

Uncertainty-Aware Spatial Screening for Suburban Infrastructure–Solar Transition: An Open-Data Study of Long Island

Muhammad Ahmad

ahmadshaukat328@gmail.com

Research Article

Keywords: spatial screening, portfolio audit, partial identification, multicriteria analysis, ecological constraints, Long Island

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11381239/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Uncertainty-Aware Spatial Screening for Suburban Infrastructure–Solar Transition: An Open-Data Study of Long Island

Abstract

Spatial transition screening often produces a single ranking despite uncertain evidence, unelicited preferences, and disputed sustainability requirements. This study presents a portfolio audit designed for reproducibility for infrastructure–solar transition in 2,136 populated block groups in Nassau and Suffolk counties. Public observations describe street connectivity, transit evidence, service intensity, classified ecological cover, and built-surface solar opportunity. Ambiguous transit components remain intervals. The primary experiment samples 20,000 policies directly over effective domain weights, targets, and solar emphasis, comparing unrestricted selection with co-equal ecological eligibility rules. Conservative rank bounds distinguish endpoint agreement from membership guaranteed for every admissible transit assignment at a sampled policy. The unrestricted rule has 58 candidates with interval-guarantee frequency at least 0.8; the 10% ecological rule has 30. Their paired portfolio overlap has median 9.3%. County-specific audits preserve distinct information regimes. An illustrative reference portfolio retains only 51.9% of its candidates under population-weighted tract aggregation. An independent municipal comparison associates the reference score negatively with completed supported solar capacity per resident ($\rho = -0.893$), but positively with capacity per land area ($\rho = 0.721$). The contribution is an integrated audit of conditional membership across evidence, preferences, compensability, and spatial support. It establishes neither a unique regional priority ranking nor comprehensive transition readiness.

Keywords: spatial screening; portfolio audit; partial identification; multicriteria analysis; ecological constraints; Long Island

1 Introduction

Renewable-energy inventories identify technical opportunity, while infrastructure assessments describe existing settlements. Turning either into a candidate portfolio requires decisions about desired changes, relative priorities, and whether strengths in one domain may offset deficits in another. A composite ranking can conceal these decisions, especially when observations have uneven geographic coverage.

Neighborhood sustainability assessment, spatial multicriteria decision analysis (MCDA), and urban-renewal scenarios already combine diverse evidence [3, 5, 6]. Recent neighborhood retrofit frameworks extend this decision-support tradition [10, 11]. Composite-indicator research treats uncertainty analysis as an element of quality assessment [15, 16]. This study builds on those capabilities rather than proposing a new sustainability index.

The original concept drew on solarpunk’s connection between energy infrastructure and social imaginaries [1], and regenerative development’s relationships among people, place, and living systems [2]. The observations cover a narrower construct: *spatial compatibility for infrastructure–solar transition screening*. They describe starting conditions and technical opportunity under declared assumptions about target shortfalls. Consent, finance, institutions, engineering suitability, and ecological function require further evidence. “Readiness” is therefore not the name of the measured construct.

The contribution is an integrated audit with archived computational materials distinguishing candidate persistence from artifacts of four modeling choices: evidence intervals, policy sampling, ecological noncompensation, and spatial support. Existing uncertainty-aware MCDA can address these choices individually. The operational contribution here is their paired treatment: each policy produces interval-qualified membership under multiple decision rules, accompanied by county-specific and aggregation diagnostics. The mathematical ingredients are established; the transferable audit procedure and its empirical demonstration constitute the novelty claim.

Nassau and Suffolk are the suburban Long Island study region; Brooklyn and Queens are outside its boundary. The questions are how evidence and policy assumptions affect membership; how ecological eligibility changes portfolios; how county information regimes and spatial aggregation change interpretation; and whether independent adoption observations support a proposed score interpretation. No regional configuration is privileged as the principal ranking.

2 Evidence and spatial processing

2.1 Sources, support, and construct coverage

Table 1: Archived evidence sources and temporal interpretation. All files and retrieval parameters are retained locally.

Source	Spatial support / information	Temporal interpretation
EPA SLD v3	2,136 block groups; form, transit, jobs	2021 release; mixed source years
Solar Roadmap	15,051 opportunity records; 17 towns/cities	2021 study; older ancillary inputs
Legacy NLCD 2021	Clipped categorical raster; land-cover fractions	2021 epoch; 2023 service product
NYSERDA / Open NY	Completed solar projects; 15 cities/towns	Cutoffs 2021 and 2025; 2026 retrieval

The EPA Smart Location Database (SLD), version 3, supplies source-native geometries, population, urban form, employment, and transit fields [17]. A complete query yields 2,136 unique populated units: 1,138 in Nassau and 998 in Suffolk. Archived population totals are 1,356,564 and 1,487,901, respectively. Source identifiers remain text; the field name `GE0ID20` alone is not interpreted as establishing a new geography vintage.

The Solar Roadmap service contributes 15,051 opportunity records from layer 2; duplicate display layers are not concatenated [18, 19]. Its minimum combined installation threshold is 250 kW DC, and some adjacent parcels are grouped by ownership. Many small residential systems consequently lie outside its scope. The legacy NLCD 2021 categorical raster supplies ecological context [20, 21]. The coverage description identifies a 2023-06-30 product served in EPSG:3857 on a 30-m grid. Retrieval in October 2026 archives mixed earlier observations rather than a contemporary infrastructure census.

