine and one implementation, and the ratio between the procedures varied between runs.

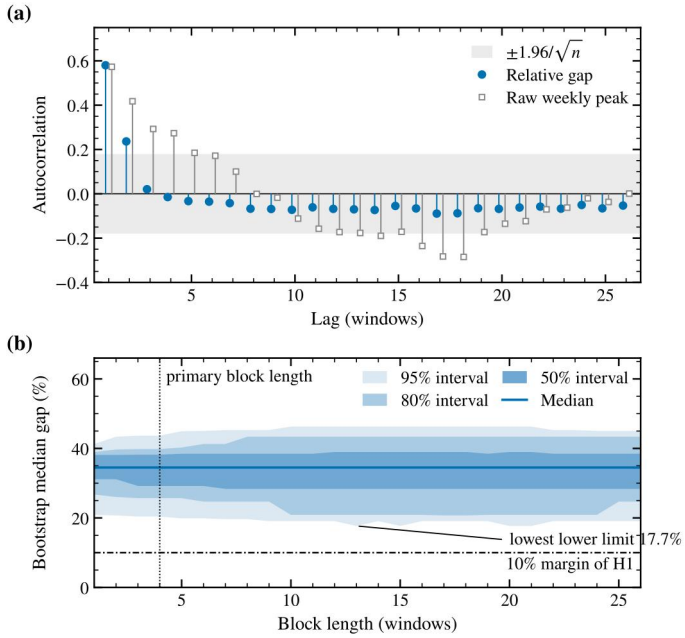


Figure 13. Dependence between evaluation windows, for a 4 kWh battery. (a) Autocorrelation of the relative gap and of the raw weekly peak at lags of 1 to 26 windows; the shaded band is ±1.96/√n with n=119"." (b) Distribution of the bootstrap median gap against the block length, from 1 to 26 windows: the shaded bands are its central 50%, 80% and 95% intervals and the line its median; the dotted line marks the primary length of four and the dash-dot line the 10% margin of H1.

9. Conclusions

On one household's record, a threshold rule set at the training 99th percentile does not meet the 10% margin set in advance for a 4 kWh battery with a 3 kW limit: its median relative gap to the offline optimum is 34.5%. Wherever the threshold is achievable the rule's peak equals it, so the gap is the threshold divided by the week's optimum, minus one; it mainly measures how the threshold is calibrated, and it changes little between 2 and 16 kWh of capaci-ty. A threshold below about 4.45 kW, the 97.5th training percentile given by (11), would have brought the median gap under the mar-gin (exploratory). The offline optimum costs a few milliseconds per week of data, over ten times less than a linear programme, so it can serve as a bound when a dispatch rule and its threshold are de-signed. In a post hoc multiverse the decision on H1 turns on the power limit and the threshold percentile alone. A rolling threshold, recomputed before each week from the preceding 180 days, low-ered the median gap to 13.5% in an exploratory analysis, with an interval that contains the margin. A confirmatory study should fix those choices in advance, compare adaptive thresholds and include a baseline without a battery.

10 Limitations and scope of the literature review

The literature review is limited. It searched two bibliographic data-bases with fixed queries, followed citations backward and forward from the closest works, and relied on abstracts; for seven cited works, by Blokhuis, Hledik, Uddin, Zheng, Chua, Leadbetter and Simonsohn with their colleagues, only the title and metadata were available, and for those we state only what the title says. It is not a full systematic review, and we do not claim priority.

The model omits battery degradation, whose cost depends on how the battery is cycled, as well as tariffs, forecast error, which affects schedules built on predicted loads, and the cost of the battery itself. The sensitivity analyses are post hoc and cover a single household. Apart from one exploratory rolling threshold, we did not test adaptive thresholds, forecast-based control or algo-rithms with worst-case guarantees, and the efficiency and power values are assumed. Public datasets with whole-house demand at finer resolution and for more homes, such as UK-DALE and RE-FIT, would allow the comparison to be repeated across house-holds.

11 Declarations

Ethics. This study analyses a publicly available dataset of electrici-ty measurements from one household, without identifiers, and involves no human participants. No ethics approval is claimed.

Data availability. The dataset is publicly available, with attribu-tion, from the UCI Machine Learning Repository at https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption. The download procedure verifies the SHA-256 hash of the text file, 4259c9d7 ece5dbee 9ab8d536 82baac68 d791c864 f0f64a52 b4043cb3 b90894b7
Code availability. The analysis code is available from the cor-responding author on request.
Funding. No funding was received for this study.
Competing interests. The author declares no competing inter-ests.
Author contributions. Luis Eduardo Muñoz Guerrero: concep-tualization, methodology, software, validation, formal analysis, investigation, data curation, visualization, writing (original draft and review and editing).

