Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Temporal portfolio stress-testing for candidate 5G upgrades under future operational burden

runzhe liu

2025082583@ynjtc.edu.cn

Yunnan Communications Vocational and Technical College

hui chen

Yunnan Communications Vocational and Technical College

yun he

Yunnan Communications Vocational and Technical College

dabin wang

Yunnan Communications Vocational and Technical College

Research Article

Keywords: 5G deployment, base-station upgrade, temporal validation, portfolio selection, multiobjective optimisation, regional equity

Posted Date: June 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10103807/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Mobile operators must prioritise 5G upgrades before future pressure is observed, while balancing future burden capture, user reach, spatial redundancy, regional equity and engineering complexity. Many scoring studies evaluate the same month used to construct the ranking and therefore cannot distinguish a deployable decision rule from a descriptive index. We developed a temporal portfolio stress test using two monthly observations from 14,129 macro sites in a provincial network. December traffic, 5G backflow, user and spatial-demand features were used for selection; January measurements were reserved for evaluation. Backflow was highly persistent (Spearman rho = 0.968), whereas high-4G/no-5G days were unstable (rho = 0.252) and were excluded from the primary target. In eight-fold city-grouped validation, gradient boosting predicted January burden with R2 = 0.847, mean absolute error = 0.071 and rank correlation = 0.911. We compared score baselines, a city-quota forecast baseline and greedy scalarised portfolios with preference, spatial, equity and cell-count terms. The key finding was score saturation rather than a large forecasting gain: at a 100-site budget, constrained portfolios changed future-burden capture by less than 0.02 percentage points relative to forecast-only ranking, but changed about half of selected sites. The spatial-equity portfolio increased the Jain index from 0.397 to 0.967 and represented all 16 cities. An engineering-aware extension reduced selected cell records from 217 to 193 while retaining 98.6% of forecast-only capture. The framework screens candidate upgrades under future operational burden; it does not estimate post-upgrade causal throughput, revenue or experience gains.

1 Introduction

The transition from fourth-generation coverage to fifth-generation capacity is not a uniform replacement problem. Operators inherit thousands of sites with heterogeneous traffic, device populations, user backflow, radio conditions and surrounding demand. The practical decision is therefore incremental: under a fixed annual or quarterly budget, which existing macro sites should be treated as candidate upgrades first? The answer affects where engineering review is concentrated, how much future operational burden is prioritised, and how investment attention is distributed across urban and rural jurisdictions. A priority list is useful only if it is honest about what it measures. A site with high future burden is a strong candidate for engineering review, but this load signal is not evidence that an upgrade will causally produce proportional throughput, revenue or experience gains.

5G network research has established the architectural importance of densification, heterogeneous radio access, backhaul and flexible resource management [1,4,5]. Techno-economic studies further show that rollout cost and coverage depend strongly on the spatial distribution of demand and the extent to which existing infrastructure can be reused [2,3]. Site-planning studies have coupled geographic information systems with heuristic optimisation [6], while performance-evaluation studies have combined expert weighting and multicriteria decision-making (MCDM) [7]. These contributions clarify what should enter a planning model, but two evaluation gaps remain common.

First, many priority scores are assessed cross-sectionally. The same traffic and user variables both construct and justify the ranking, so the evaluation measures internal consistency rather than future usefulness. A score that selects the largest December backflow values will necessarily capture much of December backflow; it is not known whether that advantage transfers to January or to a different city. Temporal separation is especially important when operational indicators have different persistence. A monthly backflow measure may be stable, whereas threshold-event counts can shift with load, maintenance or reporting rules.

Second, a single fixed weight vector conceals preference uncertainty. Network optimisation is multiobjective by construction [10]. Commercial teams may value addressable users, network teams may value traffic relief, and public-service obligations may value regional coverage. Entropy weighting or TOPSIS can make aggregation reproducible [8,9], but they do not eliminate uncertainty about whose utility is represented. Robust optimisation instead asks whether a decision remains acceptable across a set of plausible preferences [11]. For a site portfolio, robustness also has a spatial dimension: selecting several nearby high-score sites can create redundant geographic concentration, and maximising province-wide burden capture can leave low-volume cities unrepresented.

Here we treat candidate 5G upgrade selection as a temporal portfolio stress test rather than as a new radio protocol or a standalone optimiser. The individual ingredients - forecasting, MCDM, diversity penalties and equity metrics - are familiar. The contribution is the locked separation between decision-time information and future evaluation, plus a component-wise audit of what changes after predictive ranking has largely saturated. December observations construct the candidate lists; January observations evaluate them. The comparison includes raw-score baselines, entropy-TOPSIS, forecast-only ranking, city-quota allocation, spatial-equity robust selection, an engineering-aware robust variant and a hindsight oracle.

The central claim is deliberately narrower than an optimisation-theory claim. We do not argue that a new solver dominates existing operations-research methods. Instead, we ask whether a telecom planning screen that looks almost identical in future-burden capture can produce materially different and more balanced engineering shortlists. The study therefore tests six questions: which operational signals persist across months; how accurately future burden transfers across cities; whether score-based rankings have reached a practical capture plateau; how much city equity and spatial reach can be gained at that plateau; whether a cell-count complexity term changes the engineering shortlist; and whether paired city-level contrasts support the apparent trade-offs.

