



Decision-Aware Pre-Feasibility Screening for Branch-Level Battery Storage Investment Under Spatiotemporal Grid Stress: Evidence from the Tokyo Transmission Network

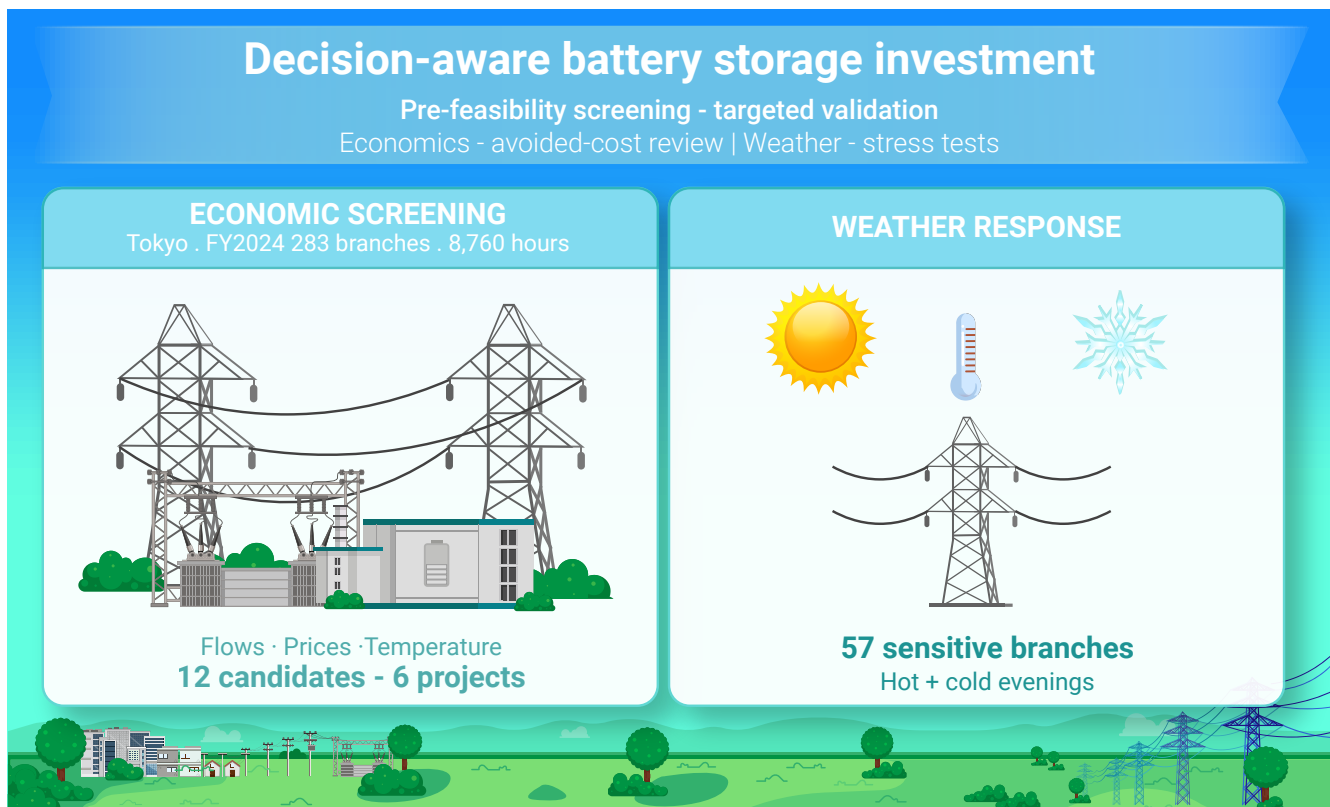
Yanxue Li,^{1,*} Tianxiao Huang,¹ and Weijun Gao²

*Correspondence: liyanxue@qut.edu.cn

Received: September 7, 2029; Accepted: September 22, 2026; Published Online: September 23, 2026; <https://doi.org/10.59717/ipj.energy-use.2026.100062>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Public grid, electricity-price and temperature records support early screening of battery storage investments in Tokyo.
- Across 283 transmission branches, the model selects six two-hour battery projects totalling 1.33 GW within a JPY 200 billion budget.
- Assumed savings from postponing grid upgrades account for 86% of modelled annual value before costs and determine project profitability.
- Fifty-seven branches show strong flow responses to hot and cold evenings; none coincides with the six selected battery locations.
- Economic and weather assessments identify distinct planning priorities, guiding targeted engineering studies and verification of potential savings.



INNOVATION
— PRESS —

Energy Use

Energy Use 3 (2026):100062

Contents lists available at INNOVATION PRESS

Decision-Aware Pre-Feasibility Screening for Branch-Level Battery Storage Investment Under Spatiotemporal Grid Stress: Evidence from the Tokyo Transmission Network

Yanxue Li,^{1,*} Tianxiao Huang,¹ and Weijun Gao²

¹Innovation Institute for Sustainable Maritime Architecture Research and Technology, Qingdao University of Technology, Qingdao 266033 China

²Faculty of Environmental Engineering, The University of Kitakyushu, Fukuoka 808-0135, Japan

*Correspondence: liyanxue@qut.edu.cn

Received: September 7, 2029; Accepted: September 22, 2026; Published Online: September 23, 2026; <https://doi.org/10.59717/ipj.energy-use.2026.100062>

© 2026 The Author(s). This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Li Y., Huang T. and Gao W. (2026). Decision-Aware Pre-Feasibility Screening for Branch-Level Battery Storage Investment Under Spatiotemporal Grid Stress: Evidence from the Tokyo Transmission Network. *Energy Use* 2:100062.

Public transmission-flow records reveal recurring operational stress, but their value for storage investment depends on a transparent link between observed flows, economic decisions, and subsequent engineering validation. This study develops a decision-aware spatiotemporal framework for branch-level battery energy storage system (BESS) pre-feasibility screening in the Tokyo transmission network. Fiscal-year 2024 TEPCO branch flows are aligned with Tokyo-area day-ahead prices and temperature observations across 283 unique branches and 8,760 hours. A composite stress screen identifies 12 candidates; branch-level linear programs co-optimize BESS power, energy, dispatch, and empirical peak reduction; and a binary portfolio model selects projects under a JPY 200 billion budget. A separate 20:00 weather model estimates branch-specific heating and cooling responses. The selected portfolio contains six two-hour projects totalling 1,328.5 MW/2,657 MWh and JPY 199.3 billion of capital, with a conditional 15-year net present value of JPY 208.0 billion. Valuation depends chiefly on the peak-deferral credit, which supplies 86.1% of gross annual operating value, and the fixed-design break-even value is JPY 14,913/kW-year. Without the credit, annualised net benefit falls to -JPY 19.8 billion/year; at the base value of JPY 30,000/kW-year it reaches JPY 20.0 billion/year. Fifty-seven of 283 branches exhibit robust two-sided temperature responses, the all-branch pressure-temperature rank correlation is weak, and none of the selected projects passes the weather screen. The framework separates economic peak-shaving candidates from weather-responsive corridors and routes each toward an appropriate validation pathway.

INTRODUCTION

Battery energy storage systems (BESSs) are becoming central assets in power systems with increasing shares of variable renewable generation, because fast response, modularity, and intertemporal flexibility allow them to support energy arbitrage, peak shaving, congestion management, and reserve provision. Their realised value is nevertheless application-specific: stationary lithium-ion system design is shaped by the intended grid service,¹ and no universally preferred siting or sizing method exists because objectives, connection locations, network models, and uncertainty treatments differ.² Capacity-expansion research shows that storage can support deep decarbonisation, with modelled system value depending on duration, operating constraints, competing flexibility, and deployment scale,³ while long-run studies attribute value to avoided generation and transmission capacity as well as operating flexibility.⁴ For long-duration storage, cost-effective design

depends on power cost, energy-capacity cost, efficiency, and discharge duration,⁵ and multi-decade hourly analysis shows that energy-capacity cost and reliability assumptions govern the cost of shaping variable generation.⁶ Market-oriented analyses show that pure arbitrage value is highly sensitive to local price volatility, efficiency, and market feedbacks and is rarely sufficient alone,⁷ and system-model sensitivities show that storage requirements depend on the cost and availability of competing flexibility.⁸ Storage planning is therefore a spatiotemporal decision problem: a technically attractive project delivers limited value if placed at the wrong branch, sized for the wrong duration, or credited for a service the evidence cannot verify.

Transmission- and distribution-level planning studies have developed detailed formulations for storage siting, sizing, and operation, coupling reserve provision, flexibility services, congestion relief, and reliability enhancement to explicit network conditions,⁹⁻¹² while market-oriented work shows how arbitrage, service stacking, and shared-storage arrangements shape realised revenue.¹³⁻¹⁶ Across this literature, storage value is conditional on some representation of line ratings, admittance, power-transfer sensitivities, nodal prices, outage states, or reliability criteria. Publicly accessible records may provide topology and measured branch flows but omit verified ratings, impedance parameters, switching states, contingency definitions, and outage consequences, so measured excursions support empirical stress screening whereas thermal-overload diagnosis and network-wide redispatch require verified electrical models. The methodological challenge is to convert observed flows into planning evidence without implying unverified electrical performance.

Decision-aware research offers a principle for this incomplete-information setting: upstream evidence should be judged by the downstream decision it can responsibly support. Chung et al. linked prediction to constrained medicine allocation and assessed the induced access decisions rather than predictive accuracy;¹⁷ Zheng et al. represented urban planning as sequential intervention on an irregular spatial graph;¹⁸ and Wessel et al. showed that regional coordination changes both the cost and the spatial allocation of low-carbon infrastructure under uncertainty.¹⁹ The same principle is formalised in the learning literature by task-based learning through stochastic optimisation,²⁰ decision-focused learning through an optimisation layer,²¹ and smart predict-then-optimize, which shows that lower predictive error need not imply better decisions;²² recent benchmarking²³ and surveys²⁴ document the problem-dependent trade-offs of embedding solvers inside learning systems. Here, decision awareness denotes a narrower planning principle: screening variables, candidate generation, portfolio selection, and uncertainty reporting are designed around the investment decision and remain traceable to it.

Weather dependence adds a second, potentially non-coincident dimension of transmission exposure. Empirical studies show nonlinear temperature–electricity relationships with substantial spatial heterogeneity, including Tokyo-scale evidence,^{25,26} and contemporary load-forecasting studies combine weather-ensemble information, semiparametric temperature-response structures, and probabilistic evaluation to represent nonlinear weather effects.^{27–29} Branch-level power flow nevertheless captures broader system operation than local demand alone: a measured corridor flow reflects spatial demand, generation dispatch, interregional transfers, network switching, and market conditions. A statistically significant temperature–flow relationship therefore evidences weather-responsive operational exposure, whereas causal attribution to temperature-controlled load or physical vulnerability requires additional data. Whether such corridors coincide with economically attractive short-duration BESS projects is a distinct planning question.

Open modelling frameworks clarify the analytical boundary of a public-data study. PyPSA provides a transparent framework for network-constrained operation and investment analysis,³⁰ PyPSA-Eur assembles public topology, generation, demand, and weather data into a reproducible transmission-level model,³¹ and synthetic-grid research offers validation criteria for public test systems when operational networks are confidential.³² Topology and branch-flow history therefore support a pre-feasibility ranking conditional on empirical proxies, whereas a security-constrained reinforcement decision requires verified electrical parameters, ratings, and operating states. Temporal scope matters equally: short-lived, system-defining events that determine back-up requirements differ from annual indicators that may appear benign,³³ and fine-scale coordination among generation, demand, storage, and transmission can matter more than aggregate capacity or mean resource quality.³⁴ A one-year, single-station branch study thus evaluates recurrent daily flexibility and indicates event-scale stress, whereas multi-year adequacy assessment requires broader data and models.

