Supporting Information for

Technoeconomic Analysis of Microgrids for AI Data Centers in the Continental United States

Alex C. Newkirk, Daniel L. Gerber, Erica R. H. Fuchs, Alex Hubbard, M. Granger Morgan, Arman Shehabi, Sarah Smith, and Constantine Samaras

Alex C. Newkirk
E-mail: acnewkirk@lbl.gov

This PDF file includes:

Supporting Information Text

1. Comprehensive SI Methods

We evaluate microgrid feasibility through technoeconomic simulation of complete power systems serving identical datacenter loads. We apply a common operational scenario: a specific number of GPUs requiring continuous power at a target location, then determine the minimum-levelized-cost configuration within each power architecture class that can meet this demand at required minimum annual uptime across the lifespan of the facility.

The analysis proceeds in three stages. First, we establish the hourly facility load that any power system must serve, combining IT demand with location-specific cooling requirements. Second, we simulate candidate power systems meeting this load profile, tracking hourly energy flows from generation through storage and conversion to final consumption. Third, we search each architecture's design space to identify configurations that minimize levelized cost while delivering required reliability.

This framework enforces scenario equivalence: all power systems face identical hourly demands across the same 25-year evaluation period at the same geographic location. We evaluate four candidate power systems: AC-coupled solar+storage, DC-coupled solar+storage, natural gas with backup diesel, and interconnection to the local electric grid as a baseline.

To establish this scenario, we accept location (lat/long coordinates), required annual facility uptime, and quantity of GPUs as specified inputs. Location-specific climate data are taken from typical meteorological year (TMY) sourced from the PVGIS public API. TMY datasets are statistically representative hourly weather sequences derived from multi-decade observations that capture seasonal and diurnal patterns while excluding extreme outliers.

We compare each system in simulation year 27, end-of-life for the PV modules (25 year lifespan), with a base-case discount rate of 7%. For scenarios with BESS, we assume one single end-of-life replacement of the battery modules occurring at the conclusion of operational year 13 (see the battery degradation subsection for more detail). All financial values are in 2022 dollars, and all discounting occurs mid year.

All software is implemented in Python, and all tools are available from the public github repository at https://github.com/acnewkirk/AI-Datacenter-Microgrid-Analysis/tree/main. Our work builds upon two existing open-source tools: NREL's BLAST-Lite optimization framework (1) and the cost calculator developed by OffGrid AI (2). Each provided critical starting points for the development of this work. To conduct the geographic analysis, we generated a hexgrid tiling of the continental United States, encompassing 4,378 individual hexagons. For each hexagon, we performed a full simulation at the center point of the hexagon for each microgrid architecture, assigning the resulting values to all points within that tile.

The following section begins with how we construct the IT facility load profile. We then detail how we design and simulate each of our candidate power system. We conclude with the techno-economic assessment and system optimization approaches.

A. Facility Modeling. Our facility model builds from the IT out, following established bottom-up datacenter modeling approaches (3, 4). We base IT demand on 8-GPU Nvidia H100 DGX nodes, drawing from empirical characterization of these systems during training workloads (5). Each node exhibits average power draw of 7.3 kW and maximum draw of 8.5 kW, inclusive of memory and interconnect components.*

We model AI training load as constant baseload demand with minimal variation (±8% random noise around unity), representing 8,760 hours of continuous operation. This scenario can be conceptualized as one continuous training run involving all facility nodes, or alternatively as a facility with mixed workloads performing jobs under optimal scheduling and saturating demand. This training load shape is an idealization of practical facility operations, but serves as a conservative limiting case. The framework accepts alternative load profiles and hardware specifications through modular inputs: node-level power parameters (average and maximum draw), hourly demand patterns, and total GPU count. While our current implementation assumes uniform load distribution across all nodes, the modular architecture readily supports extensions for heterogeneous fleet modeling.

Facilities with a more strongly diurnal usage pattern or greater time flexibility in execution have shown potential in recent work to employ demand response for cost savings and effective capacity expansion (6, 7). Our modeled scenario provides a conservative bound on microgrid viability. Systems that perform well under inflexible baseload will show improved economic performance with operational flexibility under relaxed constraints.

We allocate the chip-quantity input into 8-GPU servers, that each have the power ranges described above. We construct an expected facility IT load by taking the product of our baseload demand time series, the average node power, and the quantity of nodes. This represents the expected power that must be delivered to the IT hardware in each hour. We also define the facility modeled maximum IT power demand as the sum of the node maxima. Datacenter operators typically contract with utilities for capacity that reflects the maximum theoretical facility demand. Similarly, power delivery infrastructure, both beyond and within the fenceline, is scoped to accommodate this upper-bound value, even though average demand is much lower. This is a common feature of power systems design; infrastructure is built for peak demand, but then amortized over the typically lower level of average utilization.

The peak-provisioning principle fundamentally shapes the economics of both grid-connected and islanded facilities. To meet peak loads, operators must invest in infrastructure capable of handling maximum theoretical demand, even if this capacity sits idle during normal operations. We therefore define a facility design load representing the maximum instantaneous power the

*IT load encompasses compute nodes, embedded networking, and management systems, but excludes cooling and power conversion losses addressed in subsequent sections.

infrastructure must deliver:

$$P_{\text{design}} = N_{\text{nodes}} \times P_{\text{node,max}} \times \text{PUE}_{99} \times \alpha_{\text{contingency}} \quad [1]$$

where N_{nodes} is the number of compute nodes, $P_{\text{node,max}}$ is the maximum node power (8.5 kW), PUE_{99} is the 99th percentile PUE value, and $\alpha_{\text{contingency}} = 1.05$ represents an engineering margin for component failures. This design load, distinct from average operational load, establishes power system sizing requirements that drive capital costs across all architectures.

A.1. Cooling. Datacenter cooling systems must reject heat equal to IT power consumption plus infrastructure losses, with efficiency varying substantially by technology and climate. Lei and Masanet found that the energy consumption of a given cooling technology can vary by up to 60% location to location(8). We model location-specific cooling efficiency using hourly PUE values, capturing how ambient conditions affect the availability of free cooling and mechanical system performance throughout the year.

Cooling system performance varies with climate: mechanical-only systems show low geographic variation while free-cooling systems exhibit PUE ranges from 1.05 to 1.4 depending on ambient conditions. The rising heat density of AI chips increasingly favors liquid cooling systems—either direct-to-chip or immersion (9–12)—with future server configurations requiring such systems (13). Following industry practice for high-density deployments, we exclude configurations with insufficient cooling capacity or prohibitively costly low performance.

We consider six candidate cooling configurations for our modeled facility: (1) IT liquid cooling with waterside economizer; (2) IT liquid cooling with adiabatic-assist air-cooled chiller; (3) IT liquid cooling with dry-cooler air-cooled chiller; (4) air-cooled IT with air-cooled chiller, airside economizer, and adiabatic cooling; (5) air-cooled IT with water-cooled chiller, airside economizer, and adiabatic cooling; and (6) air-cooled IT with air-cooled chiller and airside economizer. These systems align with the cooling distribution modeled for AI-specialized facilities by (4). Each system's cooling efficiency and capacity depends on local climatological conditions, specifically dry-bulb temperature and ambient humidity.

We construct an hourly time-series of location-specific climate and weather data relevant to cooling performance (dry-bulb temperature and ambient humidity) from the TMY for the input location. Using Lei and Masanet's empirically calibrated datacenter cooling system model (8), we pre-compute PUE values for each cooling technology across the range of ambient conditions, creating lookup tables indexed by dry-bulb temperature and humidity. We then build hourly PUE performance profiles of the candidate configurations for the specified location. We assume firms select the cooling configuration with the lowest average PUE value over the year[†], as operational energy use dominates total-cost-of-ownership in cooling infrastructure (14). We then combine the hourly IT-load with this hourly PUE to construct an hourly facility load.

The integration of cooling load illustrates the core design principle of *scenario equivalence*. For a specified quantity of chips and location, the amount of power that must be delivered to the IT and cooling load, and the minimum uptime that design must deliver that power, are held constant in each modeled scenario. Differences in power delivery efficiency affect how much power must be generated at the source to meet that load, and the associated economics, but *the energy required by the IT and cooling* are constant.

Power demanded by the IT load is the hourly value defined in the previous section, and infrastructure power demand is that value multiplied by the difference between the hourly PUE value and 1. Recall that power infrastructure, both in generation and delivery, is scoped not for the annual average, but for the theoretical maximum. Once we've selected a cooling system, we select a design PUE value equal to the 99th percentile PUE produced by the selected system in the modeled year, consistent with industry practice. We then take our maximum IT load, multiply it by this design PUE, and then multiply this product by a scalar design contingency factor $\alpha_{\text{contingency}}$, 1.05 in our base case. This factor corresponds to an engineering margin accounting for unanticipated component failure or malfunctions. This facility design load is the power delivery minimum of our candidate designs for a given facility; any generating system must be able to deliver at least this much power to IT and cooling system.

A.2. Power Systems Architecture. To compare different microgrid systems, we develop candidate power flow models of all paths between generation and facility loads. We address a gap in the existing literature by modeling the explicit power distribution system losses of different facility and microgrid architectures. A scenario assumption is colocated generating assets, though the power flow models could be extended to include transmission. We model the individual voltage and current conversion steps within the facility fenceline. Candidate designs are shown in maintext Figure 4. Note that we assume that islanded natural gas facilities adopt the same power systems architecture as conventional grid-connected facilities, and that the diesel backup generators connect to the facility through the same main transformer[‡]

We model all power systems out from the main distribution bus, with the cumulative efficiency of all voltage and conversion steps necessary to each load and from each source. We calculate path efficiency as the aggregate product of all component efficiencies between the bus and the source or load. The component efficiencies, displayed in maintext Figure 4, are based on component datasheets (15) and operational estimates from industry experts (16). We discuss optimum-design levelized-cost sensitivity to our choice of component efficiency in the technical parameter sensitivity section of the SI.

Our proposed power-flow paths are consistent with conventional data center design practice (17, 18) or existing all-DC pilots (19). We assume that the grid baseline and natural gas microgrid follow identical power delivery paths, and adopt the same power-delivery architecture for each, considered from within the facility fenceline. In this baseline power-path, the backup

[†] We acknowledge that in practice, firms select cooling according to a variety of decision variables, including familiarity, vendor relationships, and compatibility with existing staff expertise and capital equipment

[‡] Sensitivity analysis shows that the power-flow pathway of the backup gensets is not significant to facility economic performance do to low annual utilization, so we adopt this simplifying assumption.

diesel generators are modeled as entering the facility through the same distribution transformer as primary power, and both cooling and IT load are supplied through double-conversion UPS to ensure consistent power quality, an architecture consistent with critical operation baseload facilities (20).

We model PV+Storage systems as using battery storage to provide both facility-scale backup power and uninterruptible power services. These architectures replace diesel backup generators and standalone UPS components with grid-scale Lithium-Iron-Phosphate (LFP) battery modules. Our power system designs follow typical microgrid models that utilize centralized generation and storage. Distributed generation and storage is also viable alternative proven in other applications (21), but is out of scope for this study. Component efficiencies are based on the averages of efficiency curve data from (22, 23).

In the AC microgrid system, we model a ground-mount solar array with outputs combined through 500 kW central inverters that output three-phase 480 V AC. The storage is a set of centralized battery banks within close proximity, AC-coupled to the switchgear though grid-forming 500 kW battery inverters. Wires are sized to minimize wire loss within reasonable cost. The AC power supply unit (PSU) is a Platinum-rated 10.2 kW 240 V AC to 48 V DC power-factor correction (PFC) rectifier, integrated into each Nvidia DGX H100 node as four plus two redundant 3300 W rectifier modules. The integrated chip-level power electronics then switch the PSU voltage down to the internal chip voltages The cooling load (HVAC, chiller, and pumps) all high performance AC induction or brushless DC motors, powered by an AC variable speed drive (VSD) or a DC electronic speed controller (ESC). Here, we only model the losses of the AC/DC PFC rectifier stage within the VSD, as the DC/AC motor inverter stage is present in all topologies. Power quality and any training induced power transients in the AC systems are managed by the UPS or the BESS components.

The DC microgrid topology mirrors the AC microgrid. Here we model the solar panels connected with high-power DC/DC power optimizers to DC switchgear. The battery's DC/DC charge controller handles grid-forming (22) through droop-voltage regulation and coordinates with the microgrid energy management system (MEMS). It distributes power on a 350 V DC bus, compliant with the emerging CurrentOS standard. Note that some DC topologies directly connect the battery to the DC bus, but these topologies are out of scope for this study. DC/DC PSUs are required to step the 350 V distribution voltage down to the 48 V of the GPU node. They do not need a rack-mount UPS, and DC VSDs can connect directly to the 350 V bus without need for DC/DC conversion.

For each pathway, we calculate cumulative efficiency as the product of all component efficiencies along that path. To determine hourly bus power requirements, we divide the IT and cooling loads by their respective bus-to-load efficiencies, accounting for conversion and distribution losses. Similarly, we calculate required generation by dividing bus power by the source-to-bus efficiency for each generation or storage asset. This power flow modeling establishes the core of the facility-specific design constraints: user-specified minimum uptime, location-aware hourly expected IT and cooling demand on the main distribution bus, facility design load for infrastructure scoping, and pathway-specific power delivery efficiencies. Note that we don't explicitly model supercapacitor banks to maintain bus voltage or any required fault management infrastructure like high efficiency circuit breakers. Full characterization of the costs and performacne characteristics of these system

B. Generation, Storage, and Degradation. Having established facility demand and power-flow paths, we next model how each generation technology supplies that demand over the project lifetime. Renewable systems must buffer resource intermittency with storage, whereas thermal plants must ride through mechanical failures and maintenance outages. All technologies degrade over time, requiring accurate modeling throughout the evaluation period.

We compare all architectures at year 27,the planned decommissioning point of the solar systems, so that capital recovery periods are aligned even though batteries undergo mid-life replacement and gas turbines operate through continuous maintenance cycles. Using location-specific TMY data, we model hourly solar irradiance and ambient temperature effects on all equipment: PV generation, battery efficiency, and gas turbine thermal derating. Solar-storage systems are simulated with hourly dispatch to capture battery cycling and multi-day weather events. We model natural-gas plants with a probabilistic availability model that reflects forced outages and scheduled maintenance. The following subsections describe the generation, storage, and degradation models in detail.

B.1. PV and BESS. We adopt the vectorized hourly energy balance approach from (2, 24), implemented with a bus-centric architecture that explicitly models conversion losses through different power-flow paths. The PV generation model fetches TMY weather data (irradiance and temperature) for the specified coordinates and feeds these into pvlib's ModelChain to calculate DC power output. The ModelChain simulates a reference PV module equipped with single-axis tracking, incorporating the CEC temperature model for thermal derating and accounting for 14% system losses from soiling, shading, mismatch, wiring, and light-induced degradation. This simulation runs across all 8760 hours of the year to produce raw DC power values. We then normalize these values by dividing each hourly output by the annual peak DC output, yielding a capacity factor profile ranging from 0 to 1. This can then be rescaled to any specified nameplate capacity.

PV module output is translated to power delivered to the distribution bus with architecture-specific efficiency chains defined in the power-systems subsection. For AC-coupled layouts the chain includes inverter clipping: the effective AC power output from the inverter, $P_{\text{effective}}$, is defined as the minimum of the efficiency-adjusted DC power produced by the PV modules and the inverter's nameplate AC output: $P_{\text{effective}} = \min(P_{\text{DC}} \times \eta_{\text{inv}}, P_{\text{inverter}})$ where P_{DC} is the instantaneous DC panel output, η_{inv} is the inverter's conversion efficiency, and P_{inverter} is the inverter's rated AC capacity. DC-coupled layouts bypass the inverter and use their own conversion stages.