2 Data and methods

2.1 Operator data and physical-site aggregation

The source workbook contained network-optimisation and market indicators for a provincial mobile network. The December and January high-4G/no-5G sheets each contained 18,591 cell-level records

corresponding to 14,129 physical macro sites in 16 cities. Multiple sector or cell records could share one physical site. We aggregated user count, 5G backflow traffic, average site traffic and high-4G/no-5G days by the median within physical site; city, county and coordinate fields used the first consistent record, and the number of distinct cells was retained as a predictor.

The two monthly physical-site sets matched one-to-one. December served as the decision month and January as the future evaluation month. Although one workbook sheet was labelled February, its note stated that the high-4G/no-5G user and traffic fields corresponded to January; we therefore refer to the second observation as January throughout. A separate February traffic field was explicitly marked by the provider as not recommended and was not used.

Three sources of external demand points were extracted: 141 property-difficulty or change points, 128 complaint/point-of-interest records and 212 competitor-coverage points. For each physical site and source, we calculated counts within 1 and 3 km and distance to the nearest point using haversine distance. The external points are sparse context features; they do not define the outcome.

Generative AI tools assisted code scaffolding, manuscript drafting and language revision under author supervision; the authors verified the analyses, claims, tables and figures in the final manuscript.

Deidentified derived tables, figure source data and the reconstruction script are supplied as Online Resource 1.

Table 1
Operator-network observations and external demand-point data.

Data element	Count	Analytical use
Cell-level records per month	18,591	Aggregated to physical sites
Physical macro sites	14,129	Matched one-to-one across months
Cities	16	Grouped outer validation
Property demand points	141	Counts and nearest distance
Complaint / POI points	128	Counts and nearest distance
Competitor points	212	Counts and nearest distance

2.2 Temporally separated burden definition

Raw users, backflow and traffic were heavy-tailed and measured in different units. For site i, metric m and month t, we used the empirical percentile

$$z_{i,t}^{m} = \frac{\operatorname{rank}\left(x_{i,t}^{m}\right)}{N} \quad (1)$$

where N = 14,129 and average ranks were used for ties. The equal-weight composite burden was

$$b_{it} = \frac{z_{i,t}^{U} + z_{i,t}^{B} + z_{i,t}^{T}}{3} \quad (2)$$

Here U, B and T denote user reach, backflow and traffic, respectively. This quantity is an ordinal operational burden, not a predicted engineering throughput gain or a causal treatment effect of upgrading the site. It gives equal influence to user reach, offloaded 5G traffic and current traffic intensity while limiting dominance by extreme raw values. The number of high-4G/no-5G days was analysed as a stability diagnostic and in the entropy-TOPSIS comparator, but excluded from the primary burden target because its month-to-month persistence was low.

The decision protocol was strict: all strategy scores used December variables only. January burden and component metrics were opened only for evaluation. The January oracle orders sites by realised January burden and is reported as an unattainable upper benchmark, not as a deployable strategy.

2.3 Cross-city forecasting of January burden

We modelled future burden as

$$\hat{b}_{i,\mathrm{Jan}} = f_{\theta}\left(\mathbf{x}_{i,\mathrm{Dec}}, \mathbf{d}_{i}\right) \quad (3)$$

where x_i,Dec contains log-transformed December users, backflow and traffic, persistence days, cell count and coordinates, and d_i contains the external demand-point features. We compared elastic net, random forest [14] and histogram gradient boosting [13]. The models were not presented as novel architectures; they test whether nonlinear relationships improve temporal transfer.

Evaluation used eight-fold GroupKFold cross-validation by city. Every site in a city was assigned to the same outer fold, so the predicted January burden for each site came from a model that had not observed January outcomes from that city. Hyperparameters were fixed before the final comparison: elastic net alpha 0.002 and L1 ratio 0.2; random forest 260 trees, maximum depth 18 and minimum leaf size 5; gradient boosting learning rate 0.06, 31 leaves, L2 regularisation 1.0 and 260 iterations. We report mean absolute error, root mean square error, R2 and Spearman rank correlation.

2.4 Deterministic comparators

Six deployable score-based strategies were constructed. **Backflow only** ranked sites by December 5G backflow. **Equal-weight December** ranked by Eq. 2. **Entropy-TOPSIS** added December persistence days to the three burden criteria, estimated entropy weights and calculated distance to positive and negative ideal solutions [8,9]. **Forecast model** ranked by the best city-held-out estimate from Eq. 3. **City-quota forecast** allocated the portfolio across cities in proportion to city-level forecast mass and then selected the highest forecasted sites within each city. **Weight-robust score** sampled preferences rather than choosing one fixed vector. The city-quota comparator is important because it separates the effect of explicit regional representation from the effect of the later spatial and preference-robust portfolio terms.

For the robust score, 1,000 vectors w = (w_user, w_backflow, w_traffic) were drawn from a symmetric Dirichlet distribution. Each vector is non-negative and sums to one. Scenario utility was

$$u_i(\mathbf{w}) = \sum_m w_m z_{i,\mathrm{Dec}}^m \quad (4)$$

We added a small external-demand term and penalised preference sensitivity:

$$q_i = \mathbb{E}_{\mathbf{w}}[u_i(\mathbf{w})] - 0.5\,\mathrm{SD}_{\mathbf{w}}[u_i(\mathbf{w})] + 0.04 d_i \quad (5)$$

where d_i is the mean percentile of the three 3-km demand-point counts. The standard-deviation term favours sites that remain useful when business priorities change. It does not represent statistical prediction uncertainty.