Footnotes

1. Andreas V. Stokke, Gerard L. Doorman, and Torgeir Ericson, "An Analysis of a Demand Charge Electricity Grid Tariff in the Residential Sector," Energy Efficiency 3, no. 3 (2010): 267–82, https://doi.org/10.1007/s12053-009-9071-9; Ryan Hledik, "Rediscovering Residential Demand Charges," The Electricity Journal 27, no. 7 (2014): 82–96, https://doi.org/10.1016/j.tej.2014.07.003; Erik Blokhuis et al., "Peak Loads and Network Investments in Sustainable Energy Transitions," Energy Policy 39, no. 10 (2011): 6220–33, https://doi.org/10.1016/j.enpol.2011.07.021; Chao Lu et al., "Optimal Sizing and Control of Battery Energy Storage System for Peak Load Shaving," Energies 7, no. 12 (2014): 8396–8410, https://doi.org/10.3390/en7128396.
2. Alexandre Oudalov, Rachid Cherkaoui, and Antoine Beguin, "Sizing and Optimal Operation of Battery Energy Storage System for Peak Shaving Application," in 2007 IEEE Lausanne Power Tech (IEEE, 2007), 621–25, https://doi.org/10.1109/pct.2007.4538388.
3. Bruce Dunn, Haresh Kamath, and Jean-Marie Tarascon, "Electrical Energy Storage for the Grid: A Battery of Choices," Science 334, no. 6058 (2011): 928–35, https://doi.org/10.1126/science.1212741; Xing Luo et al., "Overview of Current Development in Electrical Energy Storage Technologies and the Application Potential in Power System Operation," Applied Energy 137 (2015): 511–36, https://doi.org/10.1016/j.apenergy.2014.09.081.
4. Daniel D. Sleator and Robert E. Tarjan, "Amortized Efficiency of List Update and Paging Rules," Communications of the ACM 28, no. 2 (1985): 202–8, https://doi.org/10.1145/2786.2793; Anna R. Karlin et al., "Competitive Snoopy Caching," Algorithmica 3, no. 1–4 (1988): 79–119, https://doi.org/10.1007/bf01762111.
5. Amotz Bar-Noy, Matthew P. Johnson, and Ou Liu, "Peak Shaving through Resource Buffering," in Approximation and Online Algorithms, Lecture Notes in Computer Science (Springer, 2009), 147–59, https://doi.org/10.1007/978-3-540-93980-1_12.
6. Bar-Noy, Johnson, and Liu, "Peak Shaving through Resource Buffering"; Amotz Bar-Noy et al., "When to Reap and When to Sow – Lowering Peak Usage with Realistic Batteries," in Experimental Algorithms, Lecture Notes in Computer Science (Springer, 2008), 194–207, https://doi.org/10.1007/978-3-540-68552-4_15.
7. Moslem Uddin et al., "A Review on Peak Load Shaving Strategies," Renewable and Sustainable Energy Reviews 82 (2018): 3323–32, https://doi.org/10.1016/j.rser.2017.10.056.
8. Oudalov, Cherkaoui, and Beguin, "Sizing and Optimal Operation"; Lu et al., "Optimal Sizing and Control"; Jae Yong Lee and Seong Gon Choi, "Linear Programming Based Hourly Peak Load Shaving Method at Home Area," in 16th International Conference on Advanced Communication Technology (GIRI, 2014), 310–13, https://doi.org/10.1109/icact.2014.6778971; Sridhar Chouhan et al., "DER Optimization to Determine Optimum BESS Charge/Discharge Schedule Using Linear Programming," in 2016 IEEE Power and Energy Society General Meeting (PESGM) (IEEE, 2016), 1–5, https://doi.org/10.1109/pesgm.2016.7741576; Kalpesh A. Joshi and Naran M. Pindoriya, "Day-Ahead Dispatch of Battery Energy Storage System for Peak Load Shaving and Load Leveling in Low Voltage Unbalance Distribution Networks," in 2015 IEEE Power & Energy Society General Meeting (IEEE, 2015), 1–5, https://doi.org/10.1109/pesgm.2015.7285673; Menglian Zheng, Christoph J. Meinrenken, and Klaus S. Lackner, "Smart Households: Dispatch Strategies and Economic Analysis of Distributed Energy Storage for Residential Peak Shaving," Applied Energy 147 (2015): 246–57, https://doi.org/10.1016/j.apenergy.2015.02.039.
9. Nottrott, J. Kleissl, and B. Washom, "Energy Dispatch Schedule Optimization and Cost Benefit Analysis for Grid-Connected, Photovoltaic-Battery Storage Systems," Renewable Energy 55 (2013): 230–40, https://doi.org/10.1016/j.renene.2012.12.036; Alex Park and Petros Lappas, "Evaluating Demand Charge Reduction for Commercial-Scale Solar PV Coupled with Battery Storage," Renewable Energy 108 (2017): 523–32, https://doi.org/10.1016/j.renene.2017.02.060; J. Neubauer and M. Simpson, Deployment of Behind-The-Meter Energy Storage for Demand Charge Reduction, Technical Report NREL/TP-5400-63162 (Golden, CO: National Renewable Energy Laboratory, 2015), 7, 20, https://doi.org/10.2172/1168774.
10. Andrew J. Pimm, Tim T. Cockerill, and Peter G. Taylor, "The Potential for Peak Shaving on Low Voltage Distribution Networks Using Electricity Storage," Journal of Energy Storage 16 (2018): 231–42, https://doi.org/10.1016/j.est.2018.02.002.
11. Vasileios Papadopoulos et al., "Assessing the Impact of Load Forecasting Accuracy on Battery Dispatching Strategies with Respect to Peak Shaving and Time-of-Use (TOU) Applications for Industrial Consumers," in 2017 52nd International Universities Power Engineering Conference (UPEC) (IEEE, 2017), 1–5, https://doi.org/10.1109/upec.2017.8231939; Christopher J. Bennett, Rodney A. Stewart, and Jun Wei Lu, "Development of a Three-Phase Battery Energy Storage Scheduling and Operation System for Low Voltage Distribution Networks," Applied Energy 146 (2015): 122–34, https://doi.org/10.1016/j.apenergy.2015.02.012.

