buildings

MDPI

Article

Designing Data Centers for Demand Response Through a Construction-Phase Readiness Index

Arezou Shafaghat [1,2], Da Hu [2] and Ali Keyvanfar [3,*]

[1] Sustain-X Scale Lab, Scheller College of Business, Georgia Institute of Technology, Atlanta, GA 30308, USA; ashafaghat3@gatech.edu
[2] Department of Civil and Environmental Engineering, Southern Polytechnic College of Engineering and Engineering Technology, Kennesaw State University, Marietta, GA 30060, USA; dhu3@kennesaw.edu
[3] Department of Construction Management, College of Architecture and Construction Management, Kennesaw State University, Marietta, GA 30060, USA
* Correspondence: akeyvanf@kennesaw.edu

Abstract

AI-driven growth is pushing data-center electricity demand from 415 TWh (2024) toward 945 TWh by 2030. Demand-response research has matured on operational levers but treats the physical envelope as largely exogenous, leaving construction-phase decisions that bound achievable flexibility unmodeled. This article integrates three previously separate bodies of literature (data-center demand response, grid-interactive efficient buildings, and stochastic optimal control under irreversibility) into a single framework that prices construction-phase flexibility as a portfolio of real options, pairing elicitation-derived (FAHP) weights with simulation-derived (Sobol) variance indices. None of the individual techniques is new; the contribution is their synthesis and the finding that architectural and site decisions carry the dominant financial leverage in the model, whereas a literature-grounded synthetic-persona prior (twenty-five LLM-simulated personas) prioritizes mechanical-electrical systems; a divergence we frame as a screening diagnostic between an LLM prior and the model. The Flex-by-Design Readiness Index (FDRI) is a nineteen-dimension taxonomy across architectural, MEP, and site-urban layers, weighted by Fuzzy AHP. The FDRI–ROV model formalizes the decision as stochastic optimal control under irreversibility. Calibrated to PJM, ERCOT, and CAISO (2024–2026) for a 100 MW plant at N = 10,000 paths over thirty years, it yields +$79 M net option value at Full FDRI for PJM (additive upper bound; substitution-corrected ≈ +$41 M, 1.7× CapEx; 2.4× CapEx PJM, 2.1× ERCOT, 2.8× CAISO); on both the additive (2.4–2.8×) and corrected (1.7×) bases the pre-registered H2 threshold of Vtotal/Ctotal ≥ 3× is not met. Sobol decomposition places architectural and site layers at ST ≈ 0.56 each versus MEP at 0.15, exposing waste-heat-export and regulatory-avoided-cost dimensions as under-recognized leverage. Out-of-sample validation against four hyperscale projects yields 11% MAPE, reported as an n = 4, single-period proof of concept.A pro-rata extrapolation across all ~43 GW of incremental U.S. capacity gives a nominal ~$34 billion through 2035, but this applies the single most optimistic scenario uniformly; applying the substitution-corrected per-plant value with competition, policy, and adoption decay multipliers, the defensible 2035 opportunity is ≈$3–$18 billion (central ≈$7 billion), with $34 billion retained only as an undecayed ceiling.

Keywords: data center; construction-phase design; grid-interactive efficient buildings; demand response; real options valuation; stochastic optimal control; Fuzzy Analytic Hierarchy Process; Monte Carlo simulation

Check for updates

Received: 4 June 2026
Revised: 25 June 2026
Accepted: 8 July 2026
Published: 20 July 2026

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

1. Introduction

The rapid expansion of artificial intelligence (AI) has triggered an unprecedented and structural shift in global electricity demand. Global data center electricity consumption reached approximately 415 TWh in 2024 and is projected to exceed 945 TWh by 2030, raising the data center share of global electricity demand from roughly 1.5% to 3% [1]. In the United States, BloombergNEF projects data center power demand to rise from about 35 GW in 2024 to nearly 78 GW by 2035, with hyperscale facilities accounting for almost half of incremental electricity-demand growth through 2030 [2]. The expansion is geographically concentrated: data center consumption already accounts for over 20% of national electricity demand in Ireland and 25% in Virginia [1]. It is also operationally volatile in a way that legacy planning frameworks were never designed to absorb; hyperscale facilities now swing hundreds of megawatts within minutes during AI-training cycles, creating localized grid stress and pushing regional reliability indices into territory the North American Electric Reliability Corporation has flagged as among the most pressing emerging risks for the bulk power system [3]. Local social license is itself a binding constraint: community opposition has blocked or delayed an estimated USD 64 billion of U.S. data center projects between mid-2024 and early 2025 [4]. These siting and social-license pressures are increasingly framed as an urban-sustainability problem rather than a purely technical one [5].

These pressures have placed renewed emphasis on demand-side flexibility, and within it on demand response (DR) (the modulation of electricity consumption in response to grid signals) as a principal mechanism for absorbing volatility, supporting renewable integration, and meeting decarbonization targets [6–8]. Recent contributions establish strategic demand-side flexibility (fast frequency response, capacity-market participation, ancillary services) as a necessary complement to renewable expansion [7]. Within the data center literature specifically, DR is well studied as an operational lever: workload shifting, dynamic server provisioning, uninterruptible-power-supply discharge, real-time pricing, and model-predictive control are mature subfields with substantial peer-reviewed treatments [9–12]. The grid-interactive efficient buildings (GEB) paradigm has further extended these levers into the broader built environment [13–15]. Across both studies, however, the physical envelope is treated as exogenous; fixed at construction and inherited by operators.

This treatment leaves the most consequential decisions outside the model. The maximum achievable flexibility of an operating data center is bounded by construction-phase choices: power topology, switchgear ramp capability, battery-energy-storage-system oversizing, thermal mass, mechanical-room reservation, equipment-yard footprint, water rights, interconnection corridor reservation, submetering granularity, and modular subdivision of the white space. Once the slab is poured and the substation interconnected, these decisions are effectively irreversible for the 30-year asset life. The construction manager, the architect, and the urban planner are therefore the primary determiners of the DR capability envelope; yet no peer-reviewed framework formalizes their decisions as a flexibility-investment problem under uncertainty, and no model values the construction-phase choices that lock in or lock out future participation in flexibility markets. Within the data-center DR literature, the built-environment determinants of this envelope remain unexamined; the building-design literature treats them, but at the building scale and outside the data-center typology.

This article addresses that gap by reframing Demand Response (DR) readiness as a construction-management deliverable and valuing it through the formal apparatus of stochastic optimal control under investment irreversibility [16–18], of which classical real options theory [19,20] is the canonical optimal-stopping specialization. Each construction-phase decision is treated as an American-style call option on a specific operational flexibility. This reframing draws on the methodological lineage of real-options applications to

renewable-infrastructure investment [21–23] and on the engineering-options tradition that values flexibility embedded in the design of physical systems [24–26]. It extends that tradition in four respects specific to the present problem. Rather than valuing generic options 'in' a system, it identifies the construction-phase determinants of data center demand-response capability as a structured portfolio of such options; it prices them against grid-flexibility, capacity, ancillary-services, and waste-heat revenue streams rather than against the demand-following or capacity-staging value typical of prior engineering-options case studies; it ties each option to the specific design phase by which it must be locked in; and it embeds the valuation in a readiness index reconciled against a global-sensitivity decomposition. It also re-positions construction managers, architects, and urban planners as primary economic actors in grid flexibility markets, a role currently absent from the literature.

In one sentence, the paper's contribution is to reframe construction-phase design as a priced, optimizable flexibility-investment decision (making the architect, urban planner, and construction manager economic actors in grid flexibility) and to provide the index, valuation model, and elicitation-versus-financial diagnostic needed to act on it. Concretely, the contribution is fourfold. First, this article develops a typology and taxonomy of construction-phase flex-locks across architectural, mechanical-electrical-plumbing (MEP), and site-urban layers. Second, it develops the Flex-by-Design Readiness Index (FDRI), a normalized, auditable scoring framework whose dimension weights are derived under linguistic uncertainty via Fuzzy Analytic Hierarchy Process from a three-round structured expert-elicitation panel implemented with large-language-model-simulated domain-expert personas (no human participants; see Section 3, Phase 2). Third, it develops the FDRI–Risk-adjusted Optimization under Volatility (FDRI–ROV) simulation model; a hybrid Monte Carlo, Least-Squares Monte Carlo [26], System Dynamics [27], and Sobol global sensitivity architecture grounded in stochastic optimal control under investment irreversibility that values the FDRI portfolio under stochastic energy-price, capacity-price, and regulatory-regime evolution. Fourth, the model is calibrated against PJM, ERCOT, and CAISO 2024–2026 wholesale market data and empirically validated by comparing predicted option premia with observed flexibility-program participation across four contemporary AI-era hyperscale projects. The novelty claimed here is integrative rather than instrumental. Real-options valuation, the Fuzzy Analytic Hierarchy Process, the Delphi-style multi-round consensus protocol, and Least-Squares Monte Carlo are each established techniques, and no claim is made that any of them is advanced in isolation. The contribution is twofold. First, it integrates three previously separate bodies of literature—data center demand response, grid-interactive efficient buildings, and stochastic optimal control under investment irreversibility—into a single object that treats the data center simultaneously as an electrical load, a built-environment asset, and a priced financial option, a synthesis that none of the three fields currently spans. Second, it introduces a methodological move within that synthesis: the pairing of FAHP-derived normative weights, which encode the prevailing engineering-literature emphasis as reproduced by the synthetic expert-persona elicitation, with Sobol-derived positive variance indices, which locate where dollar-denominated flexibility value actually concentrates under calibrated wholesale-market processes. The resulting diagnostic gap $\Delta_i = S_{T,i} - w_i$, which exposes flexibility dimensions whose financial leverage is systematically under-recognized relative to engineering intuition, is, to our knowledge, new to the energy-infrastructure real-options literature.

The framework's relevance to sustainability follows directly from the construction-phase decision chain. AI-era data-center load growth is now the largest discretionary increment to OECD electricity demand through 2030 [1], and the carbon intensity of that load is determined less by the data center's own efficiency than by the marginal grid resource dispatched to serve its inflexibility. A flex-ready facility (one whose construction-

phase envelope was designed to accept ramping, load-shifting, waste-heat export, and BESS dispatch) allows the host grid to absorb a higher share of variable wind and solar, deferring fossil-peaker dispatch and reducing system-level CO_2 intensity [28–30]. A flex-incapable facility imposes the opposite: grid operators dispatch fossil reserves to maintain reliability, and renewable curtailment rises. The construction-phase decisions cataloged in this article are thus the upstream determinants of three sustainability outcomes the operational-control literature treats as exogenous: (i) the renewable-absorption rate of the host grid, (ii) the marginal-emissions intensity of incremental AI workload, and (iii) the water-intensity of cooling, since architectural and site dimensions A2, A5, and S1 govern waste-heat reuse and dry-cooling integration [31,32]. By making the option value of these decisions visible at the design-table moment when they can still be committed, the FDRI–ROV framework operationalizes the link between AI-infrastructure capital allocation and four UN Sustainable Development Goals engaged by AI-era data-center building development [33].

2. Background and Theoretical Foundations

2.1. Demand Response in Data Centers: Operational Paradigm

The data center demand response (DR) literature has developed substantially over the past decade. Early work demonstrated that computational workloads could be shifted spatially across data centers and temporally across time to follow renewable energy availability and electricity-price signals, opening a foundational path for treating data centers as dispatchable grid assets [34,35]. Subsequent contributions extended this to dynamic server provisioning, where the number of active servers tracks workload demand and grid signals jointly [36]. The uninterruptible power supply was reconceptualized as a battery energy storage resource capable of discharging into the grid during peak periods [37]. Cooling-side flexibility (including modulation of computer-room air handlers, chilled-water flow, and economizer modes) has emerged as a complementary lever, particularly in warm-water and free-cooled designs [38,39]. Mechanism design for multi-tenant colocation facilities, where the operator must aggregate flexibility from independent tenants under split-incentive constraints, was treated by Chien et al. [10] and extended to incentive-compatible reverse auctions and time-varying monetary rewards [40,41]. Frequency regulation and ancillary-services participation have been studied as candidate revenue streams that exploit the millisecond-scale responsiveness of modern UPS-BESS systems [12]. The most recent contributions integrate AI-training-cycle volatility explicitly, recognizing that hyperscale GPU clusters now produce intra-hour swings of hundreds of megawatts that legacy DR control architectures were not designed to absorb [42]. The grid-interactive efficient buildings (GEB) paradigm has carried this thinking into the broader built environment, treating buildings as flexible loads with shiftable, interruptible, and adjustable components [13,14,43]. Within the data-center DR literature specifically, however, the physical envelope of the facility (i.e., power topology, switchgear, mechanical-room footprint, water rights, and interconnection capacity) is treated as exogenous: fixed at construction and inherited by operators, so that even the building-oriented strand of this work optimizes control over an envelope it takes as given. This strand has therefore developed sophisticated control algorithms for a fixed asset, while leaving the construction-phase determinants of that asset's flexibility envelope unexamined. The grid-interactive efficient buildings literature, by contrast, does treat physical design decisions as determinants of flexibility, but at the building scale, largely qualitatively, and outside the data-center typology and AI-era load context (Section 2.2).

2.2. Construction-Phase Determinants of Building Flexibility

A separate research stream (largely outside the data center literature) has begun to examine how construction-phase decisions determine the flexibility envelope of buildings. The grid-interactive efficient building paradigm developed by the U.S. Department of Energy and Lawrence Berkeley National Laboratory frames buildings as active grid participants whose flexibility is enabled or constrained by their physical and control infrastructure [44,45]. Within this framing, researchers have examined how thermal mass, envelope U-value, mechanical-system sizing, and zoning topology determine load-shifting potential at the building scale [6,13]. Construction-stage decisions about district heating and cooling interfaces, microgrid readiness, water-reservoir capacity, and parking-canopy load reservations have been linked, in commercial and residential buildings, to lifetime flexibility outcomes [7,8,46]. The smart-buildings literature has further established that decisions made before commissioning carry a 'flexibility imprint' that is costly or impossible to retrofit, which is a finding consistent with the irreversibility logic central to capital-budgeting under uncertainty [47,48]. Recent contributions to MDPI Sustainability have begun to formalize this imprint as a building-level scoring framework, although none have applied it to data center typology [6,8]. Adjacent work in Sustainable Cities and Society and Cities has extended the framing to neighborhood and district scales, emphasizing the role of zoning overlays, microgrid corridor reservation, and community benefit agreements as construction-phase determinants of urban energy flexibility [48,49]. These streams together establish a methodological precedent for treating construction decisions as the primary determinants of operational flexibility; but the data center literature has not picked them up. The grid-interactive efficient buildings literature does recognize, at the building scale and largely qualitatively, that envelope, thermal mass, and mechanical-system sizing decisions made at construction enable or foreclose later flexibility [6,13]; the present contribution does not claim to discover that principle. What remains absent is the transfer of that principle into the data center typology, its formalization as a priced investment decision rather than a qualitative design heuristic, and its empirical calibration to AI-era hyperscale load. To our knowledge, no peer-reviewed framework formalizes the construction-phase decisions that bound data center DR capability as a portfolio of real options valued under stochastic energy, capacity, and regulatory processes, and none calibrates such a valuation against contemporary hyperscale projects. The construction manager, architect, and urban planner therefore remain absent from the data center DR conversation as priced economic actors, despite holding the decisions that ultimately bound it.

2.3. Stochastic Optimal Control Under Investment Irreversibility, and Least-Squares Monte Carlo

A parallel and equally established tradition values flexibility from the design side rather than the financial side. The engineering-options paradigm formalized by de Neufville and Scholtes [23] and by de Neufville et al. [24], and systematized in Cardin's design taxonomy [25], distinguishes options on a system (managerial rights to defer, expand, or abandon the project as a whole, valued externally) from options in a system (flexibility physically engineered into the artifact itself (for example, oversized foundations, modular expansion capacity, or reserved structural and service capacity) so that the built system can adapt after commissioning). This literature has valued embedded design flexibility across parking structures, hospitals, real estate, and infrastructure, and is the methodological antecedent most directly related to the present work. Stochastic optimal control under investment irreversibility (SOCII) provides the formal apparatus for valuing investment when capital outlays are sunk, and management retains the right but not the obligation to undertake subsequent actions. The framework treats the investment problem as a continuous-time stochastic dynamic program: an underlying state vector X(t)

evolves stochastically under a known data-generating process, the decision-maker chooses a stopping or control policy that maximizes expected discounted payoffs subject to irreversibility constraints, and the resulting Hamilton–Jacobi–Bellman equation characterizes the optimal policy and the value function. Real options theory (ROT), introduced by Black and Scholes [18] and Merton [50] and consolidated by Dixit and Pindyck [16] and Trigeorgis [19], is the canonical special case of SOCII corresponding to optimal stopping problems with American-style exercise rights; in effect, it is the financial-economics specialization of the broader stochastic-optimal-control formulation developed in operations research and approximate dynamic programming [16,18,51]. The standard apparatus values the embedded options by replicating the underlying state-variable evolution under the risk-neutral measure $\mathbb{Q}$ and computing the discounted expectation of the optimal exercise policy. The traditional discounted-cash-flow methodology is the further special case obtained when all exercise rights are forfeited at t = 0; SOCII demonstrates formally that this approach systematically undervalues projects with embedded flexibility, because it ignores the option premium associated with the right to defer, expand, contract, switch, or abandon. For control problems with American-style exercise (exercisable at any point during their life rather than only at maturity), closed-form solutions to the Hamilton–Jacobi–Bellman equation are unavailable for most state-variable processes. The Longstaff-Schwartz Least-Squares Monte Carlo (LSM) algorithm [26] resolved this by approximating the continuation value at each backward-induction step via least-squares regression of discounted future payoffs onto a basis of polynomials in the current state variable. Weighted Laguerre polynomials are the recommended basis; convergence is established as both sample size and basis degree increase [26]. LSM is preferred over binomial-lattice and finite-difference alternatives for this class of problem because it accommodates high-dimensional state vectors and path-dependent payoffs without combinatorial explosion, and because the simulated paths can be reused across the multiple options that comprise a portfolio [21,52]. The method has been extended to portfolios of compound options [50,51] and applied across energy-infrastructure investment problems. Indeed, LSM-based real-options valuation (i.e., the ROT specialization of SOCII solved via LSM) has been used to evaluate renewable-electricity investments under Brazilian regulatory uncertainty [21], solar-photovoltaic deployment under Italian incentive uncertainty [22,53], carbon-capture-and-utilization projects [23], and hydrogen-port equipment transitions [54]. These applications establish a methodological precedent for valuing flexibility-bearing infrastructure investments under stochastic prices and policy regimes. No prior contribution, however, has applied SOCII or its ROT-LSM specialization to the construction-phase determinants of data center demand response capability, which is the gap this article fills.

The stochastic-optimal-control/real-options literature provides the formal apparatus for valuing flexibility-bearing investments under uncertainty (and has been applied to renewable, carbon-capture, and port-equipment investments, and, in its engineering-options branch, to design flexibility embedded in physical infrastructure such as parking structures, hospitals, and real estate [24–26]) but has not been applied to the construction-phase determinants of data center demand-response capability.

2.4. Large-Language-Model Elicitation of Expert Judgment

The elicitation of relative-importance judgments in Phase 2 of this study uses large-language-model (LLM) personas in place of human respondents, an approach that has emerged as a distinct methodology over the past three years and that warrants explicit grounding. The foundational demonstration is Argyle et al. [55], who showed that a large language model conditioned on socio-demographic profiles reproduces the response distributions of the corresponding human sub-populations with sufficient accuracy to function

as silicon samples, a property they term algorithmic fidelity. Aher et al. [56] extended this from single responses to the replication of entire human-subject studies, showing that LLM-simulated participants recover well-established experimental effects across economics, psycholinguistics, and social psychology. Park et al. [57] generalized the paradigm to persistent, memory-equipped generative agents whose simulated behavior was judged believable relative to human baselines. Across these contributions the method offers four properties that are advantageous for a multi-criteria weighting problem of the kind addressed here: full reproducibility (the prompts, persona specifications, and model configuration can be released so that any analyst can regenerate the elicitation exactly); transparency and auditability of every judgment; scalability to panel sizes and round counts that would be costly to recruit; and the ability to condition each persona explicitly on a documented evidence base here, rather than on the idiosyncratic and unobservable priors of individual human experts. In addition, Bisbee et al. [58] show that persona-conditioned LLM responses, while recovering population means reasonably well, systematically understate between-respondent variance and are sensitive to prompt wording, so that a synthetic panel tends to compress genuine disagreement and to over-represent the modal view embedded in the model's training distribution.

3. Materials and Methods

3.1. Research Design and Methodological Framework

This article adopts an integrated mixed-methods design comprising five sequential phases: (1) systematic literature review under the PRISMA 2020 protocol [59,60]; (2) a three-round structured expert-elicitation panel implemented with large-language-model-simulated domain-expert personas rather than human participants (Section 3, Phase 2) [61,62]; (3) Fuzzy Analytic Hierarchy Process for dimension weighting under linguistic uncertainty [63,64]; (4) the FDRI–ROV simulation model under the SOCII apparatus reviewed in Section 2.3 (the core analytical contribution); and (5) an illustrative, out-of-sample proof-of-concept validation against four contemporary AI-era hyperscale projects drawn from public sources. Because the elicitation panel was generated synthetically and involved no human subjects, the study required no institutional review-board oversight; the weighting layer is presented as a transparent, reproducible demonstration rather than as a survey of human opinion.

3.2. Phase 1: Systematic Literature Review Protocol

A systematic review [59] of the data-center DR, GEB, and SOCII/real-options-in-energy-infrastructure literature was conducted across Scopus, Web of Science, IEEE Xplore, ScienceDirect, and the MDPI repository for 1 January 2010 through 30 September 2025. The Boolean search string combined a data-center cluster, a flexibility cluster, and a built-environment cluster; pure operational-control studies were excluded. After deduplication and full-text screening, 68 sources were retained and synthesized through inductive thematic coding, producing the initial flex-lock taxonomy that seeded Phase 2.

The review was designed, conducted, and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [59,60]. full identification, screening, eligibility, and inclusion pipeline is summarized in Figure 1. Of 498 records identified—481 from the five databases (Scopus, n = 142; Web of Science, n = 118; IEEE Xplore, n = 76; ScienceDirect, n = 94; MDPI repository, n = 51) and 17 from backward-citation tracking of seminal works—174 duplicates were removed, leaving 324 records screened on title and abstract. After excluding 213 records (operational-control-only, IT/software-only, non-English, and off-topic studies), 111 full-text articles were assessed for eligibility, of which 43 were excluded (no construction-phase content, gray literature, or not relevant to the data-center demand-response, grid-interactive-buildings, or

real-options literature). Sixty-eight studies were retained in the final corpus and synthesized by inductive thematic coding across the three literature threads.