We assume the battery system is composed of 4-hour, 60 MW LFP modules, each rated to store 240 MWh of usable energy (25). We selected LFP as our storage chemistry due to its excellent durability and ramp rate (26) and current commercial

deployment in grid-scale storage applications. While alternate chemistries provide superior energy-density, in the stationary application we model space and weight are not key performance sensitivities. We adopt a conservative LFP round-trip-efficiency (RTE) of 90% (27), and initial battery state-of-energy as 75% nameplate capacity.

Battery capacity fades over time due to calendar and cycling effects, with charge-discharge patterns and depth of discharge affecting fade severity (28–30). Battery management systems protect cell longevity by restricting operation to safe voltage ranges, typically utilizing 80-90% of theoretical cell capacity. Our model treats the manufacturer's nameplate capacity as representing this BEMS-accessible energy window; when we model a 240 MWh battery at 100% state-of-charge, this corresponds to the maximum energy the BEMS permits, not the theoretical cell maximum. This nameplate capacity degrades over time as discussed in subsequent sections.

The hourly dispatch algorithm prioritizes serving facility load from solar generation. Surplus solar charges the battery (subject to power rating and available headroom), with remainder curtailed. Deficits trigger battery discharge (subject to power rating and stored energy). The simulation implements:

$$\Delta SOE_t = \begin{cases} P_{\text{surplus},t} \cdot \Delta t / (\eta_{\text{charge}} \cdot E_{\text{nom}}) & \text{if solar exceeds load} \\ -P_{\text{deficit},t} \cdot \Delta t \cdot \eta_{\text{discharge}} / E_{\text{nom}} & \text{otherwise} \end{cases} \quad [2]$$

Subject to $0 \leq SOE_t \leq 1$ and $|P_{\text{batt}}| \leq P_{\text{batt,max}}$. The simulation tracks annual cycles, throughput, curtailment, and unmet load for economic analysis.

B.2. Natural Gas. Natural gas power plant design fundamentally differs from renewable energy system optimization due to its non-linear solution space: multiple interacting discontinuities that preclude simple scaling from reference designs. A gigawatt-scale combined cycle plant represents a fundamentally different engineering solution than what a 100 MW datacenter requires—not merely smaller, but potentially employing a different turbine class, cycle configuration, and redundancy architecture altogether. Turbine selection is inherently discrete, with aeroderivative, F-class, and H-class machines each occupying distinct niches in the capacity-efficiency-reliability design space. The number of units introduces step changes in redundancy and operational flexibility, while the choice between simple and combined cycle configurations fundamentally alters project economics: combined cycle plants approach 60% thermodynamic efficiency but require higher upfront capital investment and extended construction timelines. Our approach addresses these interdependencies by systematically generating candidate configurations from commercially available turbines, applying engineering constraints derived from industry practice, and evaluating the full lifecycle economics of each specific configuration.

Our design tool proceeds from a library of reference natural gas turbines, comprised of 16 individual turbine models, subset according to their manufacturer classification: aeroderivative, F-class, and H-class. These classifications correspond to turbine characteristics: Aeroderivative turbines are flexible but low-capacity and comparatively inefficient; class H turbines have the highest capacity and fuel-efficiency, but are the most complex to manufacture with high upfront costs; F-class turbines are medium-weight, medium capacity frames most often used in industrial applications (31). For each turbine model, we source nameplate capacity, combustion efficiency, and annual availability from manufacturer specifications and turbine data sheets.

We model the operations of these turbines according to the following formulations. Note that specific parameter values are available in the configuration file of the codebase and the SI. The effective power output and efficiency of any given gas turbine configuration are determined by its class, c, from the set $C = \{\text{aeroderivative, F-class, H-class}\}$. Performance is a function of scheduled maintenance, operational load, and ambient environmental conditions, which we model as follows:

Periodic maintenance is modeled as a deterministic reduction in a plant's availability. The Maintenance planned Outage, M_c, for a turbine of class $c \in C$ is defined as the ratio of its required annual maintenance period, H_c (in hours), to the total hours in a year, y_{year}.

$$M_c = \frac{H_c}{y_{year}} \quad [3]$$

A turbine's thermodynamic efficiency degrades when operating below its nameplate capacity. The adjusted operational efficiency, η_{op}, is the product of the turbine's rated efficiency, η_{rated}, and a class-dependent part-load multiplier, $f_{\text{part-load}}(\beta_{op}, c)$. This multiplier is a function of the plant's operational load factor, β_{op}, which is the ratio of average instantaneous power to the plant's total nameplate capacity.

$$\eta_{op} = \eta_{rated} \times f_{\text{part-load}}(\beta, c) \quad [4]$$

The part-load multiplier function is based on empirical analysis of turbine performance curves (32) and is defined as:

$$f_{\text{part-load}}(\beta, c) = \begin{cases} 1.0 - 0.10(1-\beta) & \text{if } c = \text{aeroderivative} \\ 1.0 - 0.15(1-\beta) & \text{if } c = \text{F-class} \\ 1.0 - 0.20(1-\beta) & \text{if } c = \text{H-class} \end{cases} \quad [5]$$

A turbine's maximum power output degrades with higher intake air temperature due to lower density. The available power, P_{avail}, is determined by derating the turbine's nameplate capacity, P_{rated}, based on the ambient temperature, T_{amb}. The derating is governed by a class-dependent factor (33), $r_{\text{derate}}(c)$, and applies only for temperatures above the ISO standard of $T_{ISO} = 15^{\circ}\text{C}$.

$$P_{avail} = P_{rated} \times [1 - r_{\text{derate}}(c) \times \max(0, T_{amb} - T_{ISO})] \quad [6]$$

For both capacity derating and part-load efficiency penalty, we selected conservative estimates, and varied the values in sensitivity analysis.

The total power of a combined cycle plant, P_{CC}, is calculated by augmenting the gas turbine's power, P_{GT}, with additional power generated from its waste heat. The power gain depends on the heat of the exhaust gas, which depends on the gas turbine's efficiency, η_{GT}, the effectiveness of the heat recovery steam generator, ϵ_{HRSG} (assumed to be 80%), and the efficiency of the steam turbine cycle, η_{steam} (assumed to be 34%) (34).

$$P_{CC} = P_{GT}\left[1 + \left(\frac{1}{\eta_{GT}} - 1\right) \times \epsilon_{HRSG} \times \eta_{steam}\right] \quad [7]$$

We derive plant performance according to these relationships, with a candidate configurations completely and uniquely determined by: turbine model, fleet size, and single-or-combined cycle. We restrict the design space using industry derived engineering filters to exclude infeasible and implausible configurations.

The most basic is capacity filtering: a plant must be able to meet the facility's peak demand P_{req}, defined as P_{design} scaled by cumulative power system losses, even under adverse environmental conditions. We enforce this by requiring that the plant's total capacity, after being derated for the site's 99th percentile design temperature, be greater than or equal to P_{req}.

Beyond this physical requirement, we constrain the plant's annual expected operational load factor (β_{op}) to a window of 30-90%. The 30% floor reflects a conservative minimum stable output level where turbine manufacturer specifications remain reliable. The 90% ceiling enforces an ~11% planning reserve margin, which provides the necessary headroom to handle demand fluctuations or unexpected outages. This margin is a slight relaxation of the 15% standard common in utility planning, reflecting the stable demand profile of the modeled data center.

We impose architectural constraints to ensure that designs reflect commercially viable and operationally practical power plants. First, each plant configuration is limited to a single turbine model, precluding mixed-turbine fleets to reflect the significant logistical and operational challenges of mixed-asset plants. Second, aeroderivative turbines are restricted to simple cycle configurations due to the unfavorable economies of scale of the required heat recovery systems. Finally, to reflect practical limits on plant complexity, the model constrains the number of units in a fleet. This heuristic, based on industry consultation, allows for a greater number of units in fleets of smaller turbines—such as aeroderivatives and small F-class models—than in fleets composed of large industrial frames.

Finally, a more stringent *N-1 redundancy* filter can be optionally enabled. This standard reliability constraint requires that the plant can still meet the peak demand P_{req} even with its largest single unit offline. This check is performed using the temperature-derated capacity to ensure N-1 sufficiency during hot weather conditions.

While the engineering filters ensure each plant configuration is robust, we calculate the Expected Unserved Energy (EUE) to quantify the remaining reliability risk and to appropriately size the backup power system. EUE is a standard industry metric representing the expected annual energy shortfall from a given plant design due to both random (forced) and scheduled (planned) outages.

The total EUE is the sum of the forced and planned outage components, $EUE_{total} = EUE_{forced} + EUE_{planned}$. We calculate the forced outage component, EUE_{forced}, using a binomial probability model based on the single-unit forced availability, A_f. The expected annual shortfall is the sum of all potential outage states weighted by their probabilities:

$$EUE_{forced} = H_{year} \sum_{k=0}^{N-1} \binom{N}{k} A_f^k (1 - A_f)^{N-k} \times \max(0, P_{req} - k \cdot P_{unit}) \quad [8]$$

The planned outage component, $EUE_{planned}$, is the deterministic energy shortfall that arises when units are deliberately taken out of service for maintenance. For both forced and planned outages we compute expected unserved energy using the facility's average demand. For planned outages this implicitly assumes maintenance is scheduled outside the hottest weeks.

Note that EUE is not used as a direct optimization target in our methodology. The primary application of the calculated EUE value is to inform the sizing of the backup diesel system's fuel reserves. The fuel storage capacity must be sufficient to generate the energy equivalent of the gas plant's expected annual shortfall, plus a safety margin. This ensures the integrated system can meet its overall reliability target. EUE is an annual statistical expectation, not a chronological simulation. It does not provide information on the timing, frequency, or duration of individual outage events. It serves to size the total volume of backup energy required per year, but not the power capacity of the backup system, which is sized to meet the facility's peak demand (P_{req}) .

Unlike the probabilistic treatment of gas plant reliability, diesel backup sizing follows deterministic criteria standard in mission-critical facilities. The diesel system must be able to deliver the full facility demand P_{req} regardless of the gas plant configuration, as any shortfall would violate the data center's uptime requirement. We enforce an N+1 redundancy architecture where the total installed fleet includes one additional unit for redundancy. This N+1 configuration ensures the diesel plant can meet full load even with one generator offline for maintenance or failure and is consistent with standard industry practice. Diesel backup system design rules reflect guidance for critical backup power infrastructure (35, 36), with those sources also providing our diesel system financial parameters.

While diesel generators provide the power capacity for backup, their runtime is constrained by on-site fuel storage. We size fuel reserves based on the gas plant's Expected Unserved Energy, applying a safety factor to account for uncertainty in outage patterns:

$$V_{\text{fuel}} = \max\left(\lambda \cdot \text{EUE}_{\text{total}} \cdot \text{SFC}, V_{\text{min}}\right) \quad [9]$$

where V_{fuel} is the fuel storage volume (gallons), λ is a safety factor (typically 2.0), SFC is the specific fuel consumption (gallons/MWh), and V_{min} represents the minimum code-required storage. The minimum storage ensures 72 hours of runtime at full load, a standard requirement for critical facilities. This dual criterion ensures adequate fuel for both statistical expectations and exogenous scenarios such as extended natural disasters or fuel delivery disruptions.

Note that the tool includes another optional engineering filter that we do not enforce in our base analysis: maximum diesel turbine runtime. Operators are typically only permitted to operate backup power turbines up to some annual limit (typically 200-500 hours, varying by region) per annum due to air quality standards. Above this threshold, operators need to register as prime power providers, with higher operational scrutiny. We elect to not enforce this filter in our base case as the facility in our scenario is a primary power provider, and would thus likely undergo full permitting and regulatory review. If they anticipate their design would require operating the diesel gensets above annual maximum, they would obtain permitting at that initial stage of the process. Users of the tool suite can enforce the flag to filter out designs that would exceed some specified annual max, with the practical impact of favoring multi-turbine configurations.

The diesel backup system we do not independently optimize but rather size deterministically based on the selected gas plant configuration. This coupling means that gas plant reliability directly affects total system costs: a less reliable gas configuration reduces gas plant capital costs but increases both diesel fuel storage requirements and expected fuel consumption. The life-cycle cost optimization described in design optimization subsection captures these tradeoffs, selecting the gas-diesel combination that minimizes total cost while meeting reliability constraints.

B.3. Degradation. Power-generating and energy-storage assets degrade over their useful life. Accurate degradation modeling is essential for lifecycle economic comparison, as systems that appear cost-competitive at commissioning may diverge significantly in delivered energy costs over a 25-year horizon. We implement degradation models calibrated to empirical data where available, with conservative assumptions where data is limited.

For solar PV, we apply a linear degradation model with 1% first-year loss, primarily from light-induced degradation, followed by 0.55% annual degradation thereafter (37). These degradation values were validated through consultation with a utility solar system design expert. While module-specific and climate-dependent variations exist, the total uncertainty band over 25 years has minimal impact on LCOE due to heavy discounting of later years in financial analysis.

Natural gas turbines experience both capacity and efficiency degradation, with rates differentiated by turbine class to reflect their distinct operational characteristics and maintenance requirements. Properly maintained plants see dramatically lower degradation rates. Our model assumes realistic operating environments with regular maintenance interventions such as compressor washing, resulting in annualized degradation rates that align with industry experience of 0.5-1.5% loss between major overhauls (38). We differentiate by turbine class, with aeroderivative turbines experiencing higher degradation rates due to their higher pressure ratios and sensitivity to fouling, while F-class and H-class heavy-duty frames exhibit greater durability owing to their robust design and longer service intervals. Importantly, efficiency degradations in natural gas plants affect cost performance but not the system's ability to deliver SLA-guaranteed power, and capacity degradations remain much smaller than the capacity headroom required by our design rules. This distinction allows us to adopt simplified linear calendar approaches for both PV and gas turbine degradation modeling. See the sensitivity analysis for the discussion of the sensitivity to our choice of degradation factors.

Lithium-ion battery degradation requires more sophisticated treatment. Battery capacity fade results from complex, interdependent electrochemical processes (30) that vary nonlinearly with operating conditions(39). Depth of discharge (the fraction of battery capacity used in each cycle), C-rate (the charge/discharge rate relative to battery capacity), cell temperature, and state-of-charge (the battery's current energy level as a percentage of capacity) all influence degradation trajectories. Further, batteries experience both calendar aging (degradation over time regardless of use) and cycling stress (degradation from charge-discharge cycles), and these factors can mutually interact, introducing significant non-linearity into battery fade dynamics.

Battery capacity fade modeling employs BLAST-Lite, an open-source tool implementing semi-empirical degradation models developed by Smith et al. (40). The tool's LFP model derives from multi-year cycling data of grid-scale batteries under representative duty cycles, making it directly applicable to our use case. BLAST-Lite tracks degradation through two pathways: calendar aging (time-dependent SEI growth and electrolyte decomposition) and cycle aging (mechanical stress and lithium inventory loss from charge-discharge cycles). The model accepts hourly time-series of power throughput, state-of-charge, and temperature, then computes capacity fade using stress-weighted superposition of these mechanisms.