12. Yuanyuan Shi et al., "Using Battery Storage for Peak Shaving and Frequency Regulation: Joint Optimization for Superlinear Gains," IEEE Transactions on Power Systems 33, no. 3 (2018): 2882–94, https://doi.org/10.1109/tpwrs.2017.2749512; Arash Barzkar and Seyed Mohammad Hassan Hosseini, "A Novel Peak Load Shaving Algorithm via Real-Time Battery Scheduling for Residential Distributed Energy Storage Systems," International Journal of Energy Research 42, no. 7 (2018): 2400–2416, https://doi.org/10.1002/er.4010; Kein Huat Chua, Yun Seng Lim, and Stella Morris, "A Novel Fuzzy Control Algorithm for Reducing the Peak Demands Using Energy Storage System," Energy 122 (2017): 265–73, https://doi.org/10.1016/j.energy.2017.01.063.
13. Peter M. van de Ven et al., "Optimal Control of End-User Energy Storage," IEEE Transactions on Smart Grid 4, no. 2 (2013): 789–97, https://doi.org/10.1109/tsg.2012.2232943.
14. Deepranjan Dongol, Thomas Feldmann, and Elmar Bollin, "A Model Predictive Control Based Peak Shaving Application for a Grid Connected Household with Photovoltaic and Battery Storage," in Proceedings of the 7th International Conference on Smart Cities and Green ICT Systems (SCITEPRESS, 2018), 54–63, https://doi.org/10.5220/0006685300540063; Michael Koller et al., "Defining a Degradation Cost Function for Optimal Control of a Battery Energy Storage System," in 2013 IEEE Grenoble Conference (IEEE, 2013), 1–6, https://doi.org/10.1109/ptc.2013.6652329.
15. Jason Leadbetter and Lukas Swan, "Battery Storage System for Residential Electricity Peak Demand Shaving," Energy and Buildings 55 (2012): 685–92, https://doi.org/10.1016/j.enbuild.2012.09.035; Sean Barker et al., "SmartCap: Flattening Peak Electricity Demand in Smart Homes," in 2012 IEEE International Conference on Pervasive Computing and Communications (IEEE, 2012), 67–75, https://doi.org/10.1109/percom.2012.6199851.
16. Bar-Noy, Johnson, and Liu, "Peak Shaving through Resource Buffering"; Bar-Noy et al., "When to Reap and When to Sow."
17. Ying Zhang et al., "Peak-Aware Online Economic Dispatching for Microgrids," IEEE Transactions on Smart Grid 9, no. 1 (2018): 323–35, https://doi.org/10.1109/tsg.2016.2551282; Chi-Kin Chau, Guanglin Zhang, and Minghua Chen, "Cost Minimizing Online Algorithms for Energy Storage Management with Worst-Case Guarantee," IEEE Transactions on Smart Grid 7, no. 6 (2016): 2691–2702, https://doi.org/10.1109/tsg.2016.2514412; Shizhen Zhao, Xiaojun Lin, and Minghua Chen, "Robust Online Algorithms for Peak-Minimizing EV Charging under Multistage Uncertainty," IEEE Transactions on Automatic Control 62, no. 11 (2017): 5739–54, https://doi.org/10.1109/tac.2017.2699290; Sriram Govindan, Anand Sivasubramaniam, and Bhuvan Urgaonkar, "Benefits and Limitations of Tapping into Stored Energy for Datacenters," in Proceedings of the 38th Annual International Symposium on Computer Architecture (ACM, 2011), 341–52, https://doi.org/10.1145/2000064.2000105; Rahul Urgaonkar et al., "Optimal Power Cost Management Using Stored Energy in Data Centers," in Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems (ACM, 2011), 221–32, https://doi.org/10.1145/1993744.1993766.
18. Bar-Noy et al., "When to Reap and When to Sow"; Zhang et al., "Peak-Aware Online Economic Dispatching"; Zhao, Lin, and Chen, "Robust Online Algorithms"; Nottrott, Kleissl, and Washom, "Energy Dispatch Schedule Optimization."
19. Georges Hebrail and Alice Berard, "Individual Household Electric Power Consumption," UCI Machine Learning Repository, 2012, https://archive.ics.uci.edu/ ml/ datasets/ individual +household +electric +power +consumption.
20. Alexander Loginov, Malcolm I. Heywood, and Garnett Wilson, "Benchmarking a Coevolutionary Streaming Classifier under the Individual Household Electric Power Consumption Dataset," in 2016 International Joint Conference on Neural Networks (IJCNN) (IEEE, 2016), 2834–41, https://doi.org/10.1109/ijcnn.2016.7727557; Tae-Young Kim and Sung-Bae Cho, "Predicting the Household Power Consumption Using CNN-LSTM Hybrid Networks," in Intelligent Data Engineering and Automated Learning – IDEAL 2018, Lecture Notes in Computer Science (Springer, 2018), 481–90, https://doi.org/10.1007/978-3-030-03493-1_50.
21. Donald B. Rubin, "Inference and Missing Data," Biometrika 63, no. 3 (1976): 581–92, https://doi.org/10.1093/biomet/63.3.581.
22. N. Karmarkar, "A New Polynomial-Time Algorithm for Linear Programming," Combinatorica 4, no. 4 (1984): 373–95, https://doi.org/10.1007/bf02579150.
23. Efron, "Bootstrap Methods: Another Look at the Jackknife," The Annals of Statistics 7, no. 1 (1979): 1–26, https://doi.org/10.1214/aos/1176344552.
24. Hans R. Künsch, "The Jackknife and the Bootstrap for General Stationary Observations," The Annals of Statistics 17, no. 3 (1989): 1217–41, https://doi.org/10.1214/aos/1176347265.
25. Brian A. Nosek et al., "The Preregistration Revolution," Proceedings of the National Academy of Sciences 115, no. 11 (2018): 2600–2606, https://doi.org/10.1073/pnas.1708274114.
26. Andy Georges, Dries Buytaert, and Lieven Eeckhout, "Statistically Rigorous Java Performance Evaluation," in Proceedings of the 22nd Annual ACM SIGPLAN Conference on Object-Oriented Programming Systems, Languages and Applications (ACM, 2007), 57–76, https://doi.org/10.1145/1297027.1297033; Tomas Kalibera and Richard Jones, "Rigorous Benchmarking in Reasonable Time," in Proceedings of the 2013 International Symposium on Memory Management (ACM, 2013), 63–74, https://doi.org/10.1145/2464157.2464160.
27. Q. Huangfu and J. A. J. Hall, "Parallelizing the Dual Revised Simplex Method," Mathematical Programming Computation 10, no. 1 (2018): 119–42, https://doi.org/10.1007/s12532-017-0130-5.