The literature synthesis is targeted scoping documented through primary records, available full texts, agency guidance, and bibliographic metadata. Its purpose is to connect indicator and audit choices to established methods; no exhaustive systematic-review coverage is asserted.

2.2 Indicators and opportunity allocation

Street connectivity is weighted intersection density. Transit averages reversed distance and log service-frequency density. Mixed-use/service intensity averages employment–household entropy

Table 2: Observed domains and mapped opportunity used for screening. Effective policy weights are sampled directly; indicator choice does not establish full transition feasibility.

Domain	Observed input
Street connectivity	SLD weighted intersection density (D3B).
Transit evidence	Reversed distance (D4A) and log frequency density (D4D); unavailable components remain intervals.
Service intensity	Employment–household entropy and log retail, education, and health employment per land km ² .
Ecological cover	Forest, shrub, grassland, and wetland share of classified nonwater pixels; developed-class vegetation excluded.
Built-surface solar	Log rooftop plus parking MW per land km ² ; fixed reference normalization.

and log retail, education, and health employment per land km². Ecological cover is the nonwater classified-pixel share in forest, shrub, grassland, and wetland classes (41, 42, 43, 52, 71, 90, 95). Built-surface opportunity combines rooftop and parking capacity, retaining ground capacity separately.

Urban-form evidence motivates connectivity and mixed use [4], but indicator meanings remain specific. Employment concentration is distinct from travel-time access. Classified ecological cover omits vegetation within developed classes and measures cover context rather than biodiversity condition. Eligible solar capacity is technical opportunity rather than installed generation. These distinctions delimit the screening construct before scoring.

Intersections and areas use UTM zone 18N (EPSG:32618). For record p , capacity component m , and unit i , allocation is

$$K_{im} = \sum_p K_{pm} \frac{A_{ip}}{\sum_\ell A_{\ell p}}, \quad (1)$$

where A_{ip} is intersected area. This preserves represented capacity. There are 1,058 records crossing multiple units; two geometry-less records have zero capacity. Representative-interior-point assignment tests the averaging assumption. Raster summaries use pixel-center inclusion and classified nonwater pixels, excluding nodata and water/ice. Pixel-count fractions on the served Web Mercator grid approximate ground-area fractions. Source snapshots and geometry permit alternative allocation and zonal methods to be tested.

Block groups preserve source measurement support. Subdividing their attributes into finer cells would create apparent resolution without new information. Parcel projects, mobility programs, and community initiatives may ultimately require different operational supports.

2.3 Structured transit ambiguity

SLD transit sentinels combine thresholds, unavailable feed coverage, and other source conventions. There are 999 unavailable distance observations, 914 unavailable frequency observations, and 1,044 units with either unavailable. Each unavailable normalized component remains $[0, 1]$; observed components retain their values. Component averages form domain arrays X^L and X^U . The resulting rectangular uncertainty set can be conservative where real components obey additional joint constraints.

Any transit component is unavailable in 164 Nassau units (14.4%) and 880 Suffolk units (88.2%). Within-county strata also differ in observed street and service intensity. The primary analysis therefore accompanies regional interval portfolios with county-conditional membership. Historical sentinel causes and current networks remain unresolved; upper endpoints are evidence bounds rather than assertions of uniformly strong service.

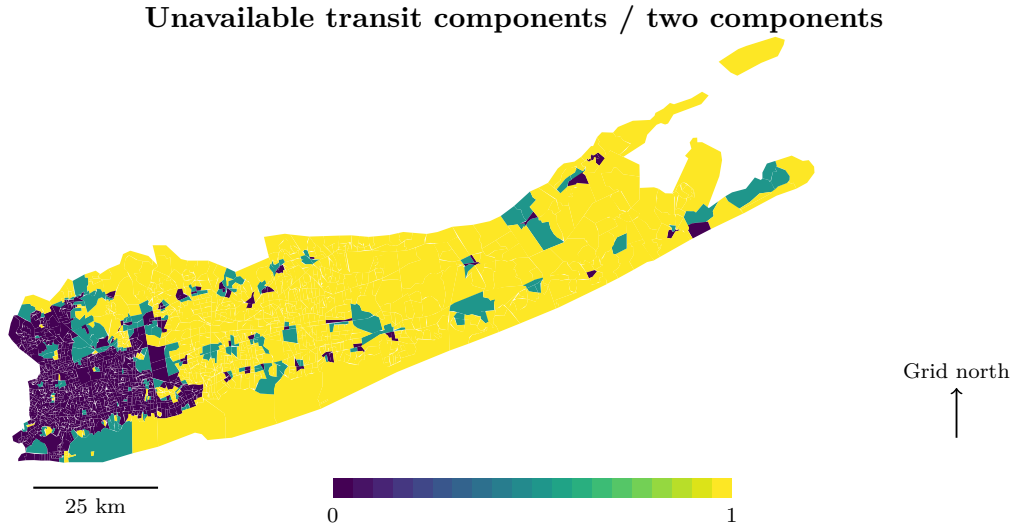


Figure 1: Spatial distribution of ambiguous SLD transit evidence. Zero means both components are observed, 0.5 means one sentinel, and one means two sentinels. This map describes archival coverage, not the contemporary transit network.