2.5 Robust portfolio selection

A ranking can still select geographically redundant sites. We therefore constructed a sequential robust portfolio. At step k, each unselected site received marginal value

$$\Delta_i(k;\eta) = q_i - 0.12 \max_{j \in S_k} \exp[-\mathrm{dist}(i,j)/3] + 0.10 h_i(k) - \eta\, c_i \quad (6)$$

where S_k is the selected set, distance is in kilometres, h_i(k) is the city shortfall capped at one, c_i is a normalised cell-count complexity proxy and eta is the complexity penalty. Eq. 6 is a weighted-sum scalarisation of four auditable planning terms: robust demand utility, local redundancy, regional shortfall and cell-count complexity. The coefficients are interpretable shadow prices rather than fitted model parameters. The expected number of sites for each city at step k was proportional to that city's total robust score. A city received a bonus when its current count was below that expectation. The exponential term reduces selection of near-duplicate sites within an approximate 3-km scale, and the cell-count term favours candidates with fewer associated cell records when two sites offer similar demand value. This greedy construction follows the general logic of diminishing-return set selection [12], but Eq. 6 is an operational scalarisation rather than a convex relaxation or a proof of global optimality.

The spatial and city coefficients were fixed at 0.12 and 0.10 before future evaluation. We did not tune them on January burden. The main robust portfolio used eta = 0 so that its equity-spatial trade-off remained directly comparable with the earlier sensitivity analysis. We then evaluated an engineering-aware extension with eta = 0.06 and a penalty sweep from 0 to 0.15. The 3-km scale is a transparent local-redundancy proxy in the absence of engineering coverage polygons, propagation simulations or interference maps; it should not be interpreted as a measured cell radius. The city shortfall term represents a provincial investment or public-service setting in which every city should remain visible in the candidate portfolio. It is not the only valid objective. A purely commercial return-on-investment screen could set the city bonus to zero. The design prioritises interpretability: an operator can see how much future burden capture is exchanged for geographic separation, city representation and cell-count complexity.

2.6 Portfolio evaluation

Strategies were evaluated at budgets of 25, 50, 100, 200, 300 and 500 sites. Primary efficiency was the share of total January composite burden contained in the selected portfolio. We also calculated shares of raw January backflow and users. Preference robustness was the fifth percentile and median capture across the 1,000 weight scenarios applied to January percentiles.

Regional equity used city service fractions. For city c, the service fraction was selected January burden divided by total January burden in that city. Equity was summarised with the Jain index

$$J = \frac{\left(\sum_c a_c\right)^2}{C \sum_c a_c^2} \quad (7)$$

where a_c is the service fraction and C = 16. J ranges from 1/C to 1; higher values indicate more even fractions across cities. Geographic reach was the number of cities represented. Spatial redundancy was summarised by the mean nearest-neighbour distance among selected sites; higher values indicate less clustering.

Uncertainty in the efficiency difference between robust and equal-weight portfolios at k = 100 was assessed by resampling cities with replacement 1,500 times. Selection sets were fixed and January burden was reaggregated within sampled cities. This cluster bootstrap preserves the heterogeneity of city scale. We also used exact paired sign-flip tests over the 16 city-level differences to compare constrained portfolios against forecast-only ranking. Tests were applied to province-wide burden-capture differences decomposed by city and to differences in mean absolute deviation of city service fractions. Finally, we calculated Jaccard overlap and turnover of the 100-site lists. These diagnostics separate numerical efficiency differences from material changes in the engineering shortlist.

Parameter sensitivity was evaluated at k = 100. We varied the redundancy radius over 1.5, 3, 5 and 8 km; the redundancy penalty over 0, 0.06, 0.12 and 0.18; the city shortfall bonus over 0, 0.05, 0.10 and 0.20; and the symmetric Dirichlet concentration over 0.5, 1, 2 and 5. For each setting, the robust portfolio was reconstructed using December information only and evaluated against the same January outcomes. Sensitivity results therefore test the stability of the decision rule, not post hoc tuning.

Because upgrade cost, propagation footprints, interference maps, backhaul headroom and spectrum assignments were unavailable, we added an engineering-complexity proxy audit rather than a false cost model. At k = 100, we recorded the number of distinct cell records attached to the selected physical sites, mean cells per selected site, the share of selected sites above the 75th percentile of cell count, burden capture per 100 selected cell records and the same equity and reach metrics used in the main portfolio evaluation. Distinct cell count is a coarse complexity proxy: it can indicate multi-sector or multi-cell sites, but it is not a monetary cost, capacity model or coverage constraint.