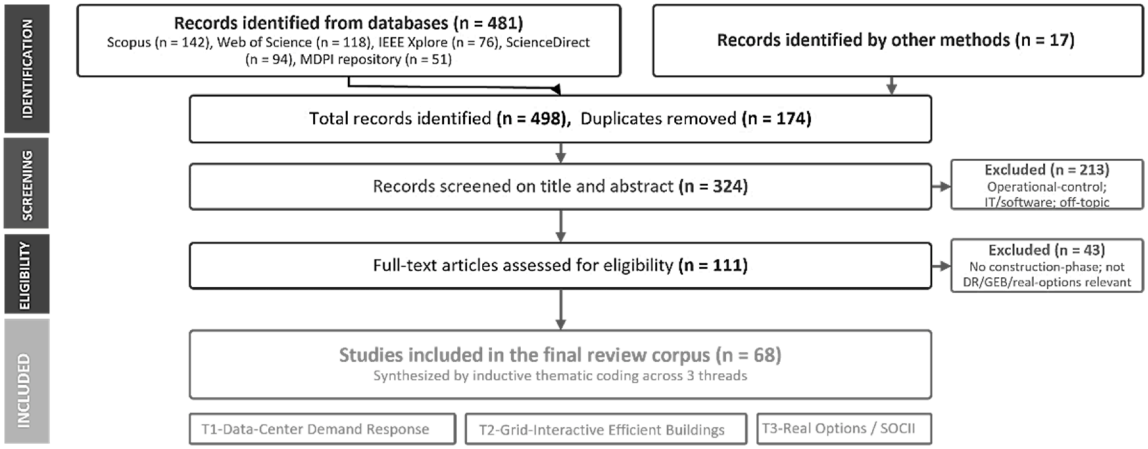


Figure 1. PRISMA 2020 flow diagram for the systematic literature review used to identify the FDRI candidate dimensions.

3.3. Phase 2: Three-Round Synthetic Expert-Elicitation Panel (LLM-Simulated Personas)

Following the LLM-elicitation methodology reviewed in Section 2.4, the relative-importance judgments required for the FDRI weighting were obtained through a structured synthetic elicitation rather than from human respondents. The three-round elicitation was conducted entirely with model-generated domain-expert personas, a design choice that yields a fully reproducible, auditable, and literature-grounded weighting prior; because no human beings participated and no personal data were collected, the procedure also does not constitute human-subjects research. The elicitation was conducted via a large language model using Anthropic Claude (Claude 3.5 Sonnet, model version claude-3-5-sonnet-20241022; accessed via the Anthropic API with the following decoding hyperparameters held constant across all twenty-five personas and all three rounds: temperature = 0.7, balancing response diversity against coherence for elicitation tasks; top-p = 0.95, nucleus sampling; maximum output tokens = 4096; no system-level frequency or presence penalties), which was conditioned on the Phase 1 literature corpus and instantiated as twenty-five distinct domain-expert personas spanning data-center engineering (n = 8), construction management (n = 6), grid system operations (n = 4), academia (n = 4), and urban planning (n = 3), each persona specified with a professional profile and an assigned regional context (North America n = 13, Europe n = 8, Asia-Pacific n = 4). Round 1 prompted each persona to propose candidate flex-locks via an open-ended instrument, organized by the architectural, MEP, and site-urban layers identified in Phase 1. Round 2 returned the consolidated dimension list for ranking on a five-point ordinal scale of design-phase irreversibility and DR leverage. Round 3 returned each persona's prior ratings together with the ensemble distribution and prompted re-rating. Dispersion across personas was summarized by Kendall's W [64] with a stopping rule of $W \geq 0.7$ or three rounds. It must be emphasized that, for a synthetic ensemble, this statistic does not measure independent inter-expert agreement in the sense it carries in a human Delphi panel: the dispersion among model-generated personas is governed by the decoding hyperparameters of the language model (in particular temperature and top-p) rather than by autonomous cognition, so a high W reflects the stability of the model's generations under the fixed configuration disclosed here, not convergence of independent judgments. Kendall's W is therefore used here strictly as a generation-stability stopping criterion and is explicitly not interpreted as evidence of expert consensus [65]. The procedure follows the multi-round structure

of the Delphi protocol [66,67] but is an entirely synthetic, model-generated elicitation. The complete persona specifications, prompts, model configuration, and round-by-round outputs are released in the reproducibility archive (Data Availability) so that the elicitation is fully auditable and reproducible. The resulting weights are accordingly to be read as a reproducible, literature-grounded prior rather than as evidence of human expert consensus; this limitation is revisited in Section 6.4.

Sensitivity to the choice of language model. Because the weighting prior is generated by a single model family, its specific values are conditional on that model and its training distribution, and this dependence is a first-order limitation rather than an implementation detail. The full decoding configuration is reported in Section 3.1 and archived for exact reproducibility, but reproducibility under a fixed model is not the same as invariance across models: a different model family, or a different version of the same family, is trained on a different corpus mixture and would in general yield a different prior, particularly for the disciplinary-anchoring effect discussed in Section 6.4, since the relative prevalence of MEP-centric versus architectural/site-centric demand-response literature in the training data is precisely what the prior reflects. We did not run the elicitation across multiple model families or versions in this study, so we do not report a cross-model sensitivity band, and we explicitly refrain from claiming that the weight vector is model-invariant. The appropriate test is repeating the three-round elicitation under at least two additional model families and one earlier and later version of the primary family and reporting the rank correlation of the resulting weight vectors. Until that test is performed, the FAHP weights should be read as the prior of one disclosed model under one disclosed configuration, which is exactly why the central FAHP–Sobol comparison is framed in this revision as a screening diagnostic between a literature-grounded model prior and the model's own financial-leverage estimates, rather than as a finding about human experts.

3.4. Phase 3: Fuzzy Analytic Hierarchy Process for Dimension Weighting

Dimension weights were derived using Fuzzy Analytic Hierarchy Process (FAHP) to accommodate linguistic uncertainty in the synthetic-expert pairwise comparisons. Classical AHP [63] is extended through the fuzzy formulation of van Laarhoven and Pedrycz [68] and Chang's [62] extent-analysis method. Following reviewer guidance, the Logarithmic Least Squares (geometric-mean) Method (LLSM) [69] is adopted as the reference weighting estimator for this study, because it derives every criterion weight from a geometric-mean aggregation of the fuzzy comparisons and therefore avoids the zero-weight pathology of extent analysis by construction. Chang's extent-analysis method is retained as a transparency cross-check rather than as the primary estimator. The two estimators yield weight vectors that are numerically equivalent to the precision reported here (Section 3.2); therefore, the dimension weights are method-invariant at the reported resolution, and are presented as the FAHP weights under the LLSM reference estimator. Each pairwise comparison is encoded as a triangular fuzzy number $\tilde{a} = (l, m, u)$ with membership

$$\mu\tilde{a}(x) = (x - l)/(m - l) \text{ for } l \le x \le m;\ (u - x)/(u - m) \text{ for } m \le x \le u;\ 0 \text{ otherwise.} \quad (1)$$

Pairwise comparisons among the n dimensions yield, for each expert $k \in \{1, \ldots, K\}$, a fuzzy reciprocal matrix $\tilde{A}(k) = [\tilde{a}ij(k)]$. The K matrices are aggregated through the fuzzy geometric mean:

$$\tilde{a}ij = (\prod k = 1K\ \tilde{a}ij(k))1/K, \text{ for all } i, j = 1, \ldots, n. \quad (2)$$

Synthetic extent values are computed under Chang's method: for row i,

$$Si = (\sum j = 1n\ \tilde{a}ij) \otimes (\sum i = 1n\ \sum j = 1n\ \tilde{a}ij) - 1, \quad (3)$$

with addition and multiplication in fuzzy-arithmetic sense. The degree of possibility that $S2 \geq S1$ is:

$$V(S2 \geq S1) = 1 \text{if } m2 \geq m1;\ (l1 - u2)/[(m2 - u2) - (m1 - l1)] \text{ if } l1 \leq u2;\ 0 \text{ otherwise.} \quad (4)$$

yielding the minimum-degree-of-possibility weight:

$$d'(Ai) = \min k \neq i\ V(Si \geq Sk);\ wi = d'(Ai)/\sum j = 1 n d'(Aj),\ \sum i\ wi = 1. \quad (5)$$

Each expert's matrix $\tilde{A}(k)$ was checked against Saaty's CR < 0.10 threshold [63]; non-compliant matrices were returned for re-elicitation. Weight robustness was further assessed through 1000-replication non-parametric bootstrap over the synthetic persona ensemble [70]: for each replicate, K personas were drawn with replacement, Equations (2)–(5) re-applied, and the 95% confidence interval reported as the empirical 2.5th–97.5th percentiles of the resulting bootstrap distribution.

A known limitation of Chang's extent-analysis method warrants explicit treatment. The method computes each weight from the minimum degree of possibility that one fuzzy synthetic extent dominates the others (Equation (5)), and when the triangular fuzzy numbers do not sufficiently overlap, this minimum can collapse to zero, so that extent analysis can assign a zero weight to a criterion that is in fact valid—a pathology documented in the decision-sciences literature [71]. We therefore checked for it directly. In the present application, none of the nineteen dimensions received a zero weight: every dimension obtained a strictly positive degree-of-possibility value $d'(Ai)$ before the normalization in Equation (5), so no valid sub-criterion was erroneously eliminated. This is a consequence of the comparison structure here; the synthetically consolidated matrices contain no dimension that is dominated by every other on all comparisons, which is the configuration that triggers the zero-weight collapse. Chang's method was nonetheless retained for three reasons: its computational transparency and wide adoption in construction and AEC multi-criteria applications make the weights directly comparable with that literature; its inputs (triangular fuzzy comparisons) map cleanly onto the linguistic synthetic-elicitation responses; and, because the zero-weight pathology did not arise, its principal documented drawback does not affect these results. Accordingly, LLSM (which is not subject to the zero-weight pathology by construction) is used as the reference weighting estimator in this study, and the Chang extent-analysis weights are reported alongside it as a transparency cross-check. The released analysis package reports both weight vectors and their rank correlation; the Chang and LLSM weights are numerically equivalent to the precision reported here and agree in rank ordering, so the FDRI composite, the dimension weights and the downstream Δi diagnostic are unaffected by the choice between the two FAHP estimators. Chang's method is retained in the exposition only for its computational transparency and its comparability with the wider construction and AEC multi-criteria literature, in which extent analysis is the most widely reported FAHP variant.

Figure 2 summarizes the four-tier FAHP decision structure: TIER 0 (overall goal (composite FDRI score)); TIER 1 (three FDRI layers (Architectural $\mathcal{A}$, MEP $\mathcal{M}$, Site/Urban $\mathcal{S}$) with FAHP-aggregated weights $\Sigma w\mathcal{A} \approx 0.33$, $\Sigma w\mathcal{M} \approx 0.40$, $\Sigma w\mathcal{S} \approx 0.27$); TIER 2 (nineteen FDRI dimensions, each annotated with its individual w_i (Equations (1)–(5))); TIER 3 (four anonymized hyperscale validation cases (see Section 5.5)).

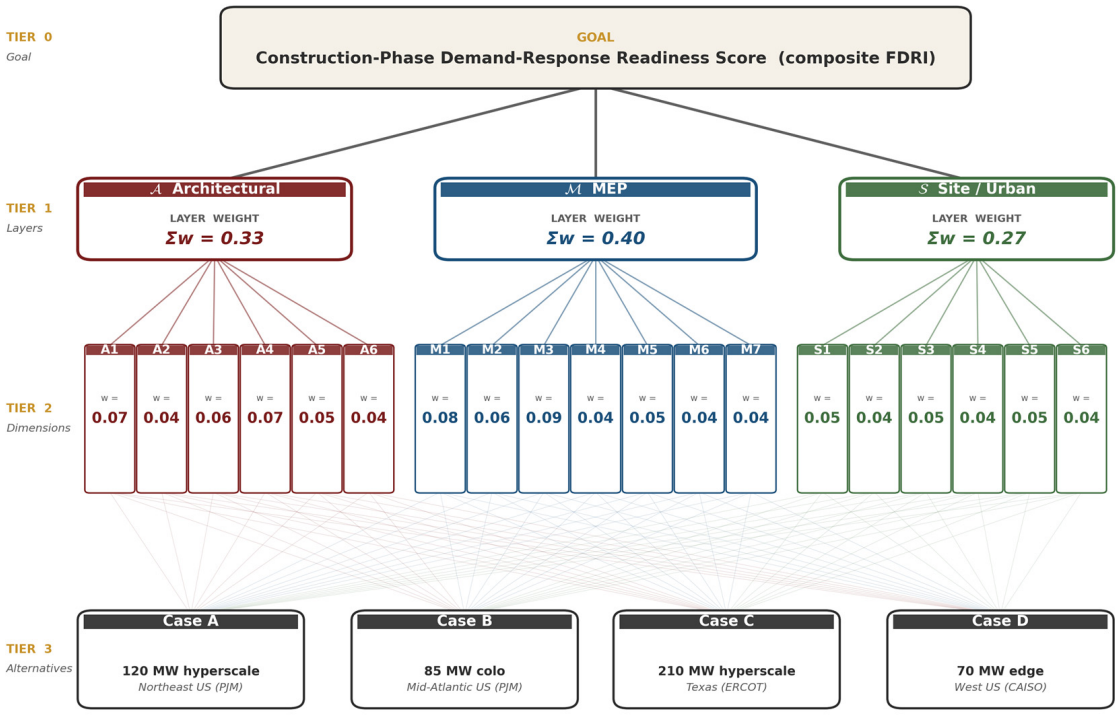


Figure 2. Fuzzy Analytic Hierarchy Process (FAHP) decision structure for FDRI weight derivation.

3.5. Phase 4: FDRI–ROV Simulation Model Architecture

The FDRI–ROV simulation model integrates the SOCII apparatus reviewed in Section 2.3 (with classical real options theory as its optimal-stopping specialization), Least-Squares Monte Carlo (LSM), Latin Hypercube Sampling (LHS), System Dynamics, and Sobol global sensitivity analysis into a single computational pipeline. The pipeline is implemented in Python (3.12.0 version); the Least-Squares Monte Carlo and stochastic simulation routines use NumPy (2.5.1) for array operations and pseudo-random path generation and SciPy (v1.18.0) for the regression and statistical functions (the Laguerre-basis continuation regression of Equation (11) is solved by ordinary least squares), market and parameter data are handled with pandas (v3.0.3), and the global sensitivity analysis uses SALib (v1.5.2) [71]; exact package versions are pinned in the requirements.txt of the released archive. Each construction-phase flex-lock is treated as an American-style call option on a specific operational flexibility, purchased by paying marginal CapEx C_i and exercisable at any time $t \in [0, T]$ over the 30-year operational life T. The state vector $X(t) \in \mathbb{R}^5$ comprises wholesale-market energy price E(t), capacity-market price K(t), ancillary-services price A(t), regulatory regime $R(t) \in \{\text{RELAXED, BASELINE, STRICT}\}$, and AI-workload index W(t), evolving under the risk-neutral measure $\mathbb{Q}$ as:

$$dEt = \mu E\, Et\, dt + \sigma E\, Et\, dWtE,\ \ dKt = \mu K\, Kt\, dt + \sigma K\, Kt\, dWtK, \tag{6}$$

$$dAt = \theta A\, (\overline{\mu} A\ -\ At)\, dt\ +\ \sigma A\, dWtA, \tag{7}$$

$$\mathbb{P}(Rt + \Delta t = j \mid Rt = i) = \pi ij\, \Delta t + o(\Delta t),\ \ i, j \in \{\text{RELAXED, BASELINE, STRICT}\}, \tag{8}$$

$$dWt = \mu W\, Wt\, dt + \nu t\, Wt\, dWtW,\ \ d\nu t^2 = \kappa\nu\, (\eta\nu\ -\ \nu t^2)\, dt + \xi\nu\, \nu t\, dZt. \tag{9}$$

Equations (6) and (9) specify Geometric Brownian Motions for energy, capacity, and AI-workload, with the latter carrying Heston-Cox-Ingersoll-Ross stochastic volatility. Equation (7) is an Ornstein-Uhlenbeck mean-reverting process for ancillary-services price [72]. Equation (8) specifies the continuous-time Markov chain governing the regu-

latory regime, with infinitesimal generator $\Pi = [\pi ij]$. Parameters are calibrated to PJM, ERCOT, or CAISO 2024–2026 statistics per Section 3.6.

The Geometric Brownian Motion and Ornstein-Uhlenbeck specifications adopted in Equations (6)–(9) are the standard continuous-path baselines for energy-price, capacity-price, and workload modeling [72], chosen here for analytical transparency and for comparability with the energy-infrastructure real-options literature against which the model is benchmarked [21,23]. They have a known limitation that is directly relevant to flexibility valuation: continuous-path diffusions do not reproduce the discrete scarcity spikes that characterize wholesale electricity markets, such as ERCOT scarcity-pricing events or PJM cold-snap excursions, which a Merton-type jump-diffusion of the form $dE/E = \mu\, dt + \sigma\, dW + (J - 1)\, dN(\lambda)$, with N a Poisson jump counter and J the jump-size multiplier, is specifically designed to capture. The direction of the resulting bias depends on how the jump component is calibrated, and we are careful not to overstate it. Because each flex-lock is valued as a convex (option-like) payoff in the underlying price, and because price spikes are precisely the states in which exercised flexibility earns its largest revenue, a specification that adds upward scarcity jumps under the physical measure (so that the jumps raise the upper tail without being offset) increases the option premia relative to the continuous-path baseline; in that case the GBM model is conservative and the reported premia are a lower bound. Under a risk-neutral, compensated jump calibration, however, the drift adjustment that holds the expected price fixed can offset the convexity gain, so the net effect on the premia is calibration-dependent and not signed a priori. This is precisely why a formal comparative evaluation is warranted rather than a qualitative assertion: the magnitude, and in the compensated case even the sign, of the GBM-versus-jump gap is an empirical question about scarcity-event frequency and amplitude in each market. A formal comparative evaluation against jump-diffusion and regime-dependent-volatility alternatives, together with parameter stress testing under extreme-scenario calibrations (elevated jump intensity and amplitude, compressed mean-reversion, and STRICT-regime persistence), is quantified for the physical-measure baseline in Section 5.2 and developed more fully as a risk-neutral robustness extension and the future-work agenda of Section 6.5; the released code package is structured so that the diffusion process can be swapped for a jump-diffusion driver without altering the valuation or sensitivity machinery.

The value of flex-lock i with payoff is the supremum over stopping times of the discounted expected payoff, evaluated under the hybrid valuation measured below:

$$Vi = \sup \tau \in \mathcal{T}\ [0,\ T]\ \mathbb{EQ}\ [e - r\tau\ \phi i(X\tau)|X0], \tag{10}$$

where, $\mathcal{T}[0, T]$ is the family of stopping times in [0, T] and r is the constant risk-adjusted real discount rate (see the measure discussion below). Following Longstaff and Schwartz [26], the continuation value at decision time t is approximated by least-squares regression of discounted future payoffs onto weighted Laguerre polynomials in the state variable:

$$\hat{C}(Xt) = \sum k = 0K\ \hat{\beta}k\ L\ k(Xt)\ \exp(-Xt/2), \tag{11}$$

where {Lk} are Laguerre polynomials of degree k, K is the basis order (default K = 3), and $\hat{\beta}$ is the OLS coefficient vector. The optimal stopping rule exercises if intrinsic payoff exceeds estimated continuation value:

$$\hat{\tau}^*(\omega) = \inf \left\{ t \in [0,\ T]\ :\ \phi t(Xt(\omega))\ \geq\ \hat{C}(Xt(\omega)) \right\}. \tag{12}$$

Equation (10) is written as an American-style optimal-stopping problem rather than a single fixed-exercise option because this is the canonical generalization that nests the simpler option types as special cases, not a narrowing of the valuation to one form. The

supremum over the full family of stopping times $\mathcal{T}[0, T]$ reduces to a European option when the admissible exercise set is the single terminal date {T}, and to a fixed-horizon deferral (timing) option when it is any single intermediate date; permitting exercise at every decision epoch is therefore the most general of the three and dominates them in value by construction (the American value is an upper envelope of the European and deferral values for the same payoff). Each flex-lock in the FDRI taxonomy is an irreversible construction-phase commitment whose optimal timing is endogenous (the owner may provision the capability at pre-design, defer it to a later design gate, or abandon it once the design freezes) so the optimal-stopping form is the type-appropriate specification rather than a convenience choice [20].

The framework consequently spans more of the real-options typology than the single American-call label suggests, and the mapping is explicit. (i) Deferral/timing options are captured directly by the endogenous stopping time $\hat{\tau}^*$ of Equation (12) and quantified by the decision-timing analysis of Section 4.4, in which the set of dimensions still under management control contracts from nineteen at pre-design to four at construction administration. (ii) Compound (staged) options arise wherever several dimensions share one revenue stream (the waste-heat cluster (A2/A5/S1), the behind-the-meter-export cluster (M1/M3/S3), and the regulatory-avoided-cost dimension S6) where the value of the downstream actuator is contingent on the upstream enabler having been exercised first; the cluster-level substitution correction of Section 5.3 and the Gamba Monte-Carlo compound-option treatment [52] address exactly this nesting. (iii) Abandonment/contraction options are present because any flex-lock not provisioned before its irreversibility gate (slab-pour, substation interconnection) is permanently forgone, which is the limiting case of a contraction option with zero salvage. (iv) Switching options (the ability to alternate operating modes once built (e.g., grid-export versus on-site consumption)) are the one archetype the present construction-phase model does not value, because its decision variable is the binary build/skip commitment rather than the post-commissioning operating policy.

Two valuation extensions therefore remain open and are flagged as such rather than claimed. First, the shared-revenue clusters are currently valued as a sum of American options with an ex-post substitution correction; a fully nested compound-option estimator [52] would price each cluster as a single option on its joint revenue and is expected to tighten, not widen, the reported premia (the additive figures are upper bounds, per Section 5.3). Second, an operating-phase switching option layered on the as-built facility would add value beyond the construction-phase commitment valued here; quantifying it requires an operating-policy state that is outside the present design-decision scope and is identified as future work in Section 6.5. These boundaries delimit the valuation; they do not reduce it to a single option type.

