

Solving Grid-Constrained Rural DC Charging with Storage



Executive Summary

Small rural communities are critical to the deployment of comprehensive electric vehicle fast charging, enabling convenient EV travel across the western United States and to its recreation destinations. Such locations can present significant and unique challenges from a power access perspective. Often rural communities have limited grid capacity, capped by substation or transmission capacity. Unlike a larger community in a more densely populated region, there may only be one substation in town, and no nearby town for many miles. Such constraints challenge regional and national deployment efforts such as the National Electric Vehicle Infrastructure program (NEVI). This study evaluates supplemental grid-supporting storage technologies and control strategies to offer high-performance charging while minimizing cost and grid impact.

This report focuses on a small southeastern Utah community whose charging demand is shaped by both local travel and regional tourism, including season-

ally varying traffic to nearby recreation and heritage destinations. That makes the site more operationally important than its local population alone would suggest, while still leaving it with relatively low utilization for much of the year. As described in Section 1, this combination creates an unusually difficult infrastructure problem: the site must be prepared for high-power service expectations during short peak periods even though a large utility upgrade is hard to justify on utilization alone. As shown in Section 3.2, projected demand can generally be served with very little waiting, so the central challenge is not chronic congestion but the electrical and economic burden of meeting NEVI-style peak service requirements to maintain driver experience during peak periods.

The key finding is that on-site storage can sharply reduce required upfront investment relative to a conventional grid-upgrade solution. The scenario analysis in Section 3.3 shows that the base grid-upgrade case requires about \$3.7M in power-infrastructure CAPEX, whereas the principal storage-supported alterna-

tives fall in a much lower range: about \$300k for the one-flywheel case, \$470k for the one-battery case, and \$790k for the two-battery case, before any incentives or negotiated discounts. Depending on the selected configuration, this corresponds to a reduction of at least about fourfold and potentially more than tenfold.

The tariff comparison in Section 3.3.6 confirms that Schedule 6A is generally the lower-cost tariff for a low-utilization site like these, which supports its use as the primary rate structure in the study. At the same time, the comparison shows that Schedule 6A weakens the economic incentive for batteries by making lightly utilized grid capacity comparatively inexpensive to operate. This is clear when comparing the base case schedule 6A with the high performance battery case on 6A; even though the peak demand is significantly reduced the bill changes only modestly compared to the major change with schedule 6. This suggests that similar sites and the utility could benefit from a lower-power tariff structure that rewards peak-demand reduction more directly.

The sensitivity analysis in Section 3.4 shows that robustness depends strongly on both storage type and the operating rules used to evaluate compliance. Without curtailment, the modeled flywheel-only configurations are too sensitive to arrival variation to provide a dependable long-run compliance strategy, while battery-supported cases are materially more robust. The two-battery case is the strongest non-curtailed option, remaining compliant through 2031 in all 25 evaluated arrival scenarios and through

2035 in 88% of them. When limited curtailment is allowed, however, every evaluated storage configuration remains compliant through 2035 at 95% confidence. In that setting, curtailment acts as a compliance buffer for lower-cost storage-supported cases: the operator can install enough storage to deliver 600kW during short peak periods, monitor actual usage, and add storage if demand grows faster than expected.

The key takeaways are therefore straightforward:

1. The primary problem is peak power cost, not queueing.
2. Storage-supported designs can cut required upfront power-infrastructure investment by roughly 4x to more than 10x relative to the conventional grid-upgrade case.
3. Battery-dominant designs are more robust than flywheel-only designs under uncertain arrival schedules.
4. If limited curtailment is acceptable within the compliance interpretation used here, much lower-cost storage configurations become viable through 2035.
5. Schedule 6A is generally cheaper, but it also reduces the financial incentive to install batteries.

The recommended decision framing follows from those results and is developed in more detail in Section 4:

1. If the site must operate with no curtailment, prioritize the two-battery configuration as the strongest storage-supported option.
2. If limited curtailment is acceptable, the one-flywheel and one-battery

cases become credible low-CAPEX pathways to compliance.

3. Reserve physical space and electrical tie-ins for modular storage expansion so the site can start with a lower-cost configuration and add equipment if observed demand warrants it.
4. Consider future tariff or control-strategy development, including plug-level power sharing as discussed in Section 3.5, to further improve compliance margin and reduce cost.

1 Site Description

The considered town, 2020 census population 240, is a small community in San Juan County, southeastern Utah, whose travel patterns are shaped not only by local activity, but also by regional tourism. The town provides access to major recreation and heritage destinations in the surrounding region, including Bears Ears National Monument, Natural Bridges National Monument, Monument Valley Navajo Tribal Park, Hovenweep National Monument, and Goosenecks State Park. In addition, the town sits at the junction of US-191, US-163, and UT-162, making it a critical linkage in long-distance regional travel for tourists visiting parks and destinations across southern Utah, the Four Corners region, and Northern Arizona, as well as travel to and from the Navajo Nation. As a result, charging demand at the site must be understood in the context of long trip lengths, limited nearby alternatives, and seasonal visitor traffic that can increase the operational importance of a single charging site beyond what would be implied by local population alone.

At the same time, the site is difficult from an infrastructure perspective. NEVI requires four ports capable of simultaneously delivering 150 kW each, implying 600 kW of peak charging output. A typical service transformer for that level of output would be on the order of 750 kVA. The existing electrical service at the site cannot support that level of demand directly. The local utility constraint becomes especially important above roughly 500 kW, so a conventional approach

based solely on grid upgrades would require substantial upstream investment. This challenge is compounded by the fact that the site is not expected to operate at high utilization for most of the year. As shown in Figure 1, the total monthly charging time remains modest, which makes a multi-million-dollar utility upgrade difficult to justify on utilization alone.

This site therefore provides a useful case study in the distinction between nominal charger capability and practical site delivery. The central question is not simply whether 600 kW can be provided in a manner that is technically feasible, economically defensible, and consistent with expected site usage. This report examines whether the constraints can be economically addressed through on-site storage, controlled curtailment, or a combination of the two.

2 Methodology

The analysis is structured to move from projected charging demand, to economic optimization, and finally to operational robustness. First, Section 2.1 describes how future charging sessions were constructed from the available data and forecasting assumptions. Section 2.2 then explains how those projected arrivals were translated into a physically feasible charging schedule at a two-cabinet, four-plug site subject to cabinet-level power-sharing constraints and vehicle charging limits.

Building on that demand representation, Section 2.3 evaluates alternative infrastructure strategies by comparing grid upgrades, electrochemical battery storage, and flywheel-supported configurations under the applicable utility cost structure. Section 2.4 then extends the analysis beyond a single example arrival pattern by testing how the conclusions change under alternative arrival scenarios and under cases that allow controlled curtailment. Taken together, these methods provide a framework for assessing whether grid constrained sites can satisfy performance and compliance requirements without relying solely on major utility upgrades.

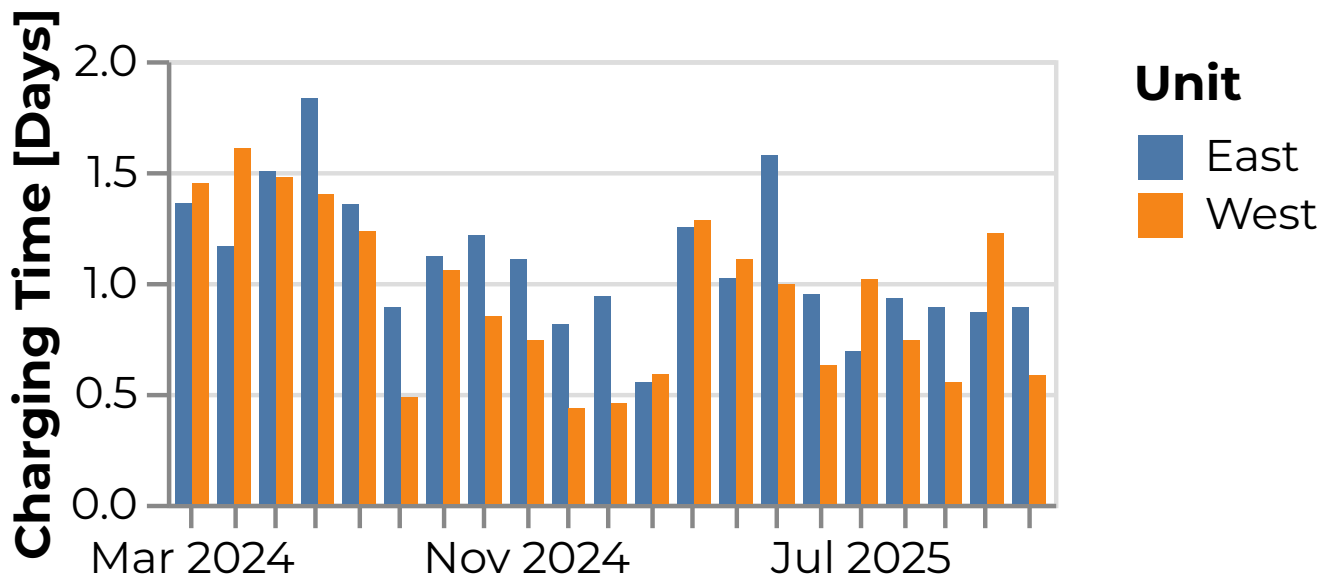


Figure 1: Monthly charging time at the existing two-plug site, illustrating the relatively low utilization that complicates justification of a major utility upgrade.

Table 1: Summary of key scenario dimensions used in the Bluff demand analysis.

Dimension	Values used
Baseline training window	One year of observed sessions (September 2024 to August 2025)
Projection years	2025 through 2035
Monte Carlo iterations	1,000 simulation draws per year
Demand-growth multipliers	2025: 1.00, 2030: 3.16, 2035: 10.00 (with intermediate annual steps)
Key outputs	Annual sessions, annual energy throughput, and hourly demand distribution

2.1 Demand Modeling

Utah Department of Transportation currently operates a two-port station in Bluff capable of 62.5 kW per port, or up to 125 kW to a single port. These stations will be replaced by the proposed higher-power station. The department shared 21 months of historical charging session data from March 2024 to November 2025, including arrival times and total energy dispensed per session. This data provided a baseline for significant variation in seasonal, day-of-week, and hourly arrival patterns in this location from which future year forecasts could be calculated.

2.1.1 Scenarios

The Bluff charging demand analysis uses a Monte Carlo scenario framework with repeated yearly simulations. Table 1 summarizes the key scenario dimensions used in the analysis.

Observed baseline context:

- 769 sessions in the baseline year.
- 30.9 MWh annual delivered energy.
- 40.2 kWh mean energy per session.

2.1.2 Approach

2.1.2.1 Data Preparation

Observed charging sessions are used to build a behavioral baseline, including session start time,

initial and final state-of-charge (SOC), session energy, and inferred battery capacity. The baseline is cleaned to include physically valid sessions and then transformed into hourly and session-level modeling frames.

2.1.2.2 Statistical Modeling Framework

The approach separates three behavior layers:

1. Session arrivals over time (hourly counts).
2. Session SOC behavior (starting SOC and fraction of remaining capacity filled).
3. Battery-capacity mix across sessions.

A Bayesian hierarchical framework is used so uncertainty in each layer is propagated into future demand realizations.

2.1.2.2.1 Hourly Arrivals Model

For each hour t , the model defines:

$$\eta_t = \alpha + \mathbf{x}_{t, \text{dow}}^\top \gamma_{\text{dow}} + \mathbf{x}_{t, \text{year}}^\top \beta_{\text{year}} + \mathbf{x}_{t, \text{hour}}^\top \beta_{\text{hour}} + \mathbf{x}_{t, \text{holiday}}^\top \gamma_{\text{holiday}}$$

$$\mu_t = \exp(\eta_t)$$

$$N_t \sim \text{NegativeBinomial}(\mu_t, \alpha)$$

where:

- N_t is hourly session count.
- $\mathbf{x}_{t, \text{dow}}$ are day-of-week indicators (one baseline category omitted).
- $\mathbf{x}_{t, \text{year}}$ are annual Fourier terms for smooth within-year seasonality.
- $\mathbf{x}_{t, \text{hour}}$ are daily Fourier terms for intraday shape.
- $\mathbf{x}_{t, \text{holiday}}$ are indicators for high-demand holiday/event windows.
- α is the overdispersion parameter.

Priors are weakly regularizing Gaussian priors on coefficients and an Exponential prior on α , which allows overdispersed counts relative to a Poisson process.

For future-year simulation, expected hourly counts are scaled by a scenario growth factor g_y :

$$\mu_{t,y}^{\text{sim}} = g_y \cdot \exp(\eta_t)$$

and sampled using the posterior predictive distribution.

2.1.2.2.2 SOC Behavior Models

Two separate Beta-regression models are fit at the session level:

1. Initial SOC, $s_i \in (0, 1)$.
2. Fraction of remaining battery filled during charging, $u_i \in (0, 1)$, where

$$u_i = \frac{\Delta s_i}{1 - s_i}$$

For each response $y_i \in (0, 1)$, the model uses:

$$\eta_i = a + \mathbf{z}_i^\top \boldsymbol{\theta}$$

$$m_i = \text{logit}^{-1}(\eta_i)$$

$$y_i \sim \text{Beta}(m_i \kappa, (1 - m_i) \kappa)$$

with κ controlling precision (Exponential prior). This preserves physical SOC bounds while allowing realistic variation.

Session SOC outcomes are then reconstructed by:

$$\Delta s_i = u_i(1 - s_i), \quad s_{f,i} = s_i + \Delta s_i$$

2.1.2.2.3 Battery-Capacity and Energy Construction

Battery capacity is sampled from the empirical observed distribution and reweighted over time toward larger batteries to anticipate the increase in both longer range and larger capacity vehicles, such as pickups and SUVs. Session energy is computed from physics-consistent SOC change:

$$E_i = C_i \cdot \Delta s_i$$

where C_i is simulated battery capacity (kWh). This ties projected energy demand directly to simulated charging behavior and battery characteristics, rather than imposing an independent energy model during production simulations.

2.1.2.3 Future-Year Simulation Logic

For each simulated year and Monte Carlo draw:

1. Hourly arrivals are sampled from the calibrated count model and scaled by the scenario growth factor.

2. Session timestamps are generated within each hour.
3. Battery capacity is sampled from the empirical distribution, with increasing weight on larger batteries in later years.
4. SOC and session energy are sampled to produce annual charging-demand realizations.

This pipeline produces uncertainty-aware annual and hourly demand outcomes rather than a single deterministic projection.

2.1.2.3.1 Scenario Growth Factor

The scenario growth factor g_y scales baseline hourly charging arrivals for year y . It is intended to represent the combined effect of two external drivers:

1. Growth in travel demand near Bluff.
2. Growth in EV adoption among vehicles traveling through the area.

Conceptually:

$$g_y \propto \text{travel-demand multiplier}_y \times \text{EV-adoption multiplier}_y$$

Travel-demand growth is anchored to long-run AADT trend behavior near Bluff (about 2% per year), which implies approximate multipliers of:

- 2030 vs 2025: 1.08
- 2035 vs 2025: 1.16

EV-adoption growth is anchored to the project's mid-range Utah adoption trajectory, which is approximately:

- 2030 vs 2025: $2.0 \times$
- 2035 vs 2025: $8.0 \times$

The implied combined multipliers are therefore:

- 2030: $1.08 \times 2.0 = 2.16$
- 2035: $1.16 \times 8.0 = 9.28$

For scenario stress testing in this case study, rounded high-growth factors are applied in simulation:

Year	Growth factor used
2025	1.00
2030	3.16

Year	Growth factor used
2035	10.00

Intermediate years use smooth progression between these anchor points.