Four gaps therefore limit the direct use of public operational data for storage screening. Network-constrained BESS models presuppose known electrical limits and sensitivities, and little attention has been paid to preserving decision usefulness when those quantities are unavailable. Public-data valuation still needs an auditable separation between revenues supported by observed prices and benefits that depend on an assumed avoided reinforcement cost. Resource adequacy and resilience require multi-year weather, outage, and recovery information, so short-term peak reduction measures routine flexibility rather than performance during rare disturbances. Finally, temperature–electricity studies usually address aggregate demand, and limited evidence exists on whether branch-level weather-responsive corridors coincide with economically preferred BESS locations.

This study therefore asks how incomplete public transmission data can inform storage-planning decisions before a load-flow-ready network model is available, through three questions. (i) Which transmission branches exhibit the strongest empirical spatiotemporal stress when peak severity, exceedance energy, ramping burden, and sustained utilisation are considered jointly? (ii) When BESS power, energy, dispatch, arbitrage revenue, and empirical peak-deferral value are co-optimised, which projects form the highest-value portfolio under a binding capital budget? (iii) To what extent do economically preferred locations coincide with branches showing robust heating- and cooling-related flow responses, and what does their convergence or divergence imply for subsequent network, adequacy, and resilience assessment?

To answer these questions we develop a decision-aware screening framework using hourly fiscal-year 2024 flows from a TEPCO 500/275/154 kV research network, Tokyo-area day-ahead prices, and Tokyo temperature observations. The framework ranks 283 unique branches by complementary indicators of empirical stress, co-optimises BESS power, energy, dispatch, and post-storage empirical peak on the 12 leading candidates, and selects investments under a JPY 200 billion capital budget; in parallel it estimates branch-specific evening temperature responses with calendar controls, Newey–West inference, false-discovery-rate correction, and cross-validation. The decision unit is a measured branch storage project, whereas the weather model uses branch-day observations as its inference unit. The contributions are an auditable bridge from public operational data to storage pre-feasibility

decisions; an integration of stress screening with investment–operation optimisation that separates price-supported arbitrage from assumption-dependent peak-deferral value; a statistically controlled branch-level weather-response screen; and the treatment of agreement or disagreement between the two priorities as an analytical result that routes branches to distinct validation pathways.

MATERIALS AND METHODS

Research design, data sources, and preprocessing

Figure 1 summarises a staged, decision-aware design that keeps data preparation, candidate generation, project optimisation, evidence integration, and downstream validation analytically distinct. Hourly inputs are aligned to a common fiscal-year clock, branch flows are cleaned and deduplicated, and peak severity, exceedance persistence, ramping burden, and sustained utilisation are combined into a transparent pressure screen. The 12 highest-scoring branches define the candidate set for independent investment–operation linear programmes, and the resulting fixed designs enter a budget-constrained binary portfolio model evaluated through value decomposition and two sensitivity modules. In parallel, the weather stream samples branch absolute flow at 20:00 and estimates piecewise heating and cooling associations under statistical controls. The two outputs are then matched by branch identifier, compared through overlap counts and rank correlations, and routed toward economic, seasonal, joint, or no-escalation review.

The flow input covers 1 April 2024 00:00 to 31 March 2025 23:00 Japan Standard Time, matching Japanese fiscal year 2024, and contains 8,760 hourly records across 472 raw line columns. The topology snapshot represents TEPCO's 500, 275, and 154 kV transmission corridors and georeferenced facilities in an author-derived research dataset. Day-ahead spot prices are the Japan Electric Power Exchange (JEPX) fiscal-year 2024 Tokyo-area spot summary, and hourly air temperature is the Japan Meteorological Agency (JMA) record for Tokyo station 47662 (35.6917° N, 139.7500° E); access links are given under Data availability. The three inputs share a common fiscal-year clock and retain their respective spatial resolutions: branch-level flows, an area-wide price, and one station-level temperature series (Table S1).

Notation is uniform throughout the methods: i indexes eligible branches, t hours ($T = 8,760$, $\Delta t = 1$ h) and d days; operators and descriptive sub- and superscripts are upright; flows and power are in MW, energy in MWh and prices in JPY/kWh; the factor 10^3 converts MW to kW and MWh to kWh; and $f_0 = 1$ MW is a unit reference that keeps logarithmic arguments dimensionless. For each raw branch column, nonnumeric values were set to missing and the initial missing fraction was recorded. A branch was excluded if more than 10% of its observations were missing, if all valid observations were zero, or if its maximum absolute flow was below 20 MW. The branch-specific 95th and 99th percentiles of absolute flow, $q_{95,i}$ and $q_{99,i}$, define the scale parameter g_i and the two-sided isolated-outlier set in Equation (1), and observations in that set were set to missing. Remaining gaps were filled by linear interpolation along the hourly index, with the nearest valid observation carried into terminal gaps. The completeness test uses the missing fraction recorded before imputation, so interpolation does not relax the exclusion rule.

$$g_i = \max \{q_{99,i} - q_{95,i}, 0.02 \max(q_{99,i}, f_0)\}, \quad (1)$$

$$\mathcal{O}_i = \{t \in \mathcal{T} : |f_{i,t}| > \max(q_{99,i} + 10g_i, 3 \max(q_{99,i}, f_0))\}.$$

Columns sharing the same parsed line identifier were treated as duplicate measurements of one investment opportunity. The representative column minimised the sum of its original missing and detected outlier counts, with ties broken in favour of the larger absolute-flow P99; nonrepresentative duplicates were removed from analysis. This deterministic rule reduced the 472 raw columns to 283 eligible, unique line representatives.

Branch-stress screening and candidate generation

Four screening metrics were calculated from hourly branch flow: exceedance energy above the absolute-flow P95 (E^{exc} , in MWh), peak gap above P95 (G , in MW), the P99 absolute hourly flow change (R , in MW), and mean absolute flow (M , in MW), as defined in Equation (2). Equation (3) gives the unit-referenced logarithmic transform, the min–max normalisation across eligible branches, and the prespecified pressure score.

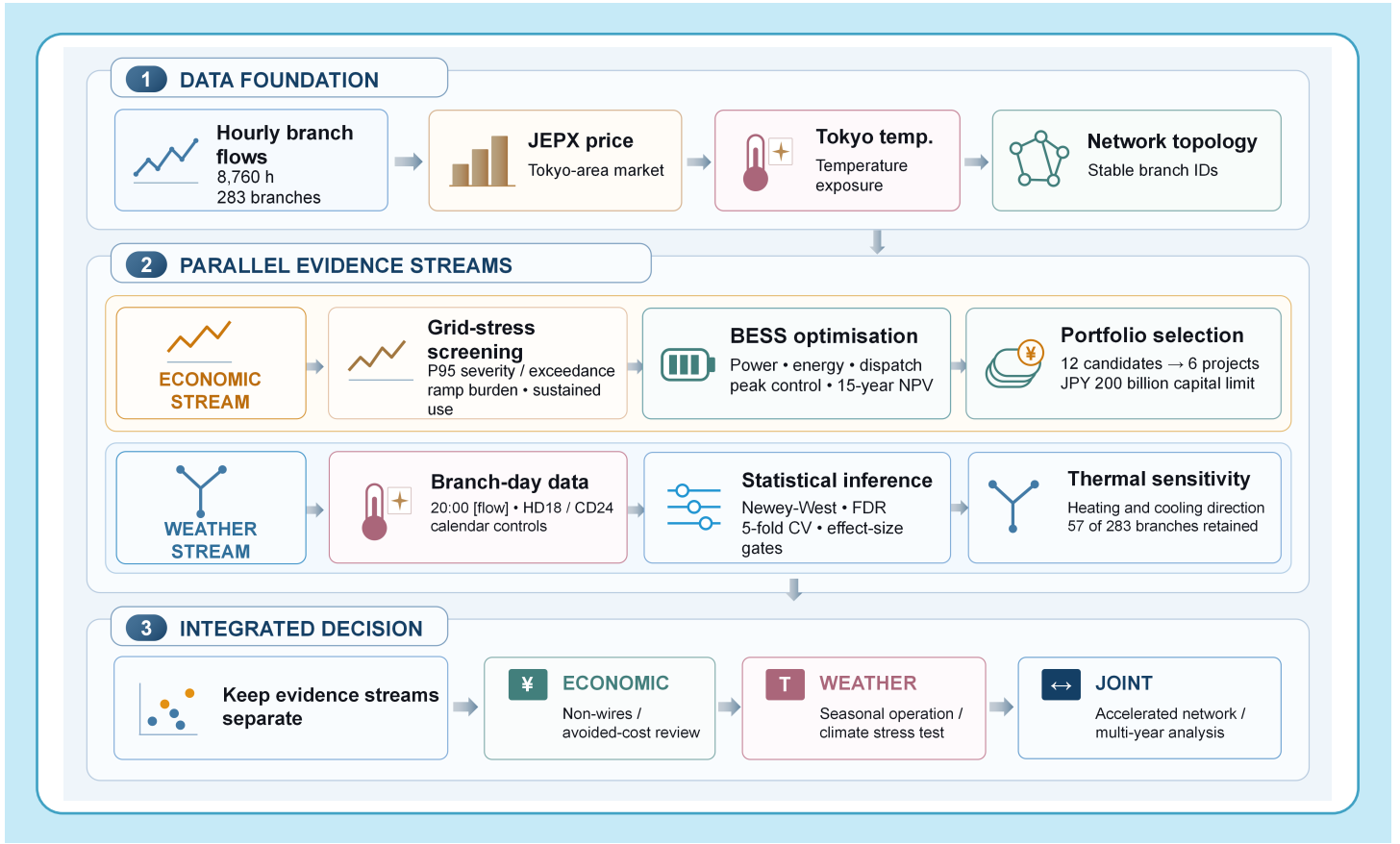


Figure 1. Research workflow.

$$E_i^{\text{exc}} = \Delta t \sum_{t \in \mathcal{T}} \max(|f_{i,t}| - q_{95,i}, 0), G_i = \max_{t \in \mathcal{T}} |f_{i,t}| - q_{95,i},$$

$$R_i = Q_{0.99}(\{|f_{i,t} - f_{i,t-1}|\}_{t=2}^T), M_i = \frac{1}{T} \sum_{t \in \mathcal{T}} |f_{i,t}|. \quad (2)$$

$$\tilde{X}_i = \frac{\ln(1 + X_i/X_0) - \min_{j \in \mathcal{N}} \ln(1 + X_j/X_0)}{\max_{j \in \mathcal{N}} \ln(1 + X_j/X_0) - \min_{j \in \mathcal{N}} \ln(1 + X_j/X_0)}, \quad X \in \{E^{\text{exc}}, G, R, M\}, \quad (3)$$

$$S_i^{\text{pres}} = 0.40\tilde{E}_i^{\text{exc}} + 0.30\tilde{G}_i + 0.20\tilde{R}_i + 0.10\tilde{M}_i \in [0, 1], \quad \forall i \in \mathcal{N}.$$