Blast-lite decomposes complex charge-discharge patterns into equivalent stress cycles. For our LFP chemistry, the model represents key empirical behaviors: minimal sensitivity to depth-of-discharge below 80%, accelerated degradation at extreme SOC ranges (>95% and <5%), and Arrhenius-type temperature dependence with roughly doubled degradation rate per 10°C increase. These path-dependent calculations produce more accurate lifetime predictions than simplified equivalent-cycle or calendar-only models.

Our 25-year facility lifespan exceeds typical 15-year battery system operating lives, making replacement timing and strategy critical design decisions. While unconstrained optimization might minimize initial capacity and add cells annually to offset

degradation, this approach misrepresents actual operational practices, as firms value operational simplicity over theoretical optimality. We therefore implement a single full-scale battery replacement at year 13, reflecting how facilities typically manage major infrastructure: as discrete, planned interventions rather than continuous adjustments. We fix replacement at year 13 to reflect practical procurement strategies where facilities execute major infrastructure upgrades as planned capital projects, not continuous optimization exercises.

Battery thermal management follows a simplified physics-based approach that tracks temperature evolution throughout the year. We model heat generation from round-trip inefficiency as $Q_{gen} = P_{throughput} \times (1 - \sqrt{\eta_{RTE}})$ where throughput power and round-trip efficiency determine waste heat. The battery's thermal mass (2.3 kWh/°C per MWh capacity) buffers these temperature fluctuations. The thermal mass parameter was derived from a representative 3.9 MWh, 38-ton containerized BESS (41), calculating the system's total thermal capacitance using mass-weighted specific heat values of its components: 57% LFP cells, 20% steel, and 23% other materials (42–44).

Active cooling engages when cell temperature exceeds 30°C, with power requirements calculated as the excess heat divided by the cooling system coefficient of performance (derived from the facility PUE). We account for this parasitic load by reducing the facility's net energy delivery by an equivalent value to the cooling energy use. Though thermal loads could theoretically be scheduled to coincide with solar curtailment periods using batteries, we naively dispatch climate control any hour outside of thermal envelope. The model includes provisions for heating when batteries are idle in cold conditions (below 15°C), this rarely engages in practice. This approach captures the essential thermal dynamics and their impact on both degradation rates and parasitic loads without requiring more detailed modeling.

We conducted initial exploratory data analysis on candidate systems using Blast-lite, revealing characteristics of our usage pattern. Batteries in systems that deliver greater than 90% annualized uptime show only modest fade at their parameterized retirement year, averaging 11% cumulative fade on average with roughly even cumulative cycle and calendar fade, as shown in S2. Because batteries are sufficient for overnight operations, and because solar panels are sufficient to both deliver facility load during daylight hours and charge the battery for expected overnight energy demand, battery state-of-energy is strongly right-skewed, with more than a quarter of typical hours spent at 100% SoE, as shown in S1. Typical depth-of-discharge is relatedly modest.

The low battery health degradation we find is consistent with underlying theoretical and empirical results on LFP cycle fade (45, 46). Islanded PV-plus-BESS systems require significant nominal power overbuild to achieve 90% annual uptime (to say nothing of 2 or 4 9s of reliability) because the sun does not shine at night. The design region is thus energy constrained, not power constrained. Beyond even the storage for a typical overnight load, the battery system must also be scoped to accommodate seasonal ebbs and outlier events.

Systems of this size charge and discharge only a small fraction of their total energy capacity in any given time step. Though the exact value varies design-to-design, the maximum nominal C-rate in a given year is approximately .2, with the average C-rate a third of that. This gentle nominal C-rate is consistent with a low-stress battery operating regime (46).This stressor pattern is also fairly stable year to year: annual throughput cycling increases on average only 11% cumulatively over the life of the battery, mean SoC reduces by 4% on average, and annual thermal load increasing less than 1% on average.

A full Blast-lite call to evaluate every candidate design in the optimization loop would have been computationally intractable. The relatively stable fade regime of our batteries allowed us to address this through gray-box surrogate distillation. The surrogate follows a physics-informed machine-learning strategy: a low-order analytic "scaffold" expresses the dominant aging mechanisms, and a data-driven term learns only the small residual that remains between the scaffold and the simulator output. We split degradation into calendar and cycle terms:

$$\Delta q_{\text{cal}} = A\, t^{\beta} \exp\!\Big(-\frac{E_{\text{cal}}}{RT}\Big)\big[1 + \gamma|\text{SOC} - 0.5|\big], \qquad \Delta q_{\text{cyc}} = B\, \text{EFC}^{\alpha} \exp\!\Big(-\frac{E_{\text{cyc}}}{RT}\Big),$$

where t is elapsed time in years, EFC is cumulative equivalent full cycles, T is mean cell temperature in kelvin, and R is the universal gas constant. We locked the Arrhenius energies at the BLAST-Lite values $E_{\text{cal}} = 5.21$ kJ mol^{-1} and $E_{\text{cyc}} = 2.19$ kJ mol^{-1}. Starting from the BLAST-Lite exponents $\beta_0 = 0.526$ and $\alpha_0 = 0.828$, we let the data move them within ± 0.25 and fitted the scale factors A, B plus the SOC-curvature coefficient γ. To capture path dependence, we simulated every design through a complete life cycle and recorded each year's incremental fade. After fitting the scaffold to those increments, we modeled the residuals

$$r = q_{\text{BLAST}} - \Delta q_{\text{scaffold}}$$

with a GP that uses a Matérn ($\nu = 2.5$) kernel and a small white-noise term. We assessed accuracy with five-fold grouped cross-validation, treating each design as one group to prevent data leakage.

We constructed a training set out of detailed Blast-lite degradation simulations of candidate systems in 13 locations over a full battery life cycle. These locations included at least one entry in every DOE climate zone in the continental United States, and 4 designs were simulated at each location . Each year's simulation provided stress features: state-of-charge statistics, temperature, and throughput metrics, creating a path dependent stressor history for each design. Group-aware 5-fold cross-validation ensured robustness across designs and allowed for hyperparameter optimization. We then evaluated the final distilled model performance on a new set of randomly generated designs from 10 locations explicitly not within the training set.

The gray-box fade proxy showed high fidelity to the underlying Blast-lite simulation: out-of-fold (OOF) mean absolute errors for calendar and cycle fade are 0.0016 and 0.0097 respectively, and each had an R^2 greater than 0.99. This proxy maintained predictive performance in our validation data, as shown in S3. The mean cumulative state-of-health error at battery end-of-life between Blast-lite and this proxy is 0.03% for our out-of-sample designs.S3

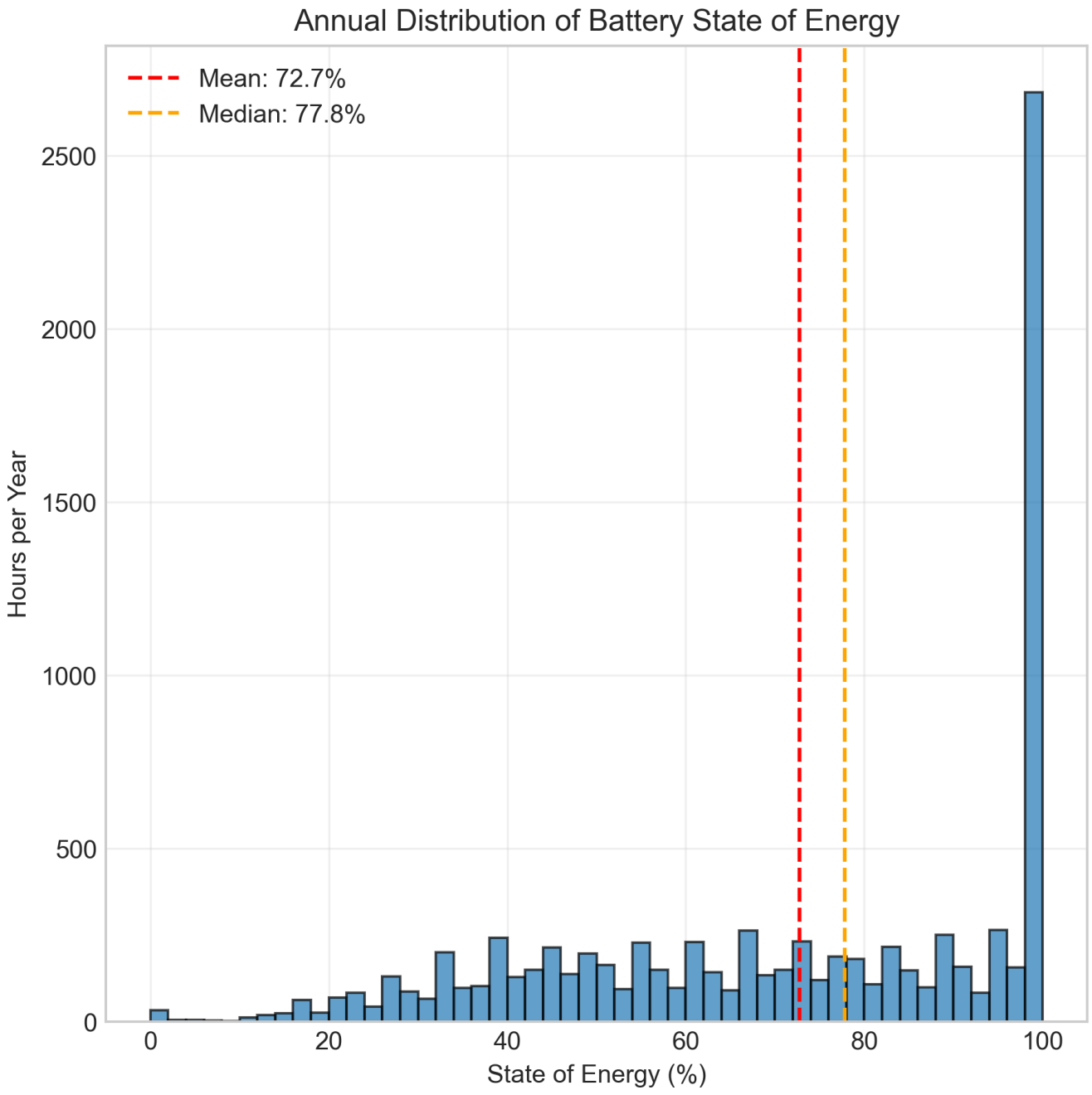

Fig. S1. Distribution of BESS state of energy over the hours in simulation year 0 for a candidate DC coupled design. The modeled design is in southern California (33.60 N, -115.14 W), would supply 75,000 GPUs (facility design load of 99.6 MW) though 515 MW solar PV and 384 MW (1536 MWh) of BESS capacity. Note the strong skew towards high states of energy, as storage systems are scoped for outlier events. In practice, alongside the practically modest discharge rate for steady overnight operations, configurations capable of high uptime exhibit comparatively low-stress cycling patterns.

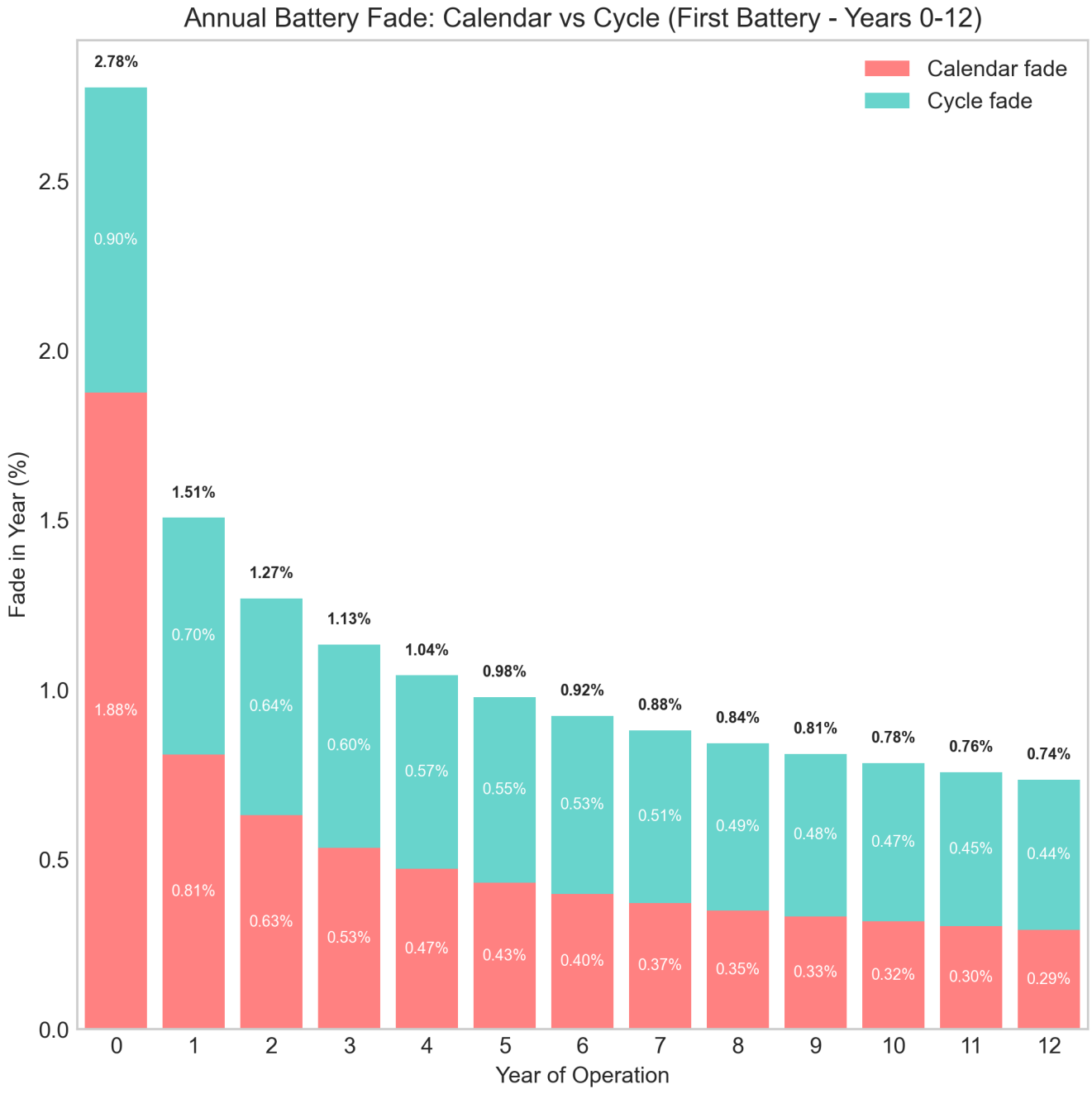

Fig. S2. Cumulative fade from cycling and calendar effects over modelled lifespan of LFP battery system. Modeled 45,000 chip faciliity, 99% uptime, eastern Colorado, DC architecture. Design load of 57 MW, 399 MW PV, 297 MW BESS.