2.7 Algorithmic complexity and implementation scope

Let N be the number of physical sites, K the portfolio budget, M the number of burden criteria, W the number of sampled preference vectors and C the number of cities. In this study, N = 14,129, M = 3, W = 1,000, C = 16 and the largest evaluated K was 500. The deterministic score baselines require O(N log N) sorting after feature construction. The city-quota forecast baseline adds only O(C + K log N) allocation and within-city selection after the forecast scores have been sorted. The preference-robust score requires O(WMN) arithmetic to evaluate scenario utilities and O(N) memory for the resulting mean and dispersion summaries. The sequential portfolio rule evaluates all currently unselected sites at each step after maintaining city shortfalls, nearest selected-site distances and the normalised cell-count complexity term, giving O(KN) marginal-value updates and O(N + C) working memory. The complexity penalty does not change the asymptotic order; it changes which nearly equivalent candidates are preferred when the forecast, diversity and city terms are close. This is deliberately simpler and more auditable than a mixed-integer or evolutionary multiobjective solver, and it returns a complete ranking that can be truncated at several budgets. The trade-off is that the greedy rule is not a proof of global optimality; it is a transparent planning heuristic evaluated out of time against January outcomes.

2.8 Confidentiality and reproducibility

The analysis does not use individual subscriber records. Site identifiers were hashed for derived tables, cities were coded C01-C16 in public outputs, and exact coordinates and operator site names are excluded from the manuscript package. Analyses were executed in Python 3.12 with pandas, scikit-learn, NumPy, SciPy and Matplotlib. The complete reconstruction script fixes random seeds and generates all anonymised tables and figures from the restricted source workbook when that workbook is available under the operator's permission.

3 Results

3.1 Operational signals had markedly different temporal persistence

Backflow traffic was the most persistent signal, with December-January Spearman rho = 0.968. Composite burden was also stable (rho = 0.888), followed by user count (0.777) and average traffic (0.715). In contrast, high-4G/no-5G days had rho = 0.252 (Table 2; Fig. 2a). The latter measure therefore captured a short-lived threshold condition rather than a stable priority target.

The hexagonal density plot of composite burden showed a strong diagonal concentration but substantial dispersion among middle- and high-burden sites (Fig. 2b). This dispersion creates room for forecasting: a December rank is informative, but it is not an oracle for January. It also justifies evaluating site portfolios out of time rather than reporting December concentration.

Table 2
Temporal rank stability of site-level operational signals.

Operational signal	December-January Spearman rho
Backflow	0.968
Users	0.777
Traffic	0.715
High-4G/no-5G days	0.252
Composite burden	0.888

3.2 Nonlinear models transferred future burden across cities

Gradient boosting provided the strongest city-held-out forecasts, with Spearman correlation 0.911, R2 0.847, mean absolute error 0.071 and RMSE 0.096 (Table 3). Random forest was close (rho = 0.907; R2 = 0.842). Elastic net remained competitive but weaker (rho = 0.884; R2 = 0.767), indicating that the major temporal signal was broadly smooth while nonlinear interactions added measurable value.

The grouped protocol is important. Predictions for each city were generated without January outcomes from that city, so the result cannot be explained by a city-specific intercept learned from the evaluation month. Nevertheless, the model estimates future operational burden, not post-upgrade gain. A site with high predicted burden is a plausible upgrade candidate for engineering review; the prediction does not imply that upgrading the site would cause a proportional throughput, revenue or experience gain.

Table 3
Eight-fold city-grouped performance for January burden forecasting.

model	MAE	RMSE	R2	Spearman
Elastic net	0.092	0.118	0.767	0.884
Random forest	0.072	0.097	0.842	0.907
Gradient boosting	0.071	0.096	0.847	0.911

3.3 A single raw metric concentrated one objective at the expense of balanced burden

At k = 100, backflow-only ranking captured 10.66% of raw January backflow, more than any multicriteria strategy. Its composite burden capture was only 1.165%, however, and it captured 1.28% of January users. The strategy represented 10 cities and had a Jain service index of 0.328. This contrast illustrates

why raw backflow capture cannot serve as the sole measure of portfolio quality: it rewards a small number of extreme sites and suppresses user reach and regional balance.

Equal-weight December ranking captured 1.342% of total January composite burden at k = 100. Entropy-TOPSIS achieved 1.341%, and the forecast model achieved 1.349%. The city-quota forecast baseline captured 1.336%, retaining 99.1% of forecast-only capture while distributing the 100 selections across all 16 cities. The hindsight oracle reached 1.384%; thus, forecast ranking retained 97.4% of the oracle's composite capture. Differences were small in absolute province-wide share because 100 sites represent only 0.71% of the 14,129-site universe, but relative efficiency provides a meaningful comparison.

3.4 The robust portfolio exchanged little future capture for large equity gains

At a 100-site budget, the spatial-equity robust portfolio captured 1.334% of total January composite burden. This was 98.9% of forecast-only capture and 96.3% of the hindsight oracle. Its fifth-percentile capture across preference scenarios was 1.322%, compared with 1.342% for forecast-only ranking. The city-block bootstrap difference from equal-weight December was centred near zero (median 0.00014 of province-wide burden capture; 95% interval − 0.00609 to 0.00587), providing no evidence of a material average efficiency gain or loss.

The distributional differences were large. Forecast-only ranking selected sites in 9 cities and produced a Jain index of 0.397. Equal-weight December covered 10 cities with J = 0.427. The city-quota forecast baseline covered all cities with J = 0.918, showing that quota allocation alone can improve equity but does not address spatial redundancy or preference sensitivity. The spatial-equity robust portfolio covered all 16 cities and reached J = 0.967. Mean nearest-neighbour distance increased from 18.1 km for forecast ranking and 17.1 km for equal weights to 28.7 km for the robust portfolio. The result is therefore a Pareto-style trade-off: almost all future composite capture was retained while geographic concentration and city imbalance were substantially reduced.