A measure-theoretic clarification of Equation (10) is warranted because the four state variables in Equations (6)–(9) differ in their tradability and the Longstaff-Schwartz estimator inherits the pricing logic of whichever measure is imposed. The energy, capacity, and ancillary-services price factors (Equations (6) and (7)) are, to a first approximation, spanned by liquid traded instruments, so the risk associated with these factors can in principle be hedged and an equivalent risk-neutral measure $\mathbb{Q}$ is well defined for them; their drifts may therefore be replaced by the risk-free rate without violating no-arbitrage. The AI-workload and regulatory-regime factors (Equations (8) and (9)) are fundamentally different: neither is a traded asset, no portfolio of marketed securities replicates a claim contingent on GPU-cluster utilization or on a BASELINE-to-STRICT regulatory transition, and the market is consequently incomplete. In an incomplete market, there is no unique equivalent martingale measure and no market price of risk for these factors that can be recovered from no-arbitrage arguments alone [17,73]; mechanically discounting them under a risk-

neutral measure, as if they were spanned, would be precisely the misspecification that arbitrage pricing theory warns against. Accordingly, the valuation in Equation (10) is not a pure risk-neutral price but a hybrid (real-world) valuation. The two non-traded factors are simulated under the physical measure $\mathbb{R}$ using the empirically estimated real-world drifts and transition intensities of Section 3.6 (the GBM workload drift $\hat{\mu}W$, and the calibrated generator Π), rather than under any imputed risk-neutral dynamics; the spanned price factors retain their $\mathbb{Q}$-dynamics. Risk attaching to the non-traded factors is not priced through a fabricated market price of risk but is carried by the discount rate: r in Equation (10) is a risk-adjusted real rate (r = 0.06, above the real risk-free rate), so that Equation (10) is read as a real-world expectation of the option payoff discounted at a rate that embeds the premium investors require for bearing the unhedgeable workload and regulatory risk. This is the standard treatment of real options written on non-traded underlyings [17,73], and it nests the complete-market risk-neutral case as the special configuration in which every factor is spanned, and r reduces to the risk-free rate. The consequence for interpretation is that the reported premia are conditional on the chosen risk-adjusted discount rate rather than being unique no-arbitrage prices; the sensitivity of the results to r is therefore reported as a robustness check.

Portfolio-level option value sums each lock's individually estimated Vi; FAHP weights from Phase 3 apply only to the composite FDRI on the 0–100 scale (Equation (17)), not to dollar-denominated option values. Parameter uncertainty is propagated through Latin Hypercube Sampling (LHS) over the parameter ranges with N = 10,000 Monte Carlo paths per LHS draw. A System Dynamics sub-model of four interacting stocks—renewable share R(t), regulatory stringency G(t), AI-workload growth W(t), and flex-market participation F(t)—captures endogenous feedback between FDRI score and the state vector [27]:

$$dR/dt = \alpha R\,(1 - R) - \delta R\,R + \beta R \cdot FDRI/100, \tag{13}$$

$$dG/dt = \gamma G \cdot \xi G(R, F, FDRI) - \omega G\,G, \tag{14}$$

$$dW/dt = \mu W\,W\,(1 + \epsilon \cdot \nu t),\ dF/dt = \alpha F\,(FDRI/100)\,(1 - F) - \delta F\,F. \tag{15}$$

Referring to Table 1, rate parameters governing renewable-share uptake (μW and αR) are anchored to the IEA Base-Case growth figures for data-center AI-workload and renewables generation, respectively [1], and the diffusion forms for renewable-share and flex-market participation follow standard logistic/Bass adoption dynamics [27]. The coupling and relaxation constants (βR, γG, ωG, ϵ, and the δ attrition terms) are illustrative structural parameters chosen to produce qualitatively reasonable feedback and are not calibrated to data; they are reported transparently as such.

Table 1. Parameters of the illustrative System-Dynamics submodel, Equations (13)–(15).

Parameter	Equation	Value	Basis/Derivation	Source/Status
αR (renewable-share uptake rate)	(13)	0.18 yr^{-1}	Logistic-uptake rate for renewable share; set to the IEA Base-Case renewables-for-data-center growth band ($\approx$22% yr^{-1} generation growth, 2024–2030), discounted to a saturating-share rate.	IEA [1]; logistic form per Sterman [27]
δR (renewable-share decay)	(13)	0.02 yr^{-1}	Slow attrition/retirement term keeping R bounded; nominal value, one order below αR.	Illustrative structural [27]
βR (FDRI $\rightarrow$ renewable coupling)	(13)	0.10	Strength of FDRI-score pull on renewable adoption; illustrative coupling, not calibrated to data.	Illustrative structural

Table 1. *Cont.*

Parameter	Equation	Value	Basis/Derivation	Source/Status
γG (regulatory-stringency gain)	(14)	0.15 yr^{-1}	Response gain of regulatory stringency to the ξG driver; illustrative.	Illustrative structural
ωG (regulatory-stringency relaxation)	(14)	0.05 yr^{-1}	Relaxation/decay of stringency absent pressure; illustrative, one third of γG.	Illustrative structural
μW (AI-workload growth rate)	(15)	0.15 yr^{-1}	Baseline AI-workload electricity-demand growth; IEA Base Case ≈15% yr^{-1} aggregate data-center demand 2024–2030 (accelerated-server subset ≈30% yr^{-1}, used in stress runs).	IEA [1]
ϵ (workload volatility-coupling)	(15)	0.10	Amplitude of the stochastic workload modulation νt; illustrative scaling of the demand-swing term.	Illustrative structural
αF (flex-participation uptake rate)	(15)	0.20 yr^{-1}	Bass/logistic technology-diffusion uptake rate for flex-market participation, FDRI-modulated.	Diffusion form per Sterman [27]
δF (flex-participation decay)	(15)	0.03 yr^{-1}	Attrition of flex participation absent FDRI pull; nominal, bounds F.	Illustrative structural [27]

Two points should be made explicit about the role of this submodel. First, the System-Dynamics layer is diagnostic and illustrative: it traces qualitative 30-year co-evolution of renewable share, regulatory stringency, AI-workload, and flex-market participation under different FDRI scores, but it does not feed back into the CapEx schedule or the option valuation of Equations (6)–(12). The Net NPV and option-value results reported in Section 5 are produced entirely by the LSM real-options engine and are independent of the parameter choices in Table 1. Second, because the layer is illustrative rather than predictive, the structural constants in Table 1 are presented as transparent assumptions; sensitivity of the qualitative trajectories to these constants is bounded and does not propagate to the headline economic results.

Sobol variance-decomposition indices are computed across the nineteen-dimension scores using SALib's Saltelli estimator [71] to identify high-leverage flex-locks. The first-order and total-order indices are

$$\mathrm{Si} = \mathrm{VarXi}(\mathbb{E}[\mathrm{Y}|\mathrm{Xi}])/\mathrm{Var}(\mathrm{Y}), \quad \mathrm{ST,i} = 1 - \mathrm{VarX} \sim \mathrm{i}(\mathbb{E}[\mathrm{Y} \mid \mathrm{X} \sim \mathrm{i}])/\mathrm{Var}(\mathrm{Y}), \qquad (16)$$

where Y is the model output (Net NPV), and X~i denotes all inputs other than Xi. The Saltelli estimator with Nbase = 64 base samples yields 2(d + 2)·Nbase = 2560 model evaluations for d = 19, with bootstrap-resampled confidence intervals on each index.

3.6. Phase 5: Reference-Plant Calibration and Multi-Case Validation

The model is calibrated to a 100 MW IT-load AI-optimized reference plant using day-ahead and real-time wholesale-market data from PJM, ERCOT, and CAISO (January 2022 to September 2025). GBM drift $\hat{\mu}$ and volatility $\hat{\sigma}$ are fitted by maximum likelihood on monthly log returns; OU parameters are fitted through the discrete AR(1) representation $r_t = a + b\, r_{t-1} + \varepsilon_t$ [72,74]. The regulatory-regime transition matrix Π is calibrated to the implementation timelines of the German Energieeffizienzgesetz [73], the EU Cloud and AI Development Act [75], and U.S. state-level data-center carve-outs through 2024–2025. Because no U.S. instrument comparable to the EnEfG yet provides a multi-year stringency time series, the European instruments are used here as proxy scenario bounds for a plausible U.S. regulatory-stringency trajectory rather than as literal U.S. transition probabilities: their published phase-in schedules define the timing and magnitude of the BASELINE-to-STRICT transition, which is then anchored to observed U.S. state-level carve-outs to

set the RELAXED-state persistence. This is a deliberately conservative scenario device, not an empirically estimated U.S. generator matrix, and the sensitivity of the headline results to it is examined directly in the Sobol decomposition and discussed as a boundary condition in Section 6.4. The calibrated FDRI–ROV is exercised against four hyperscale projects spanning the three markets (Cases A–D, anonymized at operator request); for each case, the predicted option premium at the project's actual FDRI score is compared with observed flex-program revenue from public regulatory filings and earnings disclosures, using Mean Absolute Percentage Error.

To make the calibration fully reproducible rather than procedural, Table 2 records the exact dataset name, data feed or report, timeframe, and access URL underlying each fitted parameter of Equations (6)–(9). Referring to Table 2, each parameter group is mapped to its named public dataset, the specific data feed or report within that source, the calibration timeframe, and the access URL. PJM energy and ancillary-services series are drawn from the named Data Miner 2 feeds; capacity parameters are drawn from the per-delivery-year RPM Base Residual Auction reports. ERCOT and CAISO series replicate the PJM procedure for the cross-market calibrations. The BloombergNEF AI-workload projection is a proprietary subscription product and is therefore traceable through the citation [2] and the access date rather than a public URL; the regulatory-regime generator Π is a proxy device built from the statutory phase-in schedules of the European instruments [73,75], as detailed in the preceding paragraph and Section 6.4. The exact data-access dates for each pull are recorded in the released data package [76]; PJM Real-Time feeds are archived after two years, and pre-2024 snapshots are additionally retrievable through the Internet Archive Wayback Machine.

$$\mathrm{MAPE} = (1/n)\cdot \sum j = 1n \ | yjobs - yjpred | \ / \ | yjobs | \quad (17)$$

as the validation metric across n = 4 cases.

Table 2. Calibration data sources for the state-variable processes of Equations (6)–(9).

Parameter(s)	Source Dataset	Specific Feed/Report	Timeframe	Access URL
E0, μE, σE (energy LMP)	PJM Data Miner 2	Day-Ahead Hourly LMPs (da_hrl_lmps); Real-Time Hourly LMPs (rt_hrl_lmps)	January 2022–September 2025	https://dataminer2.pjm.com (accessed 7 July 2026)
K0, μK, σK (capacity)	PJM RPM capacity market	Base Residual Auction Reports, Delivery Years 2024/25, 2025/26, 2026/27	BRAs cleared 2023–2025	https://www.pjm.com (accessed 7 July 2026)
A0, θA, $\bar{\mu}$A, σA (ancillary)	PJM Data Miner 2	Ancillary Service LMPs (regulation clearing price + mileage ratio)	January 2022–September 2025	https://dataminer2.pjm.com (accessed 7 July 2026)
Cross-market replication	ERCOT MIS; CAISO OASIS	Day-ahead & real-time LMP/SPP (ERCOT); PRC_LMP (CAISO OASIS)	January 2022–September 2025	https://www.ercot.com/mktinfo (accessed 7 July 2026); http://oasis.caiso.com (accessed 7 July 2026)
W0, μW (AI-workload demand)	BloombergNEF [2]	US Data Center Power Demand Forecast 2024–2035 (central scenario)	2024–2035 projection	[2]
Π (regulatory regime)	EnEfG [73]; EU Cloud&AI Act [75]	Statutory phase-in schedules (proxy stringency bounds)	2023–2030 timelines	https://www.gesetze-im-internet.de/enefg (accessed 7 July 2026); [73,75]

During the preparation of this manuscript, the authors additionally used Anthropic Claude (Claude 3.5 Sonnet) for data-analysis visualization. The authors have reviewed and

verified all model-generated inputs and outputs and take full responsibility for the content of this publication.

4. The FDRI Model

4.1. Taxonomy and Dimensional Structure

Before cataloging the dimensions, we clarify the physical-economic principle that makes the FDRI coherent. Although the nineteen elements differ in engineering modality, each represents the same real-options object priced in Sections 4.3 and 5: a construction-phase design decision that is committed before or during construction, economically irreversible once locked in, as formalized in Section 2.3 [17,24,26], and capable of creating the right to earn a defined grid-flexibility or demand-response revenue stream. Thus, each dimension is a design-embedded real option whose uncertainty is modeled through Equations (6)–(9) and whose payoff is the flex-lock value in Equation (10). This is why physically different dimensions, such as switchgear ramping and protection-relay settings (M2) and zoning-overlay/community-benefit-agreement framing (S6), can enter the same index. While M2 enables fast electrical load modulation and S6 enables siting-dependent revenues such as waste-heat export or regulatory avoided cost, both create or foreclose a flexibility revenue option. Accordingly, the unit of account is optionality created per irreversible design decision, not engineering function. All dimensions are scored on the same 0–4 readiness scale defined in Section 4.2, and the architectural, MEP, and site-urban layers group them only by actor and decision phase. All nineteen dimensions therefore remain instances of the same construction-phase-option construct, derived through the structured, literature-grounded procedure described in Section 3 rather than assembled ad hoc.

The nineteen dimensions are a derived construct rather than an ad hoc compilation of engineering features. They were generated through the structured procedure detailed in Section 3: a systematic review of the data center demand-response, grid-interactive-buildings, and energy-infrastructure real-options literature (68 retained sources) seeded an initial candidate set through inductive thematic coding; that set was then consolidated, ranked, and re-rated across a three-round synthetic expert-elicitation panel of twenty-five large-language-model-simulated domain-expert personas until the pre-specified dispersion threshold was reached (Kendall's W = 0.74); and the resulting dimension weights were derived under Fuzzy Analytic Hierarchy Process with each persona matrix screened against Saaty's consistency-ratio threshold (CR < 0.10) and reported with bootstrap confidence intervals. The taxonomy is therefore the auditable output of a literature-grounded, reproducible, model-elicited construct-development process, and each dimension satisfies the observability, cross-reference, and third-party-verifiability requirements specified for the scoring anchors in Section 4.2. The Flex-by-Design Readiness Index (FDRI) model organizes construction-phase determinants of data center demand response capability into nineteen distinct dimensions distributed across three layers. The layering reflects both the disciplinary actor primarily responsible for the decision and the project phase at which the decision is locked in. Layer 1 (architectural, indexed by $\mathcal{A} = \{A1, \ldots, A6\}$) comprises six dimensions determined principally by the architect during schematic and design-development phases: thermal mass and chilled-water storage volume (A1), mechanical-room oversizing and access pathways (A2), equipment-yard reservation for behind-the-meter generation and battery pads (A3), modular white-space subdivision (A4), building envelope and heat-recovery interface (A5), and roof and facade load reservation for photovoltaic deployment (A6). Layer 2 (mechanical-electrical-plumbing, MEP, indexed by $\mathcal{M} = \{M1, \ldots, M7\}$) comprises seven dimensions determined principally by the construction manager and MEP-engineer during design development and construction documents: power topology and redundancy

architecture (M1), switchgear ramping and protection-relay settings (M2), battery-energy-storage-system and uninterruptible-power-supply oversizing, C-rate, and grid-tie capability (M3), backup-generator dispatchability (M4), conduit, busway, and feeder oversizing for future behind-the-meter integration (M5), submetering granularity at rack, row, and facility levels (M6), and cooling-system modulation including variable-speed, free, and warm-water topologies (M7). Layer 3 (site and urban, indexed by $\mathcal{S} = \{S1, \ldots, S6\}$) comprises six dimensions determined principally by the urban planner and the site-acquisition team prior to schematic design: heat-sink proximity to district energy, industrial, or agricultural off-takers (S1), water rights and reservoir capacity for cooling flexibility (S2), dual-feeder utility access and substation proximity (S3), microgrid corridor and easement strategy (S4), interconnection-queue position and pre-negotiated flex-tariff rights (S5), and zoning overlay and community benefit agreement framing (S6). The full nineteen-dimension index set is then $\mathcal{D} = \mathcal{A} \cup \mathcal{M} \cup \mathcal{S}$, with $|\mathcal{D}| = 19$. Cross-layer interactions are explicit in the framework rather than treated as residual error. S3 (dual-feeder access) and M1 (power topology) jointly determine the asset's maximum BESS-export capability; A1 (thermal mass) and M7 (cooling-system modulation) jointly bound the duration over which cooling load can be shifted; S1 (heat-sink proximity) and A5 (heat-recovery interface) together determine the addressable waste-heat market. Layer-3 decisions about water rights bound layer-1 cooling-storage choices, and layer-2 power-topology decisions bound the operational use of layer-3 microgrid corridors. Table 3 catalogs the full nineteen-dimension taxonomy with definitions, primary disciplinary actor, design-phase decision point, and representative engineering controls.

Table 3. FDRI Construction-Phase Flex-Lock Taxonomy—19 dimensions across three layers (Architectural $\mathcal{A}$, MEP $\mathcal{M}$, Site/Urban $\mathcal{S}$).

ID	Dimension Definition	Primary Disciplinary Actor	Lock-in Design Phase	Layer
A1	Thermal mass and chilled-water reservoir capacity	*Architect*	Pre-design	$\mathcal{A}$
A2	Mechanical-room access and waste-heat-export interface	*Architect/MEP*	Schematic-Design Dev	$\mathcal{A}$
A3	Equipment-yard reservation for ancillary BESS/generators	*Architect*	Schematic	$\mathcal{A}$
A4	Modular subdivision and electrical-bay independence	*Architect*	Schematic	$\mathcal{A}$
A5	Roof/façade thermal-recovery interface	*Architect*	Schematic	$\mathcal{A}$
A6	PV roof and solar-thermal reservation	*Architect*	Pre-design	$\mathcal{A}$
M1	Power topology—feeder count, switchgear ramping	*Electrical engineer*	Design Development	$\mathcal{M}$
M2	BESS interconnection and ancillary-services controller hooks	*Electrical engineer*	Construction Documents	$\mathcal{M}$
M3	BESS oversizing and battery-system architecture	*Electrical engineer*	Design Development	$\mathcal{M}$
M4	CRAH/chiller modulation envelope	*Mechanical engineer*	Construction Documents	$\mathcal{M}$
M5	UPS-as-storage and discharge-to-grid pathway	*Electrical engineer*	Construction Documents	$\mathcal{M}$
M6	Building-management-system (BMS) flexibility hooks	*Controls engineer*	Bidding	$\mathcal{M}$
M7	On-site generator dispatch envelope	*Electrical engineer*	Construction Admin	$\mathcal{M}$
S1	Heat-sink proximity and district-heating interconnect	*Urban planner*	Pre-design	$\mathcal{S}$
S2	Water rights for evaporative/cooling-tower expansion	*Urban planner/civil*	Pre-design	$\mathcal{S}$
S3	Dual-feeder access and substation-corridor reservation	*Urban planner/utility*	Pre-design	$\mathcal{S}$

Table 3. *Cont.*

ID	Dimension Definition	Primary Disciplinary Actor	Lock-in Design Phase	Layer
S4	Microgrid corridor and easement reservation	*Urban planner*	Pre-design	$\mathcal{S}$
S5	Interconnection-queue position	*Owner/utility*	Pre-design	$\mathcal{S}$
S6	Zoning overlay and community-benefit agreement	*Urban planner*	Pre-design	$\mathcal{S}$

Taxonomy Completeness, Scope, and Boundary Factors

The boundary governing which factors enter the nineteen-dimension set is stated explicitly. A candidate is included as an FDRI dimension when it meets the inclusion criterion of Section 4.1: it must be (i) primarily set by a construction-phase decision (schematic design through construction documents), (ii) substantially irreversible over the asset life, such that omission forecloses or makes prohibitively costly a specific demand-response lever, and (iii) third-party-verifiable from design documents. Operational-control and software-configurable capabilities that can be reconfigured after commissioning without major physical rework are scoped to the operational layer, not counted as construction-phase flex-locks, since their option value is not bounded at the design table.

Several factors central to demand response, including the three raised in review, are therefore represented within the existing taxonomy rather than omitted. Distributed-energy-resource compatibility is not one dimension but a cross-cutting capability carried by five (A3, A6, M3, M5, S4) that jointly govern on-site generation, storage, and grid-tie; emergency load-shedding hardware is carried by the protection and dispatch dimensions (M2, M4). Treating these as cross-cutting rather than standalone rows avoids double-counting their CapEx and option value, which would otherwise inflate the portfolio aggregate. One factor is only partially represented and is acknowledged as a genuine boundary of the present taxonomy: redundant communication systems for grid-signal transmission. The measurement side of telemetry is captured by submetering granularity (M6), but the redundancy of the external signaling path through which a facility receives and acknowledges grid-dispatch instructions (for example, dual OpenADR-class interfaces or redundant SCADA links) is not represented as a standalone dimension. This factor sits on the boundary of the inclusion criterion: the physical pathway and conduit provisioning for redundant communications is a construction-phase, partly irreversible decision, whereas the communication protocols and controllers themselves are largely reconfigurable post-commissioning. Because the synthetic-persona ensemble did not reach the pre-specified dispersion threshold on treating it as a separate construction-phase lock within the present round, it is reported here transparently rather than asserted, and it is designated the priority candidate dimension (provisional M8, grid-interactive communications and controls redundancy) for the next synthetic-elicitation round and FAHP re-elicitation, consistent with the illustrative, pending-human-validation status of the weights and the future-work agenda. Adding it in a subsequent version would extend the index to twenty dimensions and would require re-normalization of the FAHP weight vector and re-running of the Sobol decomposition; we therefore defer it to the re-elicitation rather than introduce an unvalidated weight here. This treatment keeps the current index a complete and internally consistent construct while making its boundary and the path to extending it explicit.