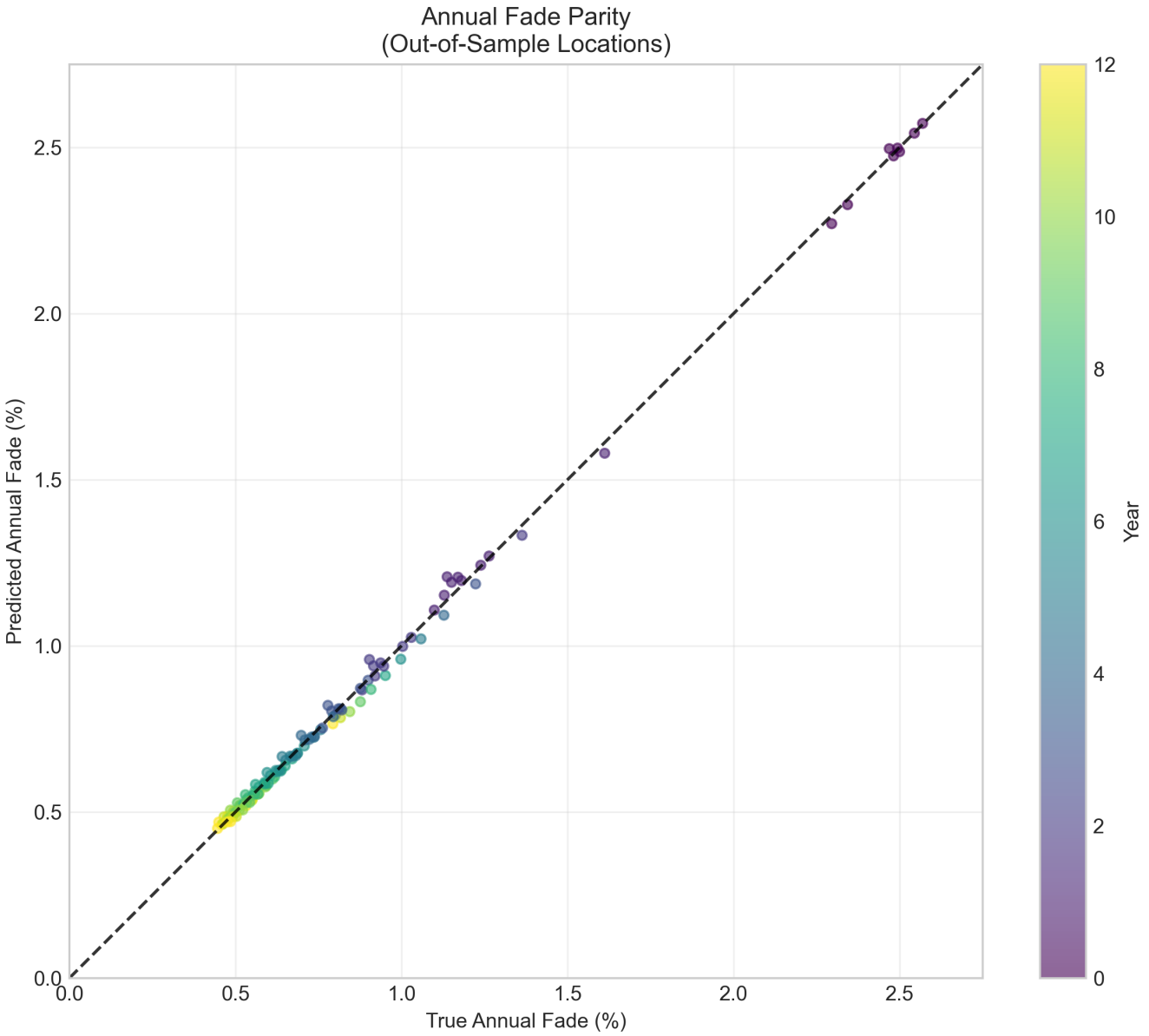

Fig. S3. Out-of-sample validation of the gray-box battery fade surrogate model. The parity plot compares predicted annual fade rates (vertical axis) against ground-truth BLAST-Lite simulations (horizontal axis) for 10 geographic locations not included in the training data. Each point represents one operational year (0-12, color-coded) for one of the validation designs. The tight clustering along the diagonal (dashed line) demonstrates that the surrogate maintains predictive accuracy when extrapolating to new locations.

We implemented battery fade using static stressor proxies with Gaussian-process regression for computational tractability. Rather than modeling the full evolution of stress factors over the battery lifecycle, we approximated degradation using simplified metrics: mean state-of-energy from year zero, and cumulative throughput calculated as year-zero equivalent full cycles multiplied by operating years. While this simplification doesn't account for evolution in stressors, validation against dynamic stressor models showed minimal accuracy loss: a mean absolute error of only 0.14% state-of-health at end of life. While we initially intended to develop a more complex stressor evolution model, exploratory testing of the y0 stressors as proxy to provide a performance baseline proved sufficient for accuracy. This static approximation captures the dominant degradation drivers while keeping the optimization computationally manageable, making it well-suited for our large-scale system design exploration.

C. Technoeconomic Analysis. We employ technoeconomic assessment (TEA) to evaluate the financial performance of candidate designs. Using the net power flow models for each system detailed in the previous sections, we associate costs with the construction and operational parameters of each system, discounting future costs into year 0 dollars. We then divide these costs by the cumulative discounted energy production of the system to calculate the levelized cost of energy (LCOE) as shown in eq xx:

$$LCOE = \sum_{t=0}^{n} \frac{\frac{Cap_t + Op_t}{(1+r)^t}}{\frac{Eng_t}{(1+r)^t}} \quad [10]$$

Where t is the year in a project, n is the lifespan of the project, Cap_t is the capital expenditure in year t, Op_t is the total operating cost in year t, Eng_t is the total output in year t, and r is the discount rate. This quantity provides us a direct point of comparison to evaluate the economic performance of different designs. Our base case discount rate is 7%, and we discount mid-year. All values are in 2022 dollars.

Note that LCOE has well-documented limitations when comparing technologies with different operational characteristics such as dispatchability (47, 48), because the metric does not completely capture how generation timing and reliability affect economic value. Our framework explicitly addresses this critique through scenario equivalence: all power systems face identical facility loads, reliability requirements, and geographic constraints, enabling direct comparison of the true cost to deliver equivalent energy services.

For PV+storage systems, capital and operating costs were derived by combining aggregate system projections from (49) with detailed cost decompositions from (50). The NREL ATB provides bundled utility-scale PV+storage system costs of $2,275/kW (2025, moderate case), which we disaggregated into component costs (PV modules, battery cells, and balance of system) using proportional shares from Ramasamy et al.'s detailed cost breakdowns. To allocate balance of system (BOS) costs between PV and storage components, we compared standalone PV systems to combined PV+storage systems in their data, attributing the full standalone PV infrastructure costs to the solar component and only incremental costs to storage.

For DC-coupled architectures, we estimated BOS cost by analyzing differences between coupling configurations in (50, 51). While electrical system costs for DC coupled architectures showed marginal increases in (51) due to required DC-DC converters, those costs are exceeded by avoidance of the inverters. In (21) the authors find that scale economies in power systems components result in DC-DC converters having lower capital costs compared to bidirectional AC-DC components. Based on this analysis, we derate DC-coupled solar BoS costs by 24% with all savings attributed to reduced PV-side inverter infrastructure while battery-specific BOS costs remained constant. Battery module replacement at year 13 is treated as a capital expenditure using NREL ATB's projected costs for that future year. Land requirements are parameterized at 4.17 acres per MW of solar capacity (52, 53).

Fixed operating and maintenance costs for PV systems were sourced from Ramasamy et al. (2023), using values for standalone PV systems to avoid double-counting battery replacement costs that are embedded in their combined system estimates. We note that these PV O&M costs include inverter replacement, which DC-coupled architectures would avoid; however, as the reference does not specifically decompose these costs, we retained the full values while acknowledging this represents a conservative overestimate for DC-coupled systems. Battery storage O&M costs were obtained from the (54) database entry for lithium iron phosphate systems.

Natural gas plant costs vary by turbine class, plant configuration, and capacity. We source baseline capital and operating costs from established references: F-class and H-class single and combined cycle configurations from the NREL ATB (49), and aeroderivative turbines from EIA's Annual Energy Outlook (55). These reference plants provide cost benchmarks at specific capacities (e.g., 237 MW for F-class simple cycle, 702 MW for H-class combined cycle).

To estimate costs for configurations that differ from these reference plants, we apply empirical scaling relationships following the well-established "six-tenths rule" (56). After reviewing component-level scaling behaviors (57), we adopt scaling exponents of 0.67 for capital costs and 0.80 for fixed O&M [§]. The scaling relationships are:

$$\text{CapEx}_{\text{scaled}} = \text{CapEx}_{\text{ref}} \times \left(\frac{P_{\text{plant}}}{P_{\text{ref}}}\right)^{0.67} \times f_{\text{complexity}} \quad [11]$$

$$\text{Fixed O\&M}_{\text{scaled}} = \text{Fixed O\&M}_{\text{ref}} \times \left(\frac{P_{\text{plant}}}{P_{\text{ref}}}\right)^{0.80} \times f_{\text{complexity}} \quad [12]$$

[§] We scale per-kW costs using exponents of -0.33 and -0.20, which is mathematically equivalent to scaling total costs with exponents 0.67 and 0.80 since per-kW cost × capacity = total cost.

where P_{plant} and P_{ref} are the candidate and reference plant capacities, and $f_{\text{complexity}}$ accounts for multi-unit configurations not represented in our reference dataset: 1.0 for 1-2 units, 1.3 for 3-4 units, and 1.8 for >4 units, reflecting increased balance-of-plant and coordination costs. Diesel system costs are linearly scaled according to system capacity from reference documentation (35). We did not vary variable O&M unit costs within a plant class.

Annual fuel costs in year t depend on the energy delivered by natural gas and diesel generators, multiplied by their respective fuel consumption rates. Natural gas generation is constrained by plant availability:

$$E_{NG,t} = \min(P_{plant,t} \cdot AF \cdot H_{year}, E_{required}) \quad [13]$$

where $P_{plant,t}$ is the nameplate capacity after degradation (MW), AF is the availability factor incorporating forced and planned outages, $H_{year} = 8760$ hours represents annual operating hours, and $E_{required}$ is the facility's annual energy demand (MWh). Any shortfall is met by diesel generation:

$$E_{diesel,t} = \max(E_{required} - E_{NG,t}, E_{test}) \quad [14]$$

where E_{test} accounts for mandatory testing requirements. The gas plant's operating efficiency depends on its operational load factor $\beta_{op} = P_{avg}/P_{plant}$, which represents the ratio of average facility demand to plant nameplate capacity. Operating efficiency in year t is adjusted for both part-load penalties and temporal degradation:

$$\eta_{op,t} = \eta_{rated} \cdot f_{part-load}(\beta, c) \cdot \eta_{deg}(t) \quad [15]$$

where η_{rated} is the rated efficiency at full load, $f_{part-load}$ is the class-dependent part-load penalty function defined previously, and $\eta_{deg}(t)$ represents efficiency degradation over time.

Total fuel cost in year t is calculated as:

$$C_{fuel,t} = E_{NG,t} \cdot \frac{\varrho}{\eta_{op,t}} \cdot p_{gas} + E_{diesel,t} \cdot \eta_{diesel} \cdot p_{diesel} \quad [16]$$

where $\varrho = 3.412$ MMBtu/MWh is the thermal energy conversion factor, p_{gas} is the natural gas price ($/MMBtu), η_{diesel} is diesel specific fuel consumption (gal/MWh), and p_{diesel} is the diesel price ($/gal). Diesel price were based on the USA national average according to the EIA circa July, 2025 (58). We price natural gas at the state level, using an Olympic scored 5-year average of each state's industrial natural gas price (59). We analogously assign each state an industrial electricity price based on the July industrial electricity prices (60) All prices are reported in y0 dollars.

Our LCOE calculation divides total system costs by delivered energy at the facility loads, not generated energy at the source. For solar+storage systems, E_t represents the hourly facility demand successfully served after accounting for all conversion losses through the power distribution chain. Natural gas plant fuel consumption and equipment sizing reflect required generation at the source, yet we amortize costs only over the energy services delivered to the load, as with the PV based systems. This approach ensures consistent comparison across architectures: all systems are evaluated based on the energy that actually reaches IT and cooling equipment.

C.1. Procurement, Construction, and Interconnection. Beyond equipment and operating costs, project economics depend critically on implementation timelines that vary substantially across generation technologies and grid interconnection pathways. Economic value from any end-use is only unlocked when construction of generation and delivery assets completes. When we refer to construction, we mean the combined site permitting, engineering design, physical construction, and testing required to begin plant operation.

We applied the following capital outlay schedule across all architectures: For projects lasting one year or less the entire investment is assumed to occur at notice to proceed (year 0); for two- and three-year builds the schedule follows industry and modeling heuristics of 40 %/60 % and 30 %/40 %/30 % in years 0–2 (55, 61). Projects longer than three years are treated as being spent uniformly: the total cost is divided equally among the full years of design and construction and prorated for any remaining partial year.

For solar and storage systems, we began with the 2 year all-in timeline from (24). To validate our assumptions, we consulted with relevant industry experts. An early stage grid storage expert estimated that the median timeline for component delivery of a large contract for LFP cathodes was 12 months using existing mature supply chain. They estimated the time to bootstrap a new production base (for example in response to export restrictions) from scratch was 18-36 months. A PV systems design expert estimated the average cumulative project timeline for an integrated utility scale PV+BESS project was 18 months, with 16-24 months as the overall range. Based on this feedback, we select 24 months as our cumulative project development timeline to bring solar+storage systems online.

Natural gas plants face more varied timeline influences. Natural gas plant construction we model as differing between single and combined cycle plants, building off estimates in (55), whit the additional 2 months for combined cycle configurations reflecting increased plant size and complexity. Natural gas plant construction is not perfectly parallelizable, especially with regard to the keystone component of the turbine. We estimate that 30% plant construction and validation must occur after turbine delivery, making turbine receipt the rate limiting factor. While reports of turbine delays can exceed five year (62), we select value of three years for all turbine models. With base construction timelines of 36 months (simple cycle) and 42 months (combined cycle), total project duration is determined by the maximum of either the base construction time or turbine delivery plus post-delivery construction share.

As a point of comparison, we also estimate the time to interconnection for baseline grid-connected facilities: this includes the initial request processing, any associated studies, and construction of marginal transmission and distribution infrastructure. The process and complexity can differ between supply-side and demand-side interconnection requests, but we lack systematic data for demand side request duration. We adopt supply side interconnection queue data as our proxy for large load interconnection timeline, using raw data data from (63). We took the 2018-2023 median initial request to commercial operation duration reported for each ISO/RTO or non-ISO region, converted it to years, and assigned that value to every state in the region. For the handful of states that straddle two regions we chose the more conservative of the two medians. Although based on supply-side projects, these medians provide an empirically grounded proxy for the timeline a large load would expect to interconnect. We examine the sensitivity of our results to these timeline assumptions sensitivity analysis section of the SI.

C.2. Design Optimization. To identify cost-optimal configurations within each architecture class, we must search design spaces while maintaining the hard constraint that all systems achieve required facility uptime. The search strategies differ fundamentally between technologies with continuous versus discrete design variables. Solar+storage optimization searches over continuous capacity ranges for both solar (typically 1-15× facility load) and battery power (typically 1-10× facility load). We employ two-stage optimization: Latin hypercube sampling with relaxed evaluation (year-0 only) identifies feasible regions, then differential evolution refines the solution using full lifecycle analysis. This approach efficiently handles the non-convex solution landscape created by the interplay between diurnal solar cycles, multi-day weather events, and battery state-of-charge dynamics

In our base case, we treat design increments in PV based systems as effectively continuous, the marginal solar cell or LFP module is minimal compared to the scale of the facility. We test sensitivity to this assumption by evaluating the affect on LCOE of jumping up to the nearest minimum PV and BESS design increment/step size. Sweeps across facility scales (10,000-100,000 GPUs), locations (Phoenix, Austin, Portland), and module sizes (1-10 MW) show cost penalties scale inversely with system size and are dominated by discrete factorization alignment rather than system performance. Median penalty across realistic configurations (2.5 MW PV, 1 MW BESS and below) is $< 0.3\%$, though the combination of small facilities ($< 25,000$ GPUs) and oversized step size (> 5 MW) can cause the penalty to exceed 2%. Given the small absolute penalty of industry representative step-sizes ($0.001 / kWh on average) as compared to the continuous optimum, we accept continuous search as a simplifying assumption.