2.2 Demand Queueing

The projected charging sessions were converted into a feasible two-cabinet, four-plug operation schedule by assigning vehicles to available plugs in arrival order and enforcing the physical charging limits at each power-sharing cabinet. In this representation, each vehicle charges as rapidly as possible subject to the lower of the cabinet power limit, the power available to the paired plug, and the vehicle charge-acceptance limit. When all plugs are occupied, an arriving vehicle remains in queue until a plug becomes available. This approach provides an estimate of site-level waiting behavior, indicative of user experience, while preserving the timing and power constraints that are most relevant for infrastructure planning.

Across all 1,000 arrival scenarios used in the analysis, the resulting queueing delays are modest. The median wait time is 0 minutes, as more than 96% of sessions begin charging immediately upon arrival. Among vehicles that wait, the average delay is approximately 5 minutes per session, and the maximum modeled wait is 45 minutes. Across nearly 35 million modeled sessions, only 105 (0.0003%) wait for longer than 30 minutes. Taken together, these results indicate that the projected demand can generally be accommodated with short queues, so the principal challenge for site design is not excessive waiting, but rather the electrical and economic difficulty of serving the high coincident charging demand.

2.3 Financial Modeling

The financial modeling framework treats the site as a co-optimization site design problem in which interconnection capacity, on-site storage sizing, and operating behavior are determined simultaneously. The site is represented as a set of energy systems connected through a common electrical bus. Within that framework, the optimizer chooses the size of the grid interconnect

and the size of the available storage technologies while also determining the time-varying energy flows needed to satisfy the projected charging demand. This formulation is especially important at grid-constrained sites because the principal design tradeoff is not between independent components, but between two competing ways of serving peak demand: increasing the amount of power that can be drawn from the grid, or installing storage that can shift a portion of that demand away from the grid connection.

In practical terms, the model determines whether it is more economical to invest in a larger interconnect, in more on-site storage, or in some combination of the two. A larger interconnect reduces reliance on storage and offers more operational certainty, but can trigger substantial utility upgrade costs. Storage can reduce the required interconnect capacity by charging during lower-demand periods and discharging during peak vehicle charging events, but this requires additional capital investment, introduces extra losses, requires physical space on site, and requires more advanced controls. Because these decisions are coupled, they must be evaluated jointly rather than sequentially.

2.3.1 Optimization Framework

The underlying optimization is formulated as a constrained cost-minimization problem. For each candidate site design, the model determines both asset sizing decisions and temporal operating decisions while enforcing the physical constraints of the system. These include the interconnect limit, the power and energy limits of the storage devices, charging-demand requirements, and the requirement that the site remain operationally feasible over the study horizon. The model therefore does not simply estimate costs for a pre-specified design; it solves for the least-cost feasible design under the assumed demand and tariff conditions.

The natural objective for this framework is net present cost. Under a net present cost formulation, the model considers not only initial capital expenses, but also future utility bills, replacement costs, and recurring operating expenses, all discounted to present value. This is the same general financial basis used elsewhere in ASPIRE's broader site-design work because it provides a

consistent way to compare capital-intensive solutions with solutions that rely more heavily on future operating expenditures. For this example, however, the project emphasis was primarily on initial capital cost rather than lifetime cost. This emphasis reflects the decision context: when viewed as a standalone site outside the strategic value of the broader charging network, this site has relatively low utilization and poor standalone return on investment, so the immediate question was whether the required service level could be achieved without a large up-front utility upgrade. In practice, that CAPEX emphasis was implemented by selecting a 1000% interest rate, which strongly suppressed the contribution of future costs while retaining the original net present cost formulation.

An additional issue is that demand grows over time, and the full optimization problem becomes computationally expensive if every month of the analysis horizon is solved simultaneously. To manage this, the financial modeling focused on the highest-demand month at the end of the study horizon. In effect, the site was sized against the most demanding monthly condition expected near the end of system life, so that a design deemed feasible under that case would be expected to handle lower-demand months as well. This approximation does not reproduce the entire year-by-year operational trajectory within a single optimization, but it provides a tractable and conservative basis for determining whether the final-year demand can be served.

2.3.2 Continuous Design Formulation

The first version of the financial analysis uses a continuous design formulation. In this setting, the optimizer is allowed to choose storage size and interconnect size as continuous variables, subject to the physical limits of the modeled technologies. This approach is useful because it reveals the underlying economic preference of the model without restricting it to preselected product offerings. If the model prefers more storage and less interconnect, or vice versa, that tendency can be observed directly. The continuous formulation is therefore best understood as a design-space exploration tool: it identifies what combination of grid capacity and storage would be preferred in principle under the modeled demand and tariff assumptions.

This continuous approach is also useful from a computational perspective because it can be represented as a linear program, without integer or binary decision variables. That makes it possible to solve the co-optimization problem efficiently even when storage sizing and dispatch are considered together over many time periods. The continuous formulation was therefore valuable both analytically and computationally: it provided a tractable way to explore the basic trade-off between interconnect capacity and storage capacity before restricting the design to specific commercial offerings.

In simplified form, the problem can be written as

$$\min \text{NPC} = C_{\text{capex}} + \sum_y \alpha_y (C_{\text{utility},y} + C_{\text{om},y} + C_{\text{replacement},y}), \quad (1)$$

where C_{capex} is the initial installed cost, $C_{\text{utility},y}$ is the annual utility cost, $C_{\text{om},y}$ is annual operating and maintenance cost, $C_{\text{replacement},y}$ represents future component replacement costs, and α_y is the discount factor for year y . The installed cost term can be represented schematically as

$$C_{\text{capex}} = c_{\text{grid}} P_{\text{grid}} + c_{\text{bat}} E_{\text{bat}} + c_{\text{fw}} E_{\text{fw}}, \quad (2)$$

where P_{grid} is the installed interconnect capacity and E_{bat} and E_{fw} are the installed battery and flywheel energy capacities. The optimizer then chooses these design variables together with the time-varying operating variables that determine when the grid is used, when storage is charged, and when storage is discharged.

The discount factor may be written as

$$\alpha_y = \frac{1}{(1+i)^y}, \quad (3)$$

where i is the annual interest rate and y is the number of years into the future. In the broader framework, this term is what distinguishes net present cost from a purely CAPEX objective. For this analysis, however, the interest rate was set to 1000%, which heavily discounts future costs and causes the original net present cost objective to behave much more like a capital-cost objective while preserving the same formal structure.

The principal operational constraints can likewise be expressed in simplified form. At each time step, energy balance must be satisfied:

$$E_{\text{grid},t} + E_{\text{bat},t}^{\text{dis}} + E_{\text{fw},t}^{\text{dis}} = E_{\text{demand},t} + E_{\text{bat},t}^{\text{ch}} + E_{\text{fw},t}^{\text{ch}}. \quad (4)$$

Grid energy in each time step is limited by the installed interconnect,

$$0 \leq E_{\text{grid},t} \leq P_{\text{grid}} \Delta t, \quad (5)$$

and the storage devices are constrained by their energy capacity, charge and discharge capability, and state of charge. The charge and discharge limits vary depending on the storage technology considered, so electrochemical storage and flywheel storage need not share the same effective rate limits. In simplified form, the charging and discharging limits can be written as

$$0 \leq E_{s,t}^{\text{ch}} \leq \rho_s^{\text{ch}} E_s \Delta t, \quad (6)$$

$$0 \leq E_{s,t}^{\text{dis}} \leq \rho_s^{\text{dis}} E_s \Delta t, \quad (7)$$

where ρ_s^{ch} and ρ_s^{dis} are the charge and discharge rate limits for storage technology s , expressed on an energy-capacity basis. The allowable depth of discharge is imposed through upper and lower limits on stored energy:

$$\underline{\sigma}_s E_s \leq E_{s,t} \leq \bar{\sigma}_s E_s, \quad (8)$$

where $\underline{\sigma}_s$ and $\bar{\sigma}_s$ define the minimum and maximum allowable state of charge. For the electrochemical storage cases, these limits restrict the usable portion of the installed energy capacity and therefore limit depth of discharge. The storage state evolution can then be written as

$$E_{s,t+1} = E_{s,t} + \eta_s^{\text{ch}} E_{s,t}^{\text{ch}} - \frac{1}{\eta_s^{\text{dis}}} E_{s,t}^{\text{dis}}. \quad (9)$$

These equations are intentionally simplified, but they capture the essential structure of the problem: the model sizes interconnect and storage together while ensuring that demand is met and all assets operate within their physical limits.

2.3.3 Curtailment Formulation

The standard financial optimization used throughout most of this analysis required all modeled charging demand to be served. In those cases, the optimizer was not allowed to drop load, and storage and interconnect capacity had to be sufficient to satisfy the full projected demand trace. Curtailment was introduced only in the subset of analyses that explicitly explored NEVI-allowable downtime¹ and curtailed-operation questions, using the federal minimum-uptime rule in 23 CFR 680.116(b)² as the reference point. It should therefore be understood as a special

¹Under 23 CFR 680.116(b), NEVI-funded stations must maintain greater than 97% average annual uptime for each charging port, calculated monthly over the previous twelve months. In round terms, that leaves less than 3% non-excluded outage time, or about 263h per port-year. This report treats that value as a formal upper bound for compliance screening, not as an operating target. An operator would not want to plan to consume the full allowance through deliberate curtailment, because the same non-excluded outage budget must preserve margin for equipment faults, network or charger failures, repair time, and any maintenance or outage events that do not qualify for exclusion. The federal uptime formula excludes certain causes when the operator can demonstrate that the port would otherwise be operational, including electric utility interruptions, vehicle-caused charging failures, scheduled maintenance, vandalism, natural disasters, and hours outside the station's identified operating schedule.

²Full text of 23 CFR 680.116(b), as displayed in the eCFR:

(b) Minimum uptime. States or other direct recipients must ensure that each charging port has an average annual uptime of greater than 97%.

1. A charging port is considered "up" when its hardware and software are both online and available for use, or in use, and the charging port successfully dispenses electricity in accordance with requirements for minimum power level (see § 680.106(d)).
2. Charging port uptime must be calculated on a monthly basis for the previous twelve months.
3. Charging port uptime percentage must be calculated using the following equation:

$$\mu = \frac{525,600 - (T_{\text{outage}} - T_{\text{excluded}})}{525,600} \times 100$$

analysis mode rather than the default site-design formulation.

In the curtailable formulation, the charging-demand model is modified so that a portion of flexible demand may be removed when the site cannot fully serve all vehicles within the imposed power limit. In simplified form, this can be written as

$$E_{\text{served},t} + E_{\text{curtailed},t} = E_{\text{demand},t}, \quad (10)$$

with

$$0 \leq E_{\text{curtailed},t} \leq E_{\text{demand},t}. \quad (11)$$

The curtailed quantity is then converted into curtailed hours and enters the objective through an explicit curtailment penalty. In simplified form, the curtailed hours can be written as

$$H_{\text{curtailed}} = \sum_t \gamma_t E_{\text{curtailed},t}, \quad (12)$$

where γ_t converts curtailed energy in interval t into an equivalent duration of curtailed service. The associated curtailment cost can then be written as

$$C_{\text{curtail}} = c_{\text{curtail}} H_{\text{curtailed}}, \quad (13)$$

where c_{curtail} is the curtailment price expressed in dollars per hour of curtailed service. In the curtailable analyses, this term is included in the monthly cost calculation, so the optimizer only accepts curtailment when doing so reduces overall cost relative to serving the demand in full. Because delivered energy is reduced when load

where μ is port uptime percentage, T_{outage} is total minutes of outage in previous year, and T_{excluded} is total minutes of outage in previous year caused by the following reasons outside the charging station operator's control, provided that the charging station operator can demonstrate that the charging port would otherwise be operational: electric utility service interruptions, failure to charge or meet the EV charging customer's expectation for power delivery due to the fault of the vehicle, scheduled maintenance, vandalism, or natural disasters. Also excluded are hours outside of the identified hours of operation of the charging station.

Source: <https://www.ecfr.gov/current/title-23/chapter-I/subchapter-G/part-680/section-680.116>

is curtailed, curtailment also lowers the modeled energy sales associated with that interval. This means that in fixed configuration modes, the optimizer is incentivized to meet all load and only curtails when meeting the load is not feasible with the given system configuration.

An important modeling simplification is that curtailed load is not replaced later. In other words, when demand is curtailed in one interval, the model does not force that unmet charging energy to be rescheduled into a later interval. Real-world operation could therefore require slightly more curtailment than this simplified formulation suggests, because some deferred charging demand may reappear later in the day. Some drivers may also divert to nearby alternatives if the site is full or power-limited. That simplification was retained because explicitly rescheduling charging demand would materially increase the complexity of the optimization and make the problem intractable.

2.3.4 Discrete Configuration Evaluation

The second version of the financial analysis moves from continuous design variables to discrete equipment configurations. In this formulation, certain aspects of the design are fixed to values that correspond to market-available modular equipment packages rather than being allowed to vary continuously. This step is important because components may not be purchased at any size and even modular options have discrete steps. A solution that is attractive in theory may require a storage size or power rating that does not correspond to an available commercial configuration, or it may imply partial deployment of equipment that would in practice be installed only in fixed increments.

The discrete analysis therefore serves a different purpose from the continuous optimization. Rather than asking what an idealized unconstrained design would look like, it asks how the site performs when limited to realistic equipment packages. In this study, this distinction matters because the project ultimately needed to compare a small number of practical storage options that could be procured and deployed. The discrete storage configurations use battery modules that each provide 270 kWh and 135 kW, corresponding to a 0.5 C device, and flywheel

modules that each provide 20 kWh and 160 kW, corresponding to an 8 C device, matching available market offerings. The continuous formulation establishes the direction of the economic tradeoff, while the discrete formulation translates that result into decision-ready cases that reflect both commercially available equipment and the threshold-driven structure of the underlying site costs.

2.3.5 Tariff Treatment and Cost Metrics

Most of the financial analysis was performed using Rocky Mountain Power Schedule 6A. That tariff structure was used as the primary basis for the site comparisons because it offers relatively favorable economics for sites with low utilization of peak demand. The analysis also considered Schedule 6 in a parallel set of scenario comparisons in order to determine whether that tariff would materially change the storage-versus-interconnect tradeoff. In practice, Schedule 6A remained the primary basis for the work, while Schedule 6 served as a sensitivity check on the tariff assumption.

The distinction between net present cost and capital cost is especially important in interpreting these results. Net present cost favors designs that reduce long-run operating costs even if they require more initial investment, because it accounts for discounted utility bills and recurring expenses over the study horizon. CAPEX-focused comparisons, by contrast, emphasize the installed cost of the initial design and therefore place more weight on minimizing the near-term budget required to make the site feasible. For this site, where the site is difficult to justify on utilization alone, that CAPEX perspective was central to the project. Accordingly, the financial results are reported in a way that preserves both views: the formal optimization framework remains grounded in net present cost, while the main project conclusions emphasize capital expense and the extent to which storage can avoid or defer expensive grid upgrades.