Table 4
Future efficiency, preference robustness and equity at a 100-site budget.

Strategy	Future capture	P05 capture	Jain index	Cities	Nearest km
Backflow only	0.012	0.010	0.328	10	12.8
Equal-weight December	0.013	0.013	0.427	10	17.1
Entropy-TOPSIS	0.013	0.013	0.521	11	20.2
Forecast model	0.013	0.013	0.397	9	18.1
City-quota forecast	0.013	0.013	0.918	16	24.4
Weight-robust score	0.013	0.013	0.419	10	16.8
January oracle	0.014	0.014	0.372	9	19.7
Robust portfolio	0.013	0.013	0.967	16	28.7
Engineering-aware robust	0.013	0.013	0.946	16	25.8

Paired city-level contrasts clarified the practical meaning of these small capture differences. Relative to forecast-only ranking, city-quota, spatial-equity robust and engineering-aware robust portfolios changed future-burden capture by only − 0.012 to -0.019 percentage points, and exact sign-flip tests did not support a detectable efficiency difference ($p = 0.947–0.963$). In contrast, service-fraction imbalance decreased consistently ($p < 0.001$), and the 100-site shortlist changed materially. The spatial-equity robust list replaced 50% of forecast-selected sites, and the engineering-aware list replaced 51%. Thus, the result is not a claim of statistically significant capture improvement. It is a score-saturation finding: once temporal prediction is strong, equity and complexity constraints can redirect the engineering shortlist without a measurable loss of future burden capture.

Table 9
Paired city-level contrasts and 100-site decision turnover.

Contrast	Capture diff. (pp)	Capture p	Imbalance diff.	Imbalance p	Jaccard	Turnover
City-quota vs Forecast	-0.012	0.963	-0.0064	< 0.001	0.575	27%
Robust portfolio vs Forecast	-0.015	0.957	-0.0076	< 0.001	0.333	50%
Eng.-aware vs Forecast	-0.019	0.947	-0.0069	< 0.001	0.325	51%
Eng.-aware vs Robust	-0.004	0.752	0.0006	0.407	0.852	8%

3.5 Equity and geographic reach were stable across practical budgets

At k = 50, the robust portfolio already represented all 16 cities, compared with 7 for forecast ranking and 8 for equal weights. Its Jain index was 0.962 and mean nearest-site separation was 42.1 km. The city-quota forecast baseline also represented all cities, but with lower separation than the robust portfolio, indicating that quota control and spatial redundancy control solve related but distinct problems. At k = 200, the robust portfolio retained all cities with J = 0.948 and captured 2.625% of January composite burden, versus 2.656% for forecast ranking. The efficiency ratio was again 98.8%.

Nearest-site distance was not monotonic with budget because adding a distant city can initially increase average separation, whereas later additions within covered cities reduce it. Figure 5 should therefore be read as a realised redundancy diagnostic rather than as a mathematical constraint curve. Across the tested budgets, the robust portfolio consistently achieved broader geographic reach and higher city equity than score-only strategies.

Table 5
Efficiency-equity trade-offs at three practical site budgets.

Strategy	Budget	Future capture	Jain index	Cities	Nearest km
Equal-weight December	50	0.007	0.314	8	19.1
Equal-weight December	100	0.013	0.427	10	17.1
Equal-weight December	200	0.026	0.541	12	67.4
Forecast model	50	0.007	0.367	7	23.9
Forecast model	100	0.013	0.397	9	18.1
Forecast model	200	0.027	0.459	11	13.8
City-quota forecast	50	0.007	0.724	12	35.8
City-quota forecast	100	0.013	0.918	16	24.4
City-quota forecast	200	0.026	0.951	16	14.8
Robust portfolio	50	0.007	0.962	16	42.1
Robust portfolio	100	0.013	0.967	16	28.7
Robust portfolio	200	0.026	0.948	16	17.9
Engineering-aware robust	50	0.007	0.962	16	43.0
Engineering-aware robust	100	0.013	0.946	16	25.8
Engineering-aware robust	200	0.026	0.948	16	17.9

3.6 The robust trade-off persisted under parameter perturbations

Parameter sensitivity supported the main interpretation rather than introducing a single fragile tuning point. Across redundancy radii from 1.5 to 8 km, the 100-site robust portfolio retained 97.8–98.9% of forecast-only January burden capture and represented all 16 cities (Table 6). Increasing the radius raised mean nearest-site separation from 26.2 to 31.9 km, as expected, with a small loss of capture.

The city shortfall bonus controlled the equity-efficiency trade-off most directly. With no city bonus, the portfolio retained 99.7% of forecast-only capture but represented 11 cities and had a Jain index of 0.521. With the default 0.10 bonus, it represented all 16 cities and reached J = 0.966. Increasing the bonus to 0.20 preserved all-city representation with similar burden capture. The result clarifies the business choice: city equity is not free, but its measured cost was small in this dataset.

Table 6
Parameter sensitivity of the 100-site robust portfolio.