Table 3: Within-county transit-evidence strata: medians of independently observed covariates. Availability is not an observed absence of transit service.

County / unavailable	n	Street	Services	Ecology	Solar
Nassau / no	974	145.057	175.077	0.000	0.369
Nassau / yes	164	102.118	46.544	0.000	0.234
Suffolk / no	118	97.814	150.533	0.000	0.481
Suffolk / yes	880	57.397	43.863	0.033	0.263

3 Portfolio audit methods

3.1 Normalization and direct effective weights

Positive indicators use valid-data 5th–95th percentile normalization,

$$x_i = \min \left[1, \max \left(0, \frac{v_i - q_{.05}}{q_{.95} - q_{.05}} \right) \right]. \quad (2)$$

Distance is reverse oriented. Solar, frequency, and service intensity are logged first. Solar opportunity E_i normalizes log rooftop-high plus parking MW per land km²; the lower-rooftop case retains the same bounds, zero and 2.5706003144. Empirical-rank normalization is a supplementary sensitivity check.

For policy $\theta = (a, t, \lambda)$, the compensatory score is

$$S_i(\theta) = (1 - \lambda) \left[1 - \frac{\sum_{j=1}^4 a_j \max(t_j - x_{ij}, 0)}{\sum_{j=1}^4 a_j t_j} \right] + \lambda E_i, \quad (3)$$

where $a_j \geq 0$, $\sum_j a_j = 1$, and $t_j > 0$. Effective weights replace the earlier multiplication of importance weights and assumed burden coefficients. Only their product enters the score and its common scale cancels; those factors are not separately identifiable. Target gaps represent assumed relative shortfalls rather than measured implementation costs.

Twenty thousand primary policies use $a \sim \text{Dirichlet}(1, 1, 1, 1)$, independent $t_j \sim U(0.25, 1)$, and $\lambda \sim U(0, 1)$, with seed 20261009. A common draw interpolates normalized solar opportunity between the low- and high-rooftop cases. Direct simplex sampling removes the earlier burden hierarchy from the primary experiment. It remains an analyst-defined exploration measure: different ranges, correlations, or measures change persistence frequencies. The ranges span materially different target ambition and energy emphasis, without empirical or literature-derived authority for their endpoints.

An illustrative reference configuration uses $t_j = 0.75$, $a = (1/3, 1/3, 2/9, 1/9)$, and $\lambda = 0.25$. It is a common control for one-at-a-time technical and aggregation checks, reported in Appendix A. Earlier local and factorized policy experiments remain in the archive as sensitivity history, separate from the primary audit.

3.2 Co-equal decision rules

Unrestricted selection takes $q = \lceil 0.1n \rceil = 214$ units per policy. Co-equal constrained rules require raw ecological cover of 5%, 10%, or 20% before selecting the highest-scoring q eligible units. Every rule uses identical policy draws, normalization, solar interpolation, and ascending-identifier tie resolution. Paired comparison therefore isolates ecological eligibility rather than changing preferences simultaneously.

Noncompensatory assessment distinguishes attribute importance from permission to substitute other strengths for its deficit [14]. Raising a weight changes trade-offs; a floor enforces a per-candidate requirement. The floors are illustrative and require local ecological justification before application. If fewer than q units qualify, the portfolio stays smaller. A supplementary all-domain veto at 0.25 or 0.50 reports certain eligibility under X^L and possible eligibility under X^U .

3.3 Interval guarantees and county competitors

Equation 3 is monotone in every domain, so X^L and X^U bound scores at a fixed policy. For an eligible set G ,

$$r_{i,\theta}^{\text{worst}} = 1 + \sum_{k \in G \setminus \{i\}} \mathbb{1}(S_k^U(\theta) \geq S_i^L(\theta)). \quad (4)$$

The bound counts competitors at upper scores and candidate i at its lower score, including ties conservatively. If this rank is at most the portfolio size, membership holds for every independently admissible transit assignment in the rectangle. Additional physical restrictions would narrow the uncertainty set and can make this sufficient guarantee conservative.

For each rule, the audit retains lower-endpoint, joint common-endpoint, and guaranteed frequencies:

$$g_i = N^{-1} \sum_b \mathbb{1}(r_{i,\theta_b}^{\text{worst}} \leq q). \quad (5)$$

Joint endpoint agreement is weaker than a full interval guarantee. The reporting threshold $g_i \geq 0.8$ summarizes frequent sufficient membership under the sampled measure, rather than a posterior or universal policy guarantee. Paired overlap measures rule-dependent substitution at each draw.

County audits keep global normalization and identical policies but restrict competitors to Nassau or Suffolk, selecting 114 or 100 units respectively. They answer within-county screening questions while preserving the differing information regimes. County restriction changes the candidate set; it neither repairs missing evidence nor establishes a fair investment quota.

3.4 Spatial support and independent comparison

A support diagnostic population-weights normalized domains and solar opportunity into 601 source-defined tracts, recomputes the reference score, and broadcasts it to constituent block groups. Equal-size overlap measures local information loss. Networks are not re-estimated at tract scale. The exercise illustrates the modifiable areal unit problem [13], rather than establishing one administrative geography as the true community unit.