2.4 Sensitivity Analysis

The sensitivity analysis was designed to answer a different question from the financial optimization. The financial modeling described in Sec-

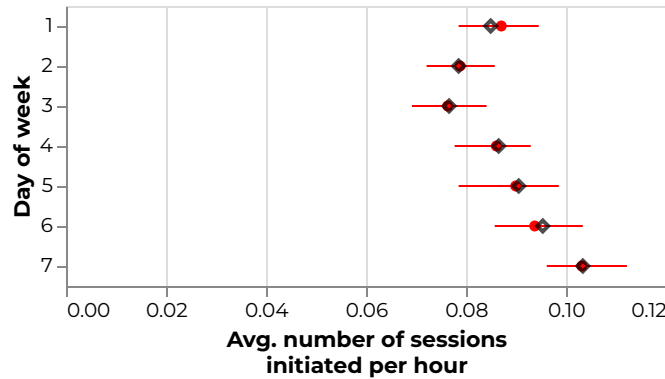


Figure 2: Baseline reproduction check for day-of-week patterning. Black diamond represents the mean of the input data. Red circle represents the mean of the means of the Monte Carlo simulations. Red line represents the interquartile range of the means of the Monte Carlo simulations.

tion 2.3 identifies attractive site designs under a given demand realization, but it does not by itself indicate how robust those designs are to variation in vehicle arrival timing. The sensitivity analysis therefore holds the physical site configuration fixed and evaluates that configuration across many alternative arrival scenarios. In this phase of the work, the quantities of interest are no longer the optimal storage and interconnect sizes, but the extent to which a selected design remains feasible when the arrival pattern changes.

For each fixed configuration, the evaluation proceeds month by month through the study horizon for each arrival scenario. Each monthly sub-problem uses the prescribed storage and interconnect configuration together with the charging schedule for that specific month and arrival scenario. The first month in which the optimization becomes infeasible is recorded as the failure point for that configuration under that arrival scenario. This produces a direct measure of robustness: rather than asking whether a design works for one representative demand trace, the analysis asks how long that same design continues to work across many plausible demand realizations.

In the non-curtailed analyses, failure occurs when the monthly problem can no longer satisfy all charging demand with the fixed site configuration. In the curtailed analyses, the same month-by-month procedure is used, but failure can also occur when cumulative curtailed hours exceed the allowable threshold for the year. This structure makes it possible to compare fixed configurations on a common basis and to distinguish

designs that are merely cost-effective in one scenario from designs that remain dependable across a wide range of future arrival patterns.

3 Results

3.1 Demand Modeling

3.1.1 Validation and Verification

3.1.1.1 Baseline Reproduction Check

The simulation reproduces baseline charging activity at the annual level:

- Observed baseline sessions: 769.
- Simulated 2025 median sessions: 770 (10th-90th percentile: 720 to 823).

For annual energy throughput:

- Observed baseline total: 30.9 MWh. Red line represents the interquartile range of the means of the Monte Carlo simulations.
- Simulated 2025 median total: 34.5 MWh (10th-90th percentile: approximately 32.1 to 37.4 MWh).

This indicates strong calibration on activity volume, with somewhat higher modeled per-session energy than observed in the baseline data.

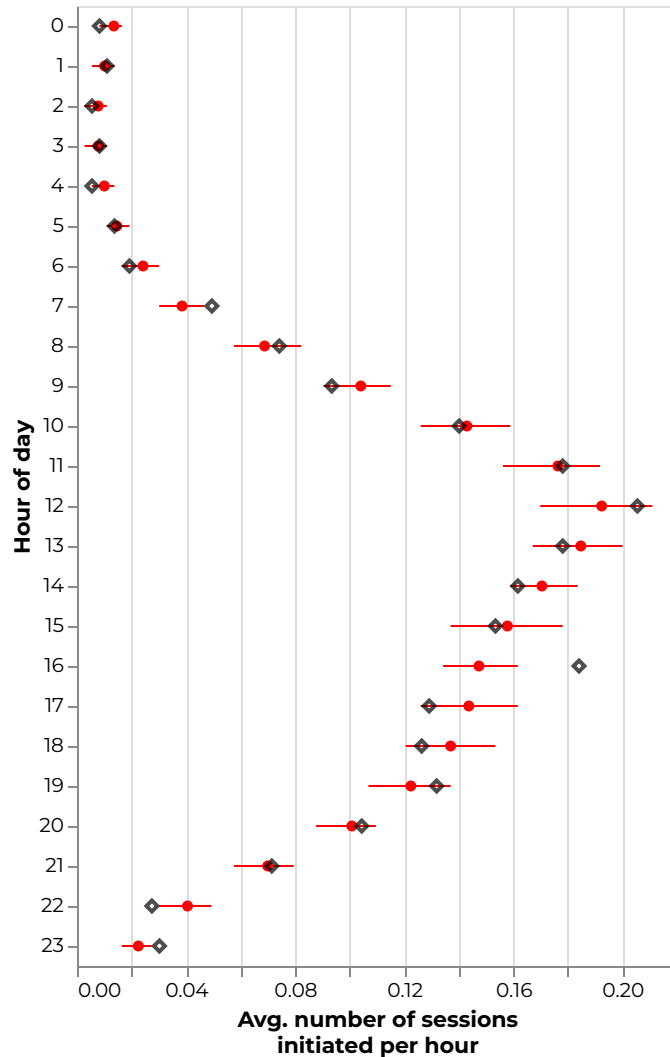


Figure 3: Baseline reproduction check for hour-of-day patterning. Black diamond represents the mean of the input data. Red circle represents the mean of the means of the Monte Carlo simulations. Red line represents the interquartile range of the means of the Monte Carlo simulations.

3.1.1.1 Detailed Baseline Reproduction Checks

3.1.1.2 Internal Consistency Checks

The simulations show expected directional behavior:

- Median annual sessions scale from 770 (2025) to 2,435 (2030) to 7,707 (2035), matching intended demand-growth structure.
- Median annual energy grows faster than sessions, reflecting increasing battery-size mix and per-session energy.
- Hourly peak pressure rises substantially across horizon years, which is critical for infrastructure sizing.

The model appears fit-for-purpose as a planning stress-test tool: it reproduces baseline activity well and generates coherent growth dynamics with explicit uncertainty bands.

3.2 Demand Queueing

The queueing results confirm that the projected demand can generally be accommodated with very short waits, even when evaluated across the full set of modeled arrival scenarios. The median wait time is 0 min, and 99.2% of sessions begin charging immediately. Averaged across all sessions, the waiting time is only 2 s. Even when considering only sessions that experience any de-

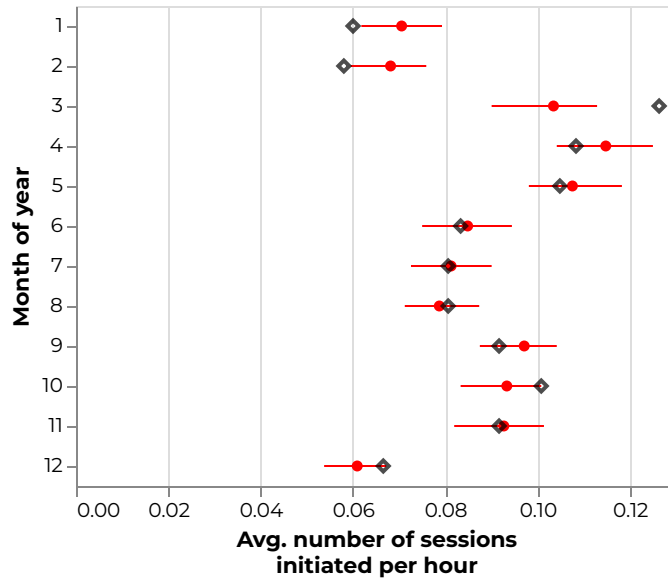


Figure 4: Baseline reproduction check for month-to-month patterning. Black diamond represents the mean of the input data. Red circle represents the mean of the means of the Monte Carlo simulations. Red line represents the interquartile range of the means of the Monte Carlo simulations.

lay, the average wait is 5 min 26 s. The maximum modeled wait is 45 min 31 s, but such events are rare and occur only during unusually concentrated periods of demand. For example, one 2035 scenario includes six arrivals within about seven minutes during the modeled Indigenous Peoples' Day/Columbus Day travel period, each waiting roughly 40-45 min; similar longer-wait events also appear around the modeled National Public Lands Day travel period.

Table 2 summarizes the most important queueing metrics. The session counts in this table are modeled session instances across the full portfolio of 1,000 simulated arrival schedules, not a forecast of how many delayed sessions would occur in a single realized future. Across that full modeled portfolio and all scenario-years, only 124,840 session instances, or 0.36%, wait longer than 5 min. That share falls to 8,684 session instances, or 0.025%, for waits longer than 15 min, and to just 105 session instances, or 0.0003%, for waits longer than 30 min. Those longer waits are concentrated late in the forecast horizon: none occur before 2030, only 4 occur before 2033, and 101 of the 105 occur during 2033-2035, including 67 in 2035. This supports interpreting the longest queues as late-horizon stress cases rather than near-term operating expectations; by that point, additional

chargers or nearby sites could also change how regional demand accumulates. It also means the exact projected dates of those events are less important than the modeled condition they represent: late-horizon growth combined with arrival bunching during holiday-related high-demand travel windows. Taken together, these results indicate that the dominant challenge is not chronic queueing, but the cost of providing sufficient electrical capacity to serve short periods of highly coincident demand.

Figure 5 shows the full waiting-time distribution, including the large mass of sessions with zero delay. The bars at the ends of the histogram include all values that fall outside the displayed x-axis range. Figure 6 then focuses only on sessions that experienced some waiting, making the right tail easier to interpret; again, the end bars include all values beyond the plotted x-axis range. Figure 7 shows how the share of delayed sessions evolves over time, and Figure 8 illustrates how queueing intensifies during the high-demand Easter period in 2035 by first summing waiting time across all sessions within each day for each arrival scenario and then averaging those daily totals across scenarios. Waiting is measured relative to each modeled driver's desired arrival time; in practice, some peak-period delay could be mitigated by

Table 2: Summary of queueing results across all modeled arrival scenarios.

Metric	Value
Median wait time	0 min
Share of sessions with zero wait	99.2%
Average wait among all sessions	2 s
Average wait among delayed sessions only	5 min 26 s
Maximum wait	45 min 31 s
Sessions above 5 min wait	124,840 (0.36%)
Sessions above 15 min wait	8,684 (0.025%)
Sessions above 30 min wait	105 (0.0003%)

driver schedule shifting or diversion, but that behavior is not assumed here. The Easter peak example should also not be read as all-day saturation: the modeled waits are concentrated in a few high-arrival periods within the day, with lower-use periods still present outside those bunches. These figures reinforce the same conclusion from different perspectives: waiting remains negligible for the overwhelming majority of sessions, and the rare longer delays are associated with a small number of holiday or travel-peak conditions rather than with persistent congestion.

3.3 Financial Modeling

The results in this section are based on a single representative arrival scenario. That framing is intentional: the goal here is to compare design concepts under a common demand trace and isolate the core tradeoff between grid expansion and on-site storage. The question of whether those same designs remain dependable across many alternative arrival realizations is taken up later in Section 3.4.

The five scenarios represent distinct design choices selected to showcase potential pathways for making the Bluff site feasible under different equipment, cost, and compliance assumptions. The base case, described in Section 3.3.1, serves all modeled charging load with a grid connection and no on-site storage. Chemical storage, described in Section 3.3.2, uses electrochemical battery storage to support the modeled 2035 load while reducing the grid interconnect requirement. The cheap battery case, described in Section 3.3.3, contrasts with chemical storage by assuming lower-cost, higher-performance battery hardware, representing a future technology-cost

case or the effect of subsidies. The flywheel case, described in Section 3.3.4, adds flywheel storage to the chemical storage case to test the value of short-duration, high-power storage. The NEVI case, described in Section 3.3.5, is a strict-compliance benchmark sized to meet NEVI requirements through 2031 rather than a full-horizon optimized design.

The aggregate figures compare those scenarios on cost, grid capacity, and storage requirements. Figure 9 shows that the storage-supported designs reduce upfront cost substantially relative to the base case. Figure 10 and Figure 11 show the mechanism behind that reduction: the storage-supported cases keep the configured grid interconnect at 500 kW and use storage to keep grid demand below that threshold. Figure 12 and Figure 13 show the energy and power requirements that distinguish the chemical storage, cheap battery, flywheel, and NEVI designs; Table 3 reports the same scenario metrics side by side.

3.3.1 Base Case

The base case is the conventional grid-upgrade solution and serves as the reference point for all of the storage-supported cases. In this single-arrival-scenario comparison, it is the only design that relies entirely on a larger grid connection rather than on shifting demand in time with storage. That simplicity comes at a high upfront cost: the site must be built around a 750kW grid interconnect, which pushes total grid capital expense to \$3.3M. Table 4 summarizes the case metrics used in this scenario comparison.

Because this case has no on-site storage, Figure 14 contains only the power-flow panel. In the March 2035 peak period, grid demand rises to

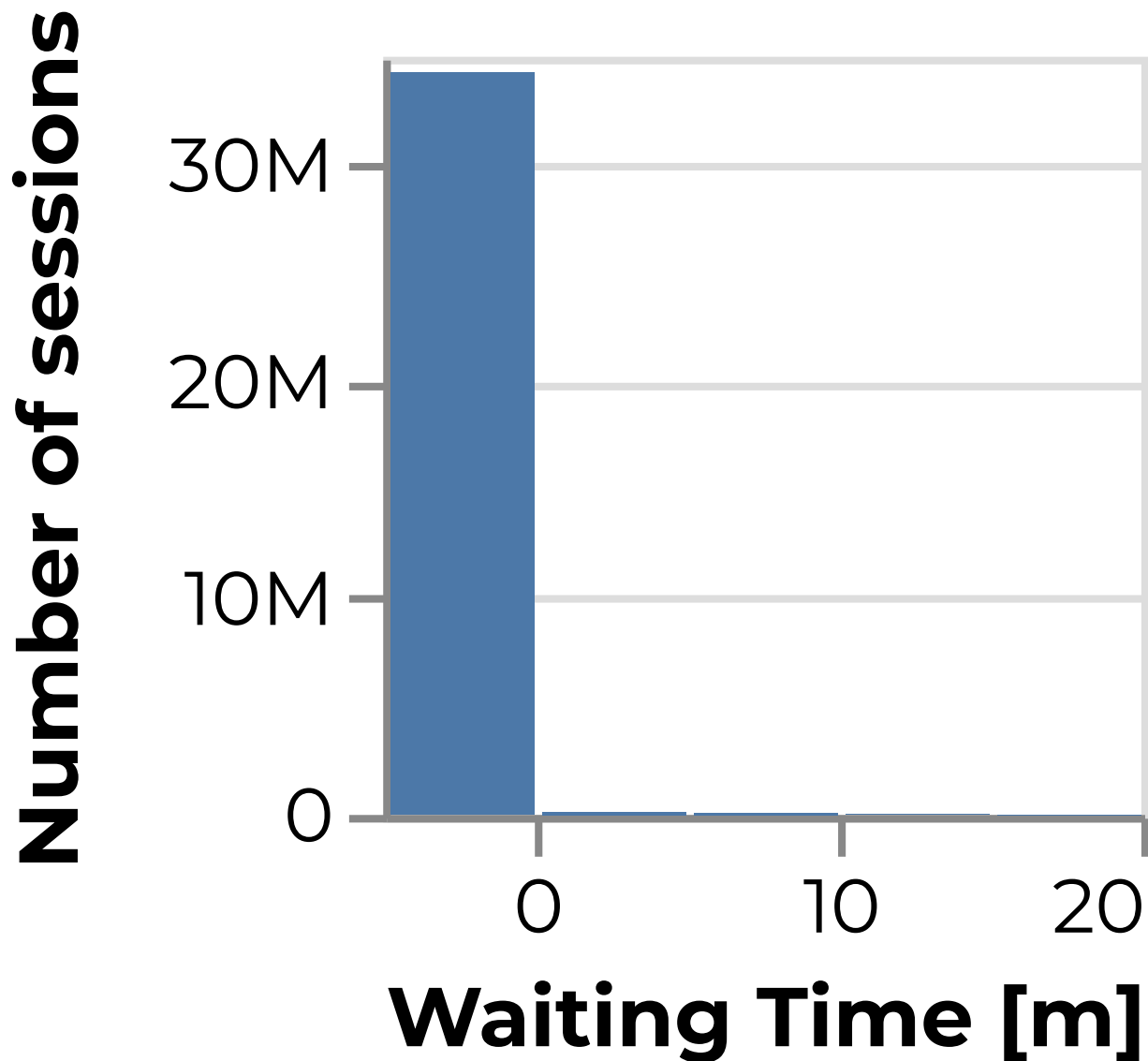


Figure 5: Distribution of session waiting time across all modeled sessions. Bars at the ends of the histogram include all values outside the displayed x-axis range.

about 720kW, showing why a pure grid solution moves the site well beyond the 500 kW threshold that drives the storage-supported designs.