The key computational challenge is avoiding full 25-year hourly simulation for each candidate design. We make two approximations. First, during search we minimize capital cost rather than LCOE for solar and storage systems —these are proportional for systems with identical operational profiles, making capital cost an exact proxy. Second, we evaluate only four anchor years that capture the system's degradation trajectory: fresh commissioning, maximum degradation before battery replacement (year 13), the discontinuity from replacement (year 14), and end-of-life (year 25). Between anchors, linear interpolation slightly underestimates the generated energy delivery. This causes our method to overestimate LCOE by 0.014% on average—acceptable given the 6-fold reduction in simulation calls.

Natural gas systems present a discrete combinatorial problem: select turbine model, determine unit count, and choose cycle configuration. The engineering filters described previously reduce the solution space sufficiently that exhaustive enumeration becomes tractable. We evaluate LCOE for each configuration passing the filters and select the minimum-cost option. Unlike the continuous solar+storage optimization requiring iterative search, natural gas selection is deterministic once the feasible set is enumerated.

C.3. Value of Energized Compute. Unlike conventional energy infrastructure that generates value only after commissioning, data centers house compute assets capable of generating revenue immediately if deployed elsewhere. Each month of construction represents not only capital carrying costs but also foregone computational output: training runs not completed, models not refined, services not deployed. This opportunity cost varies dramatically across power architectures: grid interconnection may require 4 years in congested regions, translating to 2 years of lost utility compared to faster-deploying alternatives. We term this differential the "speed premium"—the economic value of earlier compute availability. This section develops a framework for quantifying these time-dependent costs and incorporating them into our technoeconomic comparison.

Proper comparison of total system economics requires accounting for both the levelized cost of energy and the opportunity cost of delayed operations. This opportunity cost represents the value of compute services that could have been delivered during the marginal construction period of a slower system. This necessarily raises the question: what would the value of those computations have been?

The economic value of earlier operation depends on the opportunity cost of compute time. Within AI development, GPU-hours represent the fundamental unit of resource allocation, with utilization carefully tracked and optimized. The existence of liquid spot markets for H100 compute provides an empirical lower bound for this value: at $2.40 per GPU-hour (64), operators who cannot extract at least this value from their infrastructure would rationally rent capacity on the open market instead.

Infrastructure providers with internal AI development (such as major hyperscale cloud firms) must ration compute between internal workloads and external rental based on relative value. When internal workload value falls below rental rates, profit-maximizing firms should rent capacity rather than run lower-value jobs internally. The rental price thus represents the marginal decision threshold: the value of the last GPU-hour consumed internally.

We use the observed spot rental price for H100 capacity ($2.40/GPU-hour) as our baseline opportunity cost. This represents a reasonable approximation for hyperscalers who continuously optimize across diverse workloads spanning frontier model training (likely yielding much higher internal value) to routine inference and fine-tuning. This assumption likely provides a

conservative bound for specialized AI labs running exclusively proprietary training workloads. See the sensitivity analysis and Figure 3 in the results for further discussion of how GPU-hour valuation affects outcomes.

We model the value of lost compute during construction years as:

$$VLC_t = \sum_{h=1}^{H_{\text{idle},t}} N_{\text{GPU}} \times p_{\text{spot}} \quad [17]$$

where VLC_t represents the value of lost compute in year t, $H_{\text{idle},t}$ is the total idle compute hours in year t prior to operations, N_{GPU} is the total GPU count, and p_{spot} is the hourly spot-market price per GPU-hour (64). Only idle hours before the start of operations at t_{ops} are considered. We incorporate VLC_t into our total cost assessment as:

$$\text{LCOE}_{\text{adj}} = \frac{\sum_{t=0}^{n} \frac{Cap_t + Op_t + VLC_t}{(1+r)^t}}{\sum_{t=0}^{n} \frac{Eng_t}{(1+r)^t}} \quad [18]$$

This formulation treats lost compute value as functionally an additional expense during construction years, ensuring architectures with longer deployment timelines bear the full economic burden of delayed operations. The framework assumes compute assets are procured at project initiation; alternative procurement schedules can be modeled by adjusting when VLC_t becomes non-zero.

D. Sensitivity Analysis. We evaluate how parameter uncertainty affects both system costs and architectural selection through two complementary approaches: varying parameters for fixed designs to test cost estimate robustness, then re-optimizing with modified parameters to assess whether uncertainty changes which architecture is optimal. In each case, we selected a site in southeastern Colorado (37.5, -103.5) with a local gas price of $7.70 / MMBtu, and tested the 99% uptime baseline case.

D.1. Baseline LCOE sensitivity.. Solar+storage systems respond most strongly to discount rate (27% decrease at 3%, 42% increase at 12%) and construction timeline (+$0.06/kWh from 18 to 36 months). Component costs matter less: ±20% variations in batteries, modules, or balance-of-system each shift LCOE ≤$0.03/kWh. Land costs are negligible, requiring 6× increase to ∼$1M/km^2 to match even the next lowest cost parameter, battery O&M sensitivity.

Natural gas systems show strongest sensitivity to fuel prices, with continental U.S. range (Oklahoma to California) creating $0.07–0.11/kWh variation, larger for smaller turbines. Discount rate effects are comparable ($0.05–0.07/kWh across 3–12%). Construction timeline (3–5 years) and ±20% capital costs each affect LCOE ∼$0.02/kWh.

D.2. Speed-premium sensitivity.. Incorporating deployment speed dramatically amplifies temporal parameter influence. For solar PV based systems, GPU opportunity cost ($1.00–4.00/hr) and discount rate each create $0.55/kWh swings in LCOE_{adj}—larger than construction timeline effects ($0.40/kWh). Natural gas shows even stronger sensitivity ($1.25/kWh, $1.03/kWh, $0.89/kWh respectively) due to longer baseline timelines. At current GPU rental index values, deployment speed dominates nameplate energy economics.

D.3. Technical parameter sensitivity. Re-optimization with modified technical parameters shows robust architectural selection. For natural gas, 72 parameter variations across four scales triggered only five configuration changes, all from power electronics efficiency improvements reducing generation needs across turbine size thresholds. Technical parameters (part-load curves, temperature derating, conversion efficiencies) show 1–4% LCOE impacts when relevant to selected configurations. Solar systems exhibit comparable stability, with battery round-trip efficiency largest at ±3.5% (for 0.85–0.95 range) and power conversion ±1–2%. Sizing to meet end-of-life availability requirements effectively eliminates design contingency sensitivity; despite significant variation, we observe minimal impact impact on cost. In five instances during , the design optimization for natural gas systems switched to a different configuration when changing a power electronics value. In each case, high-end estimates for UPS or PSU efficiency enabled a lower-capacity design to pass minimum load engineering filters .

D.4. Optimization stability. To characterize the robustness of our microgrid designs and the stability of our optimization, we perform 30 independent runs at 14 locations representing each DOE climate zone. LCOE coefficient of variation ranges from 0.3% to 3.0% across locations (median 1.4%, mean 1.5%). Stability varies systematically with climate as shown in S7: high solar resource locations (Miami, Phoenix, Albuquerque, Las Vegas) achieve tight convergence with CV $< 1\%$, while locations with variable cloud cover or pronounced seasonality (Chicago, Baltimore, Helena) exhibit CV of 2–3%.

Component sizing exhibits substantially higher variation than LCOE, with solar capacity CV averaging 5.2% (median 4.3%) and battery capacity CV averaging 6.1% (median 5.7%), both 3–4× higher than LCOE CV. This disparity indicates near-flat objective function regions where configurations differing by 5–10% in component sizing achieve nearly identical economic performance (within ±2% LCOE). This Pareto frontier-like behavior reflects substitutability between solar and battery capacity on the margin near the boundary of the upttime constraint.

The stability differences across locations reflect search difficulty rather than fundamental problem structure. In high-solar-resource locations, the cost-optimal region represents a large target in the design space that sampling-based optimization reliably identifies. In locations with more variable solar resources, the cost-optimal region may be narrower or contain multiple local basins, reducing the probability that random initialization seeds near the global optimum. Visual inspection S7 confirms

Table S1. Base LCOE sensitivity analysis across power system architectures (excluding GPU idling costs). Percent changes are relative to each architecture's baseline LCOE (shown above result columns). Solar+storage results are scale-independent across 24k, 100k, and 500k GPU facilities. Natural gas results shown for 100k GPU facility in southeastern Colorado. Baseline assumptions: Solar construction = 2.0 yr; NG construction = 4.0 yr; discount rate = 7%; NG price = $7.70/MMBtu.

Parameter	Low ($/kWh)	High ($/kWh)	Δ% Low (%)	Δ% High (%)
AC-Coupled Solar+Storage *Baseline LCOE: $0.2226/kWh*				
Discount Rate (3–12%)	0.1605	0.3161	-27.87	+42.01
Construction Time (1.5–3.0 yr)	0.1687	0.2338	-24.19	+5.03
Solar BOS (×0.8–1.2)	0.2074	0.2377	-6.82	+6.82
Battery Cost Y0 (×0.8–1.2)	0.2103	0.2348	-5.51	+5.51
Solar Module (×0.8–1.2)	0.2154	0.2297	-3.22	+3.22
Solar Fixed O&M (×0.8–1.2)	0.2192	0.2259	-1.51	+1.51
Battery BOS (×0.8–1.2)	0.2197	0.2255	-1.31	+1.31
Battery Cost Y15 (×0.8–1.2)	0.2197	0.2254	-1.28	+1.28
Storage Fixed O&M (×0.8–1.2)	0.2218	0.2233	-0.33	+.33
Land Cost (×0.5–2.0)	0.2224	0.2230	-0.08	+0.08
DC-Coupled Solar+Storage *Baseline LCOE: $0.1852/kWh*				
Discount Rate (3–12%)	0.1348	0.2612	-27.23	+41.04
Construction Time (1.5–3.0 yr)	0.1414	0.1943	-23.66	+4.93
Battery Cost Y0 (×0.8–1.2)	0.1742	0.1961	-5.93	+5.93
Solar BOS (×0.8–1.2)	0.1750	0.1953	-5.50	+5.50
Solar Module (×0.8–1.2)	0.1783	0.1920	-3.69	+3.69
Solar Fixed O&M (×0.8–1.2)	0.1820	0.1878	-1.73	+1.73
Battery BOS (×0.8–1.2)	0.1826	0.1878	-1.40	+1.40
Battery Cost Y15 (×0.8–1.2)	0.1826	0.1878	-1.40	+1.40
Storage Fixed O&M (×0.8–1.2)	0.1845	0.1858	-0.36	+0.36
Land Cost (×0.5–2.0)	0.1850	0.1854	-0.07	+0.07
Natural Gas (100k GPU) *Baseline LCOE: $0.1626/kWh*				
Natural Gas Price ($3.01–12.75/MMBtu)	0.1235	0.2046	-24.03	+25.89
Discount Rate (3–12%)	0.1363	0.2052	-16.14	+26.20
Construction Timeline (2.0–4.8 yr)	0.1353	0.1668	-16.78	+2.59
NG CapEx (×0.8–1.2)	0.1506	0.1746	-7.37	+7.37
NG Fixed O&M (×0.8–1.2)	0.1602	0.1650	-1.48	+1.48
NG Variable O&M (×0.8–1.2)	0.1621	0.1631	-0.31	+0.31

this pattern: easy locations show tight clustering along clear trade-off curves, while challenging locations show both dense clusters (successful convergence) and scattered outliers (suboptimal convergence) incurring 2–3 cents/kWh penalties.

For challenging locations (CV $> 2\%$), running 3–5 independent optimizations and selecting the minimum provides more reliable results. Alternatively, increasing the differential evolution iteration budget from 25 to 50–100 generations improves convergence at modest computational cost for detailed investigation of specific candidate sites.

2. Additional Figures and Tables

Table S2. $\mathrm{LCOE}_{\mathrm{adj}}$ sensitivity analysis including GPU idling costs during construction. Percent changes are relative to each architecture's baseline (shown above result columns). Solar+storage results are scale-independent; NG results shown for typical facility case. Baseline assumptions: Solar construction = 2.0 yr; NG construction = 4.0 yr; discount rate = 7%; NG price = $7.70/MMBtu; GPU spot price = $2.40/chip-hr.

Parameter	Low ($/kWh)	High ($/kWh)	Δ% Low (%)	Δ% High (%)
AC-Coupled Solar+Storage				
Baseline $\mathrm{LCOE}_{\mathrm{adj}}$: $0.6668/kWh				
Discount Rate (3–12%)	0.4467	1.0074	-33.01	+51.08
GPU Idling ($1–4/chip-hr)	0.4077	0.9629	-38.86	+44.41
Construction Time (1.5–3.0 yr)	0.5057	0.9344	-24.17	+40.13
Solar BOS (×0.8–1.2)	0.6516	0.6820	-2.28	+2.28
Battery Cost Y0 (×0.8–1.2)	0.6545	0.6791	-1.84	+1.84
Solar Module (×0.8–1.2)	0.6596	0.6740	-1.07	+1.07
Solar Fixed O&M (×0.8–1.2)	0.6634	0.6701	-0.50	+0.50
Battery BOS (×0.8–1.2)	0.6639	0.6697	-0.44	+0.44
Battery Cost Y15 (×0.8–1.2)	0.6639	0.6696	-0.43	+0.43
Storage Fixed O&M (×0.8–1.2)	0.6661	0.6675	-0.11	+0.11
Land Cost (×0.5–2.0)	0.6667	0.6671	-0.02	+0.04
DC-Coupled Solar+Storage				
Baseline $\mathrm{LCOE}_{\mathrm{adj}}$: $0.6293/kWh				
GPU Idling ($1–4/chip-hr)	0.3702	0.9254	-41.17	+47.05
Discount Rate (3–12%)	0.4209	0.9254	-33.12	+51.34
Construction Time (1.5–3.0 yr)	0.4782	0.8948	-24.01	+42.18
Battery Cost Y0 (×0.8–1.2)	0.6183	0.6403	-1.74	+1.74
Solar BOS (×0.8–1.2)	0.6191	0.6395	-1.62	+1.62
Solar Module (×0.8–1.2)	0.6225	0.6362	-1.09	+1.09
Solar Fixed O&M (×0.8–1.2)	0.6261	0.6325	-0.51	+0.51
Battery BOS (×0.8–1.2)	0.6267	0.6319	-0.41	+0.41
Battery Cost Y15 (×0.8–1.2)	0.6268	0.6319	-0.40	+0.40
Storage Fixed O&M (×0.8–1.2)	0.6287	0.6300	-0.10	+0.10
Land Cost (×0.5–2.0)	0.6292	0.6296	-0.02	+0.04
Natural Gas				
Baseline $\mathrm{LCOE}_{\mathrm{adj}}$: $1.1608/kWh				
GPU Idling ($1–4/chip-hr)	0.5785	1.8263	-50.16	+57.33
Discount Rate (3–12%)	0.7689	1.8144	-33.76	+56.05
Construction Timeline (2.0–4.8 yr)	0.5848	1.4854	-49.62	+27.96
Natural Gas Price ($3.01–12.75/MMBtu)	1.1217	1.2029	-3.36	+3.62
NG CapEx (×0.8–1.2)	1.1488	1.1608	-1.03	+1.03
NG Fixed O&M (×0.8–1.2)	1.1584	1.1632	-0.21	+0.21
NG Variable O&M (×0.8–1.2)	1.1603	1.1613	-0.04	+0.04

Table S3. Technical Parameter Sensitivity Analysis Across Power System Architectures. Values show LCOE impact range (absolute $/kWh and percentage change) for parameter variations. Natural gas results shown for 100,000 GPU baseline (F-class turbines); sensitivity patterns for 24,000 GPU (aeroderivative) and 500,000 GPU (H-class) systems were similar in magnitude but differed in exact values, with AC PSU efficiency ranging +1.03% to -1.58% (24k) and +1.18% to -3.74% (500k), UPS efficiency ranging +1.13% to -6.02% (24k) and +1.30% to -3.86% (500k), and part-load efficiency ranging -0.94% to +0.98% (24k, aero) and -1.22% to +1.29% (500k, H-class).