NYSERDA’s independent Open NY dataset supplies completed supported projects [22]. Complete means interconnected, operational, and closed out. Queries aggregate projects completed before January 2022 and January 2026 by city/town. Explicit crosswalks match 15 municipalities; one reservation-assigned block group remains outside the comparison. Residential coordinates are tract centroids, so no exact-point installation allocation is attempted.

Population-weighted municipal reference scores are compared with installed capacity per archived SLD resident, projects per 1,000 residents, and capacity per land km^2 . Spearman associations, leave-one-municipality-out ranges, and within-county coefficients are descriptive external comparison/falsification diagnostics. Fifteen differently supported units, historical population estimates, incentive-program coverage, spatial dependence, and the Roadmap’s installation threshold constrain interpretation. Archived permutation calculations remain exploratory; the primary presentation emphasizes associations and denominator sensitivity.

4 Results

4.1 Paired policy-space portfolios

The unrestricted rule has 89 high lower-endpoint-frequency candidates, 59 high joint-endpoint candidates, and 58 high interval-guarantee candidates. All 58 of the latter lie in Nassau. Under the 10% ecological rule, the corresponding counts are 169, 144, and 30; the last group includes 21 Nassau and 9 Suffolk units. The large gap between joint agreement and interval guarantees under this rule demonstrates why two common endpoints cannot certify every unit-specific assignment.

For the 10% floor, paired overlap with unrestricted portfolios has 5th, median, and 95th percentiles of 1.4%, 9.3%, 68.2%. Median paired overlaps at the 5% and 20% floors are 13.6% and 3.7%. These quantiles summarize the chosen exploration measure, not confidence limits on a true portfolio. Unrestricted and constrained portfolios represent different decision specifications.

Table 4: Co-equal regional specifications across 20,000 direct-effective-weight policies. Every portfolio contains 214 eligible units. High-frequency counts use the 0.8 threshold; guarantee frequencies require selection for all admissible transit assignments at each policy.

Rule	Eligible	Joint high	Guaranteed high	Nassau / Suffolk
Unrestricted	2136	59	58	58 / 0
Ecology $\geq 5\%$	544	115	33	29 / 4
Ecology $\geq 10\%$	418	144	30	21 / 9
Ecology $\geq 20\%$	277	171	35	17 / 18

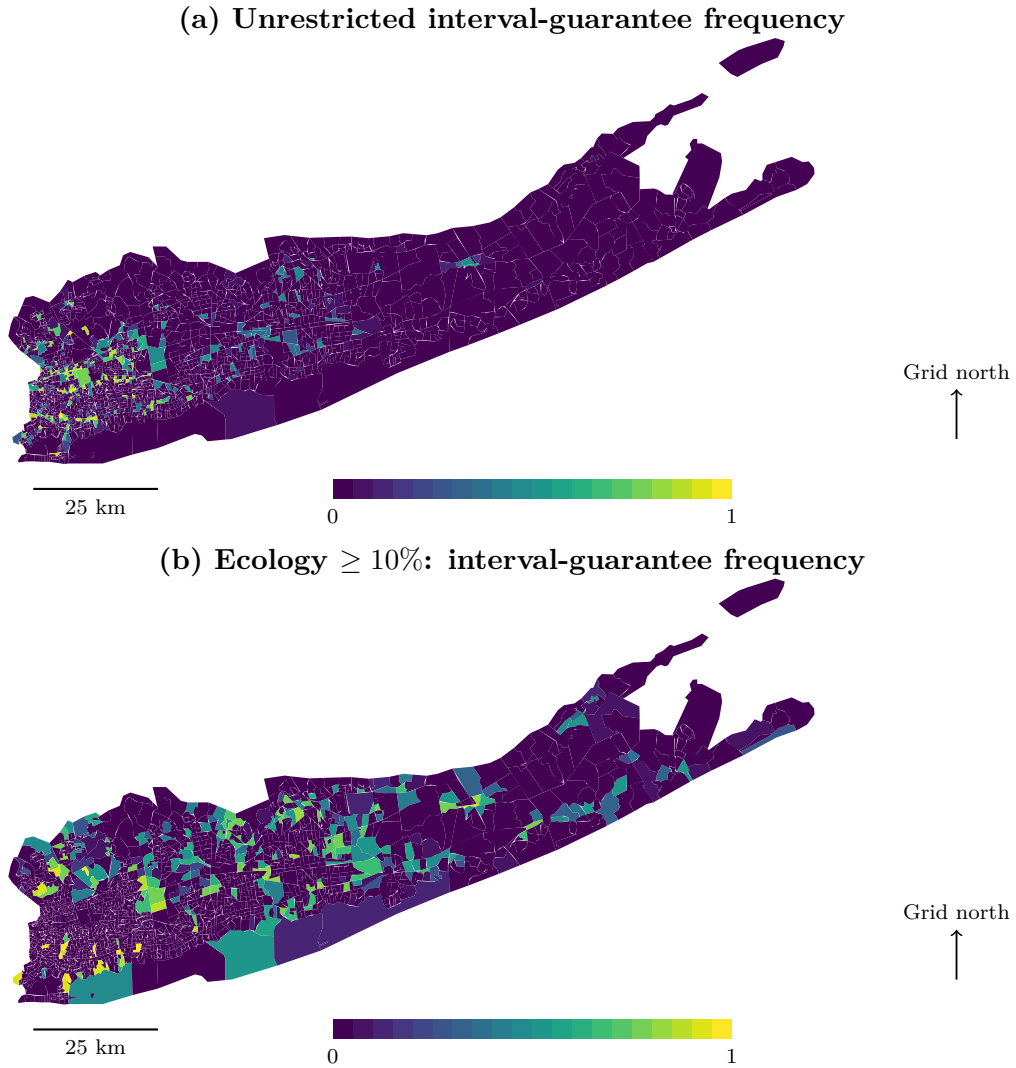


Figure 2: Co-equal policy-space portfolios. Frequencies range from zero to one across 20,000 direct-effective-weight policies. Ineligible ecological-rule units have zero frequency. Guarantees refer to the adopted transit intervals at each policy; boundaries are generalized for display only (UTM zone 18N).