3.3.2 Chemical Storage Case

The chemical storage case is the main storage-supported alternative to the base case. It pairs a 540kWh battery with a 500kW interconnect, cutting upfront cost to \$790k while still meeting the modeled charging demand in this arrival sce-

nario. The battery is represented as two 270kWh modules with 270kW total power capability at 0.5 C. Relative to the base case, the design works by trading a much smaller grid upgrade for battery dispatch during the highest coincident charging periods. Table 5 summarizes the case metrics used in this scenario comparison.

Figure 15 shows that the battery charges away from the highest-demand intervals and discharges through the sharpest spring peaks. In

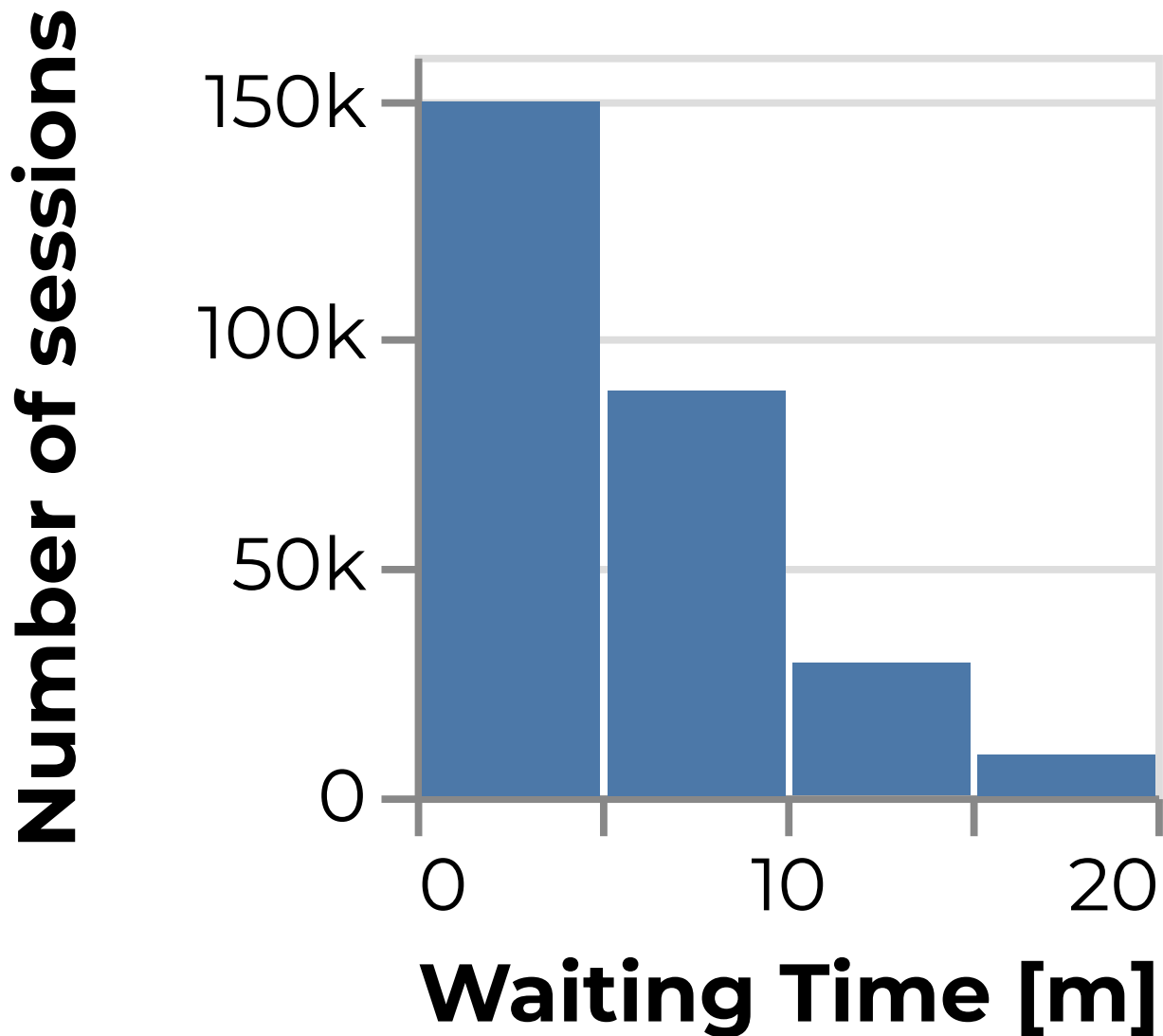


Figure 6: Distribution of session waiting time for only sessions that must wait. Bars at the ends of the histogram include all values outside the displayed x-axis range.

this case the battery is large enough to avoid charging during utility peak-demand periods, which helps keep the March 2035 grid peak at about 490kW. Section 3.4 then asks whether this same 500 kW-plus-battery architecture remains robust when arrival timing changes.

3.3.3 Cheap Battery Case

The cheap battery case keeps the same 540kWh battery energy as the chemical storage case, but assumes lower battery capital cost and a

higher one-C charge and discharge rate in order to explore the impact of technical and market advancement. That combination makes the storage-supported design materially more attractive on a first-cost basis, reducing capital expense to \$400k while preserving the same 500kW interconnect. Table 6 summarizes the case metrics used in this scenario comparison.

The important result is not just that cheaper storage lowers cost, but that it does so without shrinking the battery. Even with the higher 1 C power capability, the optimizer still keeps the

Sessions with Wait

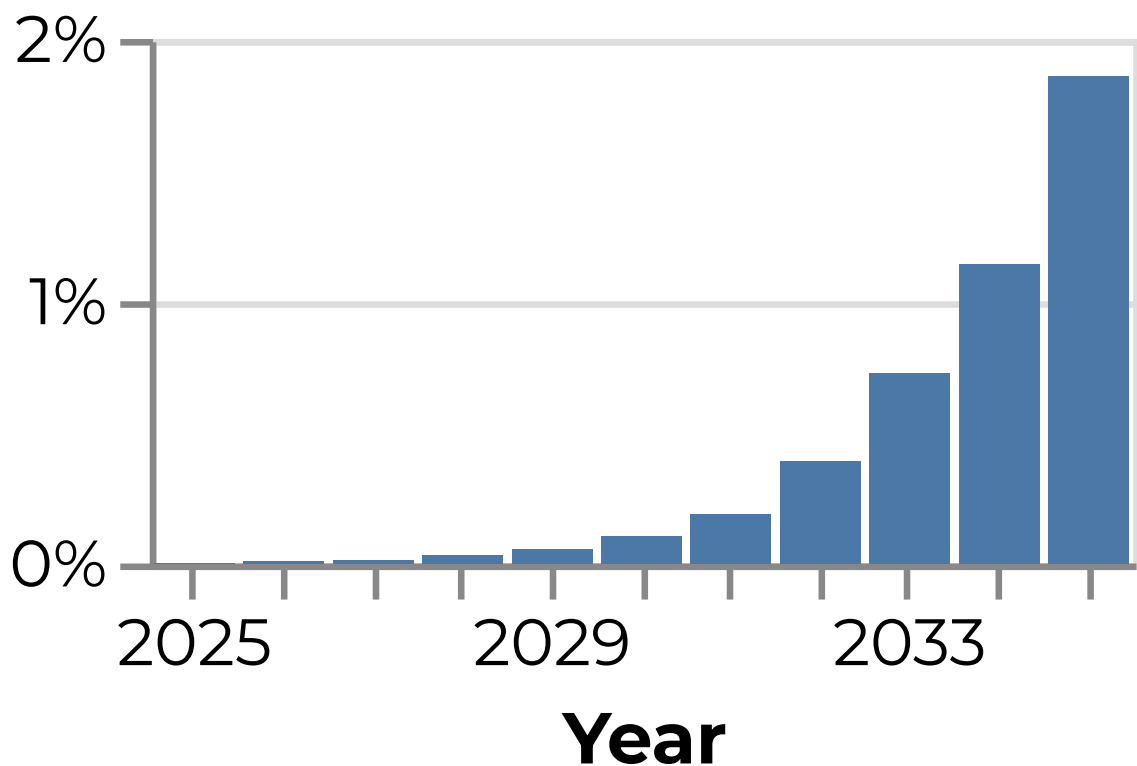


Figure 7: Share of sessions experiencing nonzero waiting time by year.

battery at 540kWh because additional energy remains cost-effective for reducing on-peak energy costs; in other words, the storage size is determined by the operational value of both energy and power, not simply by the cheapest available module count. Figure 16 shows the higher-power battery handling sharper spikes while still avoiding charging during utility peak-demand periods and pulling the March 2035 grid peak down to about 460kW. More detail on how sensitive that conclusion is to arrival timing appears in Section 3.4.

3.3.4 Flywheel Case

The flywheel case tests whether a mixed-storage architecture can manage Bluff's shortest, sharpest peaks more effectively than a battery alone. It keeps the same 540kWh battery used in the chemical storage case, adds a 20kWh flywheel, and holds the interconnect at 500kW. In

this single-arrival-scenario comparison, that produces a capital expense of \$1.0M. The modeled storage stack provides 270kW of battery power and 160kW of flywheel power. Table 7 summarizes the case metrics used in this scenario comparison.

Figure 17 shows the intended division of labor: the flywheel handles very short spikes while the battery covers more prolonged charging events. The configuration does bring the March 2035 grid peak down to about 460kW, but its additional flywheel capital cost leaves it less attractive than the battery-only options on a CAPEX basis. Whether that tradeoff looks better under broader uncertainty is addressed later in Section 3.4.

3.3.5 NEVI Case

The NEVI case is a strict-compliance benchmark rather than a full-lifetime optimized design. It asks a narrower question than the cases above:

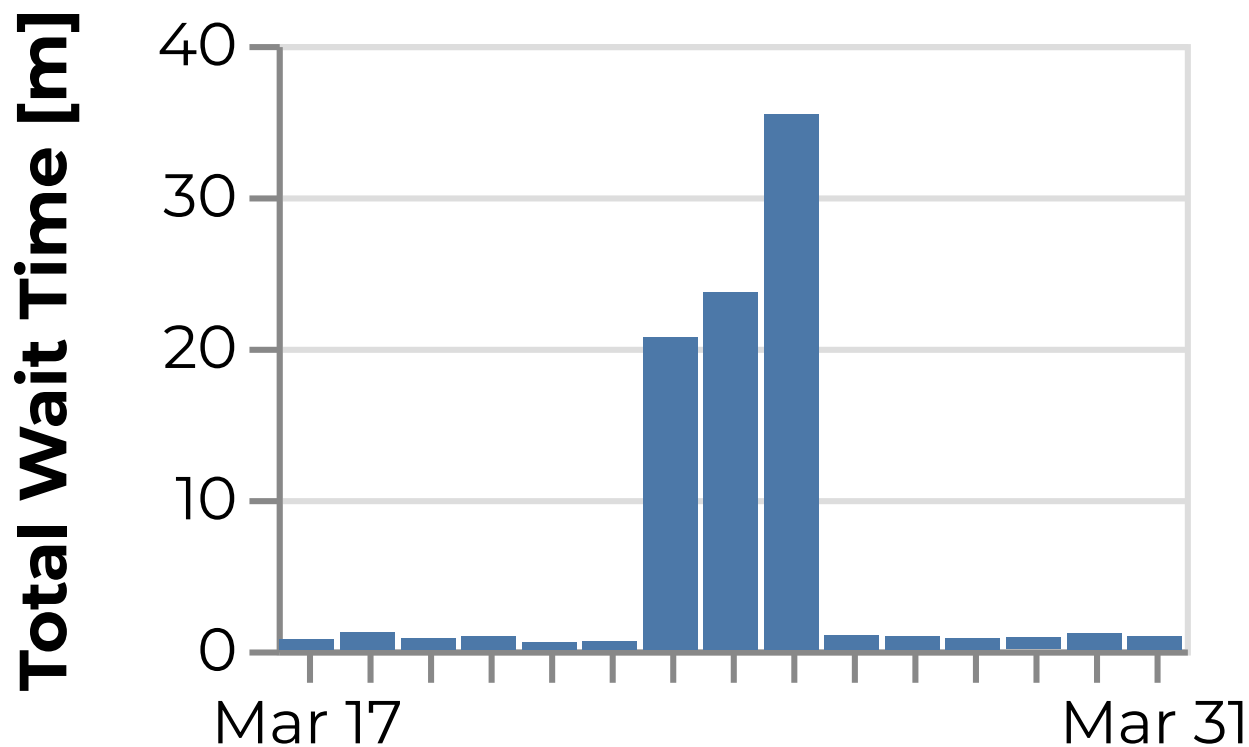


Figure 8: Average daily total waiting time during the Easter 2035 peak-demand period. For each arrival scenario, waiting time is first summed across all sessions within each day, and those daily totals are then averaged across scenarios.

how much less would the site cost if it were sized only to maintain the required NEVI service level of four ports with 150kW simultaneous output each through the end of the required compliance window, five years after opening, instead of preserving that capability through the end of system life. The other scenarios allow a port to receive up to 360kW when the paired port on the same cabinet is unused, reflecting the modeled hardware limit. The NEVI case lowers that plug-level cap to 300kW while preserving the four-port 150kW simultaneous service requirement. In this benchmark, the lower service target is represented by a 500kW grid connection paired with 40kWh of flywheel storage and no battery. Table 8 summarizes the case metrics used in this scenario comparison.

Figure 18 shows the March 2031 dispatch for that limited compliance window. This benchmark works partly because the 2031 demand forecast is materially lower than the 2035 end-of-study

demand used for the lifetime designs. It should therefore not be read as a direct replacement for those lifetime designs. Instead, it is a decision reference for the minimum upfront cost of strict NEVI compliance and could be treated as a budget-focused interim step if the site is planned with physical space and electrical tie-ins for later storage expansion. The broader robustness question is deferred to Section 3.4.