Sensitivity factor	Values tested	Forecast-ratio range	Jain range	Cities range
redundancy radius (km)	1.5, 3, 5, 8	0.978–0.989	0.956–0.973	16–16
redundancy penalty	0, 0.06, 0.12, 0.18	0.984–0.990	0.946–0.973	16–16
city shortfall bonus	0, 0.05, 0.1, 0.2	0.989–0.999	0.521–0.973	11–16
Dirichlet concentration	0.5, 1, 2, 5	0.988–0.991	0.959–0.973	16–16

3.7 A cell-count proxy made the engineering trade-off explicit

The engineering proxy audit did not overturn the main portfolio result, but it changed how the trade-off should be described. At k = 100, the spatial-equity robust portfolio selected 217 associated cell records, the same total as equal-weight December ranking and slightly more than forecast ranking (210; Table 7). Mean cells per selected site were 2.17 for the robust portfolio and 2.10 for forecast ranking. The share of high-cell-count sites was 0.17 for the robust portfolio and 0.09 for forecast ranking, indicating that the robust rule did not achieve broader city coverage by selecting obviously simpler sites.

The engineering-aware robust extension changed this frontier. With eta = 0.06, the selected portfolio used 193 associated cell records, represented all 16 cities, retained 98.6% of forecast-only future-burden capture and increased capture per 100 selected cell records to 0.00689. This exceeded the forecast-only value of 0.00642 and the spatial-equity robust value of 0.00615, while preserving a Jain index above 0.94. The gain is not a claim about true upgrade cost; it shows that a simple complexity term can redirect the greedy portfolio away from cell-record-heavy candidates without sacrificing the temporal validation result.

Table 7
Engineering-complexity proxy audit at a 100-site budget.

Strategy	Selected cells	Cells/site	High-cell share	Future capture	Capture/100 cells	Jain index	Cities
Backflow only	185	1.85	0.13	0.0117	0.00630	0.328	10
Equal-weight December	217	2.17	0.18	0.0134	0.00619	0.427	10
Forecast model	210	2.10	0.09	0.0135	0.00642	0.397	9
City-quota forecast	210	2.10	0.08	0.0134	0.00636	0.918	16
Robust portfolio	217	2.17	0.17	0.0133	0.00615	0.967	16
Engineering-aware robust	193	1.93	0.12	0.0133	0.00689	0.946	16
January oracle	203	2.03	0.08	0.0138	0.00682	0.372	9

3.8 Complexity-penalty sensitivity supported eta = 0.06 as a practical compromise

We then varied the cell-count penalty while keeping the forecast model, preference samples, spatial penalty and city bonus unchanged. The purpose was not to find a statistically optimal eta, but to test whether the engineering-aware result depended on a single arbitrary value. Penalties between 0.04 and 0.08 consistently reduced selected cell records relative to the spatial-equity robust portfolio while retaining more than 98% of forecast-only future-burden capture. The default eta = 0.06 occupied the clearest middle point: 193 selected cell records, all 16 cities represented, a Jain index of 0.946 and 98.6% of forecast-only capture.

Table 8
Complexity-penalty sensitivity for the engineering-aware robust portfolio at k = 100.

eta	Forecast ratio	Selected cells	Capture/100 cells	Jain index	Cities
0.00	0.989	217	0.00615	0.967	16
0.02	0.989	208	0.00641	0.964	16
0.04	0.987	205	0.00649	0.964	16
0.06	0.986	193	0.00689	0.946	16
0.08	0.985	191	0.00696	0.953	16
0.10	0.985	190	0.00699	0.953	16
0.15	0.979	177	0.00746	0.953	16

4 Discussion

4.1 Temporal holdout changes what a priority model is allowed to claim

The central methodological point is simple: site prioritisation should be tested on information that was unavailable when the ranking was created. Cross-sectional concentration is easy to obtain: ranking by backflow will capture backflow in the same month. The January holdout asks a more relevant question: would the December decision still locate future burden? The high persistence of backflow and composite burden explains why simple scores remained competitive, while the dispersion in Fig. 2 explains the additional value of nonlinear forecasting.

The grouped city design further separates interpolation from transfer. Mobile networks contain city-specific demand, terrain and deployment history. Random site folds would allow thousands of neighbouring sites from a test city to influence the fitted model. Grouped folds instead test whether relationships learned elsewhere carry into an unseen city. The resulting rho of 0.911 indicates strong transfer within the province, but external validation in another province or operator remains necessary.

4.2 Weight uncertainty is a governance problem as well as a statistical problem

Fixed MCDM weights are often treated as technical parameters. In practice, they encode institutional preferences. Entropy weights describe variation in the data, not the value an operator places on each objective. Our Dirichlet scenarios do not identify the correct preference vector; they expose how much a site score depends on plausible changes in emphasis. The lower-confidence score rewards sites that are jointly useful for users, traffic and backflow rather than exceptional on one criterion.

This distinction explains why backflow-only ranking appears dominant under raw backflow but weak under balanced burden and equity. No metric is universally correct. A network operation centre addressing acute offload may deliberately choose backflow-only ranking. A province-wide investment programme with public-service or geographic obligations may prefer the robust portfolio. The manuscript therefore reports the consequence of each choice instead of presenting one composite score as objective truth.

4.3 Diversity and equity can be added at modest efficiency cost

The main finding is a capture plateau rather than a new highest-capture score. At k = 100, forecast-only, city-quota, spatial-equity robust and engineering-aware robust portfolios occupied a narrow future-capture band, and paired city tests did not detect a meaningful efficiency separation. Their shortlists were nevertheless very different. This is the operational value of the framework: it shows when prediction accuracy has become less decisive than the governance constraints placed on the shortlist. In that regime, the relevant question is not only which score is largest, but which constrained list preserves the same future-burden envelope while satisfying regional and engineering requirements.