28. Torsten Hoefler and Roberto Belli, "Scientific Benchmarking of Parallel Computing Systems," in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (ACM, 2015), 1–12, https://doi.org/10.1145/2807591.2807644.
29. Travis E. Oliphant, "Python for Scientific Computing," Computing in Science & Engineering 9, no. 3 (2007): 10–20, https://doi.org/10.1109/mcse.2007.58; Stéfan van der Walt, S. Chris Colbert, and Gaël Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," Computing in Science & Engineering 13, no. 2 (2011): 22–30, https://doi.org/10.1109/mcse.2011.37.
30. John D. Hunter, "Matplotlib: A 2D Graphics Environment," Computing in Science & Engineering 9, no. 3 (2007): 90–95, https://doi.org/10.1109/mcse.2007.55.
31. Roger D. Peng, "Reproducible Research in Computational Science," Science 334, no. 6060 (2011): 1226–27, https://doi.org/10.1126/science.1213847; Geir Kjetil Sandve et al., "Ten Simple Rules for Reproducible Computational Research," PLoS Computational Biology 9, no. 10 (2013): e1003285, https://doi.org/10.1371/journal.pcbi.1003285.
32. Greg Wilson et al., "Good Enough Practices in Scientific Computing," PLOS Computational Biology 13, no. 6 (2017): e1005510, https://doi.org/10.1371/journal.pcbi.1005510.
33. Sara Steegen et al., "Increasing Transparency through a Multiverse Analysis," Perspectives on Psychological Science 11, no. 5 (2016): 702–12, https://doi.org/10.1177/1745691616658637; Uri Simonsohn, Joseph P. Simmons, and Leif D. Nelson, "Specification Curve: Descriptive and Inferential Statistics on All Reasonable Specifications" (SSRN working paper, 2015), https://doi.org/10.2139/ssrn.2694998.
34. Bar-Noy, Johnson, and Liu, "Peak Shaving through Resource Buffering"; Zhang et al., "Peak-Aware Online Economic Dispatching."
35. Neubauer and Simpson, Deployment of Behind-The-Meter Energy Storage, 3.
36. Bolun Xu et al., "Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment," IEEE Transactions on Smart Grid 9, no. 2 (2018): 1131–40, https://doi.org/10.1109/tsg.2016.2578950; Koller et al., "Degradation Cost Function."
37. Papadopoulos et al., "Impact of Load Forecasting Accuracy"; Bennett, Stewart, and Lu, "Three-Phase Battery Energy Storage Scheduling."
38. Jack Kelly and William Knottenbelt, "The UK-DALE Dataset, Domestic Appliance-Level Electricity Demand and Whole-House Demand from Five UK Homes," Scientific Data 2, no. 1 (2015): 150007, https://doi.org/10.1038/sdata.2015.7; David Murray, Lina Stankovic, and Vladimir Stankovic, "An Electrical Load Measurements Dataset of United Kingdom Households from a Two-Year Longitudinal Study," Scientific Data 4, no. 1 (2017): 160122, https://doi.org/10.1038/sdata.2016.122.