3.3.6 Schedule 6 Comparison

Schedule 6A was used for most of the project work because it is generally expected to be more favorable for sites with low utilization of grid infrastructure. That expectation should still be tested rather than assumed. This section therefore compares the main scenario designs under Schedule 6 and Schedule 6A in order to determine whether Schedule 6A is in fact the lower-cost tariff and whether the broader storage-versus-interconnect conclusions depend strongly on that choice.

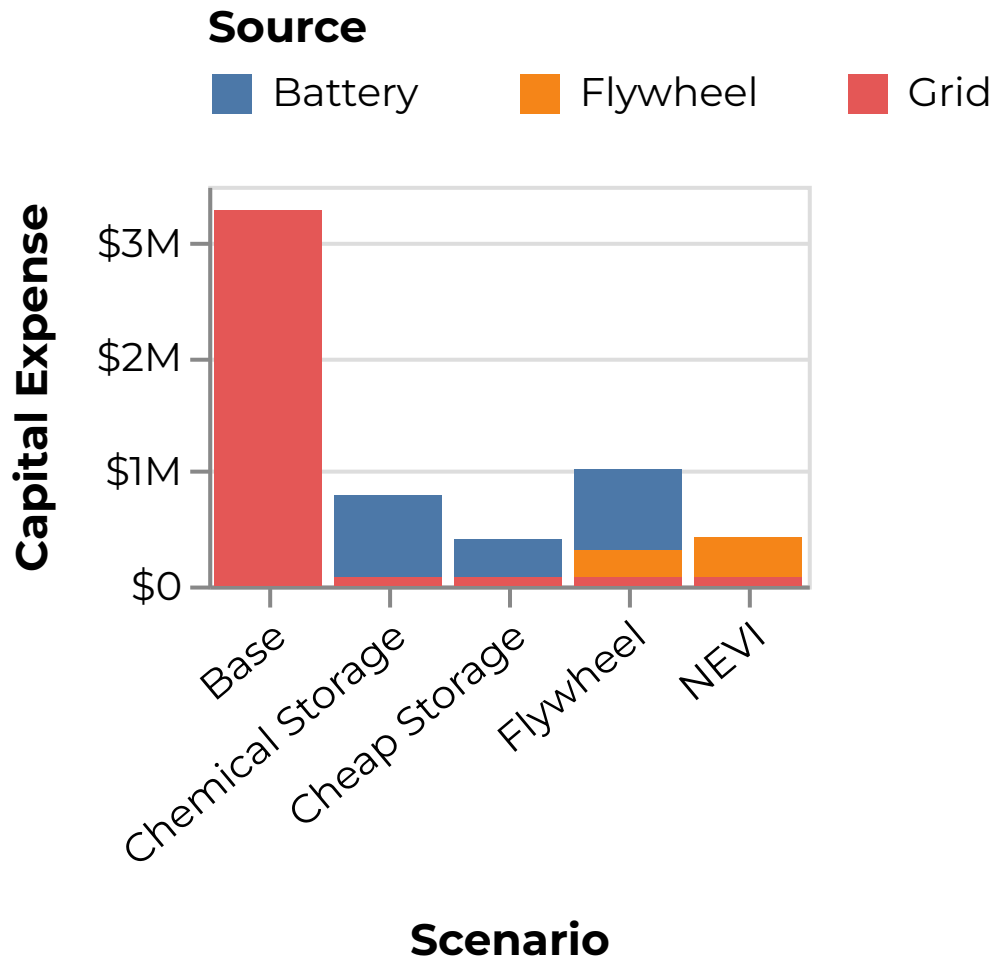


Figure 9: Capital expense by scenario, split by major cost source.

The installed equipment capacities are held fixed for each scenario, but site dispatch is reoptimized under each tariff, so the comparison reflects tariff-specific operation of the same physical design rather than identical behavior evaluated under two different rate schedules.

Because the project used an extremely high discount rate in the formal optimization, discounted operating-cost totals are not especially informative for this purpose. The more useful comparison is the non-discounted lifetime utility bill. Figure 19 presents that comparison for the main scenario set under Schedule 6 and Schedule 6A.

The comparison supports the expectation that Schedule 6A is generally the more favorable tariff for a low-utilization site. Across most of the scenario set, Schedule 6A produces the lower lifetime utility bill, which helps validate its use

as the main tariff in the broader analysis. At the same time, the figure also shows that the storage economics are shaped not only by hardware cost, but also by the rate structure used to value grid service. Schedule 6A makes low-utilization infrastructure easier to tolerate by softening the operating-cost penalty associated with carrying a larger grid connection that is used intensely only during relatively infrequent peak periods. In that sense, it reduces the operating-expense penalty for the base case and makes the underlying site problem more financially palatable even when the design relies on comparatively inefficient use of grid infrastructure.

That same feature also weakens the OPEX incentive to install batteries. One way to see this is to compare how much bill separation each tariff creates between very different hardware configurations. Under Schedule 6A, the base case and the

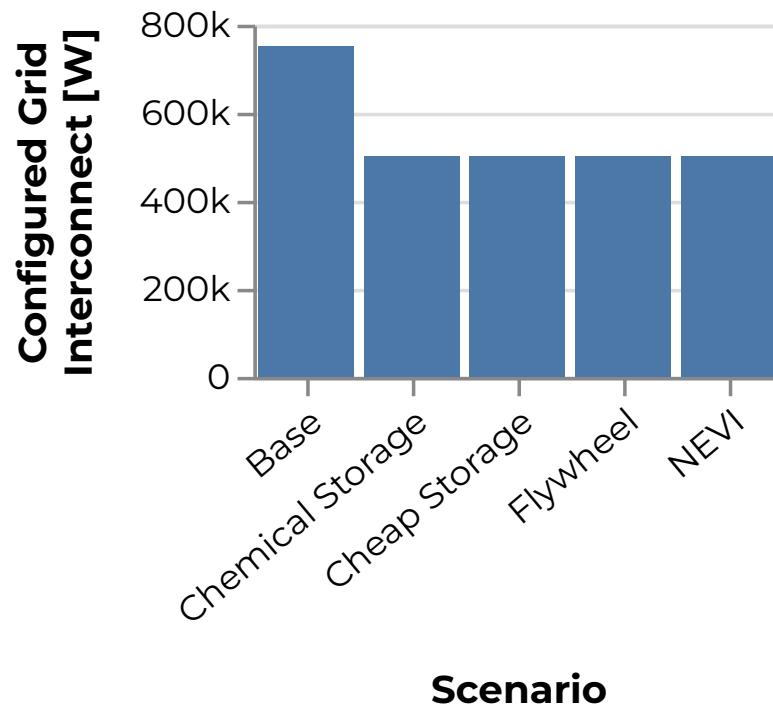


Figure 10: Grid interconnect by scenario.

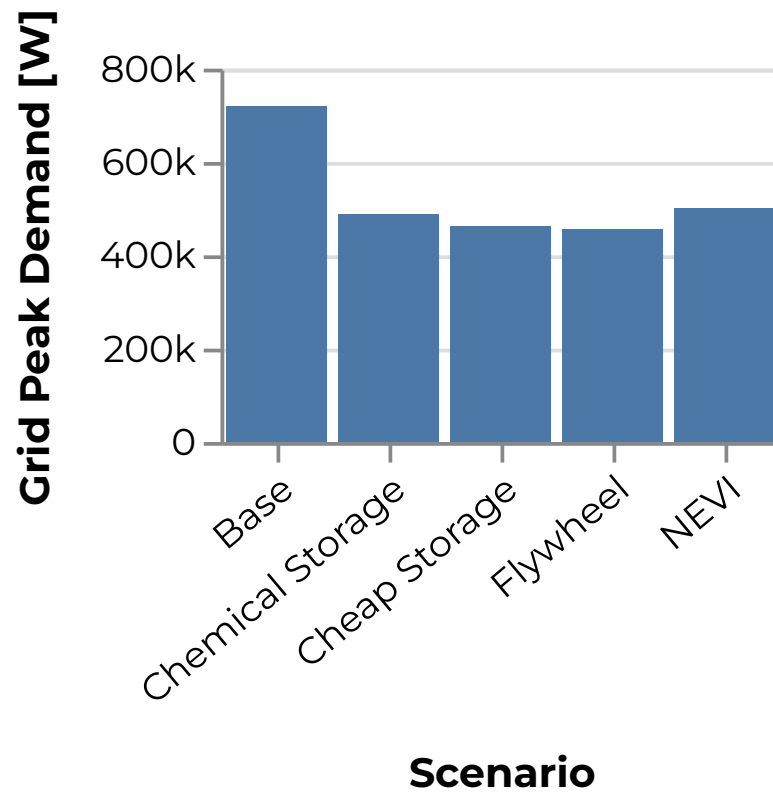
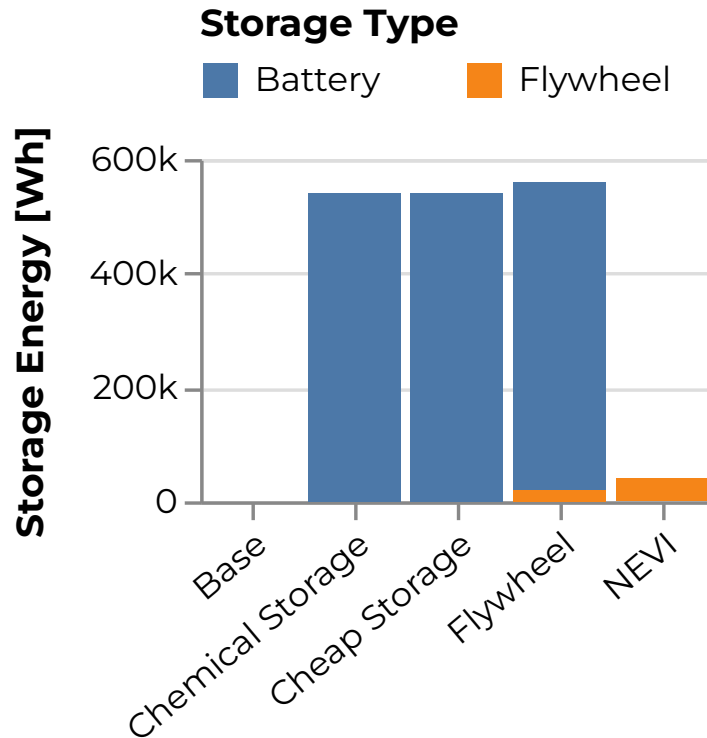


Figure 11: Grid peak demand by scenario.



Scenario

Figure 12: Installed storage energy by scenario and storage technology.

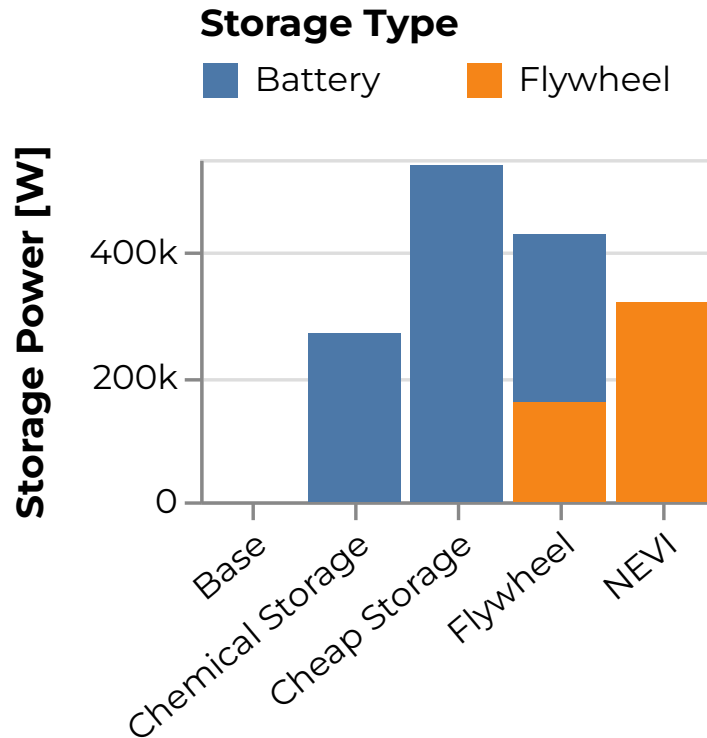
cheap storage case have relatively similar lifetime utility bills even though the cheap storage configuration uses much higher-performance storage and materially reduces grid peak demand and demand-period energy use. Under Schedule 6, the separation between those same configurations is much larger. In that sense, Schedule 6A is better from the site-cost perspective, but Schedule 6 provides a stronger reward for hardware that lowers grid impact. Put differently, Schedule 6A rewards low utilization of infrastructure in a way that can discourage storage deployment, whereas Schedule 6 sends a stronger signal in favor of batteries when those batteries are capable of materially reducing costly grid demand.

More broadly, this points toward the possible value of a new lower-power tariff designed more explicitly around peak-demand management, potentially with behavior more like Schedule 8 in the sense of lower demand charges outside true peak-demand periods. Schedule 8 is not applicable to this site as written because it applies only when a site has a peak load greater than 1

MW at least twice in an 18-month period, but it illustrates the structure of a possible medium-power, intermediate tariff concept. A structure of that kind could potentially lower customer cost, strengthen the incentive for energy storage, and reduce grid stress by rewarding designs that limit coincident peak demand and shift it away from grid peak demand periods rather than simply making lightly utilized infrastructure cheaper to operate.

3.4 Sensitivity Analysis

The sensitivity-analysis results examine how a broader set of fixed storage configurations perform when the arrival pattern is changed. Rather than reusing only the few cases highlighted in Section 3.3, this stage evaluates every fixed storage configuration except the no-storage case, spanning zero to two battery modules and zero to three flywheel modules. As noted in Section 2.3.4, each battery module provides 270 kWh and 135 kW, corresponding to a 0.5 C device, and each flywheel module provides 20 kWh and 160 kW,



Scenario

Figure 13: Installed storage power by scenario and storage technology.

Table 3: Summary of financial modeling scenario metrics.

Metric	Base	Chemical Storage	Cheap Storage	Flywheel	NEVI
Capital expense	\$3.3M	\$790k	\$400k	\$1.0M	\$420k
Grid capital expense	\$3.3M	\$76k	\$76k	\$76k	\$76k
Battery capital expense		\$720k	\$320k	\$720k	
Flywheel capital expense				\$230k	\$350k
Grid interconnect	750kW	500kW	500kW	500kW	500kW
Peak grid demand	720kW	490kW	460kW	460kW	500kW
Battery energy		540kWh	540kWh	540kWh	
Battery power		270kW	540kW	270kW	
Flywheel energy				20kWh	40kWh
Flywheel power				160kW	320kW
Compliance horizon	2035	2035	2035	2035	2031

corresponding to an 8 C device. For each fixed configuration, the monthly operational problem is rerun across 25 alternative arrival scenarios. In

the non-curtailed cases, the key output is the first month in which the configuration can no longer meet all charging demand. In the curtailed

Table 4: Base case financial metrics.