In Equation (3), X_0 is the unit reference of each metric, 1 MWh for E^{exc} and 1 MW for G , R and M , so the logarithm acts on a dimensionless argument, and the four weights sum to one, so S^{pres} lies between zero and one. The weights are prespecified planning priorities rather than estimated coefficients, emphasising the magnitude and persistence of the empirical peak burden while retaining information on ramping and sustained utilisation. Branches are sorted by pressure score with peak absolute flow as the tie-breaker, and the highest 12 define the candidate set before any storage economics are calculated. The directionality ratio D in Equation (4) measures persistence of the observed flow sign and is retained as a diagnostic and weather-screening gate, but is excluded from the pressure score.

$$D_i = \frac{\left| \sum_{t \in \mathcal{T}} f_{i,t} \right|}{\sum_{t \in \mathcal{T}} |f_{i,t}|} \in [0, 1], \quad \forall i \in \mathcal{N}. \quad (4)$$

BESS investment–operation model and portfolio selection

For each candidate branch, a deterministic annual linear programme minimises equivalent annual cost under perfect foresight of the 8,760 observed hourly flows and prices. Equation (5) states the objective, the decision set Ω – hourly charging power c , discharging power d , state of charge e , BESS power P and energy capacity E , and the post-storage annual absolute-flow peak z – and the constraint set given by Equations (6)–(8). Here c_p and c_e are the unit power and energy capital costs, π is the hourly JEPX price, c^{deg}

the throughput degradation cost, f^{M} the fixed O&M rate, v^{pk} the unit peak-deferral value, and the capital-recovery factor a converts upfront investment into an equivalent annual charge at discount rate r over asset life L .

$$\min_{\Omega_i} J_i = 10^3 \Delta t \sum_{t \in \mathcal{T}} [(\pi_t + c^{\text{deg}}) c_{i,t} + (c^{\text{deg}} - \pi_t) d_{i,t}]$$

$$+ 10^3 (a + f^{\text{M}}) (c_p P_i + c_e E_i) + 10^3 v^{\text{pk}} z_i, \quad (5)$$

s.t. Eqs. (6)–(8), $\Omega_i = \{c_{i,t}, d_{i,t}, e_{i,t} (t \in \mathcal{T}), P_i, E_i, z_i\}$,

$$a = \frac{r(1+r)^L}{(1+r)^L - 1}.$$

The modelled peak reduction is the difference between the pre-storage annual absolute-flow maximum and the optimised post-storage peak; because the pre-storage maximum is fixed, crediting this reduction at 30,000 JPY/kW-year is algebraically equivalent to penalising the optimised post-storage peak: the final term of Equation (5) and the negative peak-deferral value of Equation (9) differ by a constant. The peak-deferral value is therefore reported separately from market arbitrage and treated as a declared avoided-reinforcement assumption requiring verification against applicable tariffs, TEPCO payments, or branch-level avoided costs. Equation (6) defines adjusted branch flow, the post-storage absolute-flow bound, and the charging and discharging power limits; Equation (7) gives state-of-charge dynamics, bounds, and the cyclic boundary condition, with charging and discharging efficiencies η_c and η_d .

$$\hat{f}_{i,t} = f_{i,t} + c_{i,t} - d_{i,t}, \quad (6)$$

$$-z_i \leq \hat{f}_{i,t} \leq z_i, \quad 0 \leq c_{i,t} \leq P_i, \quad 0 \leq d_{i,t} \leq P_i, \quad \forall t \in \mathcal{T}.$$

$$e_{i,t} = e_{i,t-1} + \eta_c c_{i,t} \Delta t - \frac{d_{i,t} \Delta t}{\eta_d}, \quad \forall t \in \mathcal{T}, \quad (7)$$

$$0 \leq e_{i,t} \leq E_i, \quad \forall t \in \mathcal{T}, \quad e_{i,0} = e_{i,T}.$$

Symmetric efficiencies $\eta_c = \eta_d = \sqrt{0.90}$ give a round-trip efficiency of 0.90, and the cyclic boundary in Equation (7) links the final hour to the first so that the annual model receives no free initial or terminal energy. Equation (8)

imposes the duration range $\tau_{\min} = 2$ h to $\tau_{\max} = 6$ h, the candidate-specific power ceiling, and the post-storage peak bounds. The continuous formulation permits simultaneous charging and discharging, which efficiency losses and a positive throughput degradation cost render economically unattractive.

$$\begin{aligned} \tau_{\min} P_i &\leq E_i \leq \tau_{\max} P_i, \\ 0 &\leq P_i \leq P_i^{\max} = \min\{400 \text{ MW}, 0.05 F_i^{\max}, F_i^{\max} - q_{95,i}\}, \\ q_{95,i} &\leq z_i \leq F_i^{\max}, \quad F_i^{\max} = \max_{t \in \mathcal{T}} |f_{i,t}|. \end{aligned} \quad (8)$$

Equation (9) defines upfront investment I , annual JEPX arbitrage revenue A , annual peak-deferral value V , degradation cost C^{deg} , fixed operation and maintenance cost C^{OM} , equivalent full-cycle count N^{cyc} , and annualised net benefit NB; Equation (10) converts the recurring annual net operating cash flow and initial investment to project NPV through the present-worth factor κ_L , the reciprocal of the capital-recovery factor. Base-case parameter values are listed in Table S2.

$$\begin{aligned} I_i &= 10^3 (c_p P_i + c_E E_i), A_i = 10^3 \Delta t \sum_{t \in \mathcal{T}} \pi_t (d_{i,t} - c_{i,t}), \\ V_i &= 10^3 v^{\text{pk}} (F_i^{\max} - z_i), C_i^{\text{deg}} = 10^3 c^{\text{deg}} \Delta t \sum_{t \in \mathcal{T}} (c_{i,t} + d_{i,t}), \\ C_i^{\text{OM}} &= F_i^{\text{OM}} I_i, N_i^{\text{cyc}} = \frac{\Delta t \sum_{t \in \mathcal{T}} d_{i,t}}{E_i}, \\ \text{NB}_i &= A_i + V_i - C_i^{\text{deg}} - C_i^{\text{OM}} - a_i. \end{aligned} \quad (9)$$

$$\text{NPV}_i = -I_i + (A_i + V_i - C_i^{\text{deg}} - C_i^{\text{OM}}) \kappa_L, \quad \kappa_L = \frac{1 - (1+r)^{-L}}{r} = \frac{1}{a}. \quad (10)$$

Equation (11) states portfolio selection: the eligible set collects candidate projects whose optimised power is at least 0.10 MW, whose base-case NPV is positive, and whose individual upfront capital does not exceed the budget, and one binary variable per project maximises the sum of stand-alone project NPVs subject to total upfront investment no greater than $B = \text{JPY } 200$ billion.

$$\begin{aligned} \max_{\{x_i\}} \quad & \sum_{i \in \mathcal{C}} \text{NPV}_i x_i, \\ \text{s.t.} \quad & \sum_{i \in \mathcal{C}} I_i x_i \leq B, \quad x_i \in \{0, 1\}, \quad \forall i \in \mathcal{C}, \\ & \mathcal{C} = \{i \in \mathcal{K} : P_i \geq 0.10 \text{ MW}, \text{NPV}_i > 0, I_i \leq B\}. \end{aligned} \quad (11)$$

The additive objective combines independently optimised projects under the capital budget; joint redispatch, price feedback, competition for common network headroom, and PTDF-mediated interactions remain inputs to a subsequent security-constrained expansion analysis. Table S2 summarises the base-case assumptions.

Temperature-sensitive branch modelling and evidence gates

The temperature analysis uses one observation per day at 20:00 to reduce PV-related confounding, producing 365 branch-day records for each of the 283 retained branches. Absolute flow, denoted y , is modelled because monitored direction can reverse while the attribution question concerns branch-utilisation magnitude. Equation (12) defines the controlled piecewise model, in which the heating and cooling degree terms HD and CD are formed from the 20:00 air temperature θ and the balance points $\theta_H = 18$ °C and $\theta_C = 24$ °C.

$$\begin{aligned} y_{i,d} &\equiv |f_{i,d,20}| = \alpha_i + \beta_{H,i} \text{HD}_d + \beta_{C,i} \text{CD}_d + \gamma_i^T \mathbf{w}_d + \varepsilon_{i,d}, \quad \forall d \in \mathcal{D}, \\ \text{HD}_d &= \max(\theta_d - \theta_H, 0), \quad \text{CD}_d = \max(\theta_d - \theta_C, 0). \end{aligned} \quad (12)$$

The control vector $\mathbf{w}$ contains an intercept, month and day-of-week indicators, and a linear time trend, γ is the corresponding coefficient vector, ε the disturbance, and the separated heating and cooling balance points avoid imposing a symmetric response. A baseline model contains only the control vector, and Equation (13) defines the in-sample and cross-validated temperature partial coefficients of determination.

$$\begin{aligned} R_{\text{part},i}^2 &= 1 - \frac{\text{SSE}_{\text{part},i}^{\text{full}}}{\text{SSE}_{\text{part},i}^{\text{base}}}, R_{\text{CV},i}^2 = 1 - \frac{\text{SSE}_{\text{CV},i}^{\text{full}}}{\text{SSE}_{\text{CV},i}^{\text{base}}}, \\ \text{SSE}_i^m &= \sum_{d \in \mathcal{D}} (y_{i,d} - \hat{y}_{i,d}^m)^2, m \in \{\text{full}, \text{base}\}. \end{aligned} \quad (13)$$

Heteroskedasticity- and autocorrelation-robust (Newey–West) standard errors with a seven-day lag account for residual serial dependence, a joint test evaluates the null hypothesis that the heating and cooling slopes are both zero, and Benjamini–Hochberg adjustment controls the false-discovery rate across the 283 simultaneous branch tests. Deterministic interleaved five-fold cross-validation compares held-out squared error from the full temperature model with that from the calendar-only baseline; this design evaluates within-year held-out fit, whereas temporally blocked validation is required for out-of-year forecast skill.

Equation (14) evaluates the isolated temperature component at the cold and hot evening references $\theta_{\text{cold}} = 5$ °C and $\theta_{\text{hot}} = 30$ °C and normalises each effect by mean 20:00 absolute flow; a raw model without calendar controls is also estimated, and reversal of either slope sign after adjustment is recorded as a Simpson-direction diagnostic. Equation (15) defines the temperature-evidence score S^{temp} and the strong-evidence set, whose prespecified gates combine the joint FDR-adjusted p-value, the in-sample and cross-validated temperature partial R^2 , the maximum evening effect, directionality, and mean 20:00 absolute flow; branches passing these gates are ordered by the score, and the thresholds support statistical screening, whereas causal estimates, thermal ratings, and vulnerability limits require separate validation.

$$\begin{aligned} \delta_{H,i} &= \frac{\max(\beta_{H,i}(\theta_H - \theta_{\text{cold}}), 0)}{\bar{y}_i}, \delta_{C,i} = \frac{\max(\beta_{C,i}(\theta_{\text{hot}} - \theta_C), 0)}{\bar{y}_i}, \\ \delta_{\text{max},i} &= \max(\delta_{H,i}, \delta_{C,i}), \bar{y}_i = \frac{1}{|\mathcal{D}|} \sum_{d \in \mathcal{D}} y_{i,d}. \end{aligned} \quad (14)$$

$$\begin{aligned} S_i^{\text{temp}} &= \sqrt{R_{\text{part},i}^2 \max(R_{\text{CV},i}^2, 0) \delta_{\text{max},i} \log_{10}(1 + \bar{y}_i/f_0)}, \\ \mathcal{S} &= \{i \in \mathcal{N} : q_i^{\text{FDR}} \leq 0.05, R_{\text{part},i}^2 \geq 0.25, R_{\text{CV},i}^2 \geq 0.10, \delta_{\text{max},i} \geq 0.20, \\ & D_i \geq 0.80, \bar{y}_i \geq 50 \text{ MW}, \exists k \in \{H, C\} : \beta_{k,i} > 0 \text{ and } p_{k,i} \leq 0.05\}. \end{aligned} \quad (15)$$

Sensitivity analysis, cross-evidence alignment, and validation boundary

Economic sensitivities are calculated after solving the base-case models, with storage power, energy, dispatch, and portfolio membership held fixed. For each project and unit peak-deferral value v in JPY/kW-year, fixed-design annualised net benefit is recomputed from Equation (9) with the peak credit evaluated at v , v is evaluated from 0 to 60,000 JPY/kW-year in increments of 250 JPY/kW-year, and break-even values are the points at which the reconstructed net benefit equals zero. A second fixed-design grid varies storage-capex and JEPX price-spread multipliers independently over {0.70, 0.85, 1.00, 1.15, 1.30}; the capex multiplier changes upfront investment, annualised capex, and fixed O&M, and the price-spread multiplier scales base arbitrage revenue, while peak-deferral value remains fixed at 30,000 JPY/kW-year. Both modules are deterministic accounting decompositions, and no scenario probabilities are assigned.