References

1. Barker, Sean, Aditya Mishra, David Irwin, Prashant Shenoy, and Jeannie Albrecht. "SmartCap: Flattening Peak Electricity Demand in Smart Homes." In 2012 IEEE International Conference on Pervasive Com-puting and Communications, 67–75. IEEE, 2012. https://doi.org/10.1109/percom.2012.6199851.
2. Bar-Noy, Amotz, Yi Feng, Matthew P. Johnson, and Ou Liu. "When to Reap and When to Sow – Lowering Peak Usage with Realistic Bat-teries." In Experimental Algorithms, Lecture Notes in Computer Sci-ence, 194–207. Springer, 2008. https://doi.org/10.1007/978-3-540-68552-4_15.
3. Bar-Noy, Amotz, Matthew P. Johnson, and Ou Liu. "Peak Shaving through Resource Buffering." In Approximation and Online Algo-rithms, Lecture Notes in Computer Science, 147–59. Springer, 2009. https://doi.org/10.1007/978-3-540-93980-1_12.
4. Barzkar, Arash, and Seyed Mohammad Hassan Hosseini. "A Novel Peak Load Shaving Algorithm via Real-Time Battery Scheduling for Resi-dential Distributed Energy Storage Systems." International Journal of Energy Research 42, no. 7 (2018): 2400–2416. https://doi.org/10.1002/er.4010.
5. Bennett, Christopher J., Rodney A. Stewart, and Jun Wei Lu. "Develop-ment of a Three-Phase Battery Energy Storage Scheduling and Oper-ation System for Low Voltage Distribution Networks." Applied En-ergy 146 (2015): 122–34. https://doi.org/10.1016/j.apenergy.2015.02.012.
6. Blokhuis, Erik, Bart Brouwers, Eric van der Putten, and Wim Schaefer. "Peak Loads and Network Investments in Sustainable Energy Transi-tions." Energy Policy 39, no. 10 (2011): 6220–33. https://doi.org/10.1016/j.enpol.2011.07.021.
7. Chau, Chi-Kin, Guanglin Zhang, and Minghua Chen. "Cost Minimizing Online Algorithms for Energy Storage Management with Worst-Case Guarantee." IEEE Transactions on Smart Grid 7, no. 6 (2016): 2691–2702. https://doi.org/10.1109/tsg.2016.2514412.
8. Chouhan, Sridhar, Deepak Tiwari, Hakan Inan, Sarika Khushalani-Solanki, and Ali Feliachi. "DER Optimization to Determine Opti-mum BESS Charge/Discharge Schedule Using Linear Programming." In 2016 IEEE Power and Energy Society General Meeting (PESGM), 1–5. IEEE, 2016. https://doi.org/10.1109/pesgm.2016.7741576.
9. Chua, Kein Huat, Yun Seng Lim, and Stella Morris. "A Novel Fuzzy Control Algorithm for Reducing the Peak Demands Using Energy Storage System." Energy 122 (2017): 265–73. https://doi.org/10.1016/j.energy.2017.01.063.
10. Dongol, Deepranjan, Thomas Feldmann, and Elmar Bollin. "A Model Predictive Control Based Peak Shaving Application for a Grid Con-nected Household with Photovoltaic and Battery Storage." In Pro-ceedings of the