Metric	Value	Compared with Base
Capital expense	\$3.3M	Baseline
Total grid capital expense	\$3.3M	Baseline
Grid interconnect	750kW	Baseline
Peak grid demand	720kW	Baseline
Battery energy		Baseline
Flywheel energy		Baseline

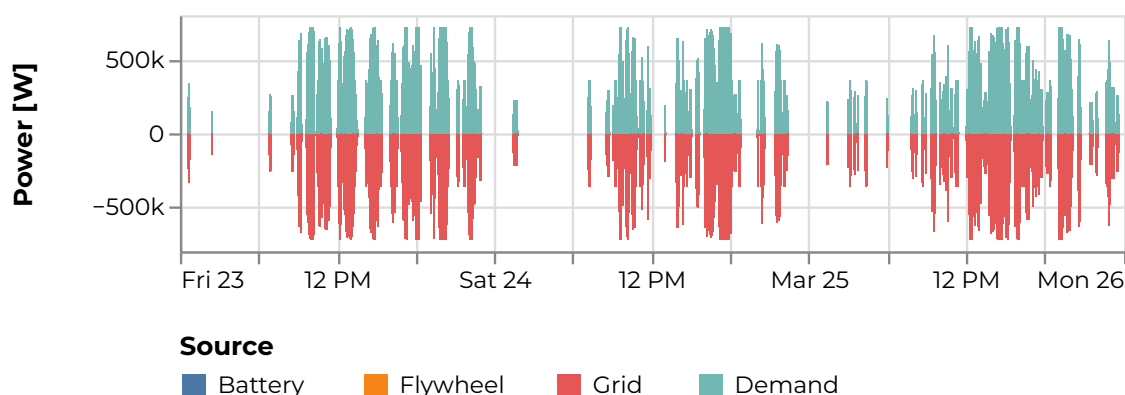


Figure 14: March 2035 peak-period energy flows for the base case. Because the site has no on-site storage, the figure contains only the power-flow panel.

Table 5: Chemical storage case financial metrics.

Metric	Value	Compared with Base
Capital expense	\$790k	76 % below base
Grid capital expense	\$76k	98 % below base
Battery capital expense	\$720k	\$720k above base
Grid interconnect	500kW	250kW lower
Peak grid demand	490kW	230kW lower
Battery energy	540kWh	
Battery power	270kW	
Flywheel energy		

cases, the same month-by-month evaluation is performed, but limited under-delivery is allowed and compliance is instead judged against the annual curtailed-hour threshold. The following subsections therefore address two related but distinct questions: first, how robust the candidate designs are if all demand must always be served, and second, how that robustness changes if a

limited amount of curtailment is treated as acceptable within the compliance framework.

3.4.1 Without Curtailment

The non-curtailed sensitivity results show that performance depends strongly on arrival cadence and that outcomes can vary substantially

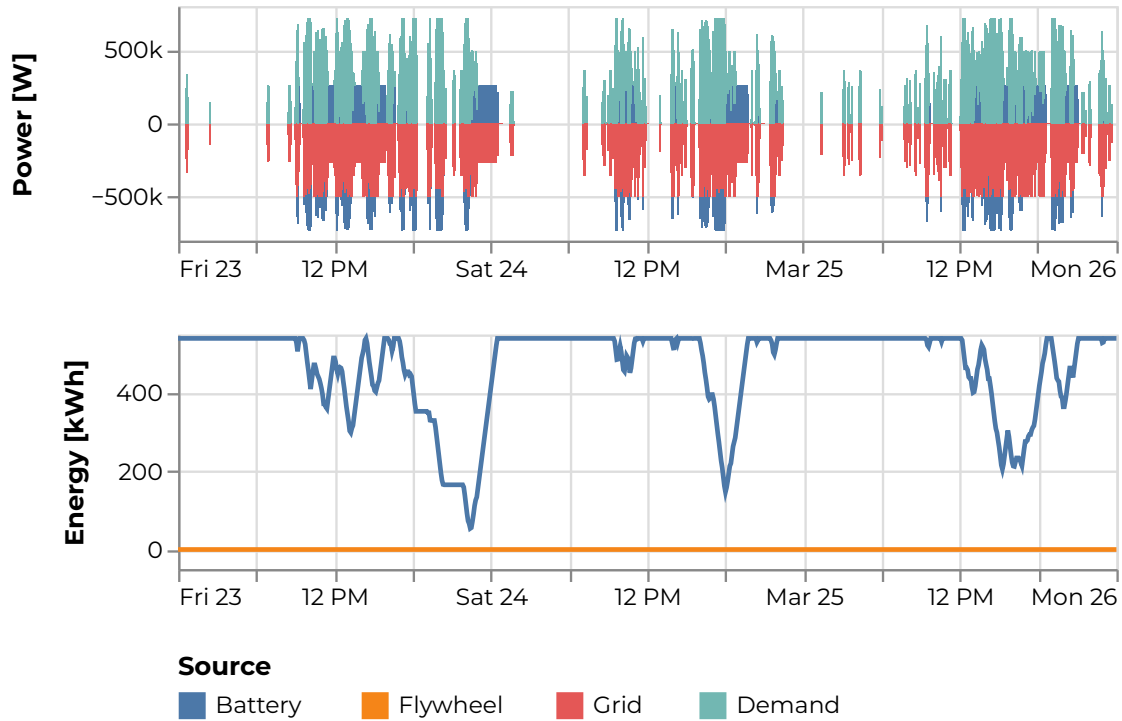


Figure 15: March 2035 peak-period energy flows for the chemical storage case.

Table 6: Cheap battery case financial metrics.

Metric	Value	Compared with Base
Capital expense	\$400k	88 % below base
Grid capital expense	\$76k	98 % below base
Battery capital expense	\$320k	\$320k above base
Grid interconnect	500kW	250kW lower
Peak grid demand	460kW	260kW lower
Battery energy	540kWh	
Battery power	540kW	
Flywheel energy		

across scenarios. Figure 20 shows the full distribution of first noncompliance year across the 25 evaluated arrival scenarios for each fixed configuration. Figure 21 presents the same information in compressed form by showing how long each configuration is expected to remain compliant at 95% confidence, meaning the month through which at least 95% of the evaluated arrival scenarios continue to satisfy the load.

The modeled flywheel-only configurations are the most sensitive to changes in arrival timing.

A single flywheel becomes noncompliant essentially immediately. Adding flywheels improves the median outcome, but not enough to make the design robust: the two-flywheel case does not survive through 2031 in any of the evaluated arrival scenarios, and the three-flywheel case survives through 2031 in only 4% of them, with a 95% confidence failure month of May 2027. This behavior should not be read as a general indictment of flywheels. Rather, it reflects this particular configuration and duty cycle: flywheels are effec-

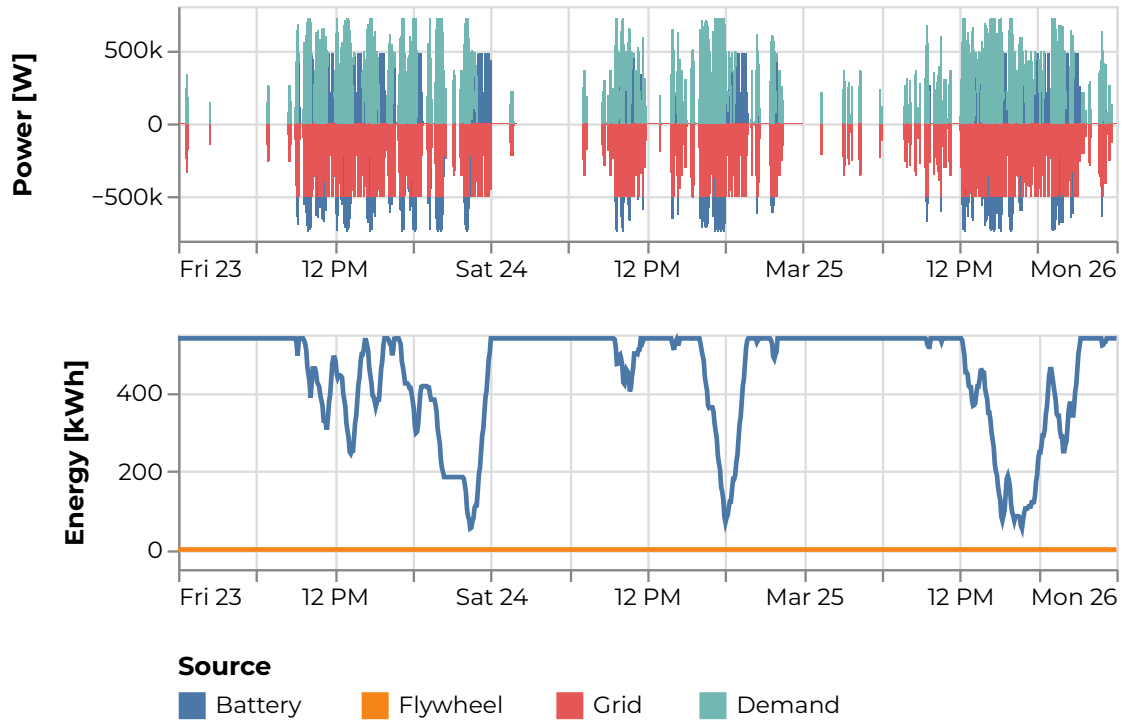


Figure 16: March 2035 peak-period energy flows for the cheap battery case.

Table 7: Flywheel case financial metrics.

Metric	Value	Compared with Base
Capital expense	\$1.0M	69 % below base
Grid capital expense	\$76k	98 % below base
Battery capital expense	\$720k	\$720k above base
Flywheel capital expense	\$230k	\$230k above base
Grid interconnect	500kW	250kW lower
Peak grid demand	460kW	260kW lower
Battery energy	540kWh	
Battery power	270kW	
Flywheel energy	20kWh	
Flywheel power	160kW	

tive at absorbing very short spikes, but difficult periods are often driven by clusters of charging demand that persist long enough for storage energy, not just storage power, to become the dominant limitation.

Electrochemical storage performs much better under the same sensitivity test. The mixed one-

battery-plus-one-flywheel case remains compliant through 2031 in 92% of the evaluated arrival scenarios, with a 95% confidence failure month of September 2031. The two-battery case remains compliant through 2031 in all 25 scenarios and through 2035 in 88% of them, with a 95% confidence failure month of October 2034. These results are important because they show that

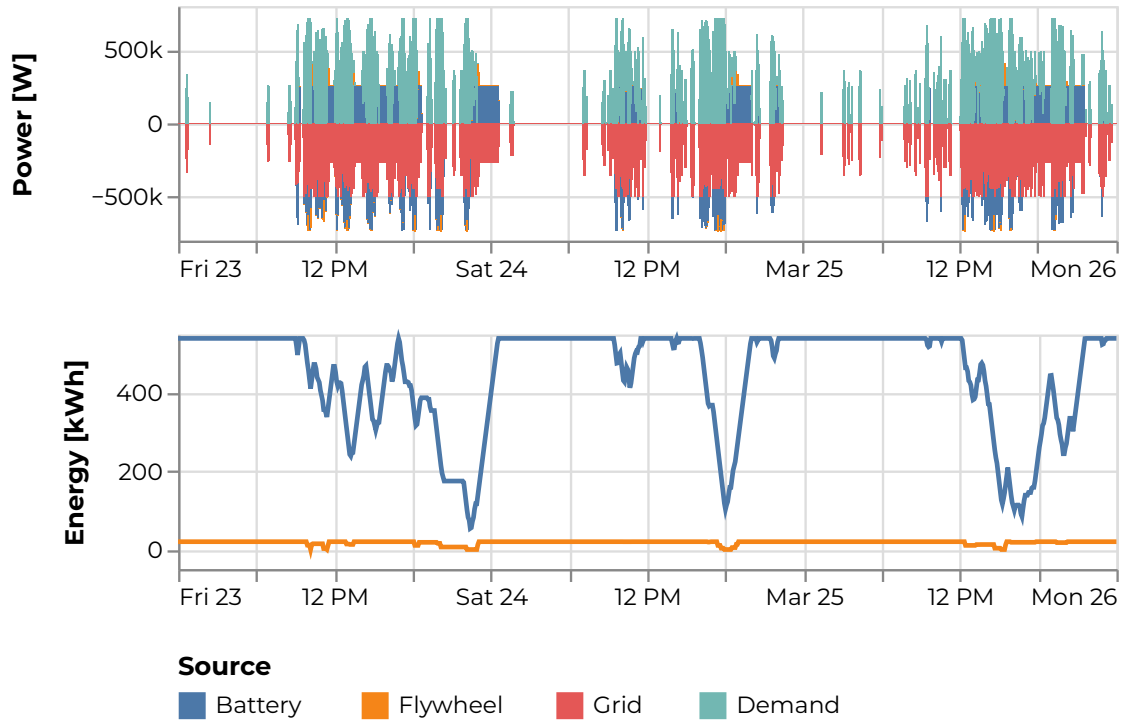


Figure 17: March 2035 peak-period energy flows for the flywheel case.

Table 8: NEVI case financial metrics.

Metric	Value	Compared with Base
Capital expense	\$420k	87 % below base
Grid capital expense	\$76k	98 % below base
Flywheel capital expense	\$350k	\$350k above base
Grid interconnect	500kW	250kW lower
Peak grid demand	500kW	220kW lower
Battery energy		
Flywheel energy	40kWh	
Flywheel power	320kW	
Compliance horizon	2031	

the main value of batteries is not only that they reduce grid peak demand in one representative scenario, but that they also provide a more dependable buffer against uncertainty in how peak demand is temporally concentrated.

Taken together, the non-curtailed results narrow the plausible design space. The modeled flywheel-only designs are too sensitive to arrival variation to serve as a dependable long-run com-

pliance strategy without curtailment. The mixed and battery-dominant cases are materially more robust, and the two-battery case provides the strongest margin. That conclusion is especially notable because the two-battery configuration is in the same general cost range as the one-battery-plus-one-flywheel case, so the improvement in robustness does not add significant project cost.

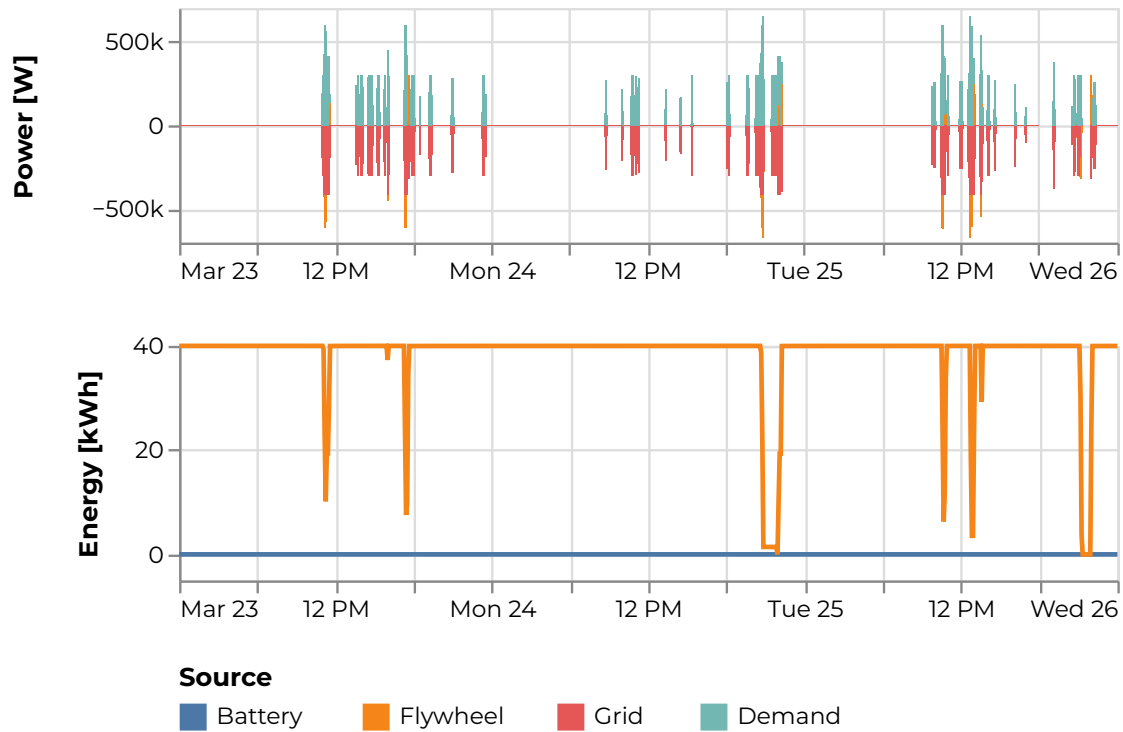


Figure 18: March 2031 peak-period energy flows for the strict NEVI-compliance benchmark.