Candidate investment results are merged with branch pressure and temperature evidence by stable source index. Association between the two bounded composite scores, S^{pres} and S^{temp} , is summarised by Spearman rank correlation for all 283 branches and, descriptively, for the 12 candidates and six selected projects, with set overlap reported separately. To quantify how the assumed peak-deferral value changes economic alignment, annualised profit is reconstructed at each value step for the 12 fixed candidate designs and the six selected projects, and Spearman correlations between profit and each evidence score are computed within each scenario. Because both evidence scores are invariant to v , the underlying pressure–temperature correlation is fixed across scenarios.

Integrated outputs are interpreted through a three-layer evidence matrix: observed exposure (branch peak, exceedance, ramp, sustained flow, and controlled 20:00 temperature response); actionable flexibility (optimised BESS power, energy, dispatch, net benefit, NPV, and portfolio membership); and validation escalation, specifying the evidence required before a candidate is assigned engineering-security, resource-adequacy, or resilience status. Because the public data contain no verified ratings, admittance model, tap states, outage schedules, PTDF, or LODF, historical P95 serves only as an empirical peak-shaving floor and storage is represented with a one-to-one effect on the monitored branch flow. Both streams use within-year observations, so inter-annual variability, forced outages, and restoration are evaluated downstream.

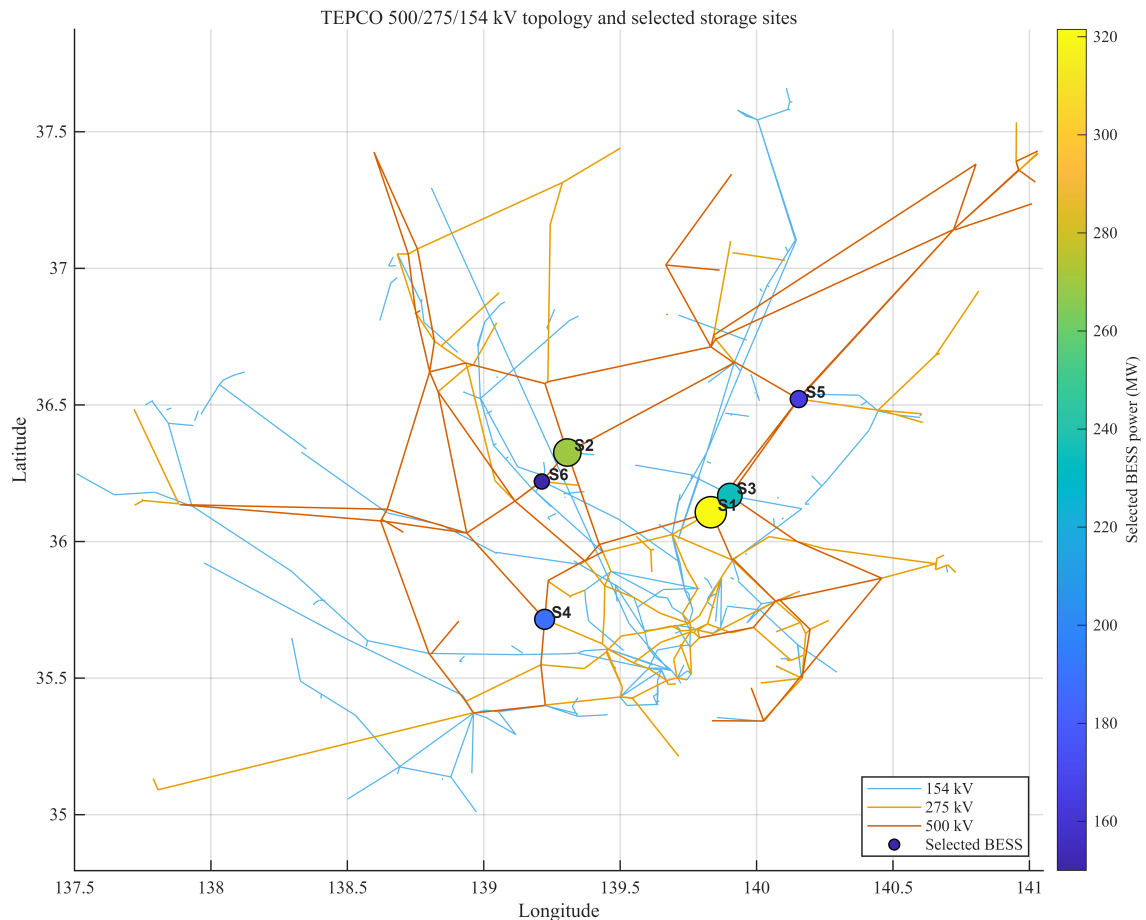


Figure 2. TEPCO 500/275/154 kV research topology and selected storage sites Marker size and color represent optimized BESS power. Site identifiers map to [Table S3](#).

RESULTS

Branch-stress screening and candidate generation

The three hourly inputs align over 8,760 timestamps. The JEPX Tokyo-area hourly price has a minimum of 0.010, mean of 13.658, median of 13.190, P95 of 20.730, and maximum of 43.575 JPY/kWh. The fiscal-year temperature series ranges from -0.9°C to 37.1°C (mean 17.62°C); at 20:00 it ranges from 0.6°C to 32.3°C and averages 17.43°C , and the only interpolated temperature observation falls outside the 20:00 sample, so all 365 branch-day observations are directly measured. Flow preprocessing interpolated 4,633 missing cells and replaced 71 isolated outliers; of 472 raw columns, nine were excluded by the quality rules and 180 were nonrepresentative duplicates. The analysis base therefore comprises 283 unique branches, 8,760 hourly flow and price observations, and 365 branch-day observations at 20:00.

The composite branch-pressure score reduces the 283 eligible branches to 12 investment candidates. The four-component metrics produce non-equivalent rankings: some candidates combine a large instantaneous peak gap with high hourly ramps, whereas others accumulate more energy above their empirical P95 or maintain higher absolute flow, so candidate generation retains information on all four dimensions rather than reproducing an annual-maximum ranking. All 12 candidate linear programmes converged and satis-

fied the eligibility filters, so the budget model could select among the complete candidate set using project value and capital requirements rather than the stress rank alone.

Storage sizing, portfolio selection, and dispatch

Under the JPY 200 billion budget, six of the 12 eligible projects are selected ([Table S3](#) and [Figure 2](#)), corresponding to pressure-screen ranks 3, 6, 7, 10, 11, and 12 and demonstrating that empirical stress and investment value produce different orders. Total installed power is 1,328.5 MW, total energy is 2,657 MWh, and capital expenditure is JPY 199.275 billion, equivalent to 99.64% of the budget. The six designs S1–S6 are listed in [Table S3](#), ranging from 321.5 MW/643 MWh at Shin-Koga Substation – Shin-Koga Line to 150 MW/300 MWh at Shin-Okabe Substation – Shin-Okabe Line, each on circuits 1 and 2, and the mapped sites span the northern and western portions of the research topology.

Each selected project reaches two design bounds: power equals 5% of the original annual absolute-flow peak, giving an aggregate modelled reduction of 1,328.5 MW, and energy equals the two-hour lower-duration bound. Equivalent cycling ranges from 274.0 to 275.0 full cycles/year across the six sites, indicating an intraday flexibility duty cycle.

Duration is not uniformly bound across the candidate set, and this distinc-

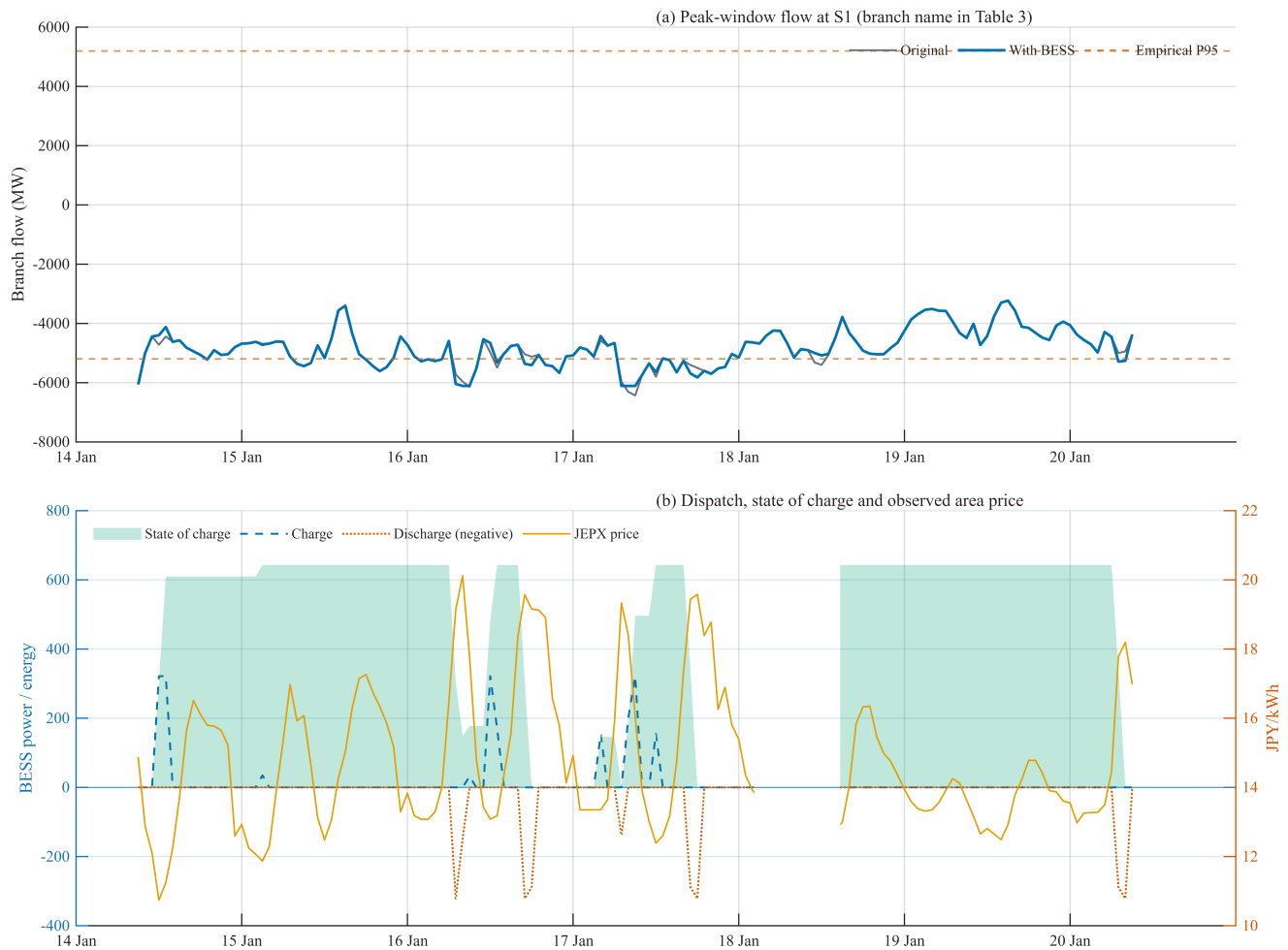


Figure 3. Peak-window dispatch for the first selected project (A) Original and post-storage branch flow relative to the empirical absolute-flow P95. (B) Charge, discharge, state of charge, and observed JEPX Tokyo-area price.

tion governs portfolio membership. Nine of the 12 candidates settle at the two-hour floor, but pressure-screen ranks 1, 2, and 8 take interior durations of 2.88, 2.71, and 2.81 h because shaving 5% of their annual peak requires sustained rather than instantaneous discharge. The additional energy capital reduces their ratio of stand-alone NPV to upfront capital to 0.620, 0.682, and 0.642, against 1.0436–1.0446 for the nine two-hour designs. Under a binding budget the portfolio model therefore removes all three interior-duration projects before comparing the remainder, so the divergence between stress rank and investment rank originates in exceedance duration rather than peak magnitude.