Parameter (Range)	AC Solar	DC Solar	Natural Gas
	Baseline: $0.2231/kWh	*Baseline: $0.1849/kWh*	*Baseline: $0.1693/kWh*
Battery Performance			
Battery RTE (0.85 – 0.95)	$-\$0.0032$ to $+\$0.0068$ -1.43% to $+3.02\%$	$-\$0.0048$ to $+\$0.0073$ -2.60% to $+3.93\%$	—
Power Electronics			
PSU Efficiency (AC: 0.92–0.96; DC: 0.96–0.99)	$-\$0.0032$ to $+\$0.0040$ -1.43% to $+1.77\%$	$-\$0.0009$ to $+\$0.0039$ -0.50% to $+2.11\%$	$-\$0.0040$ to $+\$0.0019$ -2.37% to $+1.11\%$
UPS Efficiency (0.92 – 0.96)	—	—	$-\$0.0045$ to $+\$0.0020$ -2.64% to $+1.18\%$
Inverter Efficiency (0.94 – 0.98)	$-\$0.0030$ to $+\$0.0021$ -1.35% to $+0.92\%$	—	—
Inverter Load Ratio (1.1 – 1.4)	$-\$0.0008$ to $-\$0.0003$ -0.36% to -0.14%	—	—
Turbine Performance			
Part-Load Efficiency (F-class: 0.10–0.20)	—	—	$-\$0.0018$ to $+\$0.0019$ -1.09% to $+1.15\%$
Temperature Derating (F-class: 0.004–0.012 /℃)	—	—	$+\$0.00$ to $+\$0.00$
Design Parameters			
Design Contingency (1.03 – 1.10)	$-\$0.0007$ to $+\$0.0000$ -0.31% to -0.01%	$-\$0.0003$ to $+\$0.0001$ -0.16% to $+0.04\%$	$+\$0.0000$ to $+\$0.0000$ 0.00% to 0.00%

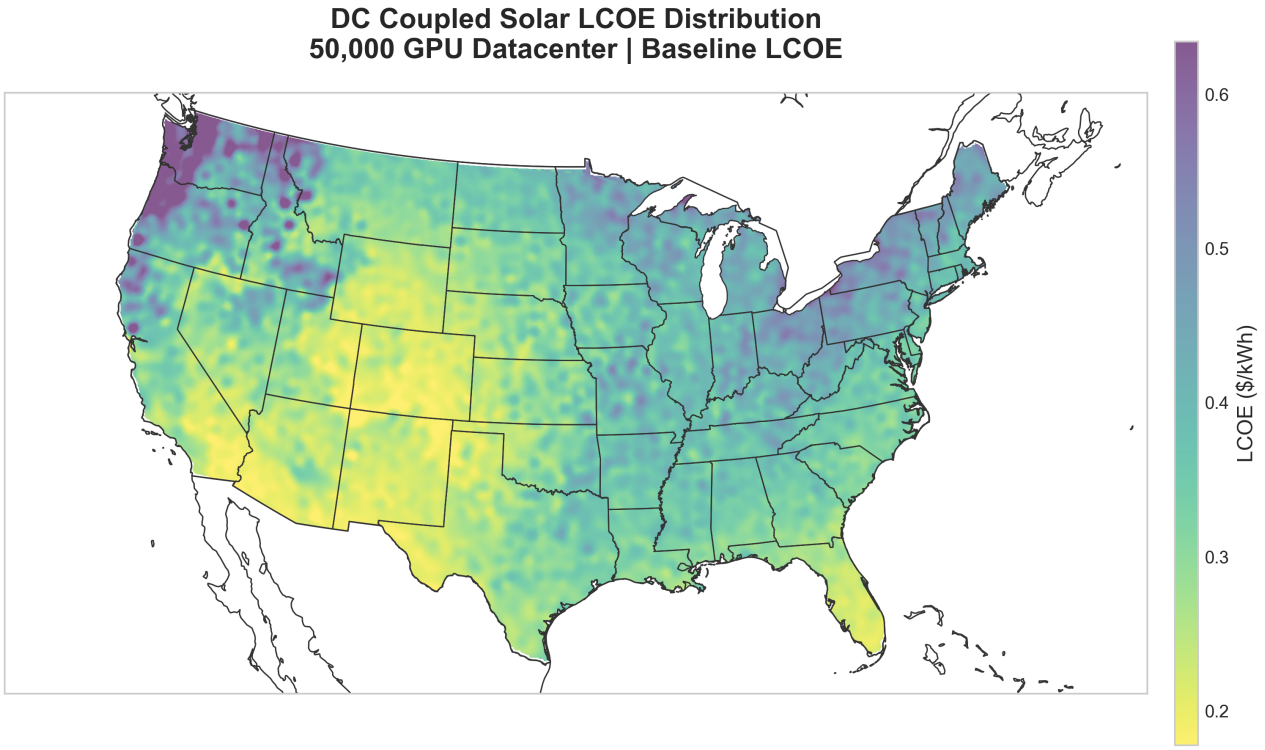

Fig. S4. LCOE for all-DC microgrids across the continental United States. Baseline LCOE ranges from $0.15/kWh in the most consistently sunny regions to $1.10/kWh in the Olympic peninsula of Washington. System cost is more sensitive to worst-case cloudcover then latitutde driven expected irradiance, see the performance delta betwee southeast and southwest Oregon.

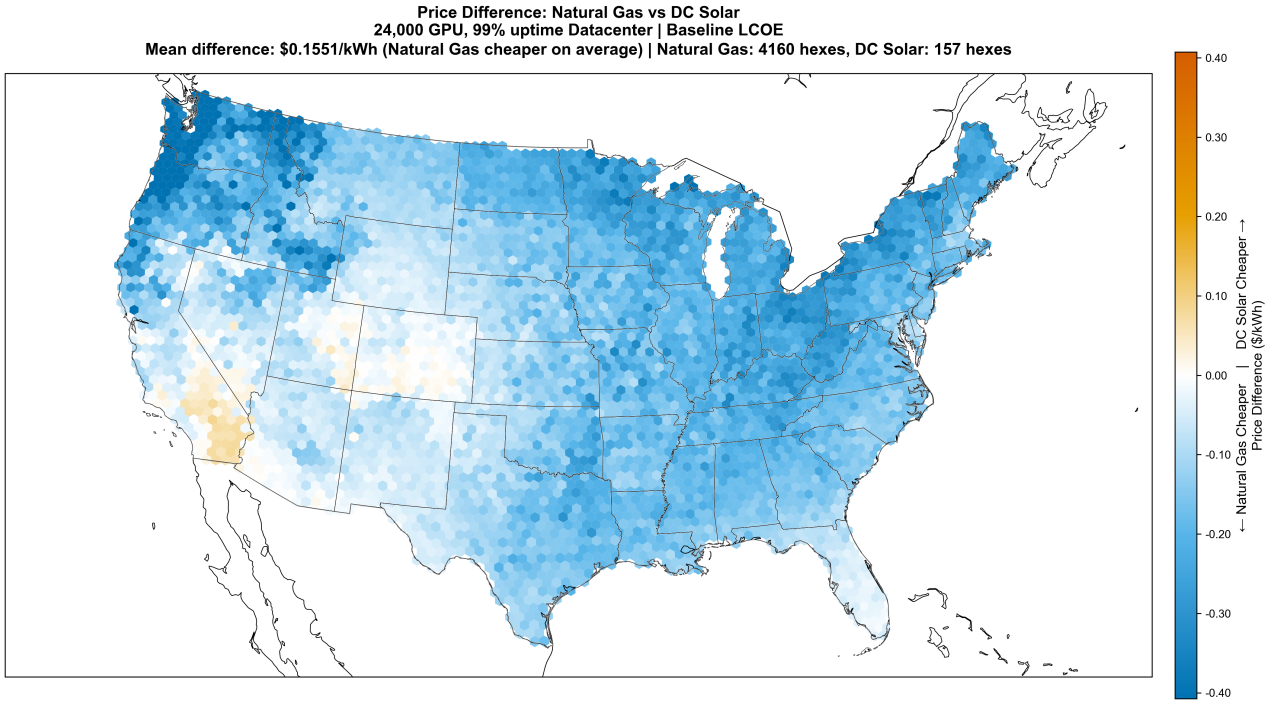

Fig. S5. Comparative LCOE of natural gas and DC Solar microgrid supplying a 24,000 GPU data center across the United States. DC solar achieves or nears cost-parity in high-insolation regions spanning southern California, Arizona, Nevada, New Mexico, and Colorado. In solar poor regions of the pacific northwest and upper midwest, natural gas significantly outperforms DC solar systems. Note the skew in performance geographically: the worst performing locations for solar are disproportionately expensive, while the spread of gas costs between the best and worst tiles is more modest.

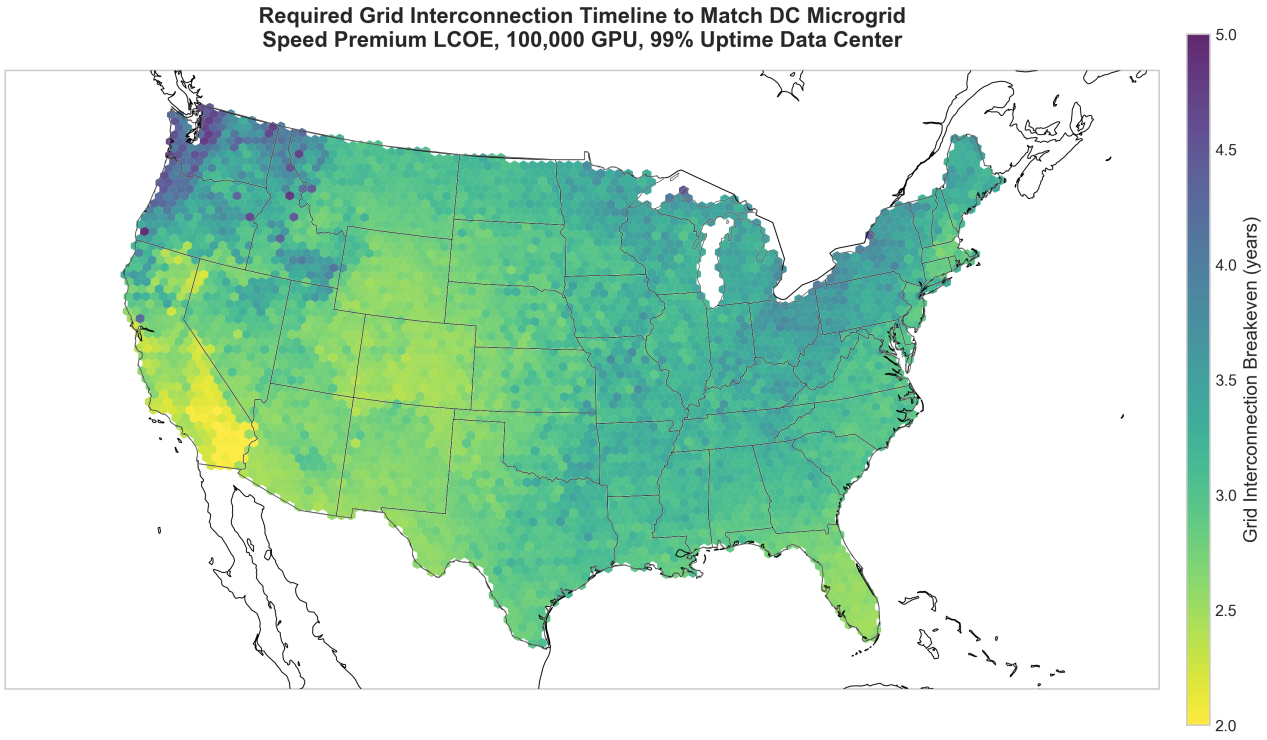

Fig. S6. Required grid interconnection timeline for grid and DC-coupled solar microgrids to achieve equal speed-premium LCOE. For a 100,000 GPU datacenter, holding GPU opportunity cost constant at $2.40/hr. Note this plot displays only the absolute interconnection time required for indifference and is not relative to regional interconnection estimates.