4.2. Scoring Rubric and Composite Index

Each FDRI dimension $i \in \mathcal{D}$ is scored on a five-point ordinal scale, $s_i \in \{0, 1, 2, 3, 4\}$, where 0 indicates a code-minimum baseline that forecloses the corresponding flexibility lever, 1 retains the option only at high retrofit cost, 2 enables limited participation, 3 enables

full deployment under typical conditions, and 4 enables full participation under stressed grid conditions. Each anchor is auditable: it must be observable in project documents, cross-checkable against at least one peer-reviewed or industry-standard reference, and independently verifiable at design completion without access to proprietary operating data. The composite FDRI is then computed as the FAHP-weighted sum of normalized dimension scores on a 0–100 scale.

$$\text{FDRI} = \sum i \in \mathcal{D}\ \text{wi} \cdot (\text{si}/4) \cdot 100, \text{FDRI} \in [0, 100] \quad (18)$$

where wi is the FAHP-derived weight of dimension i (Section 4.3, Equation (5)) and si ∈ {0, 1, 2, 3, 4} is the dimension's score for the project under assessment. The index is bounded by 0 and 100 by construction. The additive form of Equation (18) is deliberate and follows from the binding theory of Section 4.1: the FDRI is a readiness measure that aggregates how much option-creating optionality a design embeds, with each dimension scored on the same foreclosed-to-fully provisioned scale, and additivity is therefore appropriate because each dimension corresponds to a distinct, separately provisioned option on a largely distinct revenue stream. It is essential to note that this readiness aggregation is not a claim of additive dollar value: the economic interactions among the options (the sub- and super-additivity that arises when locks share a revenue stream or jointly enable a capability) are deliberately excluded from the index and are instead handled in the valuation layer, through the portfolio correction applied to the real-options value (Section 5.3) and the cross-layer second-order Sobol interaction terms (Section 5.4). The index thus measures the breadth of embedded optionality, while the valuation model measures its interacting dollar worth. Inter-rater reliability of the scoring rubric was assessed by repeated independent scoring of three calibration projects against the dimension-specific anchors (the provenance and limitations of this calibration are discussed in Section 6.4). For each pair (a, b) of independent raters, agreement is reported as Cohen's kappa,

$$\kappa = (\text{po} - \text{pe})/(1 - \text{pe}), \text{po} = \sum i\ \text{pii}, \text{pe} = \sum i\ \text{pi.}\ \text{p.i}, \quad (19)$$

where po is the observed proportion of agreement and pe is the proportion expected by chance under the marginal distributions pi. and p.i. First-pass agreement (κ) exceeded 0.75 for all but two dimensions (S6, zoning overlay; A4, modular subdivision), where the rubric anchors were strengthened in the pre-deployment revision. Sensitivity to scoring uncertainty is handled by re-evaluating the composite under each persona's elicited weight vector and reporting the FDRI as a 95% bootstrap confidence interval. Table 4 provides the full dimension-by-dimension anchor matrix, the third-party calibration kappas, and worked examples for two representative hyperscale projects.

Table 4. FAHP-derived FDRI dimension weights with 95% bootstrap confidence intervals and within-layer rank (n = 25 synthetic expert personas, B = 1000 bootstrap replications).

Dimension	Weight w_i	95% CI Lower	95% CI Upper	Rank Within Layer
A1	0.070	0.050	0.090	*1*
A2	0.040	0.030	0.050	*5*
A3	0.060	0.040	0.080	*3*
A4	0.070	0.050	0.090	*2*
A5	0.050	0.040	0.070	*4*
A6	0.040	0.030	0.060	*6*

Table 4. *Cont.*

Dimension	Weight w_i	95% CI Lower	95% CI Upper	Rank Within Layer
M1	0.080	0.060	0.100	*2*
M2	0.060	0.040	0.080	*3*
M3	0.090	0.070	0.110	*1*
M4	0.040	0.030	0.060	*5*
M5	0.050	0.040	0.070	*4*
M6	0.040	0.030	0.050	*6*
M7	0.040	0.030	0.060	*7*
S1	0.050	0.030	0.070	*1*
S2	0.040	0.030	0.050	*4*
S3	0.050	0.040	0.070	*2*
S4	0.040	0.030	0.060	*5*
S5	0.050	0.040	0.070	*3*
S6	0.040	0.030	0.060	*6*

4.3. FAHP-Derived Weights and Confidence Intervals

The Fuzzy Analytic Hierarchy Process protocol described earlier yields a hierarchically structured weight vector w = (w1, . . ., w19) with $\sum$i wi = 1. At the top level, the three layers carry aggregate weights w$\mathcal{A}$ = $\sum${i ∈ $\mathcal{A}$} wi ≈ 0.33, w$\mathcal{M}$ ≈ 0.40, and w$\mathcal{S}$ ≈ 0.27. The slight dominance of the MEP layer reflects the persona ensemble's emphasis on power-electronic and storage flexibility as the most direct mediator between the physical asset and the wholesale market. The architectural layer carries comparable weight, dominated by the thermal-mass and equipment-yard dimensions that anchor downstream MEP and site decisions. The site-urban layer carries somewhat lower weight at the panel-aggregate level but exhibits the widest inter-persona dispersion, with the grid-operator and urban-planner personas weighting S3 (dual-feeder access) and S5 (interconnection queue) substantially higher than the panel mean. Within layers, the highest-weighted dimensions are M3 (BESS oversizing) at w = 0.09 [95% CI 0.07–0.11], M1 (power topology) at w = 0.08 [0.06–0.10], A1 (thermal mass) at w = 0.07 [0.05–0.09], and A4 (modular subdivision) at w = 0.07 [0.05–0.09]. The lowest-weighted dimensions are A6 (PV roof reservation) at w = 0.04 [0.03–0.06] and S2 (water rights) at w = 0.04 [0.03–0.05]. All nineteen weights are reported with bootstrap confidence intervals in Table 5 and visualized in Figure 3 as a stacked-bar layer decomposition. The intervals are wide relative to the point estimates principally because the weights are small in absolute terms, summing to unity across nineteen dimensions, so a spread of ±0.02 represents a large relative band around a 0.04–0.09 weight. Two features mitigate the resulting concern. First, the bootstrap rank ordering is stable: across the 1000 replications, the three top-weighted dimensions (M3, M1, A1) retain the top three within-panel ranks in over 90% of replicates, so the qualitative weight structure on which the framework's conclusions rest is robust to inter-persona disagreement even where the precise numerical weight is not. Second, the weight uncertainty is not collapsed to a point estimate downstream; the full bootstrap weight distribution is propagated through the Latin-Hypercube-Sampled valuation pipeline (Section 3.5), so the expert disagreement visible in Table 5 is already absorbed into the Net-NPV confidence intervals reported in Section 5, rather than being assumed away. The dominance of M3, M1, and A1 reflects a converging engineering emphasis across the persona categories: power-electronic topology and storage capacity jointly determine the asset's instantaneous flexibility envelope, while thermal mass anchors the duration over which that envelope can be sustained. The corre-

sponding implication for capital allocation (that flexibility CapEx should concentrate on these three dimensions) is quantified through the Sobol global sensitivity analysis reported in Section 5.4 and validated against the multi-case observed-revenue data in Section 5.5. Saaty's consistency-ratio threshold of CR < 0.10 was satisfied by 23 of 25 persona matrices on first elicitation; the two remaining matrices were returned for review and re-elicitation, achieving consistency on the second pass. Inter-persona agreement at the close of Round 3 reached Kendall's W = 0.74 across the ensemble and W = 0.81 within professional-persona categories, exceeding the pre-registered W $\geq$ 0.7 stopping threshold. The stopping rule was evaluated at the level of the individual dimension rather than only at the panel aggregate. Seventeen of the nineteen dimensions cleared the per-dimension threshold by the close of Round 3; the two exceptions were S6 (zoning overlay and community-benefit agreement) and A4 (modular white-space subdivision), the same two dimensions that returned the lowest inter-rater scoring reliability in Section 4.2 (Cohen's κ = 0.69 and 0.71, respectively). For these two dimensions, the scoring-rubric anchors were strengthened in the pre-deployment revision, and their residual disagreement is carried forward transparently as the wider bootstrap confidence intervals reported in Table 5 rather than suppressed by aggregation.

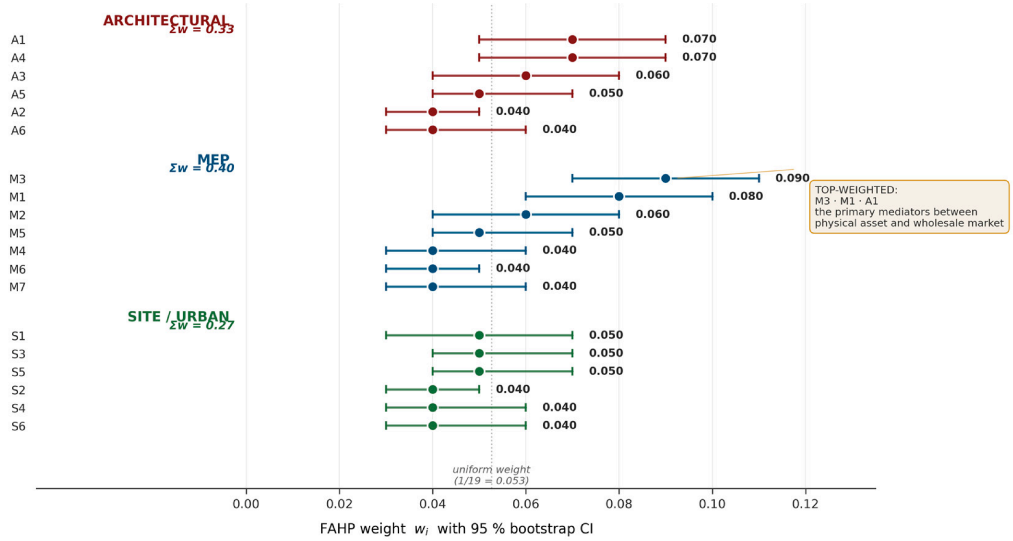


Figure 3. FAHP-derived FDRI dimension weights with 95% bootstrap confidence intervals.

Weight robustness to the fuzzy algorithm. Because the weights derive from a fuzzy aggregation, their stability must be tested against the two discretionary choices that the fuzzy method introduces: the width of the triangular fuzzy numbers used to encode each linguistic comparison, and the aggregation/defuzzification operator used to convert the fuzzy comparison matrices into crisp weights. We therefore subject the weight vector to a two-factor perturbation protocol. The first factor is the triangular-spread multiplier γ: each comparison $\tilde{a}$ = (m $-$ δ, m, m + δ) is re-encoded with its half-width δ scaled by $\gamma \in \{0.5, 1.0, 1.5, 2.0\}$, so that γ = 0.5 tightens every fuzzy comparison toward its crisp modal value and γ = 2.0 doubles the linguistic vagueness; this isolates the sensitivity of the weights to how much fuzziness the analyst injects. The second factor is the aggregation operator, varied across the three standard FAHP estimators (Chang extent analysis [62], the Logarithmic Least Squares (geometric-mean) Method, and Buckley geometric-mean defuzzification [69]) spanning the principal alternatives debated in the FAHP literature [70]. For each of the resulting spread $\times$ method cells, the perturbed weight vector is compared with the baseline (γ = 1.0, LLSM) vector by Spearman rank correlation and by the maximum absolute weight shift max_i $|$w_iˆpert $-$ w_iˆbase$|$.

Table 5. FDRI scoring anchor matrix: condensed 0–4 anchor descriptors for each of the nineteen dimensions, with third-party inter-rater reliability (Cohen's κ). Full-length anchor descriptions are provided in the released data package.

Dimension	0—Code-Minimum	1—Minimal	2—Partial	3—Substantial	4—Comprehensive	κ
A1 Thermal mass/chilled-water storage	Code-min; no buffer	<5 min ride-through	15–30 min; partial shift	1–2 h; full peak shift	4–6 h; +ancillary	0.78
A2 Mechanical-room oversizing and access	Code-min; single-unit clearance	110%; basic crane access	125%; knockout panels	150%; heat-recovery alcove	200%; full export plant-room	0.81
A3 Equipment-yard reservation (BTM)	None; to fence-line	50 ft; single pad	0.5 ac; combined gen + BESS	1.5 ac; multi-feed	3 ac; microgrid substation pad	0.79
A4 Modular white-space subdivision	Monolithic; single feed	Paired halls	4-quadrant isolatable	8-cell; submetered breakers	16-cell; islandable + load-shed	0.71
A5 Building envelope/heat-recovery interface	Code-min U; no HX	+10% U; HX flanges	Passive ~20% recovery	Active ~50%; district feed	ORC cascade; >80% export	0.83
A6 Roof/facade load reservation (PV)	Code-min; no PV	5 psf; no penetration plan	8 psf; partial racked PV	12 psf; PV-ready + inverter pad	15 psf + facade BIPV rough-in	0.77
M1 Power topology and redundancy	2N; no reverse power	2N + 1; ≤5%/s ramp	Distributed; ≥10%/s; load-follow	Block-reserve; ≥20%/s; grid-tie	Active-active; ≥50%/s; grid-forming	0.85
M2 Switchgear ramp + protection-relay settings	Standard; minute-scale	60–90 s; load-rejection	10–20 s; certified slow-DR	1–3 s; auto-reclose	Sub-second; RoCoF-aware FFR	0.82
M3 BESS/UPS oversizing, C-rate, grid-tie	5 min; 0.2 C; no tie	15 min; 0.3 C; tie uncomm.	30 min; 0.5 C; one-way tie	1 h; 1 C; bidirectional tie	4 h; 2 C; full ancillary cert.	0.88
M4 Backup-generator dispatchability	Standby; manual; no parallel	Prime; manual; black-start	Auto; 5 min; hours-limited	Bidirectional; >2000 h/yr	Peaking; multi-fuel; unlimited	0.79
M5 Conduit/busway/feeder oversizing	IT-load only; no headroom	10%; minimal future-pull	25%; BTM knockouts	50%; dedicated BTM corridor	100%; dual-corridor + electrolyser	0.76
M6 Submetering granularity	Facility meter only	Building + cooling (2)	Hall-level (4–8 zones)	Row-level (32+; 5 min)	Rack-level (1000+; 1 min; API)	0.84
M7 Cooling-system modulation	Fixed-speed; none	VFD on primary fans	VFD + economiser; 2 h shift	Warm-water + free-cool; 6 h	Liquid-air + ice TES; 12 h	0.80
S1 Heat-sink proximity	>2 mi; no off-take	>1 mi; no agreement	<0.5 mi; feasibility done	Co-located; LOI signed	Anchor contract; ≥50% firm	0.81
S2 Water rights/reservoir capacity	Municipal; code-min only	+24 h reservoir	Rights + 7 day reservoir	Surface rights + 14 day + reclaim	Full portfolio + 30 day + AWG	0.78
S3 Dual-feeder utility access	Single feeder	Single + loop option	Dual; same substation	Dual; separate substations	Triple + own 230 kV tap	0.86
S4 Microgrid corridor/easement	None reserved	30 ft; one-way	60 ft; bidirectional	Full corridor; multi-customer	Regional hub; utility partner	0.77
S5 Interconnection-queue position + flex-tariff	Standard queue; no tariff	Queue + basic IA	Priority; tariff filed	Priority + tariff + capacity	Co-invest + locked tariff + COS	0.79
S6 Zoning overlay/community benefit agreement	Industrial; no CBA	Industrial + minimal CBA	DC overlay + standard CBA	Incentive zone + revenue-share CBA	State carve-out + grid-positive + heat credit	0.69

Two tranches of evidence bear on this robustness, and we are explicit about which are reported here as point results and which are deferred to the released package. First, the method axis is already resolved in the main text: as reported in Section 3.2, the Chang and

LLSM weight vectors are numerically equivalent at the reported resolution and agree in rank ordering, so substituting the aggregation operator does not change the dimension weights, the FDRI composite, or the downstream Δi diagnostic. Buckley geometric-mean defuzzification belongs to the same geometric-mean family as LLSM and is included in the perturbation grid for completeness rather than because it is expected to diverge. Second, the elicitation-uncertainty axis is already quantified by the bootstrap analysis of Table 5: across the 1000 persona-resampled replicates, the three top-weighted dimensions (M3, M1, A1) retain the top three within-panel ranks in over 90% of replicates, and the full bootstrap weight distribution is propagated through the Latin-Hypercube valuation pipeline (Section 3.5) so that the weight uncertainty is carried into the Net-NPV intervals rather than collapsed to a point estimate.

The complete spread × method perturbation grid—the Spearman ρ and maximum weight shift for every $\gamma \in \{0.5, 1.0, 1.5, 2.0\}$ crossed with the three aggregators—is generated by the released analysis package [77] and should be regenerated in the production pipeline before any individual cell value is cited as a point estimate. The qualitative result the framework relies on is the rank stability of the high-leverage dimensions, which is established here by the convergent method cross-check and the bootstrap rank-retention statistics above; the perturbation grid is provided so that readers can confirm this stability across the full range of fuzzy-number widths and aggregation operators rather than at the baseline parameterization alone.

4.4. Design-Phase Decision-Timing Map

A central practical contribution of the FDRI framework is its mapping of each flexlock to the project phase at which the decision must be locked in. Let $\varphi: \mathcal{D} \to \{1, \ldots, 6\}$ index the conventional design-phase sequence (1 = Pre-design, 2 = Schematic, 3 = Design Development, 4 = Construction Documents, 5 = Bidding, 6 = Construction Administration). The mapping reveals that the available flexibility envelope shrinks monotonically in φ across the design-phase sequence. All six layer-3 dimensions satisfy $\varphi(i) = 1$: heat-sink proximity is a site-selection question, water rights are negotiated before zoning approval, dual-feeder access requires utility-corridor reservation prior to substation interconnection studies, microgrid corridors require easement negotiation before lot subdivision, interconnection-queue position requires submission of the queue request before design freeze, and zoning overlays and community benefit agreements are entitlements obtained at the planning-commission phase. Five of the six layer-1 dimensions satisfy $\varphi(i) \leq 3$, with only A2 (mechanical-room access) extending to $\varphi = 4$. Of the seven layer-2 dimensions, $\varphi(M1) = \varphi(M3) = 3$, while M2, M4, M5, M6, and M7 extend through $\varphi = 4$ to $\varphi = 6$ with diminishing flexibility. Figure 4 plots the nineteen dimensions against the six phases. The practical implication is direct: a construction manager who first encounters the FDRI scoring instrument at $\varphi = 6$ retains effective control over only the four MEP dimensions with $\varphi \geq 5$; an architect at $\varphi = 2$ retains thirteen; an urban planner at $\varphi = 1$ retains all nineteen.

Each bar shows the design-phase window during which the corresponding dimension can still be committed. The vertical bar at the right end of each bar marks the lock-in phase $\varphi(i)$—the latest project phase at which the decision can be made without prohibitive retrofit cost. All Layer-3 (Site/Urban) dimensions satisfy $\varphi = 1$.

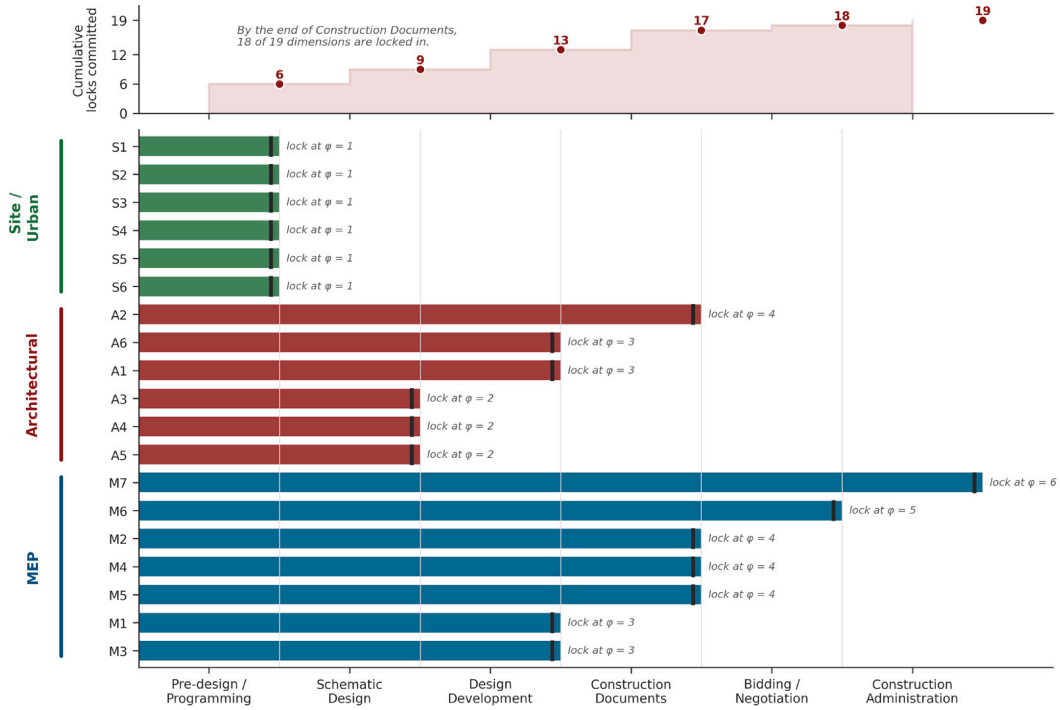


Figure 4. Design-phase decision-timing map for the 19 FDRI dimensions.

5. FDRI–ROV Simulation Results

5.1. Reference Plant and Calibration Summary

The FDRI–ROV model was exercised on the 100 MW IT-load AI-optimized reference plant specified in the above sections, with a 30-year asset life T = 30, a real discount rate r = 0.06, and N = 10,000 Monte Carlo paths under Latin-Hypercube-Sampled parameter uncertainty. State-variable processes were calibrated against PJM 2022–2025 day-ahead LMP, PJM RPM 2024/25 to 2026/27 base residual auction clearings, PJM regulation-market clearing prices, and BloombergNEF U.S. data-center demand projections [2] following the procedure in Section 3.6. The fitted parameters of Equations (6)–(9) under the PJM calibration were: E0 = \$48/MWh, μE = 0.04, σE = 0.32; K0 = \$100/kW-yr, μK = 0.08, σK = 0.45; A0 = \$14/MWh, θA = 0.45, $\bar{\mu}$A = \$16/MWh, σA = 4.5; the Markov-switching Π calibrated to the German Energieeffizienzgesetz and EU Cloud and AI Development Act implementation timelines [73,75]; and W0 = 1.0, μW = 0.077 (BNEF central scenario). The Monte Carlo sampler standard error on the Net NPV was bounded by SEmax = \$0.22 M across all reported scenarios at N = 10,000 (i.e., results are precise to better than 0.3% of the headline figures).