Among the nine two-hour designs the capital-efficiency ratio varies by only 0.1%, so the choice of six from nine is settled by how closely a subset fills the budget rather than by project merit: the selected portfolio exceeds the next-best feasible combination by only 0.18% of portfolio NPV, and candidates 4, 5, and 9 are displaced on this basis even though candidate 5 carries the highest capital-efficiency ratio. The portfolio is therefore one of several near-equivalent capital-packing solutions.

The optimised dispatch charges during lower-price intervals and discharges in higher-price or branch-peak intervals while holding adjusted absolute flow below the optimised peak. In the annual extreme-flow window of Figure 3, storage trims the measured branch peak while maintaining the cyclic state-of-charge boundary; because price and peak terms are co-optimised, the schedule differs from price-only arbitrage.

Economic performance, value composition, and sensitivity

Across the six projects, gross annual operating value is JPY 46.306 billion, of which modelled peak-deferral value contributes JPY 39.855 billion/year (86.1%) and JEPX arbitrage JPY 6.451 billion/year (13.9%). Annual degradation plus fixed O&M totals JPY 7.064 billion and annualised capex JPY 19.199 billion/year, producing a portfolio annualised net benefit of JPY 20.043 billion/year, a conditional 15-year NPV of JPY 208.041 billion, and an aggregate simple payback of 5.08 years.

Annualised net benefit is positive at all six sites and decreases from JPY 4.849 billion/year at S1 to JPY 2.264 billion/year at S6, with S2–S5 at JPY 4.065, 3.547, 2.843, and 2.475 billion/year. At every site, peak-deferral value is the largest positive component and annualised capex the largest negative component (Figure 4), and the near-proportional decline in benefit with installed power is consistent with the similar cycling intensities and common two-hour design. Because the assumed credit supplies 86.1% of gross operating value, these base-case values are conditional on independently verifying applicable TEPCO remuneration or branch-level avoided cost.

Peak-deferral value is accordingly the dominant source of profitability uncertainty. With designs and dispatch held fixed, eliminating the credit makes every selected project loss-making, from –JPY 4.796 billion/year for S1 to –JPY 2.236 billion/year for S6, and the portfolio result falls to –JPY 19.812 billion/year. Project break-even values range from 14,904.7 to 14,918.4 JPY/kW-year and the portfolio breaks even at 14,913.0 JPY/kW-

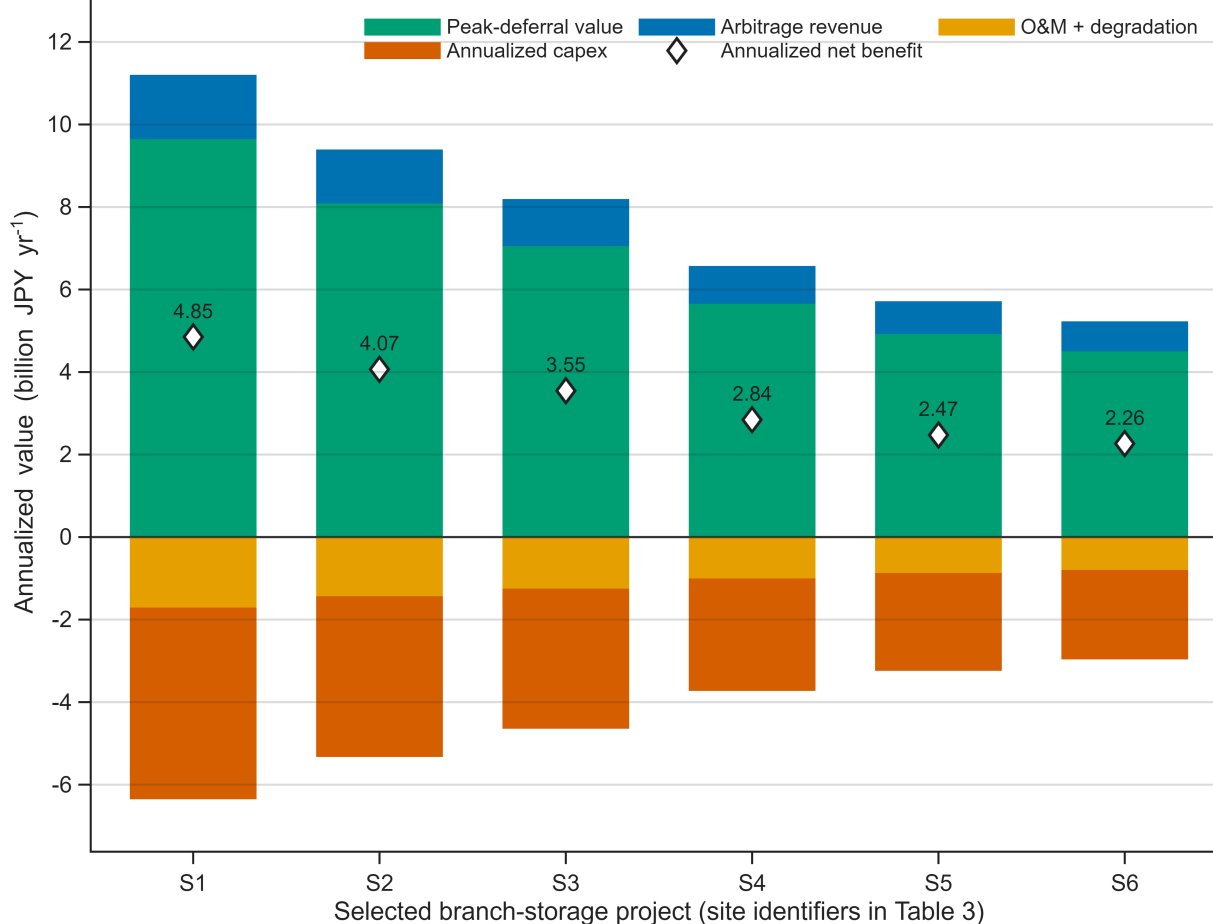


Figure 4. Annualized economics of the six selected branch-storage projects Positive stacked bars show modeled peak-deferral value and JEPX arbitrage revenue; negative stacked bars show fixed O&M plus degradation and annualized capex.

year, so all six projects are loss-making at 10,000 JPY/kW-year and profitable at 20,000 JPY/kW-year (Figure 5).

Each JPY 10,000/kW-year increase raises portfolio annualised net benefit by JPY 13.285 billion/year, reaching JPY 59.898 billion/year at 60,000 JPY/kW-year, and above the break-even region the order $S1 > S2 > S3 > S4 > S5 > S6$ is preserved because the fixed projects share two-hour duration, similar cycling, and peak reduction equal to installed power.

With the peak-deferral value fixed at its base level, portfolio NPV across a 5×5 capex and price-spread grid ranges from JPY 115.76 billion at $1.30 \times$ capex and $0.70 \times$ price spread to JPY 300.32 billion at $0.70 \times$ capex and $1.30 \times$ price spread. Moving the price-spread multiplier from 0.70 to 1.30 at base capex changes NPV by JPY 40.17 billion, whereas moving capex from $1.30 \times$ to $0.70 \times$ at the base spread changes it by JPY 144.39 billion, so accounting exposure to capex is approximately 3.6 times that to the tested price-spread range. NPV stays positive at all 25 grid points because the dominant credit is held constant and the design is not re-optimised.

Temperature-responsive branch evidence

Fifty-seven of 283 branches (20.1%) satisfy every strong-temperature-evidence criterion, and all 57 have individually significant positive heating and cooling slopes. Within this subset, temperature partial R^2 ranges from 0.254 to 0.679 (median 0.370) and cross-validated partial R^2 from 0.248 to 0.679

(median 0.363), and maximum modelled cold- or hot-evening effects range from 24.2% to 87.5% of mean 20:00 absolute flow (median 37.7%). The largest adjusted joint-temperature FDR q-value is 2.95×10^{-7} , so significance is not the operative constraint: classification is governed almost entirely by the in-sample temperature partial R^2 threshold, since relaxing it alone would admit 52 further branches and relaxing the 50 MW mean-flow gate two more, whereas the remaining gates exclude no branch that the other criteria retain.

The highest-ranked branch, Shin-Tokorozawa Substation – Minami-Sayama Line, combines mean absolute 20:00 flow of 1,573.6 MW with heating and cooling slopes of 55.50 and 94.69 MW/°C, temperature partial $R^2 = 0.564$, and a maximum modelled effect of 45.8%. The remaining leading branches vary substantially in flow scale and response shape (Table S4): Shin-Sawara – Inba reaches a cooling slope of 158.02 MW/°C and a maximum effect of 87.5%, whereas Shin-Sawara – Shin-Sodegaura combines mean flow of 3,247.2 MW with heating and cooling slopes of 113.92 and 352.82 MW/°C. Figure 6 summarises the top-15 evidence scores and the six leading responses

Only one of the 57 strong branches shows a raw-versus-controlled temperature-slope sign change, whereas 56 retain both signs after calendar adjustment. The strong-evidence set is therefore largely stable to the added controls, although causal interpretation of the estimated slopes as temperature-controlled load or thermal overload still requires additional evidence.