Their paired overlaps quantify ecological noncompensation while holding policies and evidence bounds constant. Persistence under one rule therefore cannot stand in for persistence under another. The maps attach frequencies to their sampling measure and spatial support. Continuous frequencies and paired overlaps are the primary evidence; thresholded counts provide compact summaries. Appendix D compares reporting thresholds of 0.7, 0.8, and 0.9.

4.2 County information regimes and support instability

Table 5: County-conditional direct-weight policy audit, retaining global normalization. Frequency counts use the same 0.8 threshold.

County	Portfolio size	Lower high	Joint high	Guaranteed high
Nassau	114	53	48	48
Suffolk	100	23	9	1

Nassau has 48 high interval-guarantee-frequency candidates, compared with 1 in Suffolk. Suffolk has 23 high lower-endpoint-frequency and 9 high joint-endpoint-frequency candidates. This gap concerns identifiable membership from the available observations; it does not measure absent transit or actual transition capacity. The illustrative reference further exposes the evidence imbalance: Suffolk’s regional membership changes from three under lower transit endpoints to 30 under upper endpoints and 29 when transit is omitted. County-specific lower/upper reference overlap is 97.4% in Nassau and 51.0% in Suffolk. These findings give uncertainty within each information regime greater interpretive weight than the pooled reference ranking.

Tract aggregation correlates with reference block-group scores at $\rho = 0.855$ but retains only 51.9% of the 214-unit portfolio. Overall ordering can consequently remain similar while roughly half the candidates near a decision threshold change. Operational support must follow the intervention: buildings and parcels for installation, networks for mobility, and relevant social boundaries for community programs. The study quantifies support-dependent replacement rather than identifying geographically stable block-group priorities.

4.3 Ecological requirements and resource context

At the reference policy, a 10% ecological floor retains five unrestricted candidates and selects 61 Nassau and 153 Suffolk units. Reference ecological emphasis is about 8.3% in the unsaturated regime; omitting ecology retains 97.2% of that unrestricted portfolio. The paired primary audit extends this contrast across a broad effective-weight space.

The all-domain 0.25 veto leaves 14 certainly and 35 possibly eligible units. At 0.50, one qualifies under both endpoints. Small portfolios preserve the specified integrated requirement instead of silently relaxing it to fill a quota.

High-rooftop component capacity totals 19,380.6 MW, including 6,667.5 MW on rooftops and parking. Suffolk contains 75.1% of total component capacity and 61.1% of built-surface capacity. The combined source field sums to 19,455.5 MW, while the report gives approximately 19,424 MW [18]. These distinct source quantities remain separate in Appendix C. They describe mapped opportunity rather than observed generation or deployable capacity.

4.4 External comparison and denominator sensitivity

The reference lower score correlates negatively with completed capacity per resident ($\rho = -0.893$ through 2021; -0.911 through 2025). Every leave-one-out coefficient remains negative. Within-

Table 6: External comparison across 15 municipalities. Reference score associations change sign with the capacity denominator. Leave-one-out ranges concern capacity per resident.

Cutoff	Endpoint	ρ : per resident	ρ : per land area	Leave-one-out range
2021	Lower	-0.893	0.721	-0.930 to -0.868
2021	Upper	-0.843	0.750	-0.873 to -0.820
2025	Lower	-0.911	0.686	-0.947 to -0.890
2025	Upper	-0.871	0.714	-0.903 to -0.855

county 2021 coefficients are -0.400 for Nassau’s five municipalities and -0.709 for Suffolk’s ten; corresponding 2025 values are -0.400 and -0.770 .

Installed capacity per land area instead produces positive associations (0.721 and 0.686). Opportunity density also associates positively with installed density (0.789 and 0.768), but negatively with capacity per resident. Spatial installation concentration and per-resident adoption answer different questions. This external diagnostic challenges a generic adoption-readiness interpretation while showing descriptive alignment with installation density. No model is refitted after observing the comparison.

5 Discussion

5.1 Auditing conditional selection

The central inference is that these observations and assumptions do not identify a unique infrastructure–solar transition portfolio. Effective priorities, ambiguous evidence, ecological eligibility, and spatial support each influence membership. Paired rules and interval guarantees expose those dependencies without assigning normative authority to one analyst-defined configuration.

Direct effective-weight sampling simplifies the preference model. Separating importance and assumed burden in the earlier score could not identify independent effects, because only their product matters. Removing that redundancy improves interpretation but does not make the sampling measure objectively correct. Elicited objectives and intervention-specific information would define a different applied audit space.

The resulting frequencies measure persistence under this declared exploration measure, not the probability that a location would be selected by actual stakeholders.