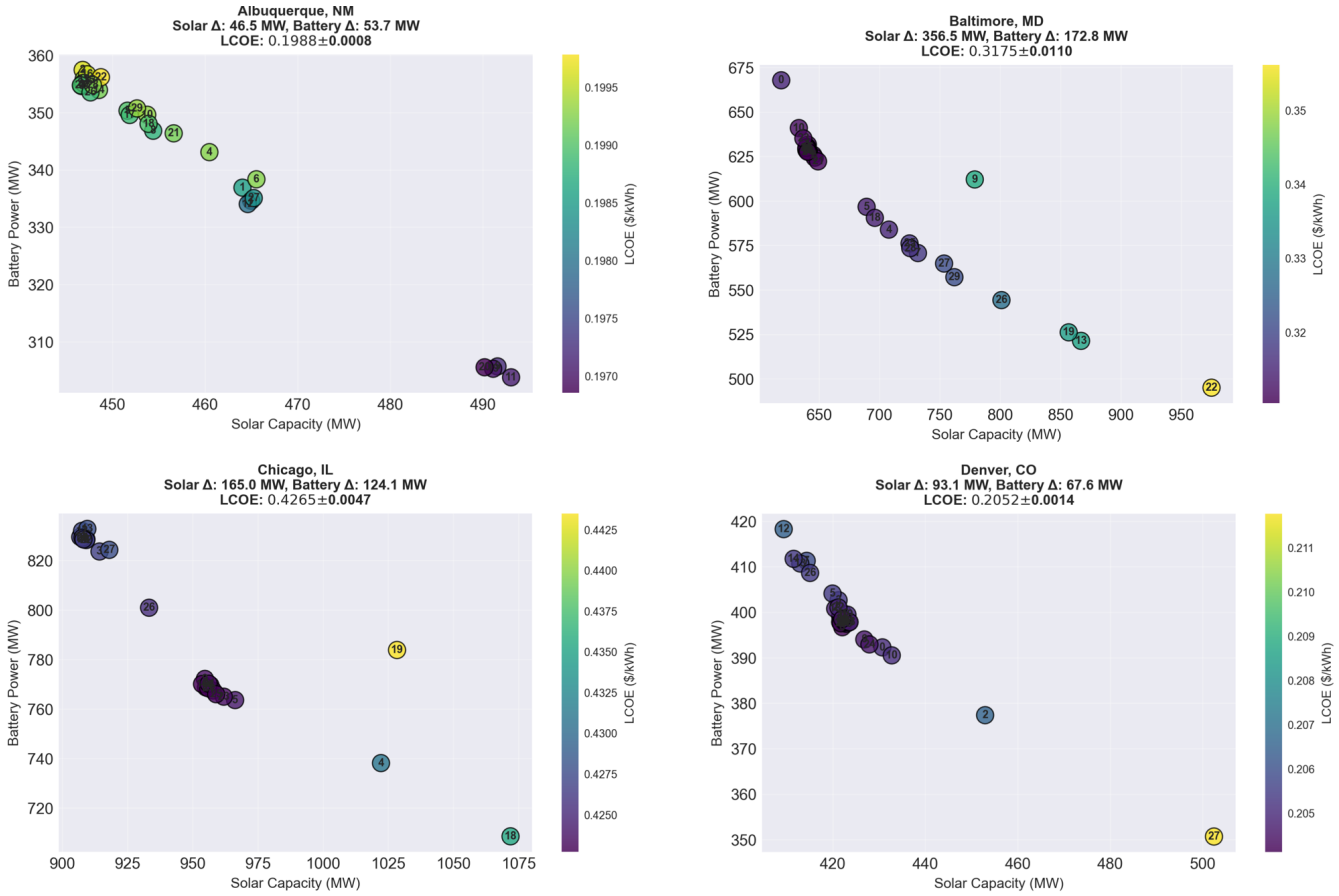

Fig. S7. Optimization stability across representative climate zones. Each point represents an independent optimization run colored by LCOE ($/kWh). Interior solutions indicate outlier runs with suboptimal convergence ($> 2\%$ LCOE penalty). Component sizing variation (5–6% solar, 6.1% battery) substantially exceeds LCOE variation (0.3–3.5%), indicating near-flat objective function regions where multiple configuration achieve nearly identical economic performance. Note the regional variation in optimal design cluster between comparatively battery favoring and compartively PV favoring. *(A)* Albuquerque, NM—high solar resource location achieving tight clustering along clear trade-off curve. *(B)* Baltimore, MD—variable cloud cover exhibiting dispersed convergence patterns with multiple local basins . *(C)* Chicago, IL—pronounced seasonality showing both dense clusters and scattered outliers. *(D)* Denver, CO—high solar resource location demonstrating reliable convergence despite variable resource timing.

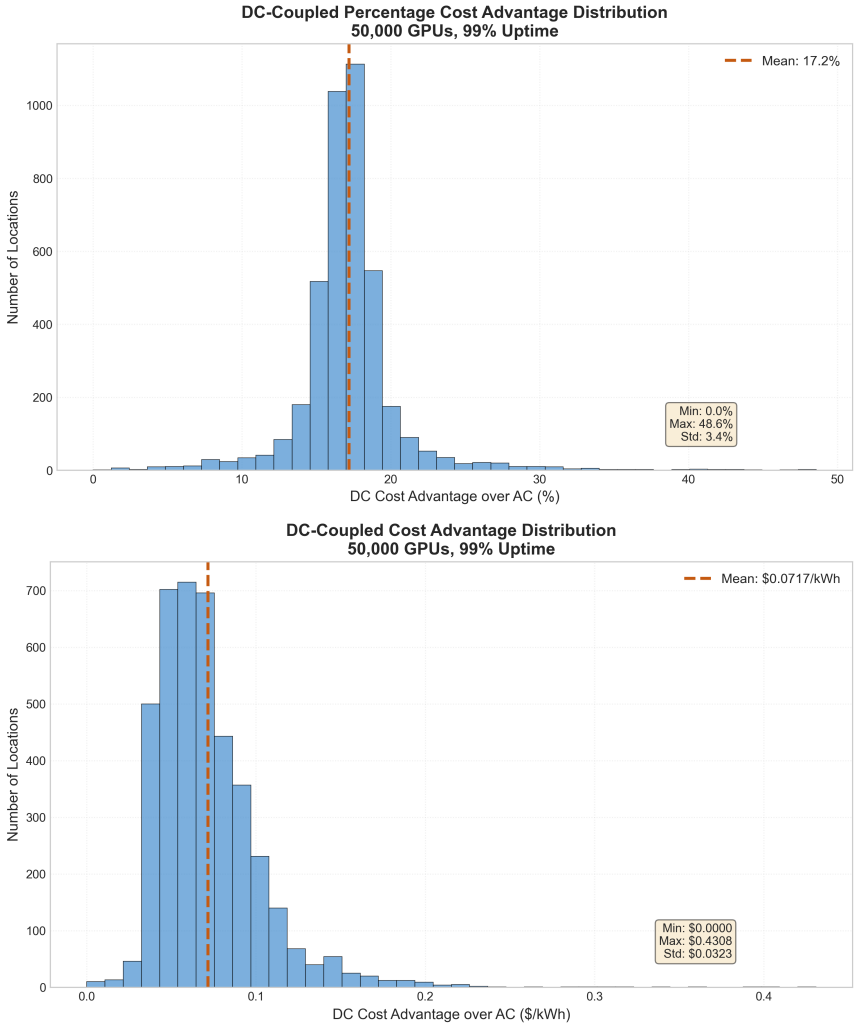

Fig. S8. Cost advantage of DC-coupled systems over AC counterparts for 50,000 GPUs at 99% uptime across all modeled continental U.S. locations. Skew in the absolute distribution reflects variation in sensitivity to solar-to-load and battery-to-load path efficiency by location. Scale of DC advantage generally increases in locations with lower quality solar resources as marginal path efficiency provides returns for highly overbuilt systems. In regions with substantial and consistent solar resources, both AC and all-DC microgrids can scope more aggressively, shrinking but not eliminating the DC cost advantage. *(A)* Percentage cost advantage distribution. Mean DC advantage is 17% of total cost of ownership for a given location. *(B)* Absolute cost advantage distribution. Mean DC advantage is $0.07 / kWh.

Table S4. Natural Gas baseline LCOE parameter sensitivity for the 24,000 and 500,000 GPU cases Modeled facilities are in southeastern Colorado. Percent changes are relative to each scale's baseline. Baseline assumptions: Discount rate = 7%; NG price = $7.70/MMBtu.

Parameter	Low ($/kWh)	High ($/kWh)	$\Delta\%$ Low (%)	$\Delta\%$ High (%)
24,000 GPUs *Baseline* LCOE*: $0.2045/kWh*				
Natural Gas Price ($3.01–12.75/MMBtu)	0.1532	0.2597	-25.08	+27.00
Discount Rate (3–12%)	0.1753	0.2518	-14.28	+23.18
NG CapEx (×0.8–1.2)	0.1929	0.2160	-5.67	+5.67
Turbine Lead Time (×0.7–1.3)	0.1903	0.2105	-6.95	+2.98
NG Construction Time (×0.5–1.2)	0.1959	0.2091	-4.19	+2.25
NG Fixed O&M (×0.8–1.2)	0.2022	0.2067	-1.11	+1.11
NG Variable O&M (×0.8–1.2)	0.2026	0.2063	-0.93	+0.93
500,000 GPUs *Baseline* LCOE*: $0.1347/kWh*				
Natural Gas Price ($3.01–12.75/MMBtu)	0.0977	0.1746	-27.49	+29.60
Discount Rate (3–12%)	0.1161	0.1649	-13.79	+22.41
NG Construction Time (×0.5–1.2)	0.1153	0.1377	-14.38	+2.21
NG CapEx (×0.8–1.2)	0.1270	0.1424	-5.72	+5.72
Turbine Lead Time (×0.7–1.3)	0.1262	0.1387	-6.35	+2.94
NG Fixed O&M (×0.8–1.2)	0.1329	0.1366	-1.38	+1.38
NG Variable O&M (×0.8–1.2)	0.1342	0.1352	-0.37	+0.37

Table S5. Natural Gas $\mathrm{LCOE}_{\mathrm{adj}}$ parameter sensitivity for the 24,000 and 500,000 GPU cases. Modeled facilities are in southeastern Colorado. Percent changes are relative to each scale's baseline. Baseline assumptions: Discount rate = 7%; NG price = $7.70/MMBtu.

Parameter	Low ($/kWh)	High ($/kWh)	$\Delta\%$ Low (%)	$\Delta\%$ High (%)
24,000 GPUs *Baseline $\mathrm{LCOE}_{\mathrm{adj}}$: $1.1882/kWh*				
GPU Idling ($1–4/chip-hr)	0.6143	1.8440	-48.30	+55.20
Discount Rate (3–12%)	0.7987	1.8341	-32.78	+54.36
NG Construction Time (×0.5–1.2)	0.6395	1.5075	-46.18	+26.88
Turbine Lead Time (×0.7–1.3)	0.7930	1.6145	-33.26	+35.88
Natural Gas Price ($3.01–12.75/MMBtu)	1.1369	1.2434	-4.32	+4.65
NG CapEx (×0.8–1.2)	1.1766	1.1998	-0.98	+0.98
NG Fixed O&M (×0.8–1.2)	1.1859	1.1905	-0.19	+0.19
NG Variable O&M (×0.8–1.2)	1.1863	1.1901	-0.16	+0.16
500,000 GPUs *Baseline $\mathrm{LCOE}_{\mathrm{adj}}$: $1.1329/kWh*				
GPU Idling ($1–4/chip-hr)	0.5506	1.7984	-51.40	+58.74
Discount Rate (3–12%)	0.7487	1.7712	-33.92	+56.33
NG Construction Time (×0.5–1.2)	0.5649	1.4563	-50.14	+28.54
Turbine Lead Time (×0.7–1.3)	0.7376	1.5653	-34.90	+38.16
Natural Gas Price ($3.01–12.75/MMBtu)	1.0959	1.1728	-3.27	+3.52
NG CapEx (×0.8–1.2)	1.1252	1.1406	-0.68	+0.68
NG Fixed O&M (×0.8–1.2)	1.1311	1.1348	-0.16	+0.16
NG Variable O&M (×0.8–1.2)	1.1325	1.1334	-0.04	+0.04

Table S6. Economic, Hardware, and Solar PV System Parameters

Parameter	Value	Source
COMPUTE HARDWARE		
GPUs per node	8	DGX H100 standard (65)
Node power (average)	7.3 kW	(5)
Node power (maximum)	8.5 kW	(5)
Design contingency	1.05	5% safety margin
ECONOMIC ASSUMPTIONS		
GPU spot price	\$2.40/GPU-hr	(64)
Discount rate	7%	Base case
Evaluation period	conclusion of year 27	Solar EOL
SOLAR PV SYSTEM		
Capital Costs (Year 0)		
Solar modules	\$402/kW_{DC}	NREL ATB 2024 (49)
Solar BOS (AC-coupled)	\$661.95/kW_{DC}	(50)
Solar BOS (DC-coupled)	\$501.98/kW_{DC}	DC coupled share of (50)
Land	\$150,000/km^2	Estimated average costs
Operations & Maintenance		
Solar fixed O&M	\$16.58/kW_{DC}-yr	(50)
Land intensity	4.17 acres/MW_{DC}	(52, 53)
Degradation		
First year loss	1.0%/yr	(37)
Subsequent years	0.55%/yr	(37)
CONSTRUCTION TIMELINES		
Solar + storage	24 months	(24)
NG simple cycle	36 months	(55)
NG combined cycle	42 months	(55)
Turbine procurement	36 months	OEM estimate
Grid interconnection	State-specific	(63)

Table S7. Battery Storage System Parameters

Parameter	Value	Source
BATTERY STORAGE SYSTEM		
Capital Costs		
Battery cells (Year 0)	$826/kWh	(49)
Battery cells (Year 13)	$508.12/kWh	(49) moderate case
Battery BOS (AC-coupled)	$385.31/kW	(50)
Battery BOS (DC-coupled)	$292.20/kW	Proportional reduction
Operations & Maintenance		
Battery fixed O&M	$4.37/kW-yr	PNNL (54)
Performance & Degradation		
Battery duration	4 hours	(49)
Round-trip efficiency (RTE)	90%	(27)
Replacement year	13	Project mid life
E_{cal} (calendar)	-5.21 kJ/mol	(40)
E_{cyc} (cycling)	$+2.19$ kJ/mol	Inverse Arrhenius (40)
β_0 (calendar exponent)	0.526	BLAST-Lite distillation (40)
α_0 (cycle exponent)	0.828	BLAST-Lite distillation (40)

Table S8. Natural Gas, Diesel, and Gas Turbine Performance Parameters

Parameter	Value	Source
NATURAL GAS TURBINES		
Simple Cycle Reference Plants		
Aero SC (211 MW) CapEx	$1,510/kW	EIA (55)
Aero SC Fixed O&M	$35/kW-yr	(55)
Aero SC Variable O&M	$8/MWh	(55)
F-class SC (233 MW) CapEx	$1,319/kW	NREL ATB (49)
F-class SC Fixed O&M	$25.70/kW-yr	(49)
F-class SC Variable O&M	$6.94/MWh	(49)
Combined Cycle Reference Plants		
F-class CC (727 MW) CapEx	$1,455/kW	(49)
F-class CC Fixed O&M	$33/kW-yr	(49)
F-class CC Variable O&M	$2.10/MWh	(49)
H-class CC (800 MW) CapEx	$1,600/kW	(49)
H-class CC Fixed O&M	$36/kW-yr	(49)
H-class CC Variable O&M	$2.10/MWh	(49)
Cost Scaling		
CapEx exponent	0.67	(56, 57)
Fixed O&M exponent	0.80	(55)
TURBINE DEGRADATION & MAINTENANCE		
Annual Degradation		
Aero capacity loss	0.20%	(38)
Aero efficiency loss	0.15%	(38)
F-class capacity loss	0.15%	(38)
F-class efficiency loss	0.13%	(38)
H-class capacity loss	0.13%	(38)
H-class efficiency loss	0.10%	(38)
Part-Load Efficiency Multiplier		
Aero	$1.0 - 0.10(1 - \beta)$	(32)
F-class	$1.0 - 0.15(1 - \beta)$	(32)
H-class	$1.0 - 0.20(1 - \beta)$	(32)
Temperature Derating (per °C above 15 °C)		
Aero	1.0%	(33)
F-class	0.8%	(33)
H-class	0.7%	(33)
DIESEL BACKUP SYSTEM		
Genset capital cost	$800/kW	(35)
Fuel storage cost	$5/gallon	(36)
Diesel fuel cost	$3.40/gallon	EIA (58)
Diesel specific fuel consumption	13.3 gal/MWh	Manufacturer data avg.
Annual testing requirement	24 hr/yr	Monthly loaded tests

Table S9. Gas Turbine Models Available for System Configuration. Aeroderivatives optimized for simple cycle (SC); F-class suitable for both SC and combined cycle (CC); H-class preferred for CC configurations.