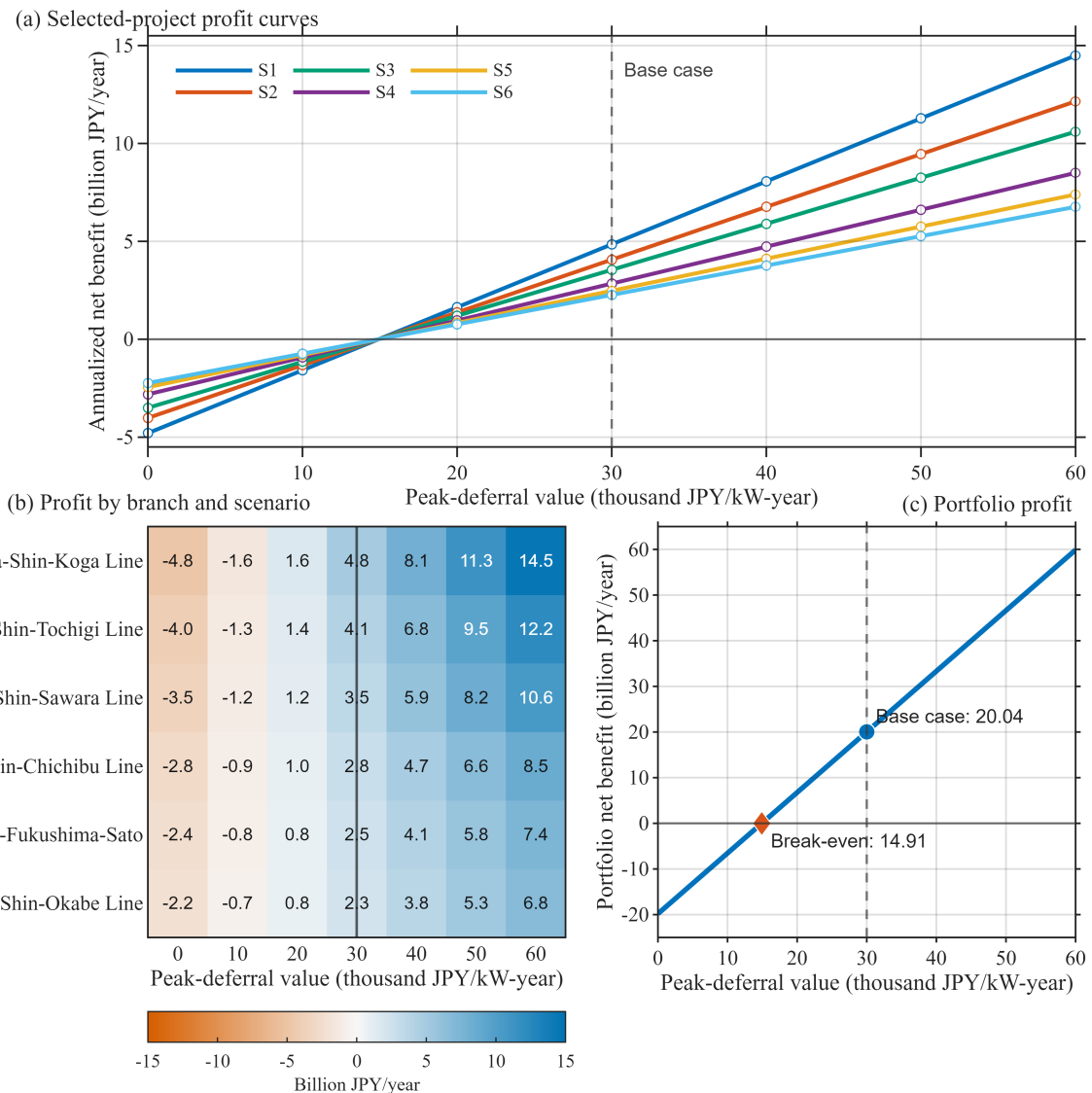


Figure 5. Fixed-design annualized-profit sensitivity to the assumed peak-deferral value (A) Annualized net benefit of the six selected branch-storage projects. (B) Branch-level profit at discrete value scenarios; the vertical line marks the 30,000 JPY/kW-year base case. (C) Aggregate portfolio annualized net benefit, with the base and break-even values identified.

Spatial alignment of economic and weather evidence

The pressure-screening and temperature-evidence scores are weakly rank-correlated across all 283 branches (Spearman $\rho = 0.142$, $p = 0.016$), and P95 flow and temperature partial R^2 are similarly weakly associated ($\rho = 0.210$). The two evidence layers therefore produce different spatial orders: pressure reflects hourly peak magnitude, exceedance persistence, ramping, and sustained flow, whereas the temperature score reflects the stability and magnitude of cold- and hot-evening associations (Figure 7).

Only two of the 12 pressure-screen candidates meet the strong-temperature rule – Shin-Sawara Substation – Shin-Sodegaura Line (pressure rank 1, temperature rank 4) and Shin-Keiyo Substation – Shin-Keiyo Line (pressure rank 2, temperature rank 17) – and neither is selected by the budget-constrained portfolio. The reason is the duration mechanism of Section 3.2 rather than weak project economics: both are interior-duration designs whose sustained peak exceedance raises energy capital and roughly halves capital efficiency. Among the six selected projects, temperature partial R^2 ranges from 0.037 to 0.188 and none passes the complete weather screen; conversely, the median pressure rank among the 57 strong-temperature

branches is 134. The selected portfolio is therefore economically prioritised but contains no joint strong-temperature project (Table S5).

Because pressure and temperature-evidence scores are fixed empirical attributes, their direct relationship is invariant to the assumed peak-deferral value. Reconstructing fixed-design annualised profit from 0 to 60,000 JPY/kW-year shows that at zero credit candidate profit is negatively associated with pressure ($\rho = -0.748$, $p = 0.007$) and temperature evidence ($\rho = -0.704$, $p = 0.011$), because larger fixed designs incur larger absolute losses. Near the break-even region the ranking changes rapidly: at 15,000 JPY/kW-year the two correlations weaken to -0.154 and -0.382 , and at 20,000 JPY/kW-year all 12 candidates are profitable but rank crossovers leave both associations close to zero (Figure 8).

At the 30,000 JPY/kW-year base value, candidate profit is positively associated with pressure ($\rho = 0.664$, $p = 0.022$) and only moderately with temperature evidence ($\rho = 0.490$, $p = 0.106$), and at 60,000 JPY/kW-year the candidate correlations rise to 0.734 and 0.620 as peak-reduction scale increasingly determines absolute profit, while the all-branch correlation remains $\rho = 0.142$. The higher correlations in the candidate pool ($\rho = 0.616$, $n = 12$) and

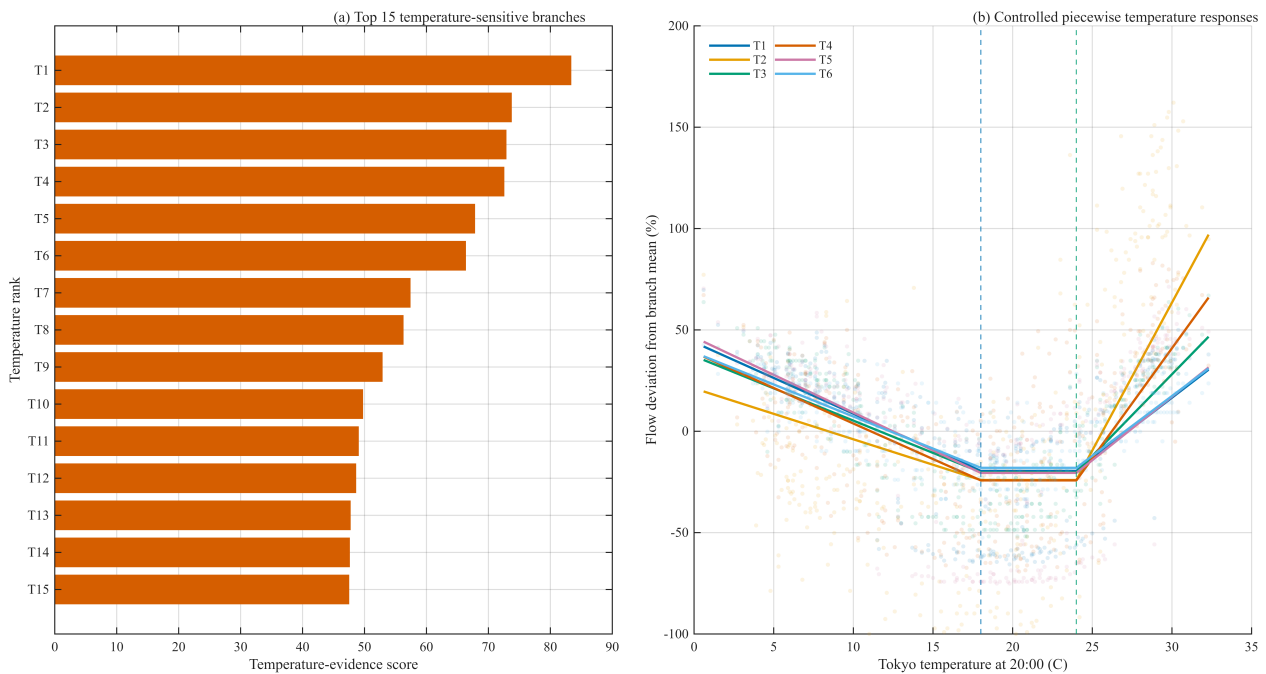


Figure 6. Temperature-sensitive branches (A) Evidence score for the top 15 branches. (B) Standardized observed 20:00 flow deviations and controlled piecewise heating/cooling components for the top six branches. T1–T5 map to Table 4; T6–T15 follow the rank order in panel (A).

selected portfolio ($p = 0.600$, $n = 6$) are selection-conditioned, so the scenario-dependent sign change is an accounting effect on profit ranks rather than a change in the underlying relationship.

DISCUSSION AND CONCLUSIONS

Decision-aware planning across spatiotemporal evidence layers

Storage-siting models can co-optimize location, size, dispatch, and grid support when they include electrical constraints, ratings, prices, or reliability criteria. This study addresses a more restrictive but common public-data setting in which branch topology and measured flows are available while the variables that convert a flow excursion into an engineering violation are not. The three-layer architecture keeps that absence visible: empirical P95 defines an observed peak-shaving target, the optimisation determines conditional actionability, and the validation layer specifies the engineering evidence still required. The architecture transfers a principle from decision-aware allocation,¹⁷ graph-based spatial intervention,¹⁸ and coordinated infrastructure planning¹⁹—intermediate evidence should be judged by the decision it changes—and extends it across temporal scales, because event-scale power deficits and annual sufficiency require different indicators³³ and spatial coordination can dominate aggregate resource quality.³⁴

This separation explains why pressure-screen ranks 3, 6, 7, 10, 11, and 12 form the budget-optimal portfolio and why none of them passes the strict temperature screen: candidate generation, investment choice, weather exposure, and reliability validation are separate decisions supported by different evidence. The contribution is a traceable planning front end that precedes security-constrained transmission analysis.

Value stacking and the dominance of peak-deferral assumptions

Arbitrage and service stacking can improve storage utilisation and revenue, although realised value depends on market feedbacks, service compatibility, and how transmission- and distribution-level benefits are represented,^{13–16} and network-oriented studies show that it also depends on placement and on the represented constraint.^{9–12} The portfolio reinforces both arguments, but the evidentiary status of its value components differs: observed JEPX prices support temporal arbitrage, whereas the peak-deferral credit is a declared planning assumption. With arbitrage providing 13.9% and peak

deferral 86.1% of gross annual value, project acceptance depends primarily on proving that storage is electrically deliverable and can defer a specified investment or reliability need.

This structure differs from the financing problem of rarely used resilience back-up, where resources selected for irregular extreme events can have low utilisation and precarious cost recovery.³³ The projects modelled here instead cycle about 274 times per year and earn value from recurrent intraday shifting plus an assumed annual peak credit; they represent operational flexibility, whereas coverage of multi-day scarcity and service continuity through outages requires separate testing.

The two-hour lower-bound solution, which holds for the six selected projects and nine of the 12 candidates, makes the temporal mechanism visible: the configured peak objective rewards power, and after energy-capital, efficiency, and degradation costs additional energy receives insufficient incremental value. Longer duration already emerges endogenously on the three highest-stress candidates, where sustained exceedance forces 2.7–2.9 h designs that the budget then excludes, and it could emerge more widely with curtailment reduction, multi-day adequacy events, reserve products, or tighter power caps. A resource-adequacy extension should therefore re-optimize duration against multi-year supply–demand trajectories rather than treating the present two-hour design as an accredited capacity resource.