The contribution to spatial MCDA is an integrated operational audit, not a mathematical capability unavailable to existing frameworks. Returning multiple membership sets, paired rule overlaps, county-qualified persistence, and support-dependent replacement makes a high score inspectable as a conditional selection artifact. This is more informative for the present evidence than a single list of allegedly most ready locations.

5.2 Ecological and distributive objectives

Ecological eligibility changes which candidates may enter, whereas an ecological weight changes trade-offs among candidates [14]. The reference portfolio’s near-complete replacement demonstrates competing definitions of the screening objective. Regenerative development requires ecological function and social relationships [2]; classified cover supplies a limited spatial safeguard. Applied floors need habitat, canopy, conservation, and intervention-specific justification.

Infrastructure advantage can also reflect historical investment advantage. Energy justice includes distributive, recognition, and procedural dimensions [12], requiring explicit objectives

and participation. A reference association of -0.130 with low-wage-worker fraction is contextual evidence. Needs-based investment and opportunity screening remain different decisions until affordability, displacement risk, participation, and access are assessed.

5.3 From archived support to operational evidence

Intervals preserve missing-information consequences without creating missing observations. Suffolk particularly requires reconciliation of historical feed coverage, sentinel causes, and contemporary services. A current network would enter as a separately documented temporal scenario. County audits locate unresolved selection questions while that reconciliation is incomplete.

Likewise, administrative aggregation diagnoses support sensitivity rather than resolving the correct intervention geography. Block groups preserve source consistency; project decisions require action-relevant evidence and boundaries. Carrying multiple supports into local evaluation follows directly from the observed portfolio replacement.

The independent comparison reinforces the same principle. Installation density can track settlement concentration, while per-resident adoption also reflects building types, roof availability, household resources, and program participation. Reporting both denominators specifies which interpretation aligns with external observations, rather than choosing the favorable one as validation of the whole construct.

6 Limitations and computational transparency

The evidence combines mixed historical years, categorical land cover, source geometries, and a thresholded solar inventory. Pixel fractions approximate area shares and omit developed-class vegetation; allocation averages heterogeneous surfaces; service and transit proxies incompletely describe access. Current transit, hazards, groundwater, biodiversity condition, structural suitability, grid capacity, finance, consent, and political feasibility remain outside the measured construct.

Policy ranges are analyst-defined, ecological floors illustrative, and the 0.8 frequency threshold a reporting convention. Interval guarantees apply within the adopted rectangle at sampled policies. Aggregation is diagnostic, and external adoption observations comprise only 15 municipalities. These boundaries delimit a screening audit rather than validated implementation feasibility.

The computational materials are supplied as Supplementary Material S1: Research Data and Reproducibility Package. The package includes processed numerical data, policy configurations, portfolio memberships, analysis scripts, software requirements, source metadata, and reconstruction instructions. Original source datasets remain available from the cited providers and subject to their distribution terms. A persistent public repository has not yet been established.

Numerical checks reconstruct prior results, compare interval ranks with brute-force competitor counts, test eligible membership and source totals, and preserve independently sourced associations. Substantive follow-up should reconcile transit, justify ecological requirements, elicit local objectives, and assess outcomes at intervention-relevant supports.

7 Conclusion

The framework audits conditional infrastructure–solar transition portfolios. Direct effective-weight sampling avoids privileging a reference hierarchy; co-equal ecological rules expose compensability; county and interval analyses retain structured evidence uncertainty; and aggregation reveals support-dependent replacement. Independent solar associations also change with outcome definition.

The principal result is quantified portfolio conditionality rather than a uniquely defensible Long Island readiness ranking. Local evidence and legitimate planning objectives are required to translate this screening audit into an operational decision.

Data and code availability

The research data and analysis code accompany this manuscript as *Supplementary Material S1: Research Data and Reproducibility Package*. It contains the processed indicator data, policy configurations, portfolio results, analysis scripts, software requirements, and instructions for reconstructing the study from the cited source datasets. Source metadata, retrieval links, and integrity checks are provided. A persistent public repository has not yet been established.