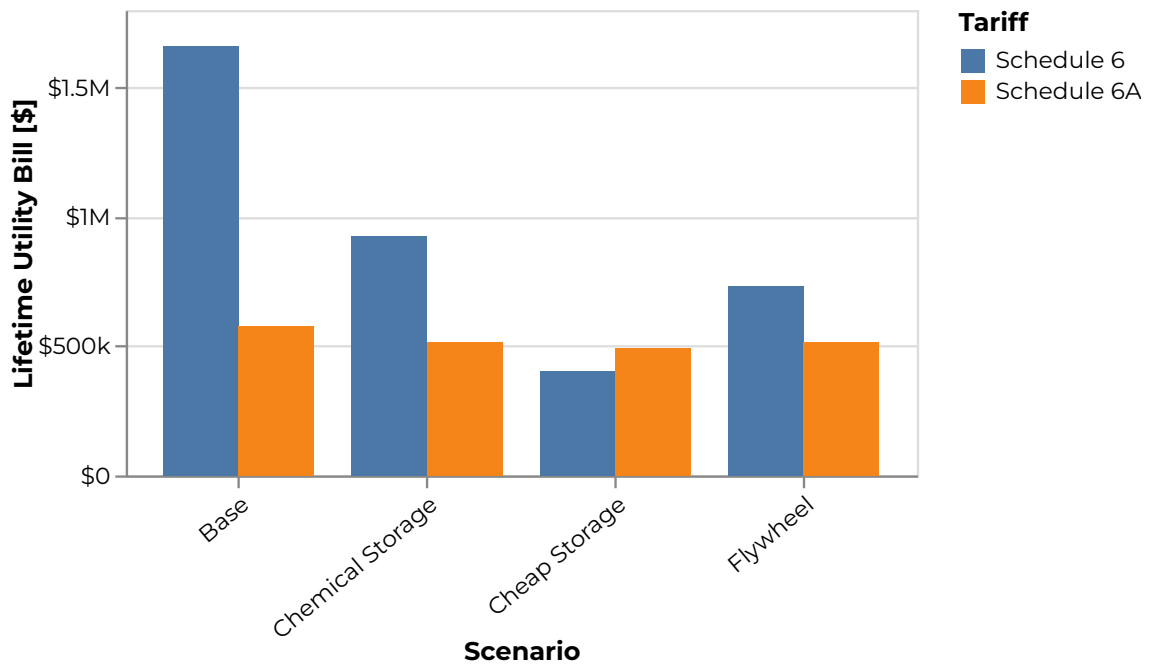


Figure 19: Comparison of non-discounted lifetime utility bills under Schedule 6 and Schedule 6A for the main scenario set.

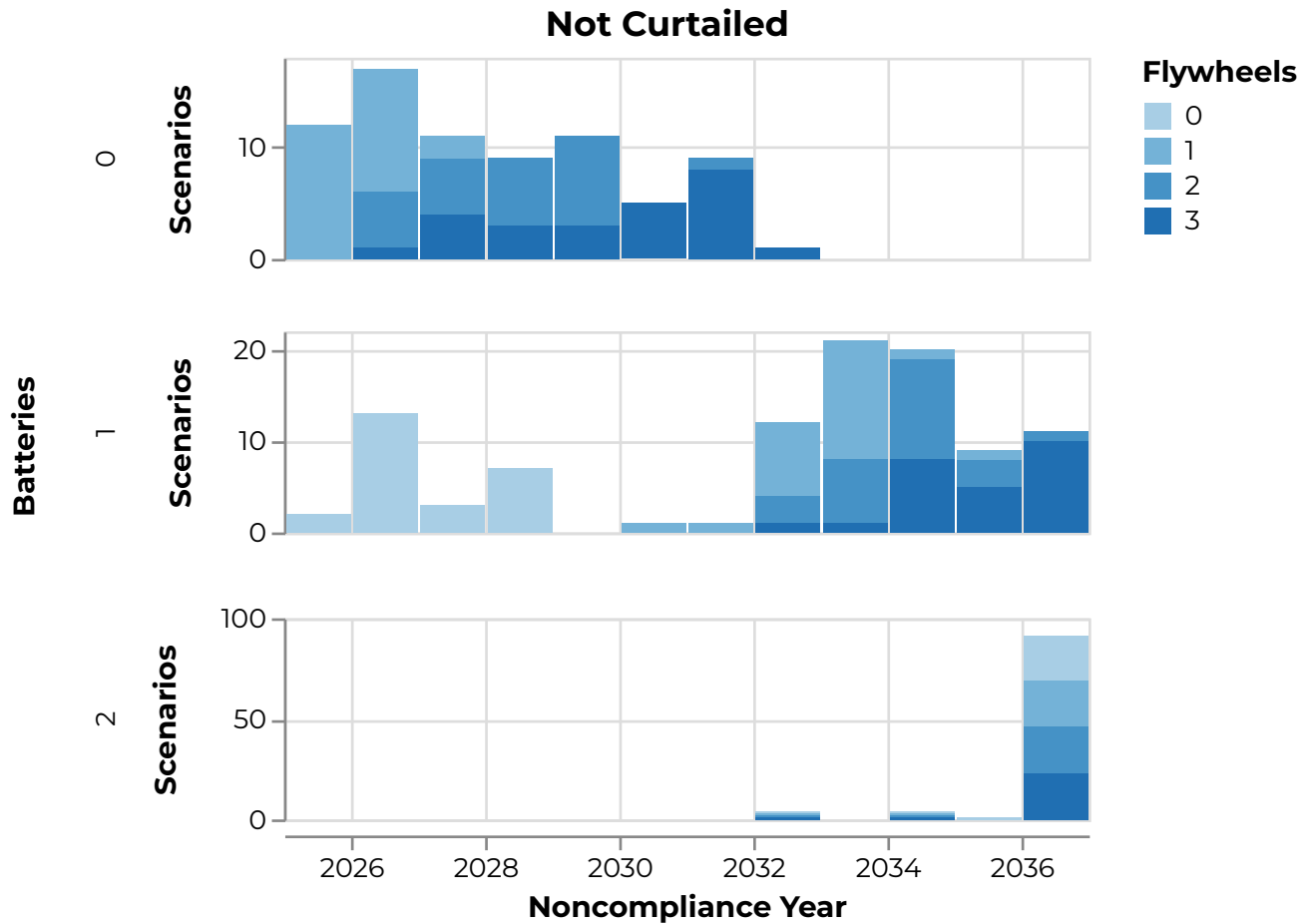


Figure 20: Distribution of first noncompliance year across the 25 evaluated arrival scenarios for fixed configurations without curtailment.

3.4.2 With Curtailment

While the sensitivity analysis without curtailment narrows the design space significantly, allowing limited curtailment changes the sensitivity picture substantially. Under the compliance interpretation used here, each plug may experience up to 3% downtime per year, which corresponds to roughly 260 h/yr. That allowance should not be interpreted as an operating target, because the same downtime budget must also absorb unscheduled outages and maintenance. It is instead a formal upper bound on how much controlled under-delivery could be used while remaining within the annual availability requirement. In this analysis, a curtailed state does not mean that the station is off. It means that storage has been depleted and the site is operating from the constrained grid connection alone; for example, a 500kW site serving four occupied plugs

would average about 125kW per plug, extending a 150kW-equivalent session by roughly 20%.

The effect of that added flexibility is shown first in the failure-timing figures. Figure 22 shows the first noncompliance year across the 25 evaluated arrival scenarios when curtailment is allowed, and Figure 23 shows the corresponding 95% confidence compliance duration. Under this criterion, all of the evaluated storage configurations remain compliant through 2035 at 95% confidence.

The curtailed-hour figures show how that improvement is achieved. Figure 24 shows the maximum annual curtailed hours, and Figure 25 shows the full annual curtailed-hour distribution. These figures are used differently from the failure-timing plots above. Rather than showing when a configuration becomes noncompliant, they show how much controlled under-delivery is

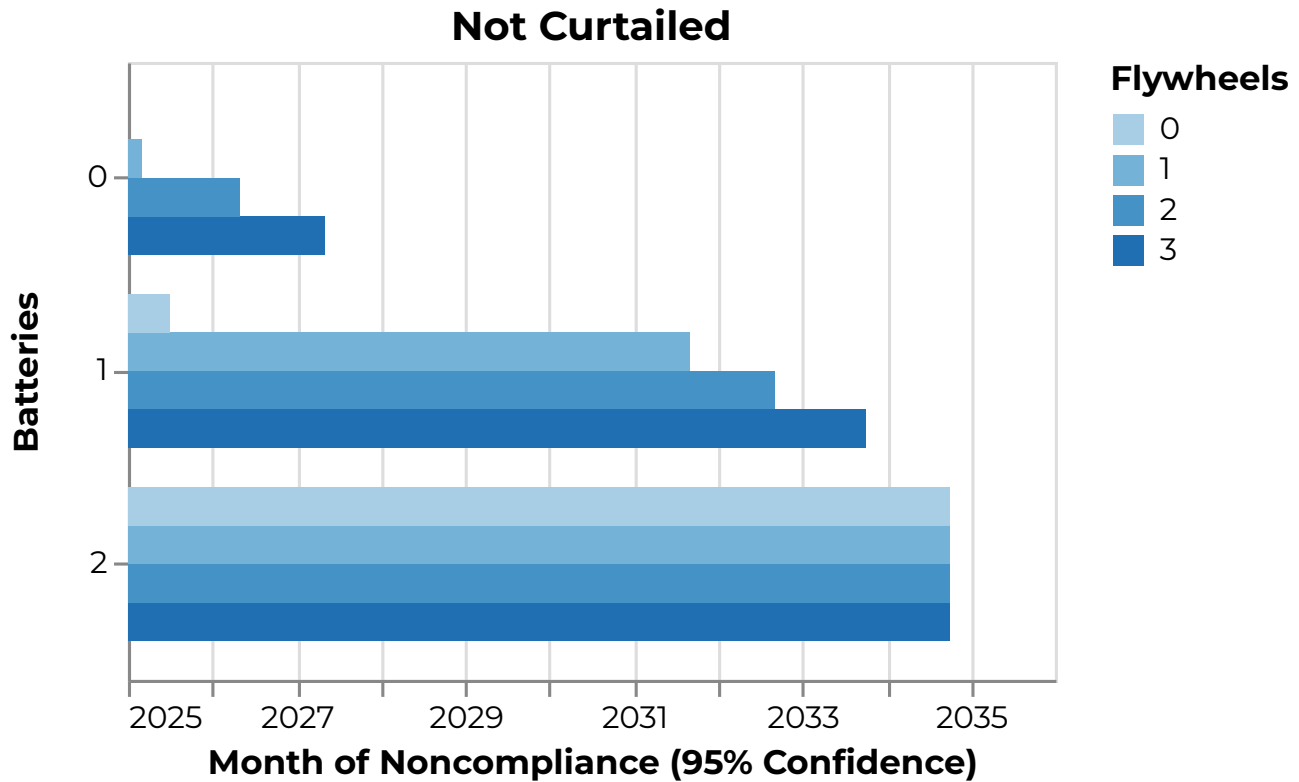


Figure 21: Month through which each fixed configuration is expected to remain compliant with 95% confidence across the 25 evaluated arrival scenarios without curtailment.

required in order to remain feasible. The no-storage reference in these two figures is calculated from the full set of modeled arrival scenarios rather than only the 25 evaluated fixed-configuration scenarios, because it is obtained directly from the complete no-storage arrival set. It is included as a reference for the value of storage under constrained-grid operation: storage can outperform a grid-only constrained design by reducing the number and duration of periods when the site falls below the target service level.

The benefit of curtailment is most obvious in the flywheel-dominant cases, which are the least robust in the non-curtailed analysis. With no curtailment allowed, the modeled flywheel-only configurations fail very early. Once modest curtailment is permitted, those same configurations remain compliant through 2035 because the amount of curtailed time required to cover the most difficult hours stays below the annual threshold. Flywheels still have advantages that may matter in this application, including rapid charge and discharge, limited capacity degradation relative to electrochemical batteries, and

lower thermal-runaway concern. The battery-supported cases also improve, but they were already comparatively robust; for batteries, curtailment operates more as an added margin than as the central mechanism making compliance possible. The mixed battery-plus-flywheel configuration keeps curtailed hours relatively low, and the two-battery case effectively eliminates modeled curtailment in the evaluated scenarios.