Weather-responsive exposure and validation routing

The 57 strong-temperature branches extend documented nonlinear temperature–electricity response functions, including Tokyo-scale evidence, to a branch-level exposure screen,^{25,26} consistent with forecasting studies that support weather ensembles, semiparametric response structures, and probabilistic evaluation.^{27–29} The double-sided associations indicate that cold and hot evenings coincide with higher absolute flow on these corridors after calendar adjustment, and their heterogeneity matches the lesson of climate-informed adequacy analysis that aggregate demand or mean resource conditions need not reveal local supply–demand–network misalignment.³⁴

Weather responsiveness identifies exposure, whereas resilience performance is assessed from disturbance and recovery outcomes. A measured branch combines demand, embedded generation, transfers, parallel-path

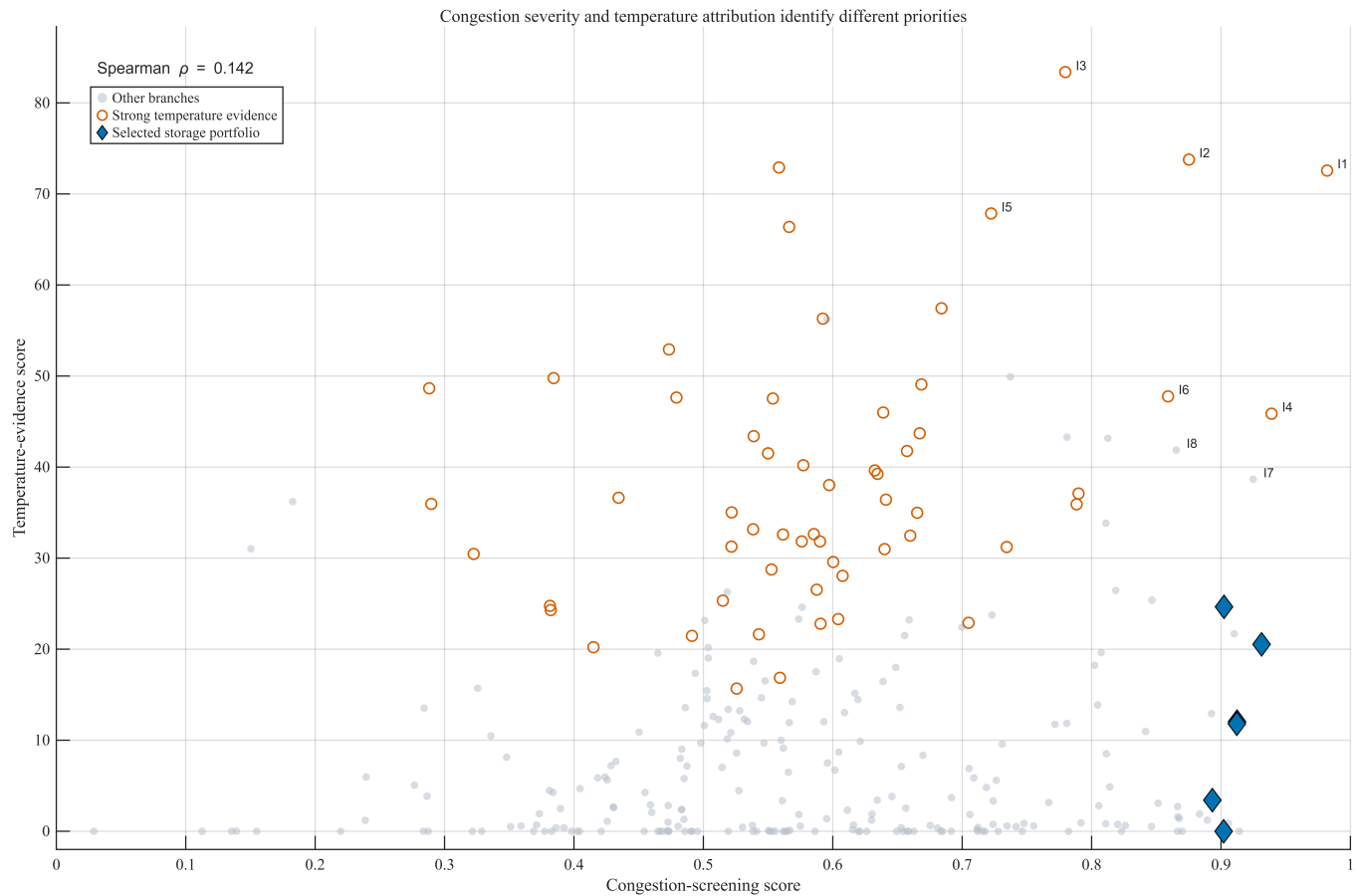


Figure 7. Integrated branch priorities. Pressure (congestion-screening) and temperature-evidence scores are weakly rank-correlated Open orange circles meet all strong temperature criteria; blue diamonds are the six projects selected by the budget-constrained NPV portfolio. I1–I8 identify descriptive integrated ranks for decision routing.

redistribution, and switching, so causal attribution of its coefficients to temperature-controlled load requires additional identification. The weak overlap therefore clarifies that economic peak shifting, and weather-responsive exposure identify different spatial priorities through different mechanisms, and a single composite score would obscure which intervention is justified. A stronger study would use multiple weather stations, all 24 hours with hour–temperature interactions, temporally blocked out-of-year validation, and multi-year histories connected to outage probabilities, storage depletion, and restoration.

For utility practice the findings imply a four-gate workflow. Gate 1 uses public flows, topology, zonal prices, and weather to produce separate economic and exposure priorities. Gate 2 replaces empirical proxies with official ratings, PTDF/LODF or AC sensitivities, outage states, and operational forecasts. Gate 3 fixes feasible capacity layouts and tests them against multi-year weather, generation, demand, forced outages, and contingencies, reporting load-loss frequency, unserved energy, storage-depletion duration, and recovery. Gate 4 values only surviving projects using utility-specific avoided reinforcement, accredited capacity or resilience benefit, degradation, and financing assumptions. The priority classes determine which gate receives attention: economic-only candidates enter non-wires-alternative review, weather-only corridors enter seasonal and climate-exposure studies, and joint candidates receive accelerated engineering and stress testing. Finally, JEPX Tokyo-area prices reward time shifting but do not reveal where storage relieves intrazonal congestion, so a nodal price or local flexibility product would align dispatch revenue with network location, and where neither exists a utility contract must compensate verified deferral value explicitly while preventing double counting with arbitrage revenue.

Study scope and limitations

The interpretation is constrained in five areas. First, the observational data cover fiscal year 2024, branch flow combines system influences beyond local load, the topology is an author-derived research representation, and representative-column selection leaves the authority of duplicate measurements unresolved. Second, the electrical model uses empirical P95 as a proxy and assumes a one-to-one storage effect on the monitored branch, whereas engineering validation requires verified ratings, PTDF/LODF, voltage, loss, protection, outage, and N–1 constraints. Third, the optimisation assumes perfect hourly foresight, continuous sizing, and additive stand-alone project NPVs; it omits cross-project network interactions and price feedback and applies prespecified screening weights and a 12-candidate cutoff without sensitivity analysis. Fourth, the economic scope excludes taxes, financing structure, construction lead time, augmentation, capacity fade, replacement, reserve revenue, site costs, and price cannibalisation, so positive NPV within the fixed-design grid establishes conditional model value rather than bankability. Fifth, the temperature component uses one weather station, fixed balance points, one 20:00 observation per day, and interleaved cross-validation, and resource-adequacy simulation would additionally require a multi-year climate ensemble, forced-outage model, and restoration process; LOLE, expected energy not served, effective load-carrying capability, and recovery duration therefore remain unestimated.

CONCLUSIONS

This study establishes a decision-aware pre-feasibility approach for using public transmission-flow data to guide branch-level BESS investment screening. The framework separates observed branch stress, conditional

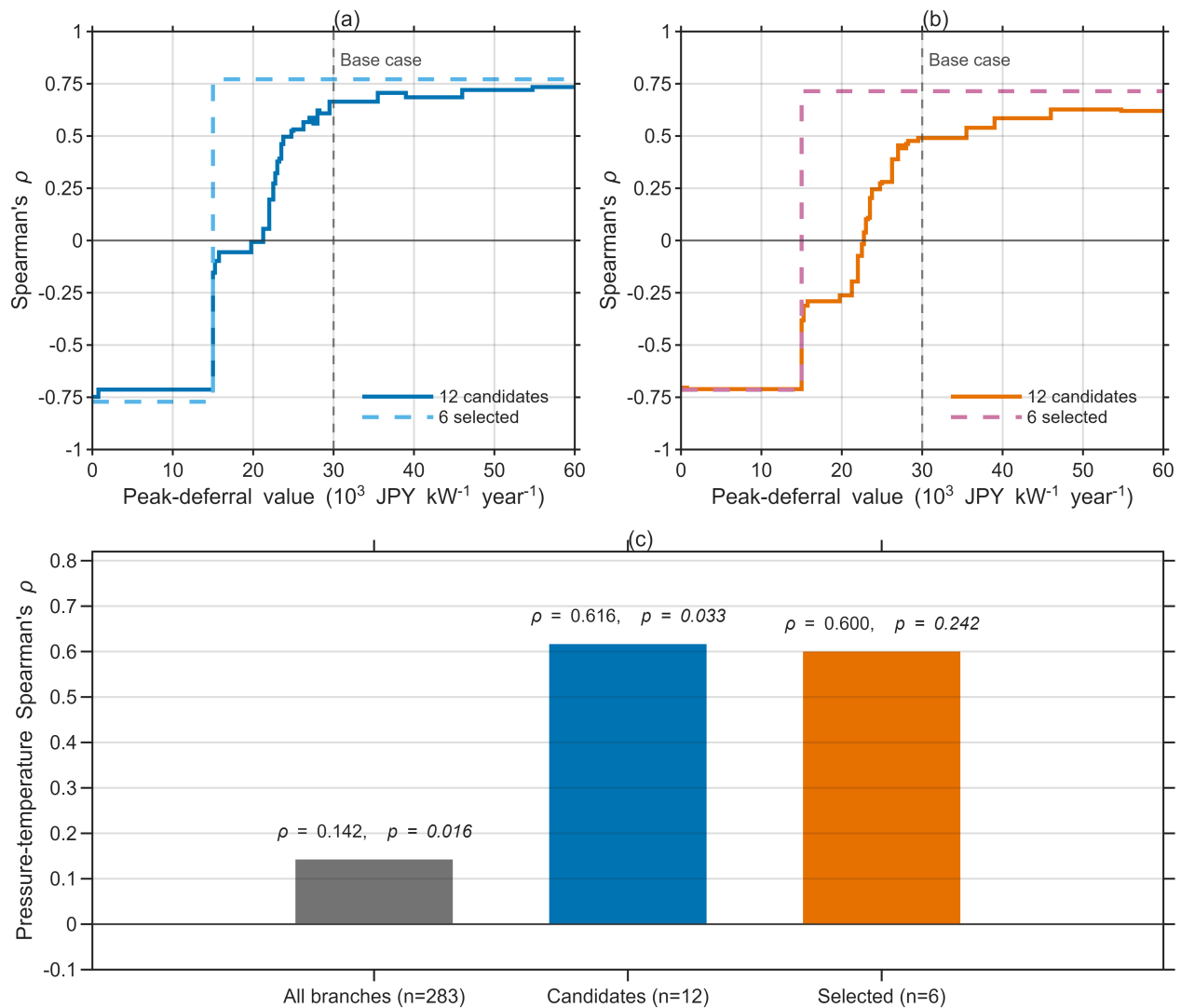


Figure 8. Correlation of fixed-design annualized profit with branch pressure and temperature evidence under alternative peak-deferral values (A) Spearman correlation between profit and pressure score. (B) Spearman correlation between profit and temperature-evidence score. Solid lines denote the 12 candidates and dashed lines the six selected projects. (C) Direct pressure-temperature correlation for all branches, candidates, and selected projects.

investment value, and weather-responsive exposure, avoiding a single ranking that would obscure distinct planning mechanisms. Using 8,760 hourly observations across 283 TEPCO research-network branches, it identified 12 candidates and selected six two-hour projects under a JPY 200 billion budget, totalling 1,328.5 MW/2,657 MWh and a conditional 15-year NPV of JPY 208.0 billion.