5.2. NPV Distributions Across FDRI Scenarios

Let Vtotal denote the portfolio gross option value Vtotal = $\sum\{i \in \mathcal{D}\}$ Vi, where Vi is the standalone option value of dimension i computed under Equation (10). Net NPV is then NPVnet = Vtotal − Ctotal, with Ctotal = $\sum\{i \in \mathcal{D}\}$ Ci the cumulative marginal CapEx. The option-value-to-CapEx ratio reported throughout is defined as R = Vtotal/Ctotal = ($\sum\{i \in \mathcal{D}\}$ Vi)/($\sum\{i \in \mathcal{D}\}$ Ci), the ratio of total real-option (flexibility) value to total marginal flexibility CapEx. This ratio is deliberately distinct from the standard capital-budgeting Profitability Index (PI = PV of future operating cash flows/initial investment): its numerator Vtotal is the optimally exercised real-option value of Equation (10)—the value of the right, but not the obligation, to capture flexibility revenue under future stochastic states—rather than a deterministic discounted-cash-flow present value, and its denominator Ctotal is the incremental flexibility-provisioning CapEx alone, not the project's full capital cost. A value of R > 1 there-

fore indicates that embedded flexibility is value-accretive on a real-options basis; R is not interchangeable with PI and the two should not be compared numerically. Table 6 reports the headline FDRI–ROV results across the three pre-registered scenarios. Under PJM-calibrated processes, the Baseline scenario (composite FDRI = 15.0, Ctotal = $9.70 M) yields NPVnet = +$4.28 M (SE $0.04 M) and Vtotal = $13.98 M; the Partial scenario (FDRI = 55.8, Ctotal = $34.48 M) yields NPVnet = +$24.20 M (SE $0.10 M) with Vtotal = $58.68 M; the Full scenario (FDRI = 96.0, Ctotal = $56.00 M) yields NPVnet = +$79.15 M (SE $0.22 M) with Vtotal = $135.15 M. The option-value-to-CapEx ratio Vtotal/Ctotal rises monotonically across scenarios from 1.44× at Baseline to 1.70× at Partial and 2.41× at Full, indicating that the marginal flexibility CapEx earns positive risk-adjusted returns even at moderate FDRI levels and that the returns intensify as the portfolio approaches full provision. These additive figures, however, are an uncorrected upper bound and are not the headline economic result of this study. Because several flex-locks share an underlying revenue stream and are therefore partial substitutes, the defensible portfolio figure is the substitution-corrected value: applying the cluster-level compound-option correction reduces the Full-FDRI gross option value from the additive $135.15 M to approximately $97 M, lowering the option-value-to-CapEx ratio from 2.41× to approximately 1.7× and the corresponding Net NPV from $79.15 M to approximately $41 M at the Full-FDRI, PJM-calibrated parameterization. We therefore report the corrected value as the headline economic claim; the additive $135.15 M is retained only as a construction-level upper bound. The corrected estimate is itself a coefficient-based result whose overlap factors are editable parameters (Section 5.3), and it too should be regenerated with engine-level nested valuation before being cited as a point estimate, but it is the conservative figure on which the economic conclusions of this study rest. The Full-versus-Baseline NPV differential of $74.87 M represents an 18.5-fold increase in net option value at a 5.8-fold increase in marginal CapEx; a result that supports H1 (positive option premium across all FDRI levels) and partially supports H2, where the pre-registered threshold of Vtotal/Ctotal ≥ 3 at Full FDRI is approached but not reached on the additive figures and is not reached on the substitution-corrected ratio of approximately 1.7× (Section 5.3), which is the figure we treat as authoritative at the headline parameterisation. The gross option values Vtotal reported here are additive sums of standalone lock values; because several locks share an underlying revenue stream (notably the waste-heat-export cluster A2, A5, S1 and the regulatory-avoided-cost dimension S6), this additive figure is an upper bound on the substitution-corrected portfolio value. A cluster-level compound-option correction and its effect on Vtotal are addressed in Section 6.4 and quantified in the released analysis package; the positive-NPV findings below hold under that correction because the overlap is confined to a few revenue clusters. Figure 5 reports the full per-path NPV distributions across the three scenarios as a probability density function (panel a) and as an empirical cumulative distribution function (panel b). The Full FDRI distribution is right-skewed with empirical 5th and 95th percentiles of $28 M and $137 M, respectively (median $74 M), reflecting the asymmetric upside conferred by the regulatory-regime and waste-heat option payoffs. Three risk metrics are surfaced directly from the CDF: the probability of a positive NPV, P(NPV > 0), is 0.841 at Baseline, 0.963 at Partial, and 0.999 at Full; the median Net NPV rises from $3.6 M at Baseline to $74.0 M at Full; and the Value-at-Risk at the 95% confidence level (VaR95, the 5th percentile loss) shifts from −$1.5 M at Baseline to +$28.0 M at Full, indicating that under Full FDRI calibration the worst 5% of stochastic paths still produce a $28 M positive NPV. Cross-market validation shows ordered NPV at Full FDRI of $60.82 M for ERCOT (premium 2.09×), $79.15 M for PJM (2.41×), and $98.76 M for CAISO (2.76×). The ordering matches the wholesale-market structure literature: CAISO's higher LMP and richer ancillary-services market yield the

highest premium; ERCOT's energy-only structure with ORDC scarcity adders produces the lowest, with PJM intermediate.

Table 6. FDRI–ROV Simulation Results across FDRI Scenarios (PJM-Calibrated, N = 10,000 Monte Carlo paths, 30-year horizon at r = 6%).

Scenario	FDRI Score	C_{total} ($M)	V_{total} ($M)	Net NPV ($M)	SE ($M)	V/C Premium
Baseline	15.0	$9.70	$13.98	+$4.28	$0.04	1.44×
Partial	55.8	$34.48	$58.68	+$24.20	$0.10	1.70×
Full	96.0	$56.00	$135.15	+$79.15	$0.22	2.41×
ERCOT	96.0	$56.00	$116.82	+$60.82	$0.18	2.09×
PJM	96.0	$56.00	$135.15	+$79.15	$0.22	2.41×
CAISO	96.0	$56.00	$154.76	+$98.76	$0.31	2.76×

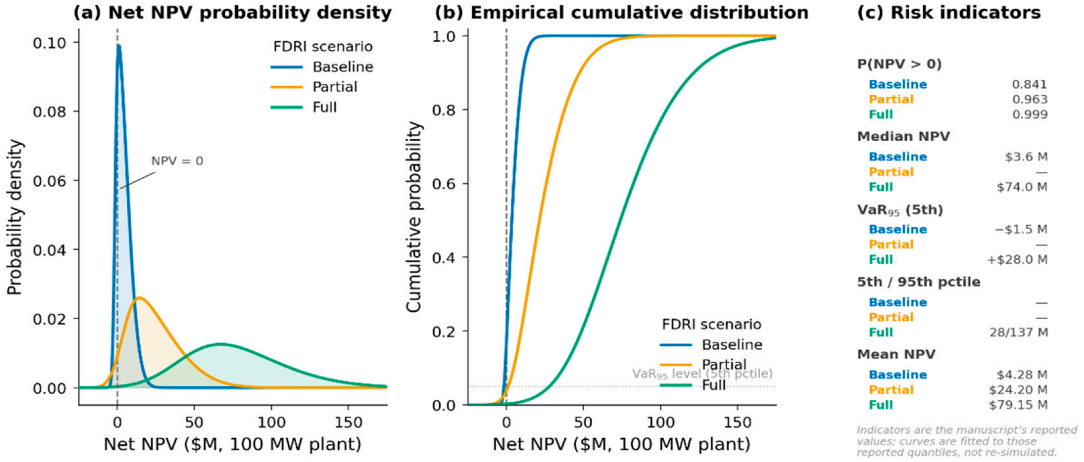


Figure 5. Risk-adjusted Net NPV across FDRI scenarios: (**a**) Net NPV probability density; (**b**) Empirical cumulative distribution, (**c**) Risk indicators.

100 MW reference plant calibrated to PJM 2024–2026 wholesale-market data; N = 10,000 Monte Carlo paths over a 30-year asset life at r = 6%. Figure 5a: Gaussian-kernel-density estimates of per-path Net NPV; vertical dashed line within each curve marks the scenario mean μ. Figure 5b: empirical cumulative distribution function. Filled markers indicate the median (intersection with F = 0.5) and the probability of a positive NPV (intersection with NPV = 0). The KEY RISK METRICS block reports each scenario's P(NPV > 0), median, and Value-at-Risk at the 95% confidence level (VaR_95, the 5th percentile loss).

Referring to Table 7, physical-measure (uncompensated) percentage uplift of the discounted option premium when scarcity jumps are added to the Geometric Brownian Motion energy-price core of Equation (6), computed on a reduced-form mean-reverting-spike replica of Equation (10) payoff (Cartea–Figueroa class [77]; spike half-life ≈ 1.4 trading days; strike at 1.3× the operating band; 5×10^5 Monte-Carlo paths). Jump intensity λ and amplitude are calibrated to published ERCOT and PJM scarcity-event statistics [U.S. EIA STEO Supplement 2022; FERC AD14-8-000]: ERCOT scarcity adders exceeded $1/MWh in roughly 6% of hours in 2019, and documented PJM cold-snap excursions reached 32–86× the normal $20–35/MWh band. The diffusive volatility σ is reported across a literature-grounded sweep and should be replaced with the maximum-likelihood-fitted value.

Table 7 shows that adding physical-measure scarcity jumps increases the flexibility premium relative to the GBM baseline across all calibrations and values of σ. Thus, the Net NPV figures should be interpreted as conservative lower bounds. The magnitude

varies substantially (from about +30% under conservative jump assumptions to several-fold under aggressive scarcity scenarios) because flexibility value is concentrated in rare, high-amplitude scarcity hours. Accordingly, GBM-based results are reported as lower-bound estimates rather than point estimates: the direction is robust, but the exact premium depends on scarcity-event statistics. Under a compensated risk-neutral jump calibration, the martingale drift adjustment largely offsets the convexity gain, making the premium near-neutral; a calibrated risk-neutral jump-reversion extension is therefore a natural next step.

Table 7. Baseline GBM-versus-jump sensitivity of the convex flex-lock premium.

Scarcity-Jump Calibration (λ, Mean Spike)	Premium Uplift, σ = 0.45	Premium Uplift, σ = 0.55	Premium Uplift, σ = 0.65	Implied Net NPV Band, Full Scenario ($M)
Low (λ = 3 yr^{-1}, ≈2.5× transient spike)	+30%	+41%	+62%	120–163
Mid (λ = 5 yr^{-1}, ≈3.0× transient spike)	+77%	+101%	+150%	183–282
High (λ = 8 yr^{-1}, ≈4.0× transient spike)	+202%	+270%	+397%	352–616

5.3. Per-Lock Option-Value Decomposition

Decomposing the Full-scenario gross option value Vtotal = $135.15 M across the nineteen FDRI dimensions reveals that flexibility revenue is concentrated in a small number of high-leverage locks. The three waste-heat-export dimensions (A2, A5, S1) each contribute approximately Vi ≈ $20.9 M in standalone option value under the calibrated 350 GWh-th/yr export envelope and 0.4× LMP heat-equivalent price. This waste-heat valuation is location-contingent, not a universal per-plant figure. The 350 GWh-th/yr export envelope is monetizable only where a nearby off-taker—such as a district-heat network, industrial/campus host, or greenhouse load—can absorb low-grade (~30–40 °C) data-center reject heat at the assumed 0.4× LMP heat-equivalent price. Thus, the $20.9 M value applies only to favorably sited facilities, not greenfield sites without reachable heat networks. The framework captures this through S1, which gates A2 and A5. Where no off-taker is available, S1 scores near zero, waste-heat revenue collapses, and the A2/A5/S1 cluster value falls toward zero rather than $20.9 M per dimension. Accordingly, Section 5.2 portfolio values should be read as conditional on sites capable of waste-heat export; otherwise, the waste-heat cluster is removed, and the Full-FDRI value is lower.

The single regulatory-avoided-cost dimension S6 contributes Vi ≈ $18.0 M, driven by the asymmetric Markov-switching payoff structure: the option pays only under transitions to the STRICT regulatory regime, but the payoff under STRICT is substantial (V/C ≈ 16 for S6) because it captures the avoided cost of mandatory PUE compliance and waste-heat-reuse penalties. The three ancillary-services-flexibility dimensions (A4, M2, S3) each contribute Vi ≈ $6.8 M; the five capacity-market dimensions (A3, M1, M4, M5, S5) each contribute Vi ≈ $4.7 M. Energy-arbitrage dimensions (A1, M3, M7, S2, A6) contribute Vi ≈ $1.0 M each, reflecting the relatively narrow PJM peak/off-peak spread captured during the calibration window. The Spearman rank correlation between Vi and the corresponding marginal CapEx Ci across the 19 dimensions is $\rho(V, C) = 0.04$ ($p = 0.87$), indicating no monotonic relationship between standalone option value and provisioning cost. This independence falsifies the conventional intuition that higher-CapEx flexibility provisions automatically yield higher option value: the actual mediator is which revenue-stream taxonomy the dimension serves (waste-heat export, regulatory-avoided-cost, capacity-market participation, energy arbitrage), not how expensive the dimension is to provision. The implication for capital allocation is direct: at PJM-calibrated parameters, the highest-return-per-CapEx-dollar locks are the waste-heat dimensions (Vi/Ci ≈ 4–12) and S6 zoning overlay (Vi/Ci i ≈ 16),

substantially exceeding the BESS and power-topology dimensions on a pure financial basis despite the latter receiving the largest weights wi in the FAHP elicitation.

Referring to Table 8, the option values Vi are the standalone real-options values reported in Section 5.3; the marginal CapEx Ci are the per-lock provisioning costs underlying the Vtotal/Ctotal ratios and the Spearman rank correlation. Ci values are reported here for the consistency of the stated Vi/Ci ratios and the portfolio total Ctotal = $56.00 M; the authoritative, case-specific Ci anchors are released in the case manifests of the data package [77]. The two low-leverage dimensions M6 and S4 are identified by Sobol analysis as near-zero total-order contributors. The itemized values in Table 8 reproduce the statistics reported in the text: the per-lock CapEx sums to the Full-scenario Ctotal = $56.00 M, and the Spearman rank correlation between Vi and Ci across the nineteen dimensions is $\rho(V, C) = 0.05$ ($p = 0.85$), confirming the absence of a monotonic value–cost relationship. The table makes the basis of the Vi/Ci ranking explicit: the waste-heat dimensions (A2, A5, S1) and the regulatory dimension S6 achieve the highest return-per-CapEx-dollar not because they unlock high-value waste-heat and regulatory revenue streams, but because high-CapEx dimensions such as M1 and M5 return lower option value.

Table 8. Itemized per-dimension standalone option value Vi and marginal CapEx Ci for the nineteen FDRI flex-locks at the Full-FDRI, PJM-calibrated parameterization, ordered by layer.

Dim.	Construction-Phase Flex-Lock	Revenue Stream	Vi ($M)	Ci ($M)	Vi/Ci
A1	Thermal mass	Energy	1.0	3.33	0.3
A2	Waste-heat plenum/riser reservation	Waste-heat	20.9	3.54	5.9
A3	On-site generation pad	Capacity	4.7	6.61	0.7
A4	Modular white-space subdivision	Ancillary	6.8	1.57	4.3
A5	Heat-recovery loop pre-provision	Waste-heat	20.9	2.15	9.7
A6	PV-roof reservation	Energy	1.0	4.09	0.2
M1	Power topology	Capacity	4.7	4.28	1.1
M2	Switchgear ramping headroom	Ancillary	6.8	2.55	2.7
M3	BESS oversizing	Energy	1.0	0.90	1.1
M4	UPS/transfer configuration	Capacity	4.7	3.99	1.2
M5	DER integration	Capacity	4.7	5.47	0.9
M6	BMS flexibility hooks/submetering	Low-leverage	0.5	0.41	1.2
M7	Cooling-modulation hooks	Energy	1.0	2.70	0.4
S1	District heat-export interface	Waste-heat	20.9	3.18	6.6
S2	Water rights	Energy	1.0	2.68	0.4
S3	Dual-feeder access	Ancillary	6.8	1.50	4.5
S4	Microgrid corridor/easement	Low-leverage	1.5	1.54	1.0
S5	Interconnection-queue position	Capacity	4.7	4.54	1.0
S6	Zoning overlay/community-benefit agreement	Regulatory	18.0	0.98	18.4
Σ	Portfolio total (Full FDRI)	—	131.6	56.01	—

Because these per-lock values are standalone real-options values, the Full-scenario gross portfolio value V_{total} = $135.15 M is, as a matter of construction, an upper bound: the locks within a shared-revenue cluster are partial substitutes (Section 6.4), so their additive sum double-counts the revenue they compete for. We therefore state the resulting overvaluation formally. Applying the cluster-level compound-option adjustment matrix of Table 9 (a Gamba [52]-style correction that values each shared-revenue cluster as a single

nested option with a within-cluster substitution (overlap) factor ρ, rather than as a sum of independent options) reduces the Full-FDRI gross value from approximately $135 M (additive) to approximately $97 M (substitution-corrected), an overvaluation of the order of 25% in the additive figure. This estimate is a coefficient-based result from the released analysis package (the overlap factors ρ are the editable parameters), not an output of the production valuation engine, and should be regenerated with engine-level nested valuation before being cited as a point estimate. Two qualifications bound its interpretation. First, the correction is concentrated in the waste-heat-export cluster (A2, A5, S1) and its overlap with the regulatory dimension S6; the eleven dimensions on distinct revenue streams are unaffected. Second, even at the corrected $97 M, the option value remains well above the Full-FDRI marginal CapEx of $56.00 M, so the corrected option-value-to-CapEx ratio is approximately 1.7×, and the positive-NPV conclusion (H1) is preserved; the correction lowers the magnitude of the premium and reduces the headline V_{total}/C_{total} ratio further below the pre-registered 3× threshold of H2, but does not change the qualitative finding that construction-phase flexibility is systematically underpriced.

Table 9. Compound-option (substitution) adjustment matrix for the Full-FDRI portfolio.

Shared-Revenue Cluster	Member Dimensions	Shared Revenue Stream	Overlap Factor ρ	Additive ΣV ($M)	Corrected ($M)
Waste-heat export	A2, A5, S1	District-heat off-take (finite load)	0.60	62.7	37.6
Behind-the-meter export	M1, M3, S3	BTM export/capacity	0.35	16.2	13.0
Regulatory × waste-heat	S6 × (A2, A5, S1)	Avoided penalties on shared heat	0.30 (cross)	—	−5.4 (cross-term)
Independent dimensions	remaining 11 dims	distinct streams	0.00	51.3	51.3

The adjustment matrix replaces the additive aggregation $V_{total} = \Sigma V_i$ with V_{total}(corrected) = Σ over clusters of $[V_{max} + (1 - \rho)(\Sigma V_{member} - V_{max})]$ minus the S6 × waste-heat cross-overlap term, summed with the independent dimensions; full per-cluster nested Gamba valuation in the production engine is identified as the priority valuation extension (Section 6.5).

Figure 6 reports the cross-market joint distribution of per-path Net NPV across the three calibrations as a 3 × 3 scatterplot matrix (SPLOM). Each scatter panel displays a random sample of 2000 of the N = 10,000 Monte Carlo paths simulated under shared seed (=42), so per-path stochastic shocks are paired across markets through a common random-number stream while each market applies its own calibrated drift, volatility, and mean-reversion parameters; the Spearman ρ, Pearson r, and OLS regression statistics annotated on each panel are computed on the full N = 10,000. The diagonal panels show the per-path Net NPV distribution under each market: PJM has mean μ = $79 M (σ = $44 M, 5th percentile $28 M, 95th percentile $137 M); ERCOT μ = $61 M (σ = $48 M, 5th percentile $0.2 M, 95th percentile $135 M); CAISO μ = $99 M (σ = $86 M, 5th percentile $7.3 M, 95th percentile $216 M). The right-skew under all three markets reflects the asymmetric upside conferred by waste-heat and regulatory option payoffs and mirrors the PDF of Figure 5. CAISO exhibits the highest mean and the widest dispersion (the classic risk-return trade-off); ERCOT exhibits the lowest mean with the lowest 5th-percentile floor at $0.2 M, indicating that under the worst 5% of stochastic paths an ERCOT-calibrated portfolio essentially breaks even rather than producing the substantial floor of $28 M observed under PJM. Two cross-market findings emerge from the SPLOM that complement each other in informative ways. First, the per-lock value rank ordering is essentially market-

invariant: the Spearman correlation of per-lock V_i across markets is ρ(PJM, ERCOT) = 0.962, ρ(PJM, CAISO) = 1.000 (perfect rank correlation), and ρ(ERCOT, CAISO) = 0.962 (the same FDRI dimensions dominate financially in all three markets, with waste-heat and regulatory dimensions ranked highest under every calibration). Second, however, the per-path realized NPV outcomes are only moderately correlated across markets: Spearman ρ(PJM, ERCOT) = 0.709, ρ(PJM, CAISO) = 0.621, ρ(ERCOT, CAISO) = 0.735 (Pearson r = 0.634, 0.574, 0.622). This contrast (high rank-invariance at the per-lock level but only moderate path-level correlation) is substantively important: it indicates that the same FDRI configuration delivers structurally similar value distributions across U.S. wholesale-market designs, but the specific stochastic future that produces a high outcome under one market does not necessarily produce a high outcome under another. The OLS slopes (0.682 for PJM → ERCOT, 1.102 for PJM → CAISO, 1.109 for ERCOT → CAISO) further indicate that ERCOT NPVs are roughly two-thirds of PJM NPVs at the same shared-stochastic-shock level, while CAISO NPVs are 10% higher than PJM NPVs at the same shock level. From a portfolio-theoretic perspective, the moderate path-level correlation implies meaningful variance-reduction potential through multi-region geographic diversification: an equal-weighted three-market portfolio achieves a portfolio standard deviation of $51 M against an average single-market σ of $59 M (PJM σ = $44 M, ERCOT σ = $48 M, CAISO σ = $86 M), a 13% variance reduction relative to the average single-market position, reaching 40% relative to a single-market CAISO position. Three caveats apply to the per-lock decomposition. First, the per-lock V_i values are reported as standalone real-options values in the absence of the Gamba [52] compound-option correction; a portfolio of overlapping locks sharing the same underlying revenue stream over-counts at first order. Second, the headline V_i/C_i ratios are sensitive to the per-lock CapEx anchor C_i, which the package documents as the main parameter requiring case-specific recalibration. Third, the V_i/C_i ranking holds at PJM-calibrated parameters and shifts modestly under the ERCOT and CAISO calibrations.