Taken together, these results show that curtailment is best understood not as a substitute for storage, but as a limited operating margin that can convert marginal high-power, low-energy, and low-cost designs into compliant ones and provide additional protection for the battery-supported cases. Without curtailment, even strong storage cases can still carry a meaningful chance of becoming noncompliant by the end of the study horizon. With modest curtailment, that risk drops sharply. The combination of storage and curtailment therefore appears to be much more effective than either tool alone: storage reduces the size and frequency of difficult peak events, and curtailment provides a controlled backstop

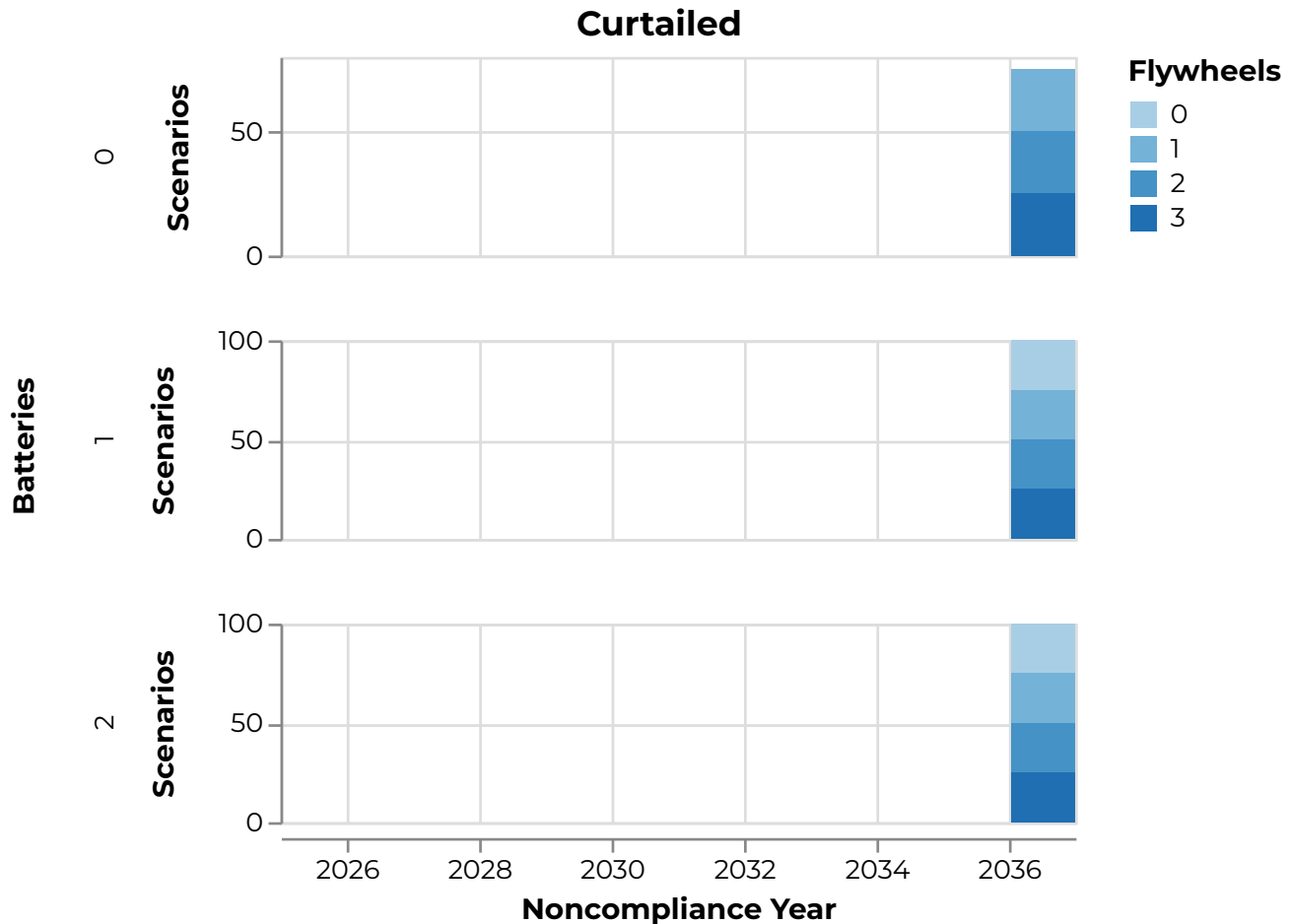


Figure 22: First noncompliance year across the 25 evaluated arrival scenarios for fixed configurations with curtailment allowed. Bars at 2036 indicate configurations that remain compliant through the end of 2035.

for the comparatively rare hours that remain problematic. A proactive version of this strategy could also limit all occupied plugs to 150kW before storage is exhausted, preserving flywheel duration instead of waiting until the site is already depleted.

3.5 Additional Compliance Strategies

The curtailed-hour accounting above is conservative in one important sense because NEVI downtime is evaluated at the plug level rather than only at the station aggregate. That creates room for intentional control strategies that distribute limited under-performance across the plugs instead of concentrating it on one connector. If the site rotates which plug is partially served during a short fully loaded interval, the same ag-

gregate power shortfall can translate into much less downtime for any one plug.

The basic idea can be illustrated with a four-hour fully loaded period in the table below. If all four plugs are uniformly curtailed at 125 kW instead of the 150 kW NEVI benchmark, that event counts as four hours of downtime per plug. If instead three plugs are held at 150 kW and the power-limited condition is concentrated at one plug, and that role is rotated evenly across the four plugs, then the effective downtime attributed to any one plug is reduced by a factor of four. Adding storage to this strategic operation can further improve performance. If a vehicle is fully curtailed for 2/3 of its curtailment rotation the storage can charge at 50 kW. The battery may then discharge at 100 kW for 1/3 the time giving the vehicle 150 kW including the 50 kW left from the grid connection. Because the level of curtailment is not penalized

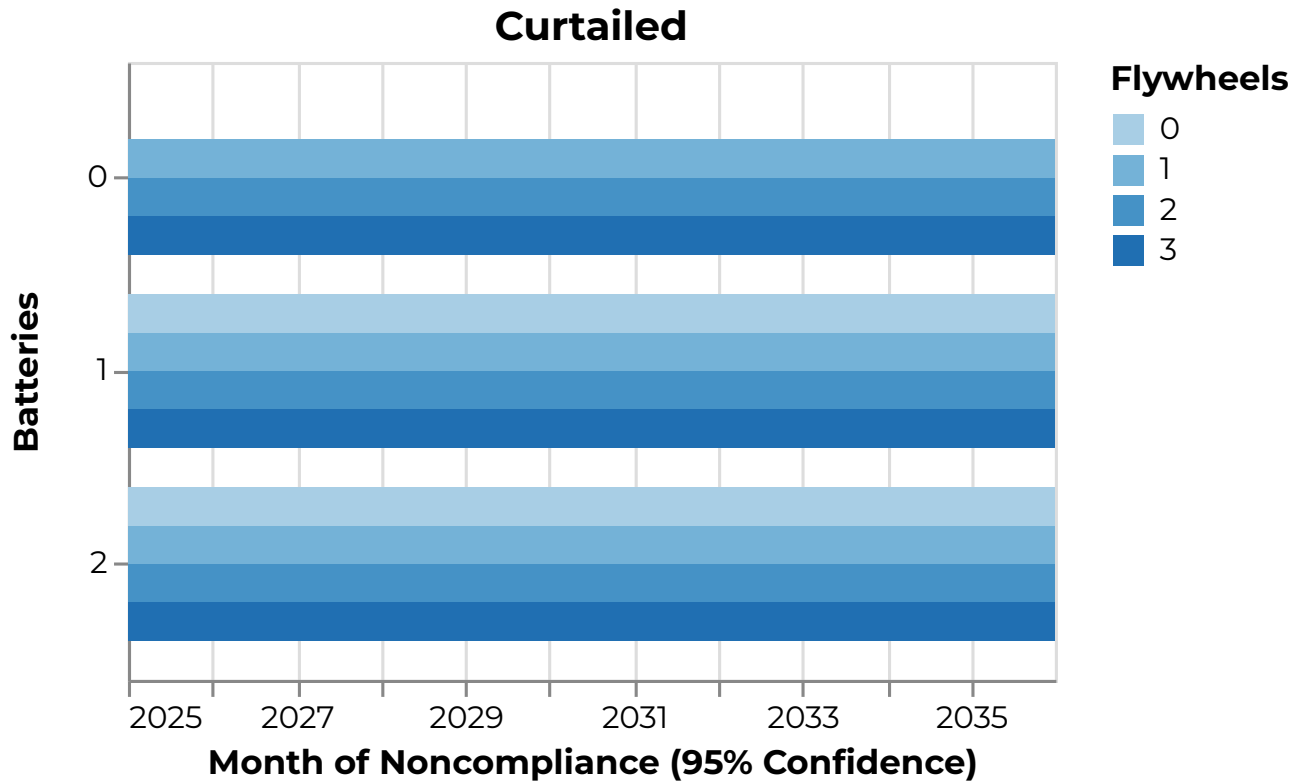


Figure 23: Month through which each fixed configuration is expected to remain compliant with 95% confidence across the 25 evaluated arrival scenarios when curtailment is allowed.

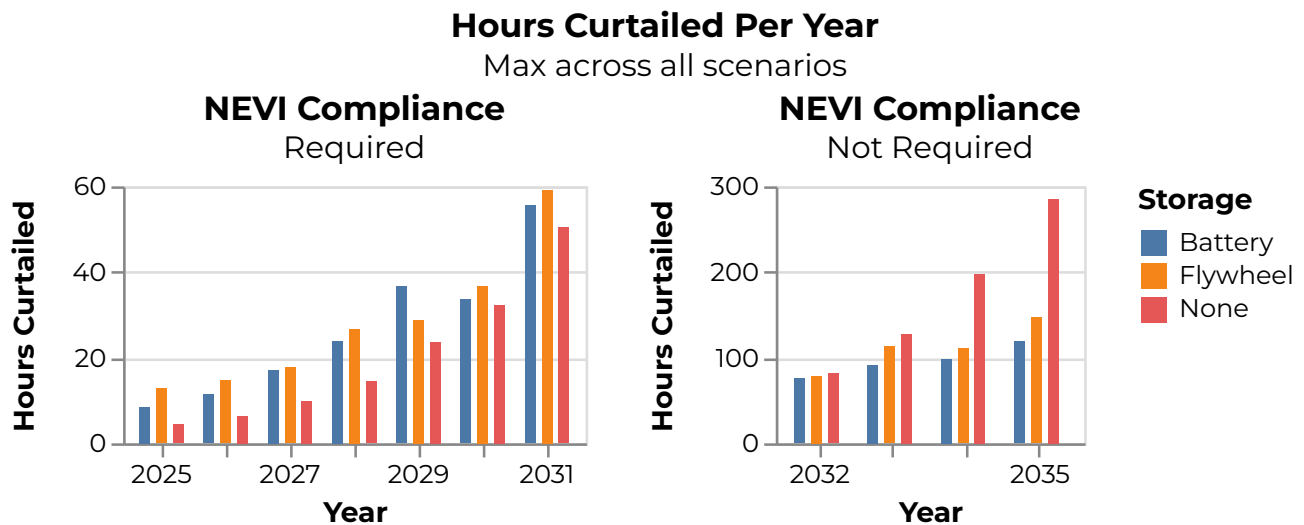


Figure 24: Maximum annual curtailed hours across the evaluated arrival scenarios, grouped by storage type and split between years when NEVI compliance is still required and years after 2031. The no-storage reference is based on the full set of modeled arrival scenarios.

by NEVI, only the binary condition of curtailed or not, this solution can reduce the plug-level downtime by 6x compared with the uniformly curtailed base case.

This kind of intentional control was not used as the basis for the main compliance conclusions in this report, and to our knowledge it does not yet exist as a market-available charger/

NEVI Compliance

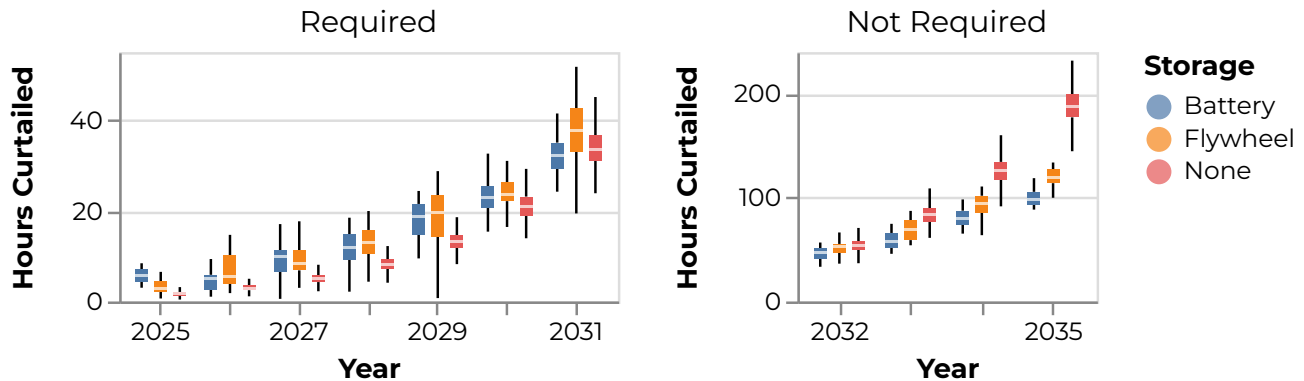


Figure 25: Distribution of annual curtailed hours across the evaluated arrival scenarios, grouped by storage type and split between years when NEVI compliance is required and years after 2031. The no-storage reference is based on the full set of modeled arrival scenarios.

Operating strategy	Example dispatch during a 4 h fully loaded period	Effective downtime per plug
Uniform curtailment	4 plugs at 125 kW	4 h
Rotating reduced-power	3 plugs at 150 kW, 1 plug at 50 kW	1 h
Rotating reduced-power with storage	3 plugs at 150 kW, 1 plug at 0 kW (67%) or 150 kW (33%)	40 min

Table 1: Illustration of how plug-level rotation and storage support reduce counted downtime per plug in the four-hour example.

storage control product. The principal sensitivity results already show that limited curtailment can preserve compliance through 2035 for the evaluated storage configurations. The importance of plug-level sharing is that it offers an additional operational margin beyond those results. If implemented carefully, it could make marginal configurations more defensible, reduce effective per-plug downtime during rare peak events, and lower the amount of dedicated storage required to achieve the same compliance outcome. The operator would still need to consider driver fairness: after the NEVI-required period, uniform allocation may be preferable unless the site can identify vehicles requesting lower power, because a guaranteed 125kW may be more useful to drivers than a rotating 0kW or 50kW condition.

4 Conclusion and Recommendations

The central conclusion of this work is that the challenge is not primarily queueing or average energy consumption, but the cost of serving short periods of highly coincident peak demand while remaining compliant with NEVI-style service expectations. Within that context, on-site storage changes the project economics dramatically. Relative to the conventional grid-upgrade solution, storage-supported designs reduce the required upfront investment by at least about fourfold and potentially more than tenfold while still providing feasible paths to compliance.

Table 2 summarizes the most decision-relevant configurations. The key contrast is between the base case and the storage-supported alternatives. The base case requires roughly \$3.7 million of power-infrastructure CAPEX, while the one-fly-wheel, one-battery, and two-battery options fall

Metric	Recommended Configurations			Base
	1x Flywheel	1x Battery	2x Battery	
Power-infrastructure CAPEX	\$300k	\$470k	\$790k	\$3.7M
Curtailed compliance	2035+	2035+	2035+	2035+
2031 curtailed hours ³	38 / 7	31 / 6	0	0
Non-curtailed compliance	2025	2025	2034	2035+

Table 2: Summary of the main decision cases.

at about \$300k, \$470k, and \$790k, respectively. These estimates do not assume incentives or negotiated discounts. On that basis alone, the potential capital-cost reduction ranges from about 4x to more than 10x.

The sensitivity results clarify how those options should be interpreted. Without curtailment, the low-CAPEX flywheel-only and one-battery cases are not robust long-run compliance strategies. The two-battery configuration is much stronger and remains the best non-curtailed option, but it still carries some late-horizon risk under unfavorable arrival timing. Once limited curtailment is allowed, however, all of the storage-supported configurations remain compliant through 2035 in the evaluated scenarios, and the flywheel and one-battery cases become much more credible low-cost pathways.

The recommended framing is therefore:

1. If the site must operate with no curtailment, the two-battery configuration is the strongest storage-supported option because it provides the most robust margin while still avoiding the very large grid-upgrade cost.
2. If limited curtailment is acceptable within the compliance interpretation used here, the one-flywheel and one-battery configurations become attractive because they preserve compliance at far lower CAPEX.
3. If the site starts with one flywheel or one battery, the site plan should reserve physical space, conduit, switchgear capacity, and control tie-ins for future modular expansion. The modular framing used here is compatible with adding up to three flywheels per flywheel control/connection block and up to four battery pods per battery control/connection block, subject to final vendor and site-engineering review.

4. The base grid-only upgrade solution should be understood primarily as a high-cost benchmark rather than as the preferred design, since it achieves performance at a capital cost that is difficult to justify relative to the storage-supported alternatives.

Taken together, these results indicate that this site is a strong case for combining grid service with supplemental storage support and operational flexibility rather than relying on a conventional grid expansion alone. The most important takeaway is not simply that storage can help, but that at this site it can reduce required upfront investment by roughly 4x to more than 10x while still supporting compliant operation when paired with an appropriate operating strategy. This grid plus supplemental storage support approach could significantly improve economics for constrained rural fast-charging sites that have strategic network value and low average utilization, but still experience periods of very high peak demand.