7th International Conference on Smart Cities and Green ICT Systems, 54–63. SCITEPRESS, 2018. https://doi.org/10.5220/0006685300540063.

11. Dunn, Bruce, Haresh Kamath, and Jean-Marie Tarascon. "Electrical Energy Storage for the Grid: A Battery of Choices." Science 334, no. 6058 (2011): 928–35. https://doi.org/10.1126/science.1212741.
12. Efron, B. "Bootstrap Methods: Another Look at the Jackknife." The Annals of Statistics 7, no. 1 (1979): 1–26. https://doi.org/10.1214/aos/1176344552.
13. Georges, Andy, Dries Buytaert, and Lieven Eeckhout. "Statistically Rigorous Java Performance Evaluation." In Proceedings of the 22nd Annual ACM SIGPLAN Conference on Object-Oriented Programming Systems, Languages and Applications, 57–76. ACM, 2007. https://doi.org/10.1145/1297027.1297033.
14. Govindan, Sriram, Anand Sivasubramaniam, and Bhuvan Urgaonkar. "Benefits and Limitations of Tapping into Stored Energy for Datacenters." In Proceedings of the 38th Annual International Symposium on Computer Architecture, 341–52. ACM, 2011. https://doi.org/10.1145/2000064.2000105.
15. Hebrail, Georges, and Alice Berard. "Individual Household Electric Power Consumption." UCI Machine Learning Repository, 2012. https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption.
16. Hledik, Ryan. "Rediscovering Residential Demand Charges." The Electric-ity Journal 27, no. 7 (2014): 82–96. https://doi.org/10.1016/j.tej.2014.07.003.
17. Hoefler, Torsten, and Roberto Belli. "Scientific Benchmarking of Parallel Computing Systems." In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 1–12. ACM, 2015. https://doi.org/10.1145/2807591.2807644.
18. Huangfu, Q., and J. A. J. Hall. "Parallelizing the Dual Revised Simplex Method." Mathematical Programming Computation 10, no. 1 (2018): 119–42. https://doi.org/10.1007/s12532-017-0130-5.
19. Hunter, John D. "Matplotlib: A 2D Graphics Environment." Computing in Science & Engineering 9, no. 3 (2007): 90–95. https://doi.org/10.1109/mcse.2007.55.
20. Joshi, Kalpesh A., and Naran M. Pindoriya. "Day-Ahead Dispatch of Battery Energy Storage System for Peak Load Shaving and Load Leveling in Low Voltage Unbalance Distribution Networks." In 2015 IEEE Power & Energy Society General Meeting, 1–5. IEEE, 2015. https://doi.org/10.1109/pesgm.2015.7285673.
21. Kalibera, Tomas, and Richard Jones. "Rigorous Benchmarking in Reason-able Time." In Proceedings of the 2013 International Symposium on Memory Management, 63–74. ACM, 2013. https://doi.org/10.1145/2464157.2464160.
22. Karlin, Anna R., Mark S. Manasse, Larry Rudolph, and Daniel D. Sleator. "Competitive Snoopy Caching." Algorithmica 3, no. 1–4 (1988): 79–119. https://doi.org/10.1007/bf01762111.
23. Karmarkar, N. "A New Polynomial-Time Algorithm for Linear Program-ming." Combinatorica 4, no. 4 (1984): 373–95. https://doi.org/10.1007/bf02579150.
24. Kelly, Jack, and William Knottenbelt. "The UK-DALE Dataset, Domestic Appliance-Level Electricity Demand and Whole-House Demand from Five UK Homes." Scientific Data 2, no. 1 (2015): 150007. https://doi.org/10.1038/sdata.2015.7.
25. Kim, Tae-Young, and Sung-Bae Cho. "Predicting the Household Power Consumption Using CNN-LSTM Hybrid Networks." In Intelligent Data Engineering and Automated Learning – IDEAL 2018, Lecture Notes in Computer Science, 481–90. Springer, 2018. https://doi.org/10.1007/978-3-030-03493-1_50.
26. Koller, Michael, Theodor Borsche, Andreas Ulbig, and Goran Andersson. "Defining a Degradation Cost Function for Optimal Control of a Bat-tery Energy Storage System." In 2013 IEEE Grenoble Conference, 1–6. IEEE, 2013. https://doi.org/10.1109/ptc.2013.6652329.
27. Künsch, Hans R. "The Jackknife and the Bootstrap for General Stationary Observations." The Annals of Statistics 17, no. 3 (1989): 1217–41. https://doi.org/10.1214/aos/1176347265.
28. Leadbetter, Jason, and Lukas Swan. "Battery Storage System for Residen-tial Electricity Peak Demand Shaving." Energy and Buildings 55 (2012): 685–92. https://doi.org/10.1016/j.enbuild.2012.09.035.
29. Lee, Jae Yong, and Seong Gon Choi. "Linear Programming Based Hourly Peak Load Shaving Method at Home Area." In 16th International Conference on Advanced Communication Technology, 310–13. GIRI, 2014. https://doi.org/10.1109/icact.2014.6778971.
30. Loginov, Alexander, Malcolm I. Heywood, and Garnett Wilson. "Bench-marking a Coevolutionary Streaming Classifier under the Individual Household Electric Power Consumption Dataset." In 2016 Interna-tional Joint Conference on Neural Networks (IJCNN), 2834–41. IEEE, 2016. https://doi.org/10.1109/ijcnn.2016.7727557.
31. Lu, Chao, Hanchen Xu, Xin Pan, and Jie Song. "Optimal Sizing and Control of Battery Energy Storage System for Peak Load Shaving." Energies 7, no. 12 (2014): 8396–8410. https://doi.org/10.3390/en7128396.
32. Luo, Xing, Jihong Wang, Mark Dooner, and Jonathan Clarke. "Overview of Current Development in Electrical Energy Storage Technologies and the Application Potential in Power System Operation." Applied Energy 137 (2015): 511–36. https://doi.org/10.1016/j.apenergy.2014.09.081.
33. Murray, David, Lina Stankovic, and Vladimir Stankovic. "An Electrical Load Measurements Dataset of United Kingdom Households from a Two-Year Longitudinal Study." Scientific Data 4, no. 1 (2017): 160122. https://doi.org/10.1038/sdata.2016.122.
34. Neubauer, J., and M. Simpson. Deployment of Behind-The-Meter Energy Storage for Demand Charge Reduction. Technical Report NREL/TP-5400-63162. Golden, CO: National Renewable Energy Laboratory, 2015. https://doi.org/10.2172/1168774.

35. Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. "The Preregistration Revolution." Proceedings of the Na-tional Academy of Sciences 115, no. 11 (2018): 2600–2606. https://doi.org/10.1073/pnas.1708274114.
36. Nottrott, A., J. Kleissl, and B. Washom. "Energy Dispatch Schedule Opti-mization and Cost Benefit Analysis for Grid-Connected, Photovolta-ic-Battery Storage Systems." Renewable Energy 55 (2013): 230–40. https://doi.org/10.1016/j.renene.2012.12.036.
37. Oliphant, Travis E. "Python for Scientific Computing." Computing in Science & Engineering 9, no. 3 (2007): 10–20. https://doi.org/10.1109/mcse.2007.58.
38. Oudalov, Alexandre, Rachid Cherkaoui, and Antoine Beguin. "Sizing and Optimal Operation of Battery Energy Storage System for Peak Shav-ing Application." In 2007 IEEE Lausanne Power Tech, 621–25. IEEE, 2007. https://doi.org/10.1109/pct.2007.4538388.
39. Papadopoulos, Vasileios, Thijs Delerue, Jurgen Van Ryckeghem, and Jan Desmet. "Assessing the Impact of Load Forecasting Accuracy on Bat-tery Dispatching Strategies with Respect to Peak Shaving and Time-of-Use (TOU) Applications for Industrial Consumers." In 2017 52nd International Universities Power Engineering Conference (UPEC), 1–5. IEEE, 2017. https://doi.org/10.1109/upec.2017.8231939.
40. Park, Alex, and Petros Lappas. "Evaluating Demand Charge Reduction for Commercial-Scale Solar PV Coupled with Battery Storage." Renew-able Energy 108 (2017): 523–32. https://doi.org/10.1016/j.renene.2017.02.060.
41. Peng, Roger D. "Reproducible Research in Computational Science." Science 334, no. 6060 (2011): 1226–27. https://doi.org/10.1126/science.1213847.
42. Pimm, Andrew J., Tim T. Cockerill, and Peter G. Taylor. "The Potential for Peak Shaving on Low Voltage Distribution Networks Using Elec-tricity Storage." Journal of Energy Storage 16 (2018): 231–42. https://doi.org/10.1016/j.est.2018.02.002.
43. Rubin, Donald B. "Inference and Missing Data." Biometrika 63, no. 3 (1976): 581–92. https://doi.org/10.1093/biomet/63.3.581.
44. Sandve, Geir Kjetil, Anton Nekrutenko, James Taylor, and Eivind Hovig. "Ten Simple Rules for Reproducible Computational Research." PLoS Computational Biology 9, no. 10 (2013): e1003285. https://doi.org/10.1371/journal.pcbi.1003285.
45. Shi, Yuanyuan, Bolun Xu, Di Wang, and Baosen Zhang. "Using Battery Storage for Peak Shaving and Frequency Regulation: Joint Optimiza-tion for Superlinear Gains." IEEE Transactions on Power Systems 33, no. 3 (2018): 2882–94. https://doi.org/10.1109/tpwrs.2017.2749512.
46. Simonsohn, Uri, Joseph P. Simmons, and Leif D. Nelson. "Specification Curve: Descriptive and Inferential Statistics on All Reasonable Speci-fications." SSRN working paper, 2015. https://doi.org/10.2139/ssrn.2694998.
47. Sleator, Daniel D., and Robert E. Tarjan. "Amortized Efficiency of List Update and Paging Rules." Communications of the ACM 28, no. 2 (1985): 202–8. https://doi.org/10.1145/2786.2793.