References

- [1] Rhys Williams (2019). ‘This Shining Confluence of Magic and Technology’: Solarpunk, Energy Imaginaries, and the Infrastructures of Solarity. *Open Library of Humanities* 5(1):60. doi:10.16995/olh.329.
- [2] Pamela Mang; Bill Reed (2012). Designing from place: a regenerative framework and methodology. *Building Research & Information* 40(1):23-38. doi:10.1080/09613218.2012.621341.
- [3] Jacek Malczewski (2006). GIS-based multicriteria decision analysis: a survey of the literature. *International Journal of Geographical Information Science* 20(7):703-726. doi:10.1080/13658810600661508.
- [4] Reid Ewing; Robert Cervero (2010). Travel and the Built Environment. *Journal of the American Planning Association* 76(3):265-294. doi:10.1080/01944361003766766.
- [5] Loghman Khodakarami; Saeid Pourmanafi; Zahra Mokhtari; Ali Reza Soffianian; Ali Lotfi (2023). Urban sustainability assessment at the neighborhood scale: Integrating spatial modellings and multi-criteria decision making approaches. *Sustainable Cities and Society* 97:104725. doi:10.1016/j.scs.2023.104725.
- [6] Maria Gracia Riera Pérez; Emmanuel Rey (2013). A multi-criteria approach to compare urban renewal scenarios for an existing neighborhood. Case study in Lausanne (Switzerland). *Building and Environment* 65:58-70. doi:10.1016/j.buildenv.2013.03.017.
- [7] Ian T. Jolliffe; Jorge Cadima (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2065):20150202. doi:10.1098/rsta.2015.0202.
- [8] Peter J. Rousseeuw (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20:53-65. doi:10.1016/0377-0427(87)90125-7.
- [9] P. A. P. Moran (1950). Notes on continuous stochastic phenomena. *Biometrika* 37(1-2):17-23. doi:10.1093/biomet/37.1-2.17.
- [10] Filippas Lygerakis; Nikolaos Kampelis; Dionysia Kolokotsa (2026). Smart, data-driven decision support for neighborhood-scale energy retrofit planning. *Energy and Buildings* 362:117524. doi:10.1016/j.enbuild.2026.117524.
- [11] Chia-Nung Li (2026). An expert-based dynamic indicator framework for sequencing net-zero neighbourhood transitions: A Taiwan expert-elicitation study. *Environmental and Sustainability Indicators* 31:101454. doi:10.1016/j.indic.2026.101454.
- [12] Kirsten Jenkins; Darren McCauley; Raphael Heffron; Hannes Stephan; Robert Rehner (2016). Energy justice: A conceptual review. *Energy Research & Social Science* 11:174-182. doi:10.1016/j.erss.2015.10.004.
- [13] A S Fotheringham; D W S Wong (1991). The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A: Economy and Space* 23(7):1025-1044. doi:10.1068/a231025.
- [14] Giuseppe Munda; Michela Nardo (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics* 41(12):1513-1523. doi:10.1080/00036840601019364.
- [15] M. Saisana; A. Saltelli; S. Tarantola (2005). Uncertainty and Sensitivity Analysis Techniques as Tools

- for the Quality Assessment of Composite Indicators. *Journal of the Royal Statistical Society Series A: Statistics in Society* 168(2):307-323. doi:10.1111/j.1467-985X.2005.00350.x.
- [16] OECD and European Commission Joint Research Centre (2008). Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD Publishing. <https://doi.org/10.1787/9789264043466-en>.
 - [17] Chapman, J.; Fox, E. H.; Bachman, W.; Frank, L. D.; Thomas, J.; Rourk Reyes, A. (2021). Smart Location Database: Technical Documentation and User Guide, Version 3.0. U.S. Environmental Protection Agency. https://www.epa.gov/system/files/documents/2023-10/epa_sld_3.0_technicaldocumentationuserguide_may2021_0.pdf.
 - [18] The Nature Conservancy and Defenders of Wildlife (2021). Long Island Solar Roadmap. Project report. https://solarroadmap.org/wp-content/uploads/2021/05/LongIslandSolarRoadmapReport_2020_LowRes.pdf.
 - [19] The Nature Conservancy and Defenders of Wildlife (2021). Long Island Solar Roadmap: Spatial Analysis Methods. Project methodology. https://solarroadmap.org/wp-content/uploads/2021/03/Long-Island-Solar-Roadmap_Spatial-Analysis-Methods.pdf.
 - [20] Dewitz, J. (2023). National Land Cover Database (NLCD) 2021 Products. U.S. Geological Survey data release; service snapshot documented in archive. <https://doi.org/10.5066/P9JZ7A03>.
 - [21] Multi-Resolution Land Characteristics Consortium (2026). National Land Cover Database class legend and description; legacy land-cover coverage service. Retrieved 7 October 2026. <https://www.mrlc.gov/data/legends/national-land-cover-database-class-legend-and-description>.
 - [22] New York State Energy Research and Development Authority (2026). Solar Electric Programs Reported by NYSERDA: Beginning 2000; dataset, overview and data dictionary. Open NY, dataset 3x8r-34rs; archived 7 October 2026. <https://data.ny.gov/Energy-Environment/Solar-Electric-Programs-Reported-by-NYSERDA-Beginn/3x8r-34rs>.

A Reference configuration and technical sensitivities

The illustrative control uses effective weights (1/3, 1/3, 2/9, 1/9), targets 0.75, and solar emphasis 0.25. These derive from earlier importance weights multiplied by assumed coefficients (3, 3, 2, 1); the reference receives no priority over the primary policy-space specifications. The reference

Table 7: Sensitivity relative to the reference block-group score. Shortlist overlap uses exactly 214 units, resolving ties by source identifier.

Alternative	Spearman ho	Top-decile overlap
Empirical-rank normalization	0.990	91.1%
Representative-point allocation	0.999	98.6%
Lower rooftop capacity	1.000	96.7%
Equal burden coefficients	0.954	92.5%
Solar opportunity removed	0.932	61.2%
Ground capacity included	0.975	92.5%
Upper-transit scenario	0.809	83.6%
Street domain omitted	0.934	80.8%
Transit domain omitted	0.906	77.6%
Services domain omitted	0.962	87.4%
Ecological domain omitted	0.993	97.2%
Population-weighted tract aggregation	0.855	51.9%

Table 8: Constrained selection under reference policy. The top 214 eligible units are selected, or all eligible units when fewer exist. Ecological floors use raw classified cover; all-domain floors use normalized domains. These are illustrative safeguards, not locally agreed requirements.