Turbine Model	Capacity (MW)	Efficiency	Availability
Aeroderivatives (Simple Cycle)			
Siemens SGT-A05 KB7S	5.8	0.323	0.92
GE LM2500+G4	35.0	0.390	0.95
Siemens SGT-A45	44.0	0.404	0.94
GE LM6000PC	51.0	0.397	0.94
Siemens SGT-800	62.0	0.411	0.92
P&W FT4000 SwiftPac 70	71.0	0.410	0.95
GE LMS100	100.0	0.395	0.93
F-Class Industrial (Simple & Combined Cycle)			
GE 6F.03	88.0	0.368	0.93
Mitsubishi M501F	185.4	0.370	0.91
GE 7F.04	202.0	0.375	0.93
Siemens SGT6-5000F	260.0	0.400	0.93
H-Class (Combined Cycle Preferred)			
GE 7HA.01	290.0	0.420	0.92
Siemens SGT-8000H	310.0	0.404	0.91
GE 7HA.02	384.0	0.426	0.91
GE 9HA.01	448.0	0.429	0.91
GE 9HA.02	571.0	0.440	0.91

Table S10. State-Level Grid and Fuel Parameters (2022 USD). Electricity and gas prices from EIA 2022 industrial rates (59, 60). Interconnection queue data from LBNL 2018–2023 median (63).

State	ISO/RTO	Queue (yr)	Elec (¢/kWh)	Gas ($/MMBtu)	Region
Connecticut	ISO-NE	2.5	15.20	7.87	New England
Maine	ISO-NE	2.5	12.01	9.47	
Massachusetts	ISO-NE	2.5	16.79	12.04	
New Hampshire	ISO-NE	2.5	15.39	12.02	
Rhode Island	ISO-NE	2.5	16.42	11.57	
Vermont	ISO-NE	2.5	11.02	5.16	
New Jersey	PJM	3.75	11.03	8.35	Mid-Atlantic
New York	NYISO	4.5	7.82	9.43	
Pennsylvania	PJM	3.75	7.26	10.67	
Illinois	PJM/MISO	3.75	7.88	6.79	E N Central
Indiana	MISO	3.75	7.49	6.17	
Michigan	MISO	3.75	7.52	7.70	
Ohio	PJM	3.75	6.39	7.02	
Wisconsin	MISO	3.75	7.53	6.22	
Iowa	MISO	3.75	5.63	6.40	W N Central
Kansas	SPP	4.0	7.02	5.04	
Minnesota	MISO	3.75	8.46	5.60	
Missouri	SPP/MISO	4.0	7.10	8.69	
Nebraska	SPP	4.0	6.67	5.54	
North Dakota	MISO/SPP	4.0	6.98	2.98	
South Dakota	SPP	4.0	7.63	5.57	
Delaware	PJM	3.75	7.92	12.43	S Atlantic
Florida	Non-ISO	3.0	8.09	6.50	
Georgia	Non-ISO	3.0	5.41	4.86	
Maryland	PJM	3.75	8.75	12.07	
North Carolina	PJM/Non-ISO	3.5	7.26	6.08	
South Carolina	Non-ISO	3.0	6.01	4.68	
Virginia	PJM	3.75	8.26	4.58	
West Virginia	PJM	3.75	7.19	3.85	
Alabama	Non-ISO	3.0	6.73	4.11	E S Central
Kentucky	PJM/MISO	3.75	6.12	4.07	
Mississippi	MISO	3.75	6.38	5.04	
Tennessee	Non-ISO/MISO	3.5	5.99	4.99	
Arkansas	SPP/MISO	4.0	5.85	8.57	W S Central
Louisiana	MISO/SPP	4.0	5.98	3.49	
Oklahoma	SPP	4.0	4.68	3.01	
Texas	ERCOT	3.0	5.48	3.24	
Arizona	Non-ISO	4.0	6.84	6.29	Mountain
Colorado	Non-ISO	4.0	7.75	7.70	
Idaho	Non-ISO	4.0	6.15	6.03	
Montana	Non-ISO	4.0	6.67	7.14	
Nevada	CAISO/Non-ISO	4.5	7.45	10.64	
New Mexico	SPP/Non-ISO	4.0	5.39	4.58	
Utah	Non-ISO	4.0	6.11	9.06	
Wyoming	Non-ISO	4.0	7.26	5.92	
California	CAISO	4.75	18.50	12.75	Pacific
Oregon	Non-ISO	4.0	7.18	7.23	
Washington	Non-ISO	4.0	5.87	10.32	
Alaska	Non-ISO	4.0	20.09	5.92	
Hawaii	Non-ISO	4.0	33.71	25.82	
US Average		4.0	7.25	4.11	National Average

References

1. P Gasper, N Prakash, K Smith, NREL/BLAST-Lite (2025) original-date: 2023-04-06T19:51:44Z.
2. B James, offgridai-us/cost-calculator (2025) original-date: 2024-12-27T15:48:36Z.
3. JG Koomey, Estimating Total Power Consumption by Servers in the US and the World. (2007).
4. A Shehabi, et al., 2024 United States Data Center Energy Usage Report, National Lab Report (2024).
5. AC Newkirk, et al., Empirically-Calibrated H100 Node Power Models for Reducing Uncertainty in AI Training Energy Estimation (2025) arXiv:2506.14551 [cs].
6. T Norris, T Profeta, D Patino-Echeverri, A Cowie-Haskell, Rethinking load growth: assessing the potential for integration of large flexible loads in US power systems. (2025) Publisher: Nicholas Institute for Energy, Environment & Sustainability.
7. P Colangelo, et al., Turning AI Data Centers into Grid-Interactive Assets: Results from a Field Demonstration in Phoenix, Arizona (2025) arXiv:2507.00909 [cs].
8. N Lei, E Masanet, Climate- and technology-specific PUE and WUE estimations for U.S. data centers using a hybrid statistical and thermodynamics-based approach. *Resour. Conserv. Recycl.* **182**, 106323 (2022).
9. A Narayanan, Q Wang, S Ozguc, RW Bonner, Investigation of Server Level Direct-To-Chip Two Phase Cooling Solution for High Power GPUs. (American Society of Mechanical Engineers Digital Collection), (2024).
10. R Kong, et al., Enhancing data center cooling efficiency and ability: A comprehensive review of direct liquid cooling technologies. *Energy* **308**, 132846 (2024).
11. D Beberide, D Regany, J Camarasa, J Barrau, M Vilarrubí, Performance assessment of an advanced direct-to-chip liquid cooling solution in real conditions inside a data centre in *2024 30th International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*. pp. 1–5 (2024) ISSN: 2474-1523.
12. SMHy say, Meta details AI data center redesign that led to facilities being scrapped (2023).
13. M Connatser, Nvidia CEO admits next gen DGX systems necessitate liquid cooling - and the new systems are coming soon (2024).
14. K Cho, H Chang, Y Jung, Y Yoon, Economic analysis of data center cooling strategies. *Sustain. Cities Soc.* **31**, 234–243 (2017).
15. M Beheshti, Wide-bandgap semiconductors: Performance and benefits of GaN versus SiC. *Analog. Des. J.* (2020).
16. VF Pires, A Pires, A Cordeiro, DC Microgrids: Benefits, Architectures, Perspectives and Challenges. *Energies* **16**, 1217 (2023) Publisher: Multidisciplinary Digital Publishing Institute.
17. S Chalise, et al., Data center energy systems: Current technology and future direction in *2015 IEEE Power & Energy Society General Meeting*. pp. 1–5 (2015) ISSN: 1932-5517.
18. LA Barroso, U Hölzle, P Ranganathan, *The datacenter as a computer: Designing warehouse-scale machines*. (Springer Nature), (2019).
19. W Tschudi, B Fortenbery, M Ton, DC power for improved data center efficiency. *Lawrence Berkeley Natl. Lab.* **10** (2008) Publisher: Citeseer.
20. D Sartor, et al., USER GUIDE FOR IMPLEMENTING ECBC IN DATA CENTERS, (Lawrence Berkeley National Lab), Technical report (2021).
21. DL Gerber, B Nordman, R Brown, J Poon, Cost analysis of distributed storage in AC and DC microgrids. *Appl. Energy* **344**, 121218 (2023).
22. DL Gerber, et al., A Comprehensive Loss Model and Comparison of AC and DC Boost Converters. *Energies* **14**, 3131 (2021) Publisher: Multidisciplinary Digital Publishing Institute.
23. DL Gerber, et al., A simulation-based efficiency comparison of AC and DC power distribution networks in commercial buildings. *Appl. Energy* **210**, 1167–1187 (2018).
24. K Baranko, D Cambell, Z Hausfather, J McWalter, N Ransohoff, Fast, scalable, clean, and cheap enough, Technical report (2024).
25. V Ramasamy, et al., U.S. Solar Photovoltaic System and Energy Storage Cost Benchmarks, With Minimum Sustainable Price Analysis: Q1 2022, (National Renewable Energy Laboratory), Technical Report NREL/TP-7A40-83586September 2022 (2022).
26. JB Goodenough, KS Park, The Li-Ion Rechargeable Battery: A Perspective. *J. Am. Chem. Soc.* **135**, 1167–1176 (2013) Publisher: American Chemical Society.
27. M Rehm, et al., Comparing the electrical performance of commercial sodium-ion and lithium-iron-phosphate batteries. *J. Power Sources* **633**, 236290 (2025).
28. P Gasper, et al., Degradation and modeling of large-format commercial lithium-ion cells as a function of chemistry, design, and aging conditions. *J. Energy Storage* **73**, 109042 (2023).
29. AJ Crawford, et al., Lifecycle comparison of selected Li-ion battery chemistries under grid and electric vehicle duty cycle combinations. *J. Power Sources* **380**, 185–193 (2018).
30. R Xiong, P Wang, Y Jia, W Shen, F Sun, Multi-factor aging in Lithium Iron phosphate batteries: Mechanisms and insights. *Appl. Energy* **382**, 125250 (2025).
31. K Darrow, R Tidball, J Wang, A Hampson, Catalogue of CHP Technologies, (U.S. Environmental Protection Agency), Technical report (2017).
32. JM Muñoz de Escalona, D Sánchez, R Chacartegui, T Sánchez, Part-load analysis of gas turbine & ORC combined cycles. *Appl. Therm. Eng.* **36**, 63–72 (2012).

33. A De Sa, S Al Zubaidy, Gas turbine performance at varying ambient temperature. *Appl. Therm. Eng.* **31**, 2735–2739 (2011).
34. V Ganapathy, Heat-recovery steam generators: Understand the basics. *Chem. Eng. Prog.* **92** (1996).
35. SJ Ericson, DR Olis, A Comparison of Fuel Choice for Backup Generators, Technical Report NREL/TP–6A50-72509, 1505554 (2019).
36. A Mott, Best Practices for Standby Generator Operations and Maintenance (2021) Publisher: Pacific Northwest National Laboratory.
37. H Yang, Y Yin, A Abu-Siada, A Comprehensive Review of Solar Panel Performance Degradation and Adaptive Mitigation Strategies. *IET Control. Theory & Appl.* **19**, e70040 (2025) _eprint: https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/cth2.70040.
38. R Kurz, K Brun, Degradation in Gas Turbine Systems in *Volume 2: Coal, Biomass and Alternative Fuels; Combustion and Fuels; Oil and Gas Applications; Cycle Innovations*. (American Society of Mechanical Engineers, Munich, Germany), p. V002T03A003 (2000).
39. HR Morin, JF Whitacre, J Michalek, Quantifying the degradation cost of frequent fast charging across multiple electric vehicle battery chemistries. *J. Power Sources* **652**, 237552 (2025).
40. K Smith, P Gasper, AM Colclasure, Y Shimonishi, S Yoshida, Lithium-Ion Battery Life Model with Electrode Cracking and Early-Life Break-in Processes. *J. The Electrochem. Soc.* **168**, 100530 (2021) Publisher: The Electrochemical Society.
41. D Reber, SR Jarvis, MP Marshak, The role of energy density for grid-scale batteries (2022).
42. Y Tang, T Li, X Cheng, Review of Specific Heat Capacity Determination of Lithium-Ion Battery. *Energy Procedia* **158**, 4967–4973 (2019).
43. S Neupane, M Alipanah, D Barnes, X Li, Heat Generation Characteristics of LiFePO4 Pouch Cells with Passive Thermal Management. *Energies* **11**, 1243 (2018) Publisher: Multidisciplinary Digital Publishing Institute.
44. MV Morganti, S Longo, M Tirovic, CY Blaise, G Forostovsky, Multi-Scale, Electro-Thermal Model of NMC Battery Cell. *IEEE Transactions on Veh. Technol.* **68**, 10594–10606 (2019).
45. VN Lam, et al., A decade of insights: Delving into calendar aging trends and implications. *Joule* **9**, 101796 (2025).
46. SB Peterson, J Apt, JF Whitacre, Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization. *J. Power Sources* **195**, 2385–2392 (2010).
47. PL Joskow, Comparing the Costs of Intermittent and Dispatchable Electricity Generating Technologies. *Am. Econ. Rev.* **101**, 238–241 (2011).
48. F Ueckerdt, L Hirth, G Luderer, O Edenhofer, System LCOE: What are the costs of variable renewables? *Energy* **63**, 61–75 (2013).
49. B Mirletz, et al., Annual Technology Baseline: The 2024 Electricity Update, (National Renewable Energy Laboratory (NREL), Golden, CO (United States)), Technical Report NREL/PR-7A40-89960 (2024).
50. V Ramasamy, et al., U.S. Solar Photovoltaic System and Energy Storage Cost Benchmarks, With Minimum Sustainable Price Analysis: Q1 2023, (National Renewable Energy Laboratory), Technical report (2023).
51. V Ramasamy, D Feldman, J Desai, R Margolis, U.S. Solar Photovoltaic System and Energy Storage Cost Benchmarks: Q1 2021, Technical Report NREL/TP-7A40-80694, 1829460, MainId:77478 (2021).
52. S Ong, C Campbell, P Denholm, R Margolis, G Heath, Land-Use Requirements for Solar Power Plants in the United States, (National Renewable Energy Laboratory (NREL), Golden, CO (United States)), Technical Report NREL/TP–6A20-56290 (2013).
53. M Bolinger, G Bolinger, Land Requirements for Utility-Scale PV: An Empirical Update on Power and Energy Density. *IEEE J. Photovoltaics* **12**, 589–594 (2022).
54. Pacific Northwest National Lab, ESGC Cost and Performance Database: Lithium-ion Battery (LFP and NMC) (2024) Publisher: Pacific Northwest National Laboratory.
55. A LaRose, J Diefenderfer, R Bowers, N Vincent, Capital Cost and Performance Characteristics for Utility-Scale Electric Power Generating Technologies, (Energy Information Administration), Technical report (2024).
56. MA Tribe, RLW Alpine, Scale economies and the "0.6 rule". *Eng. Costs Prod. Econ.* **10**, 271–278 (1986).
57. R James, S Leptinsky, M Turner, Quality Guidelines for Energy System Studies - Capital Cost Scaling Methodology: Revision 4a Report, Technical Report DOE/NETL–2022/3340, 1893821 (2022).
58. Gasoline and Diesel Fuel Update (2025).
59. Natural Gas Industrial Price (2025).
60. Electric Power Monthly - U.S. Energy Information Administration (EIA) (2025).
61. T Schmitt, et al., Cost and Performance Baseline for Fossil Energy Plants Volume 1: Bituminous Coal and Natural Gas to Electricity, (National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States)), Technical Report DOE/NETL-2023/4320 (2022).
62. P Gerke, NextEra CEO warns against scorning renewable generation amidst long lead times for gas and nuclear development (2025).
63. J Rand, et al., Queued Up: 2024 Edition, Characteristics of Power Plants Seeking Transmission Interconnection As of the End of 2023. (2024).
64. Bloomberg L.P., SDH100 Price Index (2025) Accessed via Bloomberg Terminal.
65. Nvidia, NVIDIA DGX H100/H200 User Guide (year?).