48. Steegen, Sara, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. "Increasing Transparency through a Multiverse Analysis." Perspec-tives on Psychological Science 11, no. 5 (2016): 702–12. https://doi.org/10.1177/1745691616658637.
49. Stokke, Andreas V., Gerard L. Doorman, and Torgeir Ericson. "An Analy-sis of a Demand Charge Electricity Grid Tariff in the Residential Sec-tor." Energy Efficiency 3, no. 3 (2010): 267–82. https://doi.org/10.1007/s12053-009-9071-9.
50. Uddin, Moslem, Mohd Fakhizan Romlie, Mohd Faris Abdullah, Syahirah Abd Halim, Ab Halim Abu Bakar, and Tan Chia Kwang. "A Review on Peak Load Shaving Strategies." Renewable and Sustainable Ener-gy Reviews 82 (2018): 3323–32. https://doi.org/10.1016/j.rser.2017.10.056.
51. Urgaonkar, Rahul, Bhuvan Urgaonkar, Michael J. Neely, and Anand Si-vasubramaniam. "Optimal Power Cost Management Using Stored Energy in Data Centers." In Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems, 221–32. ACM, 2011. https://doi.org/10.1145/1993744.1993766.
52. van der Walt, Stéfan, S. Chris Colbert, and Gaël Varoquaux. "The NumPy Array: A Structure for Efficient Numerical Computation." Compu-ting in Science & Engineering 13, no. 2 (2011): 22–30. https://doi.org/10.1109/mcse.2011.37.
53. van de Ven, Peter M., Nidhi Hegde, Laurent Massoulié, and Theodoros Salonidis. "Optimal Control of End-User Energy Storage." IEEE Transactions on Smart Grid 4, no. 2 (2013): 789–97. https://doi.org/10.1109/tsg.2012.2232943.
54. Wilson, Greg, Jennifer Bryan, Karen Cranston, Justin Kitzes, Lex Neder-bragt, and Tracy K. Teal. "Good Enough Practices in Scientific Computing." PLOS Computational Biology 13, no. 6 (2017): e1005510. https://doi.org/10.1371/journal.pcbi.1005510.
55. Xu, Bolun, Alexandre Oudalov, Andreas Ulbig, Goran Andersson, and Daniel S. Kirschen. "Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment." IEEE Transactions on Smart Grid 9, no. 2 (2018): 1131–40. https://doi.org/10.1109/tsg.2016.2578950.
56. Zhang, Ying, Mohammad H. Hajiesmaili, Sinan Cai, Minghua Chen, and Qi Zhu. "Peak-Aware Online Economic Dispatching for Microgrids." IEEE Transactions on Smart Grid 9, no. 1 (2018): 323–35. https://doi.org/10.1109/tsg.2016.2551282.
57. Zhao, Shizhen, Xiaojun Lin, and Minghua Chen. "Robust Online Algo-rithms for Peak-Minimizing EV Charging under Multistage Uncer-tainty." IEEE Transactions on Automatic Control 62, no. 11 (2017): 5739–54. https://doi.org/10.1109/tac.2017.2699290.

58. Zheng, Menglian, Christoph J. Meinrenken, and Klaus S. Lackner. "Smart Households: Dispatch Strategies and Economic Analysis of Distribut-ed Energy Storage for Residential Peak Shaving." Applied Energy 147 (2015): 246–57. https://doi.org/10.1016/j.apenergy.2015.02.039.