Spatial separation also reduced redundant concentration. The exponential penalty does not model radio interference or overlapping propagation footprints; it is a planning proxy. Its practical role is to prevent several nearby high-score sites from exhausting a limited portfolio when a slightly lower-score site would expand geographic reach. The sensitivity analysis showed that the qualitative trade-off persisted across 1.5-8 km radii, but the radius still remains an engineering assumption. The cell-count analysis adds a second guardrail. Without a complexity term, the robust portfolio used a similar or slightly heavier cell-count envelope than score-only strategies. With the engineering-aware term, the portfolio moved toward a more favourable efficiency-complexity frontier while keeping all-city representation. A production system should replace both proxies with coverage polygons, backhaul constraints, spectrum constraints and estimated upgrade cost where available.

4.4 Relation to network planning and MCDM literature

Prior GIS-heuristic planning optimises site location under spatial coverage and network constraints [6], whereas the present problem is a brownfield upgrade portfolio over existing sites. Prior hybrid MCDM work formalises indicator systems and expert preferences [7]. Our framework complements both by introducing a future month as the evaluation target and treating preference, redundancy and equity as separate measurable dimensions.

The approach is deliberately simpler than evolutionary multiobjective optimisation such as NSGA-II [15]. A greedy sequence is easier to audit, produces a complete ordering for multiple budgets and has an O(KN) portfolio stage that is practical at 14,129 sites. It does not guarantee the global optimum of Eq. 6. The city-quota forecast baseline shows what can be gained by a classical allocation rule; the engineering-aware variant shows how a lightweight complexity proxy can be added without changing the

evaluation design. Energy-aware cloud-RAN studies have additionally shown how consolidation and base-station sleeping can be incorporated into operational control [16]; those actions are complementary to, rather than substitutes for, capital-upgrade prioritisation. If detailed cost, capacity and coverage simulations become available, the same temporally held-out evaluation can be applied to mixed-integer, evolutionary or submodular solvers. The evaluation protocol is independent of the optimiser.

4.5 Limitations

Five limitations define the scope. First, only two monthly snapshots were available; the model captures one-step persistence rather than seasonal dynamics. Second, the target is a percentile composite of operational burden, not observed post-upgrade throughput, retention or revenue. Third, exact engineering coverage, interference, backhaul capacity and upgrade cost were unavailable, so the portfolio uses an equal site budget, a distance proxy and a cell-count complexity proxy; this proxy improves auditability but is not a substitute for an engineering bill of materials. Fourth, external demand points were sparse and may reflect reporting intensity. Fifth, the data come from one provincial operator network and cannot establish national transfer.

The next validation should follow selected and unselected sites after deployment, estimate realised incremental outcomes and include at least 12 months to model seasonality and structural breaks. Cost and coverage models can then replace equal budgets and distance penalties. A prospective trial in which candidate sites are randomised within priority strata would provide the strongest evidence of causal upgrade benefit.

5 Conclusion

A credible candidate-upgrade portfolio requires more than a high composite score. In 14,129 physical sites, December backflow, users and traffic contained substantial information about January burden, and a city-held-out gradient-boosting model achieved R2 = 0.847 and rank correlation 0.911. The decisive result was not a statistically significant capture gain. It was a score-saturation result: constrained portfolios stayed within 0.02 percentage points of forecast-only future-burden capture, while the spatial-equity robust list replaced 50% of forecast-selected sites, increased city equity from 0.397 to 0.967 and represented every city. The engineering-aware extension reduced selected cell records to 193 and retained 98.6% of forecast-only capture. Parameter sensitivity showed that this result was not driven by one isolated tuning setting, although fairness and complexity weights remain business choices. Temporal separation prevents descriptive rankings from being mistaken for forecasts, and multi-dimensional reporting makes the efficiency-equity-complexity trade-off explicit. The framework is suitable as a transparent candidate-screening layer until richer cost, coverage and post-upgrade outcome data support a full network optimisation model and a causal evaluation of realised upgrade benefits.

Declarations

Competing interests:

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval:

Not applicable; the analyses used administrative, aggregate or site-level operational data and did not involve human participants or animals.

Consent to participate:

Not applicable.

Consent for publication:

Not applicable.

Use of generative AI tools

Generative AI tools assisted code scaffolding, manuscript drafting and language revision under author supervision. The authors verified the analyses, claims and final submitted text.

Funding:

This work was supported by the Scientific Research Fund Project of Yunnan Provincial Department of Education.

Author Contribution

Author contributionsRunzhe Liu: conceptualisation, methodology, software, data curation, formal analysis, validation, visualisation, writing - original draft, and writing - review and editing. Hui Chen: resources, data verification, investigation, domain interpretation, and writing - review and editing. Yun He: supervision, project administration, funding acquisition, methodology oversight, writing - review and editing, and correspondence.Dabin Wang: resources, engineering interpretation, validation, and writing - review and editing.All authors read and approved the final manuscript.

Data Availability

Deidentified derived tables, figure source data and the Python reconstruction script supporting this study are provided as Online Resource 1. Raw site-level operator data are commercially restricted and are not publicly redistributed. Access to the raw operator data is subject to approval by the data-owning operator.