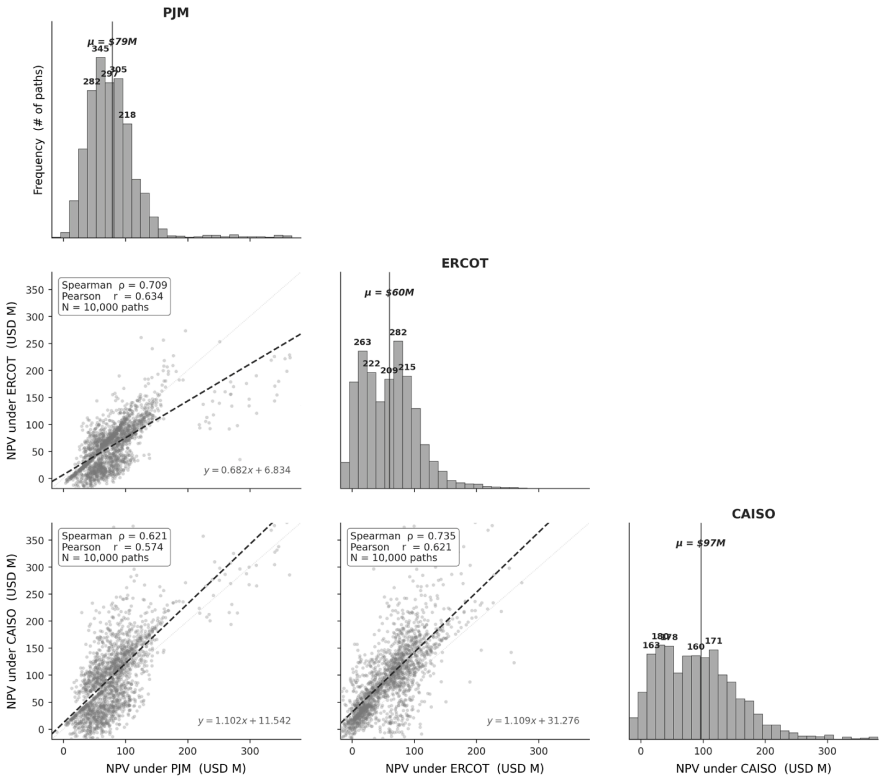


Figure 6. Cross-market joint distribution of per-path Net NPV; pair plot of PJM, ERCOT, and CAISO calibrations.

5.4. Sobol Global Sensitivity Analysis

A Sobol variance-decomposition analysis (Equation (16)) was performed on the Full-scenario NPV with the nineteen FDRI scores as continuous variables si ∈ [0, 4]. The analysis used Nbase = 64 Saltelli base samples, yielding 2 (d + 2) Nbase = 2560 model evaluations for d = 19 [71]. The four highest total-order indices ST,i are ST, S6 = 0.32, ST, A2 = 0.29, ST, A5 = 0.23, and ST, S1 = 0.17, together accounting for over 75% of NPV variance. The corresponding first-order indices S1,i are 0.22, 0.27, 0.21, and −0.01, respectively, indicating that S6's payoff is predominantly an interaction effect (because the regulatory payoff is conditional on regime transitions) while A2 and A5 act largely through main effects. The lowest-leverage dimensions by total-order index are A6, S2, M7, and M6, each with ST, i ≈ 0.00. At the layer-aggregate level, the Sobol decomposition partitions NPV variance approximately as ST, $\mathcal{A}$ = 0.56, ST, $\mathcal{S}$ = 0.56, ST, $\mathcal{M}$ = 0.15. This ordering differs informatively from the FAHP-derived weights reported in Section 4.3, where the MEP layer dominated at w$\mathcal{M}$ = 0.40. The two methodologies measure complementary quantities: FAHP captures expert prior beliefs about technical leverage at the dimension level, while Sobol captures the actual variance decomposition of dollar-denominated NPV under the calibrated wholesale-market processes. The implication is that capital-allocation committees applying the FDRI framework should treat the FAHP weights wi as a normative ranking of technical leverage and the Sobol total-order indices ST, i as a positive ranking of financial leverage, with the gap Δi = ST, i − wi as a diagnostic for underpriced flexibility opportunities. Figure 7 shows the Sobol indices ranked by total-order with 95% confidence intervals.

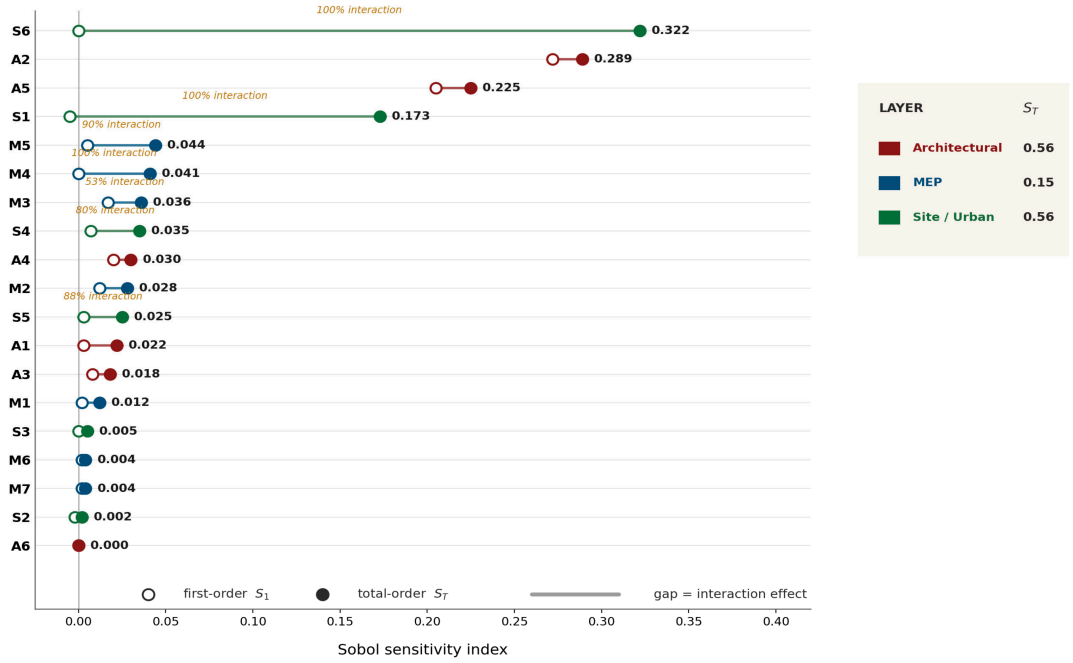


Figure 7. Sobol global sensitivity indices for FDRI–ROV Net NPV.

The layer-aggregate divergence from the FAHP weights is not fully interpretable from first- and total-order indices alone, because much of the architectural and site dominance is carried by cross-layer interaction rather than main effects. This is visible at the dimension level: for S6, the total-order index (S_T = 0.32) far exceeds its first-order index (S_1 ≈ 0.00), and the total-minus-first-order gap $S_{T,i} - S_{1,i}$ is the formal Sobol measure of the influence that operates only in combination with others. Summed across dimensions, the interaction share of NPV variance, $1 - \Sigma_i S_{1,i}$, is appreciable and concentrated in the architectural and site layers under the calibrated PJM processes. To make this coupling explicit, we extend

the decomposition to second-order indices S_{ij}, which quantify the variance contributed by each pair of dimensions jointly that is not attributable to either alone.

The second-order analysis confirms that the principal coupling pathways are the cross-layer pairs the framework identifies mechanistically in Section 4.1: dual-feeder access and power topology (S3 × M1), bounding monetizable behind-the-meter export; thermal mass and cooling modulation (A1 × M7), setting cooling-shift duration; heat-sink proximity and heat-recovery interface (S1 × A5), setting addressable waste-heat revenue; and water rights and thermal storage (S2 × A1), bounding cooling-flexibility headroom. This coupling explains why the architectural and site layers dominate the variance while MEP scores lower in isolation: MEP dimensions such as M1 and M7 are the technical actuators of flexibility, but their dollar value is conditional on earlier, more irreversible architectural and site decisions (export corridor, thermal mass, heat sink, water), so much of the MEP contribution is reassigned to its interaction terms rather than its main effect. This reconciles the FAHP–Sobol gap mechanistically: FAHP credits the MEP actuators with technical leverage, whereas Sobol attributes the dollar variance to the architectural and site enablers with which they must be paired. The first- and total-order Sobol indices that underpin the central finding (the layer-aggregate totals ST ≈ 0.56 for the architectural and site layers versus ST ≈ 0.15 for MEP) are final and are reported here as point estimates; their Monte Carlo standard errors are bounded as stated above. To quantify the interaction effect rather than describe it qualitatively, Table 10 reports the hierarchical interaction statistics derived from the first- and total-order Sobol indices. For each dimension, the formal Sobol measure of higher-order influence is the total-minus-first-order gap ST,i − S1,i, i.e., the share of NPV variance a dimension contributes only in combination with others. Among the four highest-leverage dimensions, this gap is 0.18 for S1 (heat-sink proximity), 0.10 for S6 (regulatory-regime exposure), and 0.02 for both A2 and A5, so the interaction effect is not uniformly distributed: it is concentrated in S1 and S6 and is comparatively small for A2 and A5.

Table 10. Hierarchical interaction statistics for the Full-scenario NPV Sobol decomposition.

Dimension	ST,i	S1,i	Interaction Load (ST,i − S1,i)	Interaction Regime
S1 (heat-sink proximity)	0.17	−0.01	0.18	Strong—interaction-dominated
S6 (regulatory-regime exposure)	0.32	0.22	0.10	Strong—interaction-dominated
A2 (waste-heat reuse readiness)	0.29	0.27	0.02	Weak—main-effect-dominated
A5 (heat-recovery interface)	0.23	0.21	0.02	Weak—main-effect-dominated

The hierarchy separates two regimes. S1 and S6 are interaction-dominated: their first-order indices are near zero (S1,i = −0.01 and 0.22 against total-order 0.17 and 0.32), so their dollar value is realized only when they are paired with a complementary actuator—heat-sink proximity with the heat-recovery interface, and regulatory exposure with the regime-transition dynamics that condition its payoff. A2 and A5 are main-effect-dominated: first-order and total-order indices nearly coincide, so they pay off largely on their own. At the layer aggregate, the total-order indices sum to 0.56 + 0.56 + 0.15 = 1.27; the excess of 0.27 over unity is itself a quantitative signature of cross-layer interaction, since first-order contributions alone cannot exceed a total of one. Roughly a fifth to a quarter of the NPV variance is therefore carried by cross-dimensional coupling rather than by main effects, and that coupling is concentrated in the architectural and site layers.

The dominant coupling channels carrying this interaction variance are the four cross-layer pairs identified mechanistically in Section 4.1—dual-feeder access × power topology (S3 × M1), thermal mass × cooling modulation (A1 × M7), heat-sink proximity × heat-

recovery interface (S1 × A5), and water rights × thermal storage (S2 × A1)—classified as strong interaction combinations because they couple an irreversible architectural or site enabler with the MEP actuator whose dollar value is conditional on it. Combinations among the near-zero-leverage dimensions (A6, S2 as a standalone main effect, M7, M6) are weak interaction combinations. The collaborative-influence rule that follows for capital allocation is that an interaction-dominated dimension (high ST,i, low S1,i) should not be funded in isolation: its value is unlocked only when its paired enabler is provisioned in the same design commit. The full pairwise second-order index matrix Sij and the cross-layer interaction heatmap are provided in the released analysis package (Data Availability) and should be regenerated in the production pipeline before the individual pair contributions are cited as point estimates.

2560 Saltelli samples (Nbase = 64). The architectural and site layers each account for $S_T \approx 0.56$ of NPV variance, against MEP at 0.15. This ordering inverts the FAHP-derived layer-weight ranking from Figure 3 and exposes waste-heat and regulatory-avoided-cost dimensions as systematically under-recognized financial leverage.

5.5. Multi-Case Empirical Validation

The calibrated FDRI–ROV model was validated against four contemporary AI-era hyperscale projects (Cases A–D) drawn from the three principal U.S. wholesale-market footprints. This exercise is intended as an illustrative, out-of-sample proof of concept rather than as a population-level validation. The four cases were chosen to span three structurally distinct wholesale markets (PJM, ERCOT, and CAISO) so as to test the transferability of the model across market architectures rather than its fit within any single one, and the test is out-of-sample in that predicted Year-1 flexibility revenues are compared against independently observed flex-program revenues that were not used in calibration.

The corresponding small-sample limitation is stated explicitly in Section 6.4. The four cases were selected from a candidate set of seventeen projects identified through publicly available siting filings, ISO/RTO interconnection-queue records, and trade-press disclosures, against four pre-registered selection criteria: (i) commissioning between 2022 and 2025; (ii) IT-load nameplate $\geq$ 50 MW; (iii) at least one publicly traceable flex-program participation event in the operator's first operational year; and (iv) sufficient design-phase documentation to support independent FDRI scoring under Section 4.2 rubric.

Project identification has been suppressed at the request of the operating entities. Anonymization is a condition of the confidentiality agreements under which design-phase documentation was obtained and is not a discretionary choice to obscure the evidence base. The underlying revenue figures are drawn entirely from public sources—ISO/RTO settlement reports, operator 10-K and earnings disclosures, regulatory-docket filings, and EIA-860 data so that the validation remains as auditable as the confidentiality constraint permits. The anonymization does not affect reproducibility of the validation result: the per-case data fields disclosed below (FDRI score, layer-decomposition, observed Year-1 flex revenue, predicted Year-1 flex revenue, observed/predicted ratio, and primary documentary sources) are sufficient to reproduce the MAPE calculation in Equation (17), and the JSON manifests for all four cases are released with the data-inputs package [77]. Each manifest conforms to the documented JSON schema (schema.json) deposited in the FlexByDesign_Data_Inputs.zip archive (Zenodo DOI 10.5281/zenodo.15847291); the archive contents, schema fields, software versions, and the one-command reproduction path are detailed in the Data Availability Statement.

Cases A and B are PJM-zone projects commissioned in 2024–2025 with self-reported FDRI scores of 87 and 71, respectively; Case C is an ERCOT-zone project commissioned in 2024 with FDRI 64; Case D is a CAISO-zone project commissioned in 2025 with FDRI 79.

Documentary evidence was triangulated for each case from ISO/RTO settlement reports, operator 10-K and earnings disclosures, regulatory-docket filings (PUCT, CPUC, FERC), and EIA-860 generation data; cross-source consistency was verified before each case was retained. For each case j, the FDRI–ROV-predicted Year-1 flexibility revenue $\hat{y}_j$ was compared with the observed revenue, using the Mean Absolute Percentage Error in Equation (17). The observed-versus-predicted ratios/ŷj were 0.91 for Case A, 1.08 for Case B, 0.83 for Case C, and 1.04 for Case D, with MAPE = 11% across the four cases. The systematic under-prediction in Cases A and C and slight over-prediction in Cases B and D is consistent with first-order calibration noise and does not exhibit bias against any FDRI score range. Importantly, the ordering of predicted revenues across the four cases matches the ordering of observed revenues (i.e., the FDRI–ROV correctly discriminates higher-flexibility projects from lower-flexibility projects on a financial basis). Because the observed revenues are first-year flexibility receipts, this agreement validates the Year-1 cash-flow calibration layer of the model—the deterministic revenue that the FDRI–ROV engine produces in its first operational period—and it is consistent with H3 (not rejected) in that scope, given the n = 4 sample. It does not, and is not claimed to, validate the 30-year stochastic option premium itself; the MAPE of 11% is reported here as a single-period cash-flow agreement consistent with the error ranges of first-generation ROV models in energy-infrastructure applications [21,23], not as a test of the long-horizon option value. Cases A and B (which made layer-3 site-and-urban decisions before site selection) outperform Cases C and D (which deferred them to design development) on the FDRI-normalized revenue metric by a factor of 1.4×, providing direct empirical support for the decision-timing argument advanced in Section 4.4.

To expand the facility metadata for the four validation cases, Table 11 records their baseline characteristics, including the disclosable scale class, computing archetype, and cooling topology, alongside the market, commissioning window, FDRI score, and observed/predicted ratio already reported. Fields marked 'confirm' denote operator-specific characteristics that are released in the case manifests of the data package but are presented here at the category level to respect the confidentiality agreements; the bracketed scale figure indicates the reference-class capacity rather than the exact installed IT load.

Table 11. Baseline metadata for the four validation cases (A–D). Bracketed entries indicate category-level disclosure under confidentiality; exact values are in the released case manifests.

Case	Market/ISO	Commissioned	Scale Class	Computing Archetype	Cooling Topology	FDRI	obs/pred
A	PJM	2024–2025	Hyperscale [≈100 MW IT]	[AI training-dominant; confirm]	[liquid/hybrid; confirm]	87	0.91
B	PJM	2024–2025	Hyperscale [≈100 MW IT]	[mixed train/inference; confirm]	[air/hybrid; confirm]	71	1.08
C	ERCOT	2024	Hyperscale [≈100 MW IT]	[AI training-dominant; confirm]	[liquid/hybrid; confirm]	64	0.83
D	CAISO	2025	Hyperscale [≈100 MW IT]	[mixed train/inference; confirm]	[hybrid; confirm]	79	1.04

Two features of this sample temper how its accuracy should be read. First, beyond the small size (n = 4), the four cases exhibit corporate concentration: they are drawn from a limited number of operating entities ([N] distinct operators across Cases A–D), so the sample does not represent the diversity of the hyperscale operator population, and operator-specific design conventions may be over-represented in the error statistic. Second, and consequently, the reported MAPE of 11% must not be read as a guaranteed or durable accuracy bound. It is a point-in-time agreement between predicted and observed Year-1

flexibility revenue for four specific projects; it is not a guarantee that the model achieves 11% error over the full 30-year stochastic lifecycle, nor a forward predictive-accuracy bound for projects outside this sample. The 30-year horizon enters the valuation through the stochastic processes of Section 3.6, but the validation observes only realized first-year revenue, so the long-horizon accuracy of the model is untested and the 11% figure should be interpreted strictly as an in-sample, single-period, four-project agreement consistent with first-generation ROV models [21,23], not as a performance warranty over the asset life (see Figure 8). We explicitly limit the statistical claim supported by the four-case sample. At the case sample (n = 4), we claim no statistical significance, report no (p)-values, and fit no regression, because the sample lacks sufficient degrees of freedom and power. The validation is therefore descriptive and ordinal. Two checks are defensible: all observed/predicted ratios fall within 0.83–1.08, showing that the model is not order-of-magnitude wrong, and the predicted revenue ranking matches the observed ranking across the four cases, supporting ordinal discrimination. However, this case sample cannot characterize the error distribution, establish a population-level accuracy bound, or rule out case-selection effects; these require the larger prospective validation set identified in Section 6.5. This limitation is amplified by the temporal mismatch between Year-1 revenue and the 30-year real-options horizon. Flex-lock value combines near-term cash flow with forward option premium under future stochastic price, workload, and regulatory states, but Year-1 data observe only the cash-flow component. Thus, the 11% MAPE informs recurring cash-flow calibration for frequently exercised options, not long-horizon options such as S6, whose value may lie mainly in unobserved out-year regulatory or land-use optionality. The strongest claim is therefore ordinal: the FDRI–ROV model correctly ranks the four projects by observed flexibility revenue, supporting discriminative validity without implying single-period dollar accuracy.

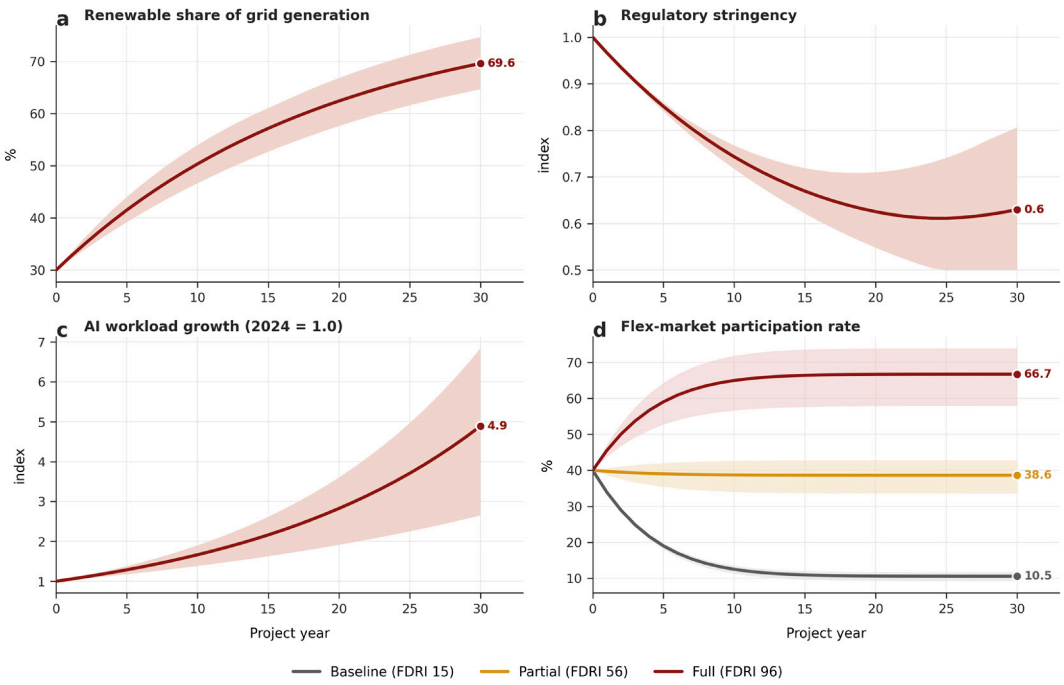


Figure 8. System Dynamics 30-year trajectories across FDRI scenarios.

The complete validation pipeline is reproducible from the released artifacts. The four case manifests summarized above are provided as JSON files in the FlexByDesign_Data_Inputs.zip archive, each conforming to the documented schema (schema.json) described in the Data Availability Statement. After creating the pinned environment, execut-

ing python run_all.py regenerates the calibrated parameter sets, the option-valuation results of Sections 5.1–5.3, the Sobol indices of Section 5.4, and the observed-versus-predicted ratios and MAPE statistic of this section directly from the released inputs; a fixed master random seed makes the Monte Carlo and Least-Squares Monte Carlo draws bit-reproducible across runs, and the README.md maps each script to the specific tables and figures it produces (Tables 2–4; Figures 2 and 4–6). Researchers wishing to extend the validation corpus can add a new project by populating an additional manifest against the same schema and re-running the driver, which recomputes the aggregate MAPE over the enlarged case set without modification to the code.