Rule	Eligible L / U	Selected	Nassau	Reference retained
Ecological 0.05	544 / 544	214	77	8
Ecological 0.1	418 / 418	214	61	5
Ecological 0.2	277 / 277	214	36	3
All domains 0.25	14 / 35	14	12	5
All domains 0.5	1 / 1	1	1	1

lower/upper full-ranking correlation is 0.809, equal-size overlap 83.6%, and median interval rank width 693. There are 179 interval-guaranteed units at this fixed policy. Earlier experiments identified 130 high joint-endpoint-frequency units in the local policy space and 74 high interval-guarantee-frequency units in the broader factorized space. Those values belong to different sampling measures from the direct-effective-weight primary audit.

B Supplementary profiles and spatial dependence

PCA standardizes the four lower domains and solar opportunity [7]; its first two variance shares are 45.4% and 28.3%. Its first component correlates 0.944 with the reference score through shared inputs. K-means examines $k = 2, \dots, 6$ with 30 starts, and silhouette selects five profiles at 0.302 [8]. Alternative seed agreement has minimum adjusted Rand index 0.990. These summaries describe observations and algorithmic stability rather than independent construct validity. Profile

Table 9: Descriptive profiles selected by silhouette. Entries are means of normalized lower-scenario domains and solar opportunity. Profiles are not externally validated neighborhood categories.

Profile	n	Street	Transit	Services	Ecology	Solar	S
1	397	0.694	0.709	0.794	0.014	0.708	0.765
2	494	0.692	0.651	0.479	0.005	0.125	0.556
3	519	0.318	0.090	0.719	0.084	0.595	0.444
4	287	0.057	0.020	0.471	0.779	0.211	0.256
5	439	0.377	0.056	0.311	0.063	0.134	0.252

1 combines infrastructure and solar strengths with very low classified ecological cover; profile 4 emphasizes ecology with weaker infrastructure evidence. The contrast explains decision-rule differences. Original exploratory figures remain archived; the principal figures emphasize portfolio persistence and evidence availability.

Moran’s $I = 0.604$ for row-standardized eight-nearest-centroid neighbors, with 999 permutations giving $p = 0.001$ [9]. It describes concentration for the specified graph and cautions against interpreting block groups as independent regional replicates.

C Distributions and source reconciliation

Table 10: Raw indicator distributions across valid archived observations. Sentinel counts are preserved separately. Quantiles here summarize raw rather than normalized or log-transformed values.

Indicator	Valid n	Sentinel n	Median	$q_{.05}$	$q_{.95}$
Street intersections / mi ²	2136	0	103.490	17.437	214.963
Transit distance (m)	1137	999	470.730	136.790	1053.048
Transit service / mi ²	1222	914	66.274	2.923	470.745
Employment–housing entropy	2136	0	0.553	0.177	0.852
Service jobs / km ²	2136	0	90.162	0.273	1306.641
Ecological fraction	2136	0	0.000	0.000	0.447

Transform order is raw values, declared logs, valid-data quantiles, clipping, orientation, averaging, target gaps, and policy combination. Transit sentinels remain intervals after exclusion from valid distributions. Raster class fractions sum to one among retained nonwater pixels. Identifiers remain text and ascending order resolves selection ties. Distinct capacity totals are

Table 11: Capacity reconciliation: separate archived quantities are not forced to agree. The component-based high total sums rooftop-high, parking, and ground fields.

Quantity	MW
Archived rooftop-low component	2,369.2
Archived rooftop-high component	3,173.6
Archived parking component	3,493.9
Archived ground component	12,713.1
Component-based high total used here	19,380.6
Archived combined <code>TotalHigh_MW</code> field	19,455.5
Published Roadmap total (approximately)	19,424
Built-surface high total used here	6,667.5

preserved rather than forced to agree. Reproduction requires versioned source snapshots because fresh service responses can change. Environment versions and hashes are archived; the generated LaTeX source is self-contained.

D Reporting-threshold sensitivity

Continuous membership frequencies and paired portfolio overlaps underpin the interpretation. The 0.8 threshold offers a concise summary of candidates meeting the sufficient interval-membership condition in at least four-fifths of sampled policies; it is a reporting convention rather than an empirically identified decision cutoff. Table 12 compares thresholds 0.7, 0.8, and 0.9 without changing any policy draws or portfolio rules.

Table 12: Candidates with interval-guarantee frequency at or above the stated reporting threshold. Each column uses the same 20,000 policies; changing the reporting threshold does not change portfolio construction.

Decision rule	$g_i \geq 0.7$	$g_i \geq 0.8$	$g_i \geq 0.9$
Unrestricted	94	58	15
Ecology $\geq 5\%$	70	33	16
Ecology $\geq 10\%$	67	30	17
Ecology $\geq 20\%$	76	35	12
Nassau only	66	48	26
Suffolk only	4	1	0

The unrestricted count declines from 94 to 58 to 15, while the 10% ecological-rule count declines from 67 to 30 to 17. These differences caution against treating the size of a thresholded set as an intrinsic property of a location or as a statistically certified priority list. The paired median overlap of 9.3% is calculated from all policy draws and is independent of this reporting threshold. Complete lower-endpoint, joint-endpoint, and guarantee-frequency counts, including county breakdowns, are provided in Supplementary Material S1.