Portfolio economics are governed by the peak-deferral assumption: JEPX arbitrage contributes JPY 6.451 billion/year against JPY 39.855 billion/year from the assumed credit, or 86.1% of gross annual operating value. With fixed designs and dispatch, removing the credit reduces annualised net benefit to -JPY 19.812 billion/year, and the portfolio break-even value is approximately JPY 14,913/kW-year. The two-hour duration floor and approximately 274 equivalent cycles/year show that the selected designs target recurrent intraday peak shifting.

The spatiotemporal evidence produces two distinct spatial priority sets. Fifty-seven branches satisfy controlled, false-discovery-rate-screened, and cross-validated heating-and-cooling criteria; two enter the 12-branch candidate set and none enters the selected portfolio, and the all-branch rank correlation ($\rho = 0.142$) confirms that observed branch stress and temperature-responsive exposure operate through different spatial mechanisms. Changing the peak-deferral credit modifies profit-evidence alignment but not the

underlying association, so disagreement between the evidence streams becomes actionable information rather than a reason to collapse them into one rank. The central contribution is therefore a transparent decision-routing architecture for BESS planning under incomplete public network information: economic-only candidates proceed to non-wires-alternative and avoided-cost assessment, weather-responsive corridors to seasonal and climate-exposure analysis, and joint priorities to accelerated network and multi-year stress testing, with procurement appraisal reserved for verified ratings, transfer sensitivities, contingencies, and scenario-specific re-optimisation.

REFERENCES

- Hesse H., Schimpe M., Kucevic D., et al. (2017). Lithium-Ion Battery Storage for the Grid-A Review of Stationary Battery Storage System Design Tailored for Applications in Modern Power Grids. *Energies* **10**, 2107, DOI:10.3390/en10122107
- Liu Z., Su T., Quan Z., et al. (2023). Review on the Optimal Configuration of Distributed Energy Storage. *Energies* **16**, 5426. DOI:10.3390/en16145426
- Levin T., Bistline J., Sioshansi R., et al. (2023). Energy storage solutions to decarbonize electricity through enhanced capacity expansion modelling. *Nat. Energy* **8**, 1199–1208. DOI:10.1038/s41560-023-01340-6
- Mallapragada D., Sepulveda N. and Jenkins J. (2020). Long-run system value of battery energy storage in future grids with increasing wind and solar generation. *Appl. Energy* **275**, 115390. DOI:10.1016/j.apenergy.2020.115390

5. Sepulveda N., Jenkins J., Edington A., et al. (2021). The design space for long-duration energy storage in decarbonized power systems. *Nat. Energy* :506–516. DOI:10.1038/s41560-021-00796-8
6. Ziegler M., Mueller J., Pereira G., et al. (2019). Storage Requirements and Costs of Shaping Renewable Energy Toward Grid Decarbonization. *Joule* **3**:2134–2153. DOI:10.1016/j.joule.2019.06.012
7. Antweiler W. (2021). Microeconomic models of electricity storage: Price Forecasting, arbitrage limits, curtailment insurance, and transmission line utilization. *Energy Eco.* **101**:105390. DOI:10.1016/j.eneco.2021.105390
8. Schill W. and Zerrahn A. (2018). Long-run power storage requirements for high shares of renewables: Results and sensitivities. *Renew. Sustain. Energy Rev.* **83**:156–171. DOI:10.1016/j.rser.2017.05.205
9. Massucco S., Pongiglione P., Silvestro F., et al. (2021). Siting and Sizing of Energy Storage Systems: Towards a Unified Approach for Transmission and Distribution System Operators for Reserve Provision and Grid Support. *Elect. Power Sys. Res.* **190**:106660. DOI:10.1016/j.epsr.2020.106660
10. Khajeh H., Parthasarathy C., Doroudchi E., et al. (2023). Optimized siting and sizing of distribution-network-connected battery energy storage system providing flexibility services for system operators. *Energy* **285**:129490. DOI:10.1016/j.energy.2023.129490
11. Zhao Y., Zhang Y., Li Y., et al. (2024). Optimal configuration of energy storage for alleviating transmission congestion in renewable energy enrichment region. *J. Energy Stor.* **82**:110398. DOI:10.1016/j.est.2023.110398
12. Inaolaji A., Wu X., Roychowdhury R., et al. (2024). Optimal allocation of battery energy storage systems for peak shaving and reliability enhancement in distribution systems. *J. Energy Stor.* **95**:112305. DOI:10.1016/j.est.2024.112305
13. Lamp S. and Samano M. (2022). Large-scale battery storage, short-term market outcomes, and arbitrage. *Energy Eco.* **107**:105786. DOI:10.1016/j.eneco.2021.105786
14. Hjalmarsson J., Thomas K. and Boström C. (2023). Service stacking using energy storage systems for grid applications – A review. *J. Energy Stor.* **60**:106639. DOI:10.1016/j.est.2023.106639
15. Marnell K., Obi M. and Bass R. (2019). Transmission-Scale Battery Energy Storage Systems: A Systematic Literature Review. *Energies* **12**:4603. DOI:10.3390/en12234603
16. Ma M., Huang H., Song X., et al. (2022). Optimal sizing and operations of shared energy storage systems in distribution networks: A bi-level programming approach. *Appl. Energy* **307**:118170. DOI:10.1016/j.apenergy.2021.118170
17. Chung A., Abdulai J., Bayoh P., et al. (2026). Improving access to essential medicines via decision-aware machine learning. *Nature* :1178–1183. DOI:10.1038/s41586-026-10433-7
18. Zheng Y., Lin Y., Zhao L., et al. (2023). Spatial planning of urban communities via deep reinforcement learning. *Nat. Comput. Sci.* :748–762. DOI:10.1038/s43588-023-00503-5.
19. Wessel J., Chowdhury A., Wild T., et al. (2025). Regional coordination can alleviate the cost burden of a low-carbon electricity system. *Nat. Commun.* **16**:9033. DOI:10.1038/s41467-025-64093-8
20. Donti P. L., Amos B. and Kolter J. Z. (2017). Task-based End-to-end Model Learning in Stochastic Optimization. *Adv. Neural Inform. Proces. Sys.* **30**:5484–5494. DOI:10.48550/arXiv.1703.04529
21. Wilder B., Dilkina B. and Tambe M. (2019). Melding the Data-Decisions Pipeline: Decision-Focused Learning for Combinatorial Optimization. *Proceed. AAAI Conf. A.I.* **33**:1658–1665. DOI:10.1609/aaai.v33i01.33011658
22. Elmachoub A. and Grigas P. (2022). Smart "Predict, then Optimize". *Manag. Sci.* **68**:9–26. DOI:10.1287/mnsc.2020.3922
23. Mandi J., Kotary J., Berden S., et al. (2024). Decision-Focused Learning: Foundations, State of the Art, Benchmark and Future Opportunities. *J. A.I. Res.* **81**:1623–1701. DOI:10.1613/jair.1.15320
24. Kotary J., Fioretto F., Van Hentenryck P., et al. (2021). End-to-End Constrained Optimization Learning: A Survey. *Proceed. 13 Int. Joint Conf. A.I.* :4475–4482. DOI:10.24963/ijcai.2021/610
25. Hiruta Y., Gao L. and Ashina S. (2022). A novel method for acquiring rigorous temperature response functions for electricity demand at a regional scale. *Sci. Total Environ.* **819**:152893. DOI:10.1016/j.scitotenv.2021.152893
26. Takane Y., Nakajima K., Yamaguchi K., et al. (2023). Decarbonisation technologies can halve the nonlinear increase in electricity demand in densely populated areas due to climate change. *Sustain. Cit. Soc.* **99**:104966. DOI:10.1016/j.scs.2023.104966
27. Ludwig N., Arora S. and Taylor J. W. (2023). Probabilistic load forecasting using post-processed weather ensemble predictions. *J. Operat. Res. Soc.* **74**:1008–1020. DOI:10.1080/01605682.2022.2115411
28. Miller J. I. and Nam K. (2022). Modeling peak electricity demand: A semiparametric approach using weather-driven cross-temperature response functions. *Energy Econ.* **114**:106291. DOI:10.1016/j.eneco.2022.106291
29. Hou H., Liu C., Wang Q., et al. (2022). Review of load forecasting based on artificial intelligence methodologies, models, and challenges. *Elec. Power Sys. Res.* **210**:108067. DOI:10.1016/j.epsr.2022.108067
30. Brown T., Hörsch J. and Schlachtberger D. (2018). PyPSA: Python for Power System Analysis. *J. Open Res. Soft.* **6**:4. DOI:10.5334/jors.188
31. Hörsch J., Hofmann F., Schlachtberger D., et al. (2018). PyPSA-Eur: An open optimisation model of the European transmission system. *Energy Strat. Rev.* **22**:207–215. DOI:10.1016/j.esr.2018.08.012
32. Birchfield A., Xu T., Gegner K., et al. (2017). Grid Structural Characteristics as Validation Criteria for Synthetic Networks. *IEEE TPS* **32**:3258–3265. DOI:10.1109/tpwrs.2016.2616385
33. Grochowicz A., Bloomfield H. and Victoria M. (2026). Resilience metrics to guide back-up investments in the power system during extreme weather. *Nat. Sustain.* DOI:10.1038/s41893-026-01897-6
34. Qiu L., Khorramfar R., Wang S., et al. (2026). Climate change reshapes resource adequacy risks and optimal renewable energy siting in wind and solar energy systems. *Nat. Energy*. DOI:10.1038/s41560-026-02109-3

FUNDING AND ACKNOWLEDGMENTS

This study was supported by Shandong Provincial Youth Innovation Team Project for Higher Education Institutions (No. 2024KJH025).

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND CODE AVAILABILITY

The replication package accompanying this manuscript contains the processed 8,760-hour branch-flow matrix used in the study and the author-derived topology snapshot. The topology repository is available at <https://github.com/xiaowangzi777/tepco-pg-transmission-topology-data>. The branch-flow input is a station-grouped compilation of TEPCO Power Grid transmission line-flow records published through the company's system-information service (<https://www.tepco.co.jp/pg/consignment/system/>); it is a research compilation and not an official operational network model. Day-ahead prices are the FY2024 Tokyo-area spot summary published by JEPX (<https://www.jepx.jp/electricpower/market-data/spot/>), aggregated to hourly resolution as the arithmetic mean of each adjacent pair of half-hour delivery periods. Temperature observations are the JMA hourly record for Tokyo station (<https://www.data.jma.go.jp/risk/obsdl/index.php>).

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/ijp.energy-use.2026.100062>