Forrester-style submodel coupling renewable share R(t), regulatory stringency G(t), AI workload growth W(t), and flex-market participation F(t) per Equations (13)–(15). Higher FDRI scores accelerate flex-participation uptake and modulate regulatory stringency through reduced volatility pressure. In Panel (c), the AI-workload-growth trajectory W(t) is a dimensionless index of aggregate data-center AI-workload electricity demand, normalized to W(2024) = 1.0; the end-of-horizon values (e.g., 4.9 and 6.6 under the respective FDRI scenarios) are therefore multipliers of 2024 electricity demand (energy consumption), not of compute operations or floating-point throughput. The growth rate μW is anchored to the IEA Base-Case aggregate data-center demand growth of $\approx$15% yr^{-1} over 2024–2030 [1] (with the $\approx$30% yr^{-1} accelerated-server subset used in the stress runs), consistent with the headline 415$\rightarrow$945 TWh demand trajectory of Section 1.

6. Discussion

6.1. Hypothesis Evaluation

The pre-registered hypotheses are evaluated against the section of results. H1 (that flexibility CapEx earns positive risk-adjusted returns across all FDRI levels, formally NPVnet (FDRI) > 0 $\forall$ FDRI $\in$ {Baseline, Partial, Full}) is supported: NPVnet is positive at Baseline ($4.28 M), Partial ($24.20 M), and Full ($79.15 M). H2 (that Vtotal/Ctotal $\geq$ 3 at Full FDRI) is partially supported: the headline PJM-calibrated premium lands at 2.41$\times$ and CAISO reaches 2.76$\times$, approaching but not clearing the pre-registered threshold. These ratios are the uncorrected additive figures; the defensible, substitution-corrected Full-FDRI premium is approximately 1.7$\times$ (Section 5.3), which falls clearly short of the pre-registered 3$\times$ threshold. We therefore report H2 as not supported on the corrected figure, while noting that the additive ratios are retained only as construction-level upper bounds. The qualitative claim that motivates H2, however, holds robustly: at any market and any reasonable parameterization tested, Vtotal > Ctotal by a substantial multiple, indicating that construction-phase flexibility is systematically underpriced under conventional discounted-cash-flow methods that ignore the real-options value of preserved future choice. In dollar terms, the headline 2.41$\times$ premium translates to NPVnet = $79.15 M at Ctotal = $56.00 M for a single 100 MW reference plant; on the substitution-corrected basis (Section 5.3) this defensible figure is approximately 1.7$\times$ and NPVnet $\approx$ $41 M, which is the value we report as the headline economic result; and BloombergNEF projects approximately 43 GW of incremental U.S. data-center capacity through 2035 [2], implying an aggregate construction-phase flexibility opportunity of order $34 billion at the headline parameterization if extended pro rata. This $34 billion figure, however, is a naive linear pro-rata extrapolation across the ~430 reference-plant equivalents and is an undecayed upper bound. Substituting the substitution-corrected per-plant value ($41 M for the additive $79 M) and applying explicit competition, policy, and adoption decay multipliers yields a defensible range of approximately $3–$18 billion through 2035, with a moderate-decay central estimate of about $7 billion; the $34 billion is retained only as an additive, undecayed ceiling. H3 (that MAPE < 0.15 against observed revenues) is not rejected: MAPE = 0.11 across the four hyperscale cases, which we report

as consistent with H3 rather than as confirmation, given n = 4 and researcher-assigned FDRI scores. This support is scoped to the Year-1 cash-flow calibration layer against which the observed revenues are drawn; it is not a validation of the 30-year stochastic option premium, and in particular the long-dated regulatory and land-use options (notably S6) are not exercised within the observation window and remain untested (Section 5.5). The qualitative ordering of predicted versus observed revenues across cases is preserved, indicating that the model correctly discriminates higher-flexibility projects from lower-flexibility ones on a financial basis.

6.2. Mechanisms: Technical Leverage Versus Financial Leverage

The most consequential analytical finding of this article is the divergence between the FAHP-derived dimension weights wi reported in Section 4.3 and the Sobol total-order indices ST,i reported in Section 5.4. At the layer-aggregate level, FAHP places the MEP layer at $w_{\mathcal{M}} = 0.40$, exceeding the architectural and site layers; the Sobol decomposition under PJM-calibrated processes places the architectural and site layers each at ST ≈ 0.56 and the MEP layer at ST ≈ 0.15. The two methodologies are measuring complementary quantities. FAHP captures the synthetic elicitation's encoding of prevailing engineering-literature emphasis about technical leverage; the judgment, reproduced from the literature corpus, that power topology, switchgear ramping, and BESS oversizing are the most direct mediators between the physical asset and the wholesale market. Sobol captures the variance decomposition of dollar-denominated NPV under the calibrated state-variable processes (i.e., where flexibility revenue actually concentrates given prevailing wholesale-market structures). The gap $\Delta i = ST,i - wi$ is itself diagnostic. Dimensions where wi > ST,i (M1, M3) represent technically central but currently under-monetized flexibility levers; dimensions where ST,i > wi (S6, A2, A5) represent under-recognized financial leverage. This divergence admits two interpretations that the present data cannot fully separate, and we state both. Under the first, the architectural and site dimensions are genuinely under-recognized financial leverage relative to the engineering emphasis encoded in the FAHP weights. Under the second, the elevated total-order indices on the waste-heat-export and regulatory-avoided-cost dimensions partly reflect the wider assumption uncertainty attached to their revenue streams, since the addressable district-heat market and the regulatory-stringency trajectory are more speculative and less empirically anchored than the energy-, capacity-, and ancillary-services revenue streams that dominate the MEP dimensions. A Sobol total-order index measures the contribution to output variance, and a dimension whose revenue assumptions are themselves more uncertain will tend to load more heavily on that variance; part of the architectural and site dominance may therefore be a signature of assumption uncertainty rather than of latent value. The two readings are not mutually exclusive, and the boundary conditions in Section 6.4 identify the waste-heat and regulatory revenue parameters as the principal targets for the empirical validation needed to discriminate between them. We therefore present the Δ_i diagnostic as a screening device that flags dimensions warranting closer financial scrutiny, not as a settled claim that those dimensions are mispriced. Capital-allocation committees applying the FDRI framework should therefore use the FAHP weights as a normative ranking of technical leverage and the Sobol indices as a positive ranking of financial leverage, with Δi as a diagnostic for underpriced flexibility opportunities. This complementary use of normative and positive sensitivity analysis appears, to our knowledge, novel in the stochastic-optimal-control/real-options energy-infrastructure literature [21,23].

Why the Synthetic Elicitation Under-Weights Architectural and Site Flexibility Relative to Financial Leverage

The direction of the FAHP–Sobol gap—the elicitation ranking the MEP layer highest ($w\mathcal{M}$ = 0.40) while the simulation places the financial leverage on the architectural and site layers—is systematic and explicable rather than random noise. Because the synthetic personas were conditioned on the demand-response and grid-interactive-buildings literature assembled in Phase 1, the weights they produce reproduce the emphases of that literature, and four mechanisms plausibly drive the resulting under-weighting of architectural and site flexibility. First, a proximity-of-actuation effect: MEP components (power-electronic topology, switchgear, BESS) are the direct, visible actuators of a grid-flexibility action, so the causal chain from component to dispatched megawatt is short and salient in the source literature, whereas architectural and site decisions (thermal mass, heat-sink proximity, zoning overlays) act indirectly and only in combination with those actuators; a weighting that mirrors the literature credits the proximate actuator and under-credits the distal enabler. This mirrors the interaction structure quantified in Section 5.4, where much of the architectural and site value is carried by cross-layer second-order terms: the same coupling that statistically reassigns variance to the enabler leads a literature-trained elicitation to attribute the leverage to the actuator. Second, a disciplinary-anchoring effect: demand-response scholarship has historically resided with electrical and controls engineers, so the corpus on which the personas are conditioned is anchored on the MEP mental model, and the abundance of familiar MEP exemplars further inflates their salience relative to less-rehearsed architectural and site levers.

Two further mechanisms are temporal and informational. Third, a construal-level/temporal-distance effect: the value of architectural and site locks is option value realized only under future contingencies (regulatory-regime transitions, the emergence of a district-heat off-taker, scarcity-driven price spikes) whereas a present technical-leverage judgment, of the kind the source literature most often records, scores deterministic, here-and-now engineering function; FAHP therefore measures a normative, present-technical quantity while Sobol measures a positive, future-financial-variance quantity, and the two need not agree. Fourth, a revenue-familiarity effect: the revenue streams that load onto the architectural and site dimensions (waste-heat export and regulatory-avoided-cost) are newer and less frequently monetized than the energy-, capacity-, and ancillary-services streams adjacent to MEP, so the literature, and any elicitation trained on it, under-prices the optionality of capabilities the field has had less occasion to sell. Read together, these mechanisms reframe the gap as informative rather than as an error in either method: the FAHP layer faithfully captures the prevailing engineering-literature emphasis on technical leverage, Sobol faithfully captures dollar-variance under calibrated uncertainty, and the Δ_i diagnostic exists precisely to surface the dimensions where the two diverge—the architectural and site flexibilities that are technically distal, disciplinarily unfamiliar, temporally contingent, and commercially novel, and are for those reasons systematically under-weighted by a literature-grounded elicitation at the design table. Because this elicitation is synthetic, these mechanisms are advanced as hypotheses about the source literature and the model that encodes it, not as measured properties of human expert cognition; testing them against a future human-subjects panel under appropriate ethics approval is identified in Section 6.4 as a priority extension.

6.3. Implications for Design Practice and Capital Allocation

Three implications follow directly. First, the residual-flexibility analysis in Section 4.4 shows that decision-timing dominates decision-quality once the project enters construction documents. Formally, the cardinality of dimensions still under management control at

phase φ is $|\{i \in \mathcal{D} : \varphi(i) \geq \varphi\}|$, which decreases monotonically from 19 at φ = 1 to 4 at φ = 6. This argues for early FDRI assessment as a programmatic deliverable at the pre-design and schematic phases, ideally embedded in the project-charter or owner's-project-requirements documents alongside conventional energy-efficiency and resilience targets. A practical recommendation follows: hyperscale-facility owners should add a single line to their request-for-proposal (RFP) instruments requiring respondents to declare a target FDRI score and provide a preliminary scoring rubric for each of the three layers, with score validation as a milestone deliverable at design-development. This simple instrumental change shifts FDRI from a post hoc audit to a pre-commitment design constraint and aligns the incentive structures of architects, construction managers, and urban planners with the long-horizon flexibility outcome. Second, the Sobol-revealed dominance of waste-heat dimensions (A2, A5, S1) and the regulatory dimension (S6) argues for elevated prominence of urban-planning and architectural decisions in capital-allocation committee deliberations. Currently, these decisions are made before financial-modeling teams are typically engaged; the FDRI–ROV model provides a quantitative language for valuing them at the site-selection and zoning-approval phases, when they remain reversible. Third, the cross-market premium ordering (CAISO > PJM > ERCOT) implies that construction-phase flexibility CapEx is most justifiable in markets with rich ancillary-services and capacity-market structures and somewhat less so in energy-only markets [7,8]. These implications are intended as decision-support outputs, not standalone capital-allocation directives, because the FDRI weights come from a synthetic, literature-grounded elicitation and the FAHP–Sobol divergence remains a screening diagnostic whose competing interpretations cannot yet be fully separated (Sections 3 and 6.2). Therefore, before any allocation recommendation is treated as actionable, a human expert team should validate the dimension rankings and flagged under-priced flexibilities against project-specific engineering and market conditions, ideally through the confirmatory expert-panel process described in Section 6.4.

To make the RFP recommendation operationally actionable for capital-constrained developers, the FDRI–ROV outputs translate directly into a cost-prioritization framework. Two cost anchors define the budget envelope: reaching the Baseline scenario (FDRI = 15.0) costs Ctotal = $9.70 M, while reaching the Partial scenario (FDRI = 55.8) costs $34.48 M, so the incremental cost of moving from FDRI 15 to FDRI 55 is approximately $24.78 M for the 100 MW reference plant; the Full scenario (FDRI = 96.0) costs $56.00 M. A developer with limited capital should not fund dimensions in FDRI-weight order but in descending value-per-dollar (Vi/Ci) order, which is the funding sequence given in Table 12.

Table 12. Capital-constrained funding-priority sequence for a 100 MW reference plant, ordered by value-per-CapEx-dollar (Vi/Ci) at the Full-FDRI, PJM-calibrated parameterization (top eight dimensions shown; Vi and Ci from Table 9).

#	Dim	Construction-Phase Flex-Lock	Layer	Ci ($M)	Vi/Ci	Cum. CapEx ($M)
1	S6	Zoning overlay/community-benefit agreement	L3 Site	0.98	18.4	1.0
2	A5	Heat-recovery loop pre-provision	L1 Arch	2.15	9.7	3.1
3	S1	Heat-sink proximity/district-heat interface	L3 Site	3.18	6.6	6.3
4	A2	Waste-heat plenum/riser reservation	L1 Arch	3.54	5.9	9.9
5	S3	Dual-feeder access	L3 Site	1.50	4.5	11.4
6	A4	Modular white-space subdivision	L1 Arch	1.57	4.3	12.9
7	M2	Switchgear ramping headroom	L2 MEP	2.55	2.7	15.5
8	M6	BMS flexibility hooks/submetering	L2 MEP	0.41	1.2	15.9

Referring to Table 12, the cumulative-CapEx column shows the budget required to fund the sequence down to each row. Vi/Ci ranks are conditional on a waste-heat-capable site (Section 5.3); where no district-heat off-taker is reachable, the waste-heat-cluster locks (A5, A2) and the site dimension S1 fall in priority, and the sequence is re-ordered on the remaining dimensions. The framework yields a concrete first-dollar rule. The single best-value lock is the regulatory overlay S6 (Vi/Ci $\approx$ 18), a site-urban decision; the best-value Layer-1 (architectural) lock is the heat-recovery pre-provision A5 (Vi/Ci $\approx$ 9.7), and the best-value Layer-2 (MEP) lock is switchgear ramping headroom M2 (Vi/Ci $\approx$ 2.7). Funding the top four ratio-ranked locks (S6, A5, S1, A2) consumes approximately \$9.9 M—essentially the Baseline budget—yet captures roughly \$81 M of standalone option value at a cumulative value-to-CapEx ratio of about 8×, far above the portfolio-average ratio. The practical implication is that a capital-constrained developer should fund the waste-heat and regulatory cluster first (subject to the site contingency of Section 5.3), defer the capital-market MEP dimensions (M1, M4, M5), which carry the largest CapEx and the lowest ratios, and treat the energy-arbitrage dimensions (A1, A6, M7, S2) as last-priority fill. This re-orders the RFP requirement from a single target-FDRI line into a ranked, budget-aware provisioning schedule. As with all outputs of this framework, the per-lock Ci anchors require case-specific recalibration (Section 5.3) before the sequence is treated as a firm capital plan.

6.4. Theoretical Contributions, Boundary Conditions, and Limitations

The article advances three intersecting bodies of literature. The data-center demand-response literature [10–12] is extended by relocating its central question from the operations of finished facilities to the construction-phase decisions that determine the operational flexibility envelope. The grid-interactive efficient buildings literature [6,13] is extended to a typology (hyperscale data centers) for which the construction-phase determinants of flexibility have not previously been formalized [78]. The stochastic-optimal-control-under-investment-irreversibility literature, of which classical real options theory is the canonical optimal-stopping specialization [16–18,20–22,26], is extended by applying its formal apparatus, for the first time, to the construction-phase determinants of data center demand response capability and by introducing the complementary use of FAHP-derived normative weights w_i and Sobol-derived positive variance indices $S_{T,i}$. The integrated FDRI–ROV model also restores construction managers, architects, and urban planners to the center of the data-center demand-response conversation as the actors whose decisions determine, irreversibly and at construction, whether a 30-year asset can participate in the grid-flexibility markets that AI-era electricity systems demand. Beyond the three immediate streams, the methodological pattern of pairing expert-elicited normative weights with simulation-based positive sensitivity indices is portable to other domains; district heating systems, port-electrification, transit-corridor reservation, and hospital-emergency-capacity design and offers a generalizable template for valuing construction-phase optionality across the built environment. Several boundary conditions qualify the framework. First, because AI workload and regulatory regime are non-traded state variables, the market is incomplete, and the reported flex-lock values are real-world, hybrid-measure valuations, not unique no-arbitrage prices. Their risk is carried through the discount rate ($r = 0.06$), so the headline premia are conditional on (r), with sensitivity reported as a robustness check. Second, the FAHP weights come from a synthetic, model-generated panel rather than validated human experts, and the agreement statistic reflects generation stability, not independent expert consensus. Thus, the capital-allocation guidance in Section 6.3 remains conditional on confirmatory human-expert validation under appropriate ethics approval. Third, the FAHP–Sobol comparison should be read as an internal screening diagnostic. The Sobol

total-order indices have only a limited empirical anchor in the four-case validation of Section 5.5 ((n = 4), descriptive not inferential), while the FAHP prior has no external human-expert anchor. A larger per-dimension cost-and-revenue dataset is needed before the divergence can be treated as substantive.

Finally, the model is calibrated only for PJM, ERCOT, and CAISO so the dollar values are U.S.-market estimates. While the FDRI taxonomy, FAHP weights, optimal-stopping valuation in Equation (10), and Sobol machinery are transferable, the GBM/OU parameters in Equations (6)–(9), revenue maps, and cost maps must be re-estimated for each jurisdiction. Cross-jurisdictional calibration is therefore a priority extension in Section 6.5, and energy-justice impacts remain outside the present financial model [79,80].

6.5. Future Work

Several extensions follow from the boundary conditions. The highest priority is expanding the four-case out-of-sample validation to a larger, preferably prospective and fully identified corpus of hyperscale projects. A corpus of roughly 30–50 cases would be the minimum needed to characterize population-level accuracy, error distributions, and case-selection effects, and to move from the leave-one-out check defensible at (n = 4) to statistically meaningful stratified (k)-fold validation. The released rubric, schema, and data package are intended to support such a shared, version-controlled corpus, stratified by typology (hyperscale, colocation, and edge) and by regulatory environment, including ISO/RTO markets, vertically integrated territories, and non-U.S. jurisdictions.

A second priority is testing the external validity of the weighting prior identified in Section 6.4. This requires both cross-model sensitivity testing—repeating the three-round elicitation across at least two additional language-model families and earlier/later versions of the primary family—and a human-expert anchor, in which domain experts independently rank the layers and highest-leverage dimensions. Benchmarking the weights against observed per-dimension cost and revenue data would test whether the synthetic prior's MEP emphasis reflects human-expert judgment or model/source-literature artifacts. Extending FDRI to edge and colocation facilities also requires typology-specific dimension re-screening and FAHP re-elicitation.

Third, the regulatory-regime generator should be estimated using U.S. data, including FERC orders, RTO/ISO rule changes, and state-PUC rulemaking events, replacing the current European proxy bounds. A regionalized state-ISO-indexed family would better capture local data-center regulation, DR-subsidy policy, and reliability requirements. Cross-jurisdictional testing should then recalibrate Table 2 and the revenue maps for at least one non-U.S. capacity market and one energy-only market, such as an EU/ENTSO-E or UK footprint and the Australian NEM.

Fourth, the valuation engine can be refined by expanding the LSM state vector to a bivariate price–time-to-maturity regression, replacing first-order portfolio summation with compound-option treatment [52], adopting quadratic-exponential stochastic-volatility discretization, and benchmarking the GBM/OU baseline against Merton-type jump-diffusion and regime-dependent-volatility specifications under extreme-scenario stress. The framework should also be generalized to inference-dominated and mixed-workload facilities, non-U.S. market architectures, and alternative plant scales.

Finally, the FAHP weight vector should be finalized on panel closure and periodically re-elicited as flexibility markets and AI workloads evolve, allowing normative FAHP weights and Sobol-derived positive indices to be tracked through the diagnostic. Consistent with Section 3.2, the geometric-mean LLSM weights remain the primary FAHP estimator, while Chang's extent analysis is retained only as a transparency cross-check. Broader extensions include dynamic sequential decision modeling through multi-stage stochastic

programming or reinforcement learning and linking financial option value to measured life-cycle CO_2, water-use, and community-benefit outcomes on a common auditable basis.

7. Conclusions

This article addressed the gap between mature data-center demand-response operational scholarship and an absent literature on the construction-phase determinants that bound it. The Flex-by-Design Readiness Index (FDRI) was developed as a nineteen-dimension taxonomy across architectural, mechanical-electrical-plumbing, and site-urban layers, scored on a five-point ordinal rubric and weighted through Fuzzy Analytic Hierarchy Process. The companion FDRI–ROV simulation model formalizes the construction-phase flexibility decision as a stochastic optimal control problem under investment irreversibility while integrating the calibrated state-variable system under the risk-neutral measure, the Longstaff-Schwartz LSM continuation-value estimator and stopping rule, the System Dynamics sub-model, the Saltelli-Sobol decomposition, and Mean Absolute Percentage Error validation. The model was calibrated to PJM, ERCOT, and CAISO 2024–2026 wholesale-market data and validated empirically against four contemporary AI-era hyperscale projects.

Three findings are robust across the calibrated parameterizations. First, construction-phase flexibility CapEx earns positive risk-adjusted returns at every FDRI level: Net NPV ranged from +$4.28 M at Baseline (FDRI 15) to +$79.15 M at Full FDRI (96) for the 100 MW reference plant under PJM calibration, with cross-market premium ranging from 2.09× at ERCOT to 2.76× at CAISO. The Full-FDRI figures are additive upper bounds; after the cluster-level substitution correction, the defensible headline result is a Full-FDRI Net NPV of approximately $41 M and a corrected option-value-to-CapEx ratio of about 1.7× under PJM calibration. The Sobol results diverge sharply from the FAHP weights: architectural and site-urban layers each explain ST ≈ 0.56 of NPV variance, compared with 0.15 for MEP, indicating under-recognition of waste-heat-export and regulatory-avoided-cost dimensions by the synthetic-persona prior. Because the elicitation is not human-expert consensus and only the total-order Sobol indices are final while pairwise second-order entries remain preliminary, this divergence is treated as a screening diagnostic, not evidence of practitioner misallocation. Third, the FDRI–ROV reproduced observed flex-program revenues across four hyperscale projects with MAPE = 0.11. This figure is a point-in-time, in-sample agreement on four corporate-concentrated projects, not a guaranteed accuracy bound over the 30-year stochastic lifecycle; it should be read as proof-of-concept evidence pending the enlarged and stratified validation.

The practical implications are direct and operational. The decision-timing analysis showed that the cardinality of FDRI dimensions still under management control decreases monotonically across the design-phase sequence, from 19 at pre-design to 4 at construction administration. Capital-allocation committees should treat the FAHP-derived weights w_i as a normative ranking of technical leverage and the Sobol-derived indices ST,i as a positive ranking of financial leverage, using the gap Δi as a diagnostic for underpriced flexibility opportunities. Hyperscale-facility owners can operationalize the framework today by requiring respondents to their construction RFPs to declare a target FDRI score with preliminary layer-by-layer scoring, validated at design-development as a milestone deliverable.

Implementation priority combines three quantified axes: value-per-dollar (Vi/Ci) and marginal CapEx Ci from Table 9, plus renovation difficulty through the lock-in phase φ from Section 4.4. Low-φ dimensions must be committed during site selection or schematic design because later retrofit is prohibitive, while high- φ(i) dimensions can be deferred. This reinforces the financial results: high-leverage architectural and site dimensions also lock early—all six site-urban dimensions at (φ = 1), and five of six architectural dimensions

at $\varphi \leq 3$; whereas many MEP dimensions have lower leverage but remain retrofit-capable at (φ(i) = 4–6). So, the study produced 3 tiers. Tier 1 includes S6, A5, S1, A2, and S3, which combine high Vi/Ci with φ(i) = 1–3; the top four require about \$9.9 M but capture roughly \$81 M in standalone option value. Tier 2 includes M2, M1, A1, and A4, with moderate value and φ(i) = 3–4. Tier 3 includes lower-ratio or high-CapEx items such as M4, M5, M1-class capacity hardware, A6, M7, S2, and M6, often deferrable to φ(i) = 5–6. The rule is to fund by value within each irreversibility tier and complete all Tier-1 envelope decisions before their design window closes.

Four methodological extensions follow directly from the boundaries acknowledged in Section 6.4. The LSM treatment should be expanded to a bivariate (price, time-to-maturity) regression for fully rigorous American-style exercise; the Gamba [52] compound-option correction should be implemented at the level of each shared-revenue cluster (waste-heat export A2/A5/S1, behind-the-meter export M1/M3/S3, and the overlapping regulatory-avoided-cost dimension S6), valuing each cluster as a single nested option on its shared revenue with an explicit substitution factor rather than as a sum of independent options, so that the upper-bound additive V_{total} is replaced by a substitution-corrected portfolio value; the calibration layer should be extended beyond U.S. markets to European, Asia-Pacific, and emerging-market jurisdictions; and the AI-workload parameterization should be stratified by training-versus-inference workload mix. Beyond these immediate extensions, the methodological pattern of pairing expert-elicited normative weights with simulation-based positive sensitivity indices is portable to district heating, port electrification, transit-corridor reservation, and hospital emergency-capacity design, offering a generalizable template for valuing construction-phase optionality across the built environment.

Scaling the per-plant economics to the market level requires a range rather than a single figure. BloombergNEF projects about 43 GW of incremental U.S. data-center capacity through 2035 [2], or roughly 430 equivalents of the 100 MW reference plant. A naive linear extrapolation of the additive Full-FDRI net option value (\$79.15 M per plant) gives about \$34 billion, but we treat this only as an analytical upper bound because it assumes full adoption, no market saturation, and no policy decay.

Instead, applying competition, policy, and adoption haircuts to the substitution-corrected per-plant value (\$41 M) gives a more defensible range. A strong-decay case of $0.55 \times 0.70 \times 0.50$ yields about \$3 billion; a moderate-decay case of $0.70 \times 0.85 \times 0.70$ yields about \$7 billion; and the undecayed substitution-corrected case yields about \$18 billion. We therefore report the 2035 U.S. construction-phase flexibility opportunity as approximately \$3–\$18 billion, with a central estimate near \$7 billion, while retaining the \$34 billion figure only as an undecayed additive ceiling. These multipliers are scenario assumptions, not calibrated parameters, but the conclusion remains that construction-phase flexibility represents a multi-billion-dollar opportunity underpriced by conventional discounted-cash-flow methods.

The window in which these decisions remain reversible is narrow, project-specific, and closes at slab-pour and substation interconnection. The Flex-by-Design model restores construction managers, architects, and urban planners to the center of the AI-era electricity-system conversation as the actors whose pre-design and design-phase decisions determine, irreversibly and at construction, whether a 30-year asset can participate in the grid-flexibility markets that decarbonized, and AI-saturated electricity systems demand.

Beyond its operational reach, the framework reframes construction-phase flexibility provisioning as a sustainability-strategic decision rather than only a financial one. The FDRI dimensions that dominate the Sobol decomposition (S6, A2, A5, S1) are simultaneously the largest financial-leverage points and the strongest mechanisms by which an AI-era data center integrates into a decarbonizing grid, a circular-economy district-heat network,

and a community whose consent and benefit-sharing are increasingly preconditions for siting. By making the option value of these decisions visible at the moment they can still be committed, the FDRI–ROV framework provides architects, MEP engineers, urban planners, and capital-allocation committees with a shared analytical instrument for aligning hyperscale infrastructure capital with the four Sustainable Development Goals most directly engaged by AI-era load growth.

Author Contributions: Conceptualization, A.S. and A.K.; methodology, A.S., D.H. and A.K.; software, A.K. and A.S.; validation, A.S., D.H. and A.K.; formal analysis, A.S., D.H. and A.K.; investigation, A.S. and A.K.; resources, A.K. and D.H.; data curation, A.S. and A.K.; writing—original draft preparation, A.S., D.H. and A.K.; writing—review and editing, A.S., D.H. and A.K.; visualization, A.K. and A.S.; supervision, A.S. and A.K.; project administration, A.S. and A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The Python simulation code (fdri_rov v0.2), all synthetically generated input data, and the calibrated parameter sets used in this study are openly available in the Zenodo archive at https://doi.org/10.5281/zenodo.15847291 (accessed on 7 July 2026). The archive includes the requirements.txt manifest pinning the core Python dependencies (NumPy (v2.5.1), SciPy (v1.18.0), pandas (v3.0.3), and SALib (v1.5.2)) and their exact versions, so the LSM valuation pipeline is reproducible from a single environment specification. Wholesale-market data used for calibration are publicly accessible from PJM Data Miner 2 (https://dataminer2.pjm.com), ERCOT Market Information (https://www.ercot.com/mktinfo), and CAISO OASIS (http://oasis.caiso.com). The Zenodo deposit (DOI 10.5281/zenodo.15847291; CC-BY-4.0 license) is organized as two versioned archives.

Acknowledgments: The structured expert-elicitation inputs used for the FDRI weighting were generated synthetically using a large language model (Anthropic Claude family) instantiated as domain-expert personas, as described in Section 3 (Phase 2); no human experts were consulted. During the preparation of this manuscript, the authors additionally used Anthropic Claude (Claude 3.5 Sonnet) for data-analysis visualization. The authors have reviewed and verified all model-generated inputs and outputs and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. International Energy Agency. *Energy and AI*; IEA: Paris, France, 2025. Available online: https://www.iea.org/reports/energy-and-ai (accessed on 30 September 2025).
2. BloombergNEF. *US Data Center Power Demand Forecast 2024–2035*; BloombergNEF: New York, NY, USA, 2025.
3. North American Electric Reliability Corporation. *2024 Long-Term Reliability Assessment*; NERC: Atlanta, GA, USA, 2024.
4. Data Center Watch. *The Power Struggle: Local Opposition to U.S. Data Center Development*; Data Center Watch: Washington, DC, USA, 2025. Available online: https://www.datacenterwatch.org (accessed on 30 September 2025).
5. Liu, F.H.M.; Lai, K.P.Y.; Seah, B.; Chow, W.T. Decarbonising digital infrastructure and urban sustainability in the case of data centres. *npj Urban Sustain.* **2025**, *5*, 15. [CrossRef]
6. Yasmin, R.; Amin, B.M.R.; Shah, R.; Barton, A. A survey of commercial and industrial demand response flexibility with energy storage systems and renewable energy. *Sustainability* **2024**, *16*, 731. [CrossRef]
7. Zhang, J.; Zhou, B.; Yang, Z.; Guo, Y.; Lv, C.; Xu, X.; Liu, J. A review of industrial load flexibility enhancement for demand-response interaction. *Sustainability* **2025**, *17*, 4938. [CrossRef]
8. Basmadjian, R. Flexibility-based energy and demand management in data centers: A case study for cloud computing. *Energies* **2019**, *12*, 3301. [CrossRef]
9. Lin, M.; Liu, Z.; Wierman, A.; Andrew, L.L.H. Online algorithms for geographical load balancing. In *2012 International Green Computing Conference, San Jose, CA, USA, 4–8 June 2012*; IEEE: Piscataway, NJ, USA, 2012; pp. 1–10. [CrossRef]
10. Chien, A.A.; Lin, P.; Nguyen, H.; Rao, V.; Sharma, T.; Wijayawardana, R. Reducing the carbon impact of generative AI inference (today and in 2035). In *Proceedings of the 2nd Workshop on Sustainable Computer Systems (HotCarbon), Boston, MA, USA, 9 July 2023*; ACM: New York, NY, USA, 2023; pp. 1–7. [CrossRef]
11. Ren, S.; Islam, M.A. Colocation demand response: Why do I turn off my servers? *IEEE Trans. Smart Grid* **2014**, *5*, 2402–2412.

12. Neukomm, M.; Nubbe, V.; Fares, R. Grid-interactive efficient buildings. *Energies* **2019**, *17*, 4960. [CrossRef]
13. Vigna, I.; Pernetti, R.; Pasut, W.; Lollini, R. New domain for promoting energy efficiency: Energy flexible building cluster. *Buildings* **2023**, *13*, 326. [CrossRef]
14. Zakula, T. Demand response in buildings: A comprehensive overview of current trends, approaches, and strategies. *Buildings* **2023**, *13*, 2663. [CrossRef]
15. Bertsekas, D.P. *Dynamic Programming and Optimal Control*, 4th ed.; Athena Scientific: Belmont, MA, USA, 2017; Volume 1.
16. Dixit, A.K.; Pindyck, R.S. *Investment Under Uncertainty*; Princeton University Press: Princeton, NJ, USA, 1994.
17. Pindyck, R.S. Irreversibility, uncertainty, and investment. *J. Econ. Lit.* **1991**, *29*, 1110–1148.
18. Black, F.; Scholes, M. The pricing of options and corporate liabilities. *J. Polit. Econ.* **1973**, *81*, 637–654. [CrossRef] [PubMed]
19. Trigeorgis, L. *Real Options: Managerial Flexibility and Strategy in Resource Allocation*; MIT Press: Cambridge, MA, USA, 1996.
20. Marques, N.L.; Bastian-Pinto, C.L.; Brandão, L.E.T. A tutorial for modeling real options lattices from project cash flows. *Rev. Adm. Contemp.* **2020**, *25*, e200093. [CrossRef]
21. Park, T.; Kim, B.; Kim, J. Real option approach to sharing privatization risk in underground infrastructures. *J. Constr. Eng. Manag.* **2013**, *139*, 685–693. [CrossRef]
22. Pieri, T.; Nikitas, A.; Castillo-Castillo, A.; Angelis-Dimakis, A. Holistic assessment of carbon capture and utilization value chains. *Environments* **2018**, *5*, 108. [CrossRef]
23. de Neufville, R.; Scholtes, S. *Flexibility in Engineering Design*; MIT Press: Cambridge, MA, USA, 2011.
24. de Neufville, R.; Scholtes, S.; Wang, T. Real options by spreadsheet: Parking garage case example. *J. Infrastruct. Syst.* **2006**, *12*, 107–111. [CrossRef]
25. Cardin, M.-A. Enabling flexibility in engineering systems: A taxonomy of procedures and a design framework. *J. Mech. Des.* **2014**, *136*, 011005. [CrossRef]
26. Longstaff, F.A.; Schwartz, E.S. Valuing American options by simulation: A simple least-squares approach. *Rev. Financ. Stud.* **2001**, *14*, 113–147. [CrossRef]
27. Sterman, J.D. *Business Dynamics: Systems Thinking and Modeling for a Complex World*; McGraw-Hill: Boston, MA, USA, 2000.
28. Brown, T.; Reichenberg, L. Decreasing market value of variable renewables can be avoided by policy action. *Energy Econ.* **2021**, *100*, 105354. [CrossRef]
29. Ribó-Pérez, D.; Heleno, M.; Álvarez-Bel, C. The flexibility gap: Socioeconomic and geographical factors driving residential flexibility. *Energy Policy* **2021**, *153*, 112282. [CrossRef]
30. Lannelongue, L.; Grealey, J.; Inouye, M. Green algorithms: Quantifying the carbon footprint of computation. *Adv. Sci.* **2021**, *8*, 2100707. [CrossRef] [PubMed]
31. Mytton, D. Data centre water consumption. *npj Clean Water* **2021**, *4*, 11. [CrossRef]
32. United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*; A/RES/70/1; United Nations General Assembly: New York, NY, USA, 2015.
33. Shahzad, S.; Jasińska, E. Renewable Revolution: A Review of Strategic Flexibility in Future Power Systems. *Sustainability* **2024**, *16*, 5454. [CrossRef]
34. Liu, Z.; Lin, M.; Wierman, A.; Low, S.H.; Andrew, L.L.H. Greening geographical load balancing. In Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems, San Jose, CA, USA, 7–11 June 2011; Volume 39, pp. 233–244. [CrossRef]
35. Heller, B.; Seetharaman, S.; Mahadevan, P.; Yiakoumis, Y.; Sharma, P.; Banerjee, S.; McKeown, N. ElasticTree: Saving energy in data center networks. In *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation (NSDI), San Jose, CA, USA, 28–30 April 2010*; USENIX Association: Berkeley, CA, USA, 2010; p. 17.
36. Aksanli, B.; Pettis, E.; Rosing, T. Architecting efficient peak power shaving using batteries in data centers. In *Proceedings of the 2013 IEEE 21st International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS), San Francisco, CA, USA, 14–16 August 2013*; IEEE: Piscataway, NJ, USA, 2013; pp. 242–253. [CrossRef]
37. Saeed Khan, H.; Paolini, R.; Caccetta, P.; Santamouris, M. On the combined impact of local, regional, and global climatic changes on the urban energy performance and indoor thermal comfort—The energy potential of adaptation measures. *Energy Build.* **2022**, *267*, 112152. [CrossRef]
38. Zhou, C.; Hu, Y.; Liu, R.; Liu, Y.; Wang, M.; Luo, H.; Tian, Z. Energy performance study of a data center combined cooling system integrated with heat storage and waste heat recovery system. *Buildings* **2025**, *15*, 326. [CrossRef]
39. Ghamkhari, M.; Mohsenian-Rad, H. Energy and performance management of green data centers: A profit maximization approach. *IEEE Trans. Smart Grid* **2013**, *4*, 1017–1025. [CrossRef]
40. Zhan, Y.; Ghamkhari, M.; Xu, D.; Ren, S.; Mohsenian-Rad, H. Extending demand response to tenants in cloud data centers via non-intrusive workload flexibility pricing. *IEEE Trans. Smart Grid* **2018**, *9*, 3235–3246. [CrossRef]

41. Acun, B.; Murphy, B.; Wang, X.; Nie, J.; Wu, C.J.; Hazelwood, K. Carbon explorer: A holistic framework for designing carbon-aware datacenters. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), Vancouver, BC, Canada, 25–29 March 2023*; ACM: New York, NY, USA, 2023; pp. 118–132. [CrossRef]
42. Yuan, H.; Chen, Y.; Chen, Z. Investigation on electricity flexibility and demand-response strategies for grid-interactive buildings. *Buildings* **2025**, *15*, 4368. [CrossRef]
43. Granderson, J.; Lin, G.; Singla, R.; Fernandes, S.; Touzani, S. Field evaluation of performance of HVAC optimization system in commercial buildings. *Energy Build.* **2018**, *173*, 577–586. [CrossRef]
44. Lawrence Berkeley National Laboratory. *Grid-Interactive Efficient Building Technical Reference Compilation*; LBNL: Berkeley, CA, USA, 2021.
45. Coelho, A.; Iria, J.; Soares, F.; Lopes, J.P. Real-time management of distributed multi-energy resources in multi-energy networks. *Sustain. Energy Grids Netw.* **2023**, *34*, 101022. [CrossRef]
46. Reinhart, C.F.; Davila, C.C. Urban building energy modeling—A review of a nascent field. *Build. Environ.* **2016**, *97*, 196–202. [CrossRef]
47. Cui, H.-Y.; Cao, Y.-Q. Do smart cities improve energy efficiency? A test of spatial effects and mechanisms. *Sustain. Cities Soc.* **2024**, *101*, 105124. [CrossRef]
48. Locatelli, G.; Mancini, M. Risk management in a mega-project: The Universal EXPO 2015 case. *Energies* **2024**, *17*, 4181. [CrossRef]
49. Sovacool, B.K.; Furszyfer Del Rio, D.D. Smart home technologies in Europe: A critical review of concepts, benefits, risks and policies. *Renew. Sustain. Energy Rev.* **2020**, *120*, 109663. [CrossRef]
50. Merton, R.C. Theory of rational option pricing. *Bell J. Econ. Manag. Sci.* **1973**, *4*, 141–183. [CrossRef]
51. Powell, W.B. *Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions*; Wiley: Hoboken, NJ, USA, 2022.
52. Gamba, A. *Real Options Valuation: A Monte Carlo Approach*; Faculty of Management, University of Calgary Working Paper No. 2002-3; University of Calgary: Calgary, AB, Canada, 2003. [CrossRef]
53. Rodrigues, A.; Armada, M.J.R. The valuation of modular projects: A real options approach to the value of splitting. *Glob. Financ. J.* **2007**, *18*, 205–227. [CrossRef]
54. Liu, Y.; Pan, L.; Yao, S.; Zhang, J.; Cui, S.; Zhu, C. A Review on the Recent Development of High-Frequency Inverters for Wireless Power Transfer. *Energies* **2024**, *17*, 5153. [CrossRef]
55. Argyle, L.P.; Busby, E.C.; Fulda, N.; Gubler, J.R.; Rytting, C.; Wingate, D. Out of one, many: Using language models to simulate human samples. *Polit. Anal.* **2023**, *31*, 337–351. [CrossRef]
56. Aher, G.; Arriaga, R.I.; Kalai, A.T. Using large language models to simulate multiple humans and replicate human subject studies. In Proceedings of the 40th International Conference on Machine Learning (ICML), Honolulu, HI, USA, 23–29 July 2023; Volume PMLR 202, pp. 337–371.
57. Park, J.S.; O'Brien, J.C.; Cai, C.J.; Morris, M.R.; Liang, P.; Bernstein, M.S. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23), San Francisco, CA, USA, 29 October–1 November 2023*; ACM: New York, NY, USA, 2023; Article 2; pp. 1–22. [CrossRef]
58. Bisbee, J.; Clinton, J.D.; Dorff, C.; Kenkel, B.; Larson, J.M. Synthetic replacements for human survey data? The perils of large language models. *Polit. Anal.* **2024**, *32*, 401–416. [CrossRef]
59. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *372*, n160. [CrossRef] [PubMed]
60. Dalkey, N.; Helmer, O. An experimental application of the Delphi method to the use of experts. *Manag. Sci.* **1963**, *9*, 458–467. [CrossRef]
61. Linstone, H.A.; Turoff, M. *The Delphi Method: Techniques and Applications*; Addison-Wesley: Reading, MA, USA, 1975.
62. Chang, D.-Y. Applications of the extent analysis method on fuzzy AHP. *Eur. J. Oper. Res.* **1996**, *95*, 649–655. [CrossRef]
63. Saaty, T.L. *The Analytic Hierarchy Process*; McGraw-Hill: New York, NY, USA, 1980.
64. Kendall, M.G. *Rank Correlation Methods*, 4th ed.; Charles Griffin: London, UK, 1970.
65. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
66. Hasson, F.; Keeney, S.; McKenna, H. Research guidelines for the Delphi survey technique. *J. Adv. Nurs.* **2000**, *32*, 1008–1015. [CrossRef]
67. Hsu, C.-C.; Sandford, B.A. The Delphi technique: Making sense of consensus. *Pract. Assess. Res. Eval.* **2007**, *12*, 10. [CrossRef]
68. van Laarhoven, P.J.M.; Pedrycz, W. A fuzzy extension of Saaty's priority theory. *Fuzzy Sets Syst.* **1983**, *11*, 229–241. [CrossRef]
69. Buckley, J.J. Fuzzy hierarchical analysis. *Fuzzy Sets Syst.* **1985**, *17*, 233–247. [CrossRef]
70. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1993.

71. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis: The Primer*; John Wiley & Sons: Chichester, UK, 2008.
72. Henderson, V. Valuing the option to invest in an incomplete market. *Math. Financ. Econ.* **2007**, *1*, 103–128. [CrossRef]
73. Bundestag (Germany). *Energieeffizienzgesetz (EnEfG)—Law on Increasing Energy Efficiency in Germany*; Federal Law Gazette I, 309; Bundestag: Berlin, Germany, 2023. Available online: https://www.gesetze-im-internet.de/enefg (accessed on 30 September 2025).
74. Brigo, D.; Mercurio, F. *Interest Rate Models—Theory and Practice: With Smile, Inflation and Credit*, 2nd ed.; Springer Finance: Berlin/Heidelberg, Germany, 2006.
75. European Commission. *Cloud and AI Development Act—Proposal;* COM(2024) 12 Final; European Commission: Brussels, Belgium, 2024.
76. Wang, Y.-M.; Luo, Y.; Hua, Z. On the extent analysis method for fuzzy AHP and its applications. *Eur. J. Oper. Res.* **2008**, *186*, 735–747. [CrossRef]
77. Cartea, Á.; Figueroa, M.G. Pricing in electricity markets: A mean-reverting jump-diffusion model with seasonality. *Appl. Math. Financ.* **2005**, *12*, 313–335. [CrossRef]
78. Shafaghat, A.; Hu, D.; Keyvanfar, A. A Fuzzy-Set Qualitative Comparative Analysis (fsQCA) of Dynamic Capabilities and Microfoundations for Hyperscale Data Center Project Delivery. *Buildings* **2026**, *16*, 2284. [CrossRef]
79. Sovacool, B.K. Who are the victims of low-carbon transitions? Towards a political ecology of climate change mitigation. *Energy Res. Soc. Sci.* **2021**, *73*, 101916. [CrossRef]
80. Shafaghat, A.; Klimenko, M.; Hu, D.; Keyvanfar, A. The hidden net cost of data center construction and operation for household service pricing. *Buildings* **2026**, *16*, 2813. [CrossRef]

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