Code availability:

The Python reconstruction script used to reproduce the reported tables and figures is available in Online Resource 1.

References

1. Andrews JG, Buzzi S, Choi W, et al. What will 5G be? IEEE Journal on Selected Areas in Communications. 2014;32:1065–1082. https://doi.org/10.1109/JSAC.2014.2328098
2. Oughton EJ, Frias Z. The cost, coverage and rollout implications of 5G infrastructure in Britain. Telecommunications Policy. 2018;42:636–652. https://doi.org/10.1016/j.telpol.2017.07.009
3. Oughton EJ, Lehr W. Surveying 5G techno-economic research to inform the evaluation of 6G wireless technologies. IEEE Access. 2022;10:25237–25257. https://doi.org/10.1109/ACCESS.2022.3153046
4. Tezergil B, Onur E. Wireless backhaul in 5G and beyond: issues, challenges and opportunities. IEEE Communications Surveys & Tutorials. 2022;24:2579–2632. https://doi.org/10.1109/COMST.2022.3203578
5. Agarwal B, Togou MA, Ruffini M, Muntean GM. A comprehensive survey on radio resource management in 5G HetNets. IEEE Communications Surveys & Tutorials. 2022;24:2495–2534. https://doi.org/10.1109/COMST.2022.3207967
6. Wang Q, Zhao X, Lv Z, Ma X, Zhang R, Lin Y. Optimizing the ultra-dense 5G base stations in urban outdoor areas: coupling GIS and heuristic optimization. Sustainable Cities and Society. 2020;63:102445. https://doi.org/10.1016/j.scs.2020.102445
7. Liang M, Li W, Ji J, et al. Evaluating the comprehensive performance of 5G base station: a hybrid MCDM model based on Bayesian best-worst method and DQ-GRA technique. Mathematical Problems in Engineering. 2022;2022:4038369. https://doi.org/10.1155/2022/4038369
8. Hwang CL, Yoon K. Multiple Attribute Decision Making. Springer; 1981. https://doi.org/10.1007/978-3-642-48318-9
9. Shannon CE. A mathematical theory of communication. Bell System Technical Journal. 1948;27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

10. Miettinen K. Nonlinear Multiobjective Optimization. Springer; 1998. https://doi.org/10.1007/978-1-4615-5563-6
11. Bertsimas D, Sim M. The price of robustness. Operations Research. 2004;52:35–53. https://doi.org/10.1287/opre.1030.0065
12. Nemhauser GL, Wolsey LA, Fisher ML. An analysis of approximations for maximizing submodular set functions I. Mathematical Programming. 1978;14:265–294. https://doi.org/10.1007/BF01588971
13. Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of Statistics. 2001;29:1189–1232. https://doi.org/10.1214/aos/1013203451
14. Breiman L. Random forests. Machine Learning. 2001;45:5–32. https://doi.org/10.1023/A:1010933404324
15. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation. 2002;6:182–197. https://doi.org/10.1109/4235.996017
16. Sigwele T, Hu YF, Susanto M. Energy-efficient 5G cloud RAN with virtual BBU server consolidation and base station sleeping. Computer Networks. 2020;177:107302. https://doi.org/10.1016/j.comnet.2020.107302

Figures

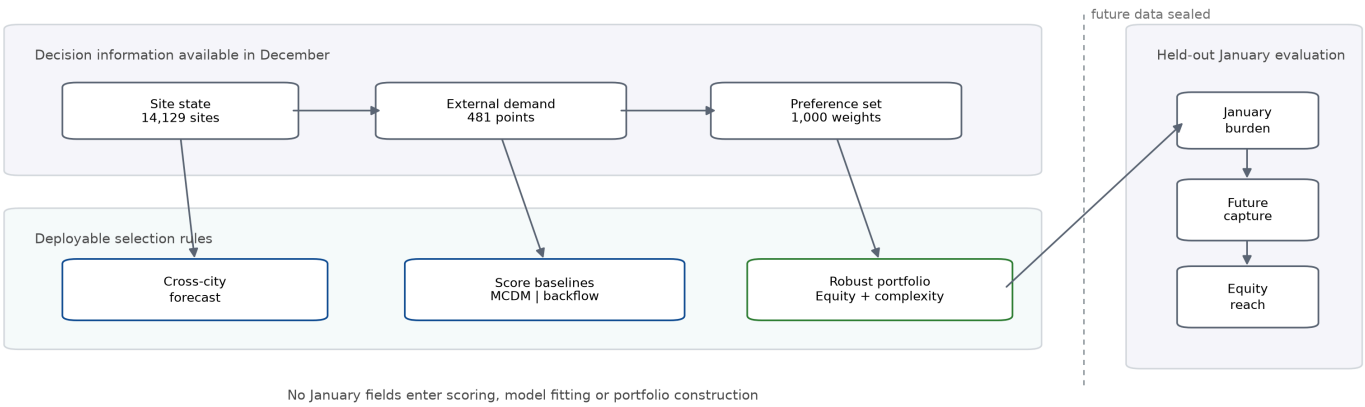


Figure 1

Temporally separated decision and evaluation framework. December observations and external demand features define deployable strategies; January outcomes are reserved for held-out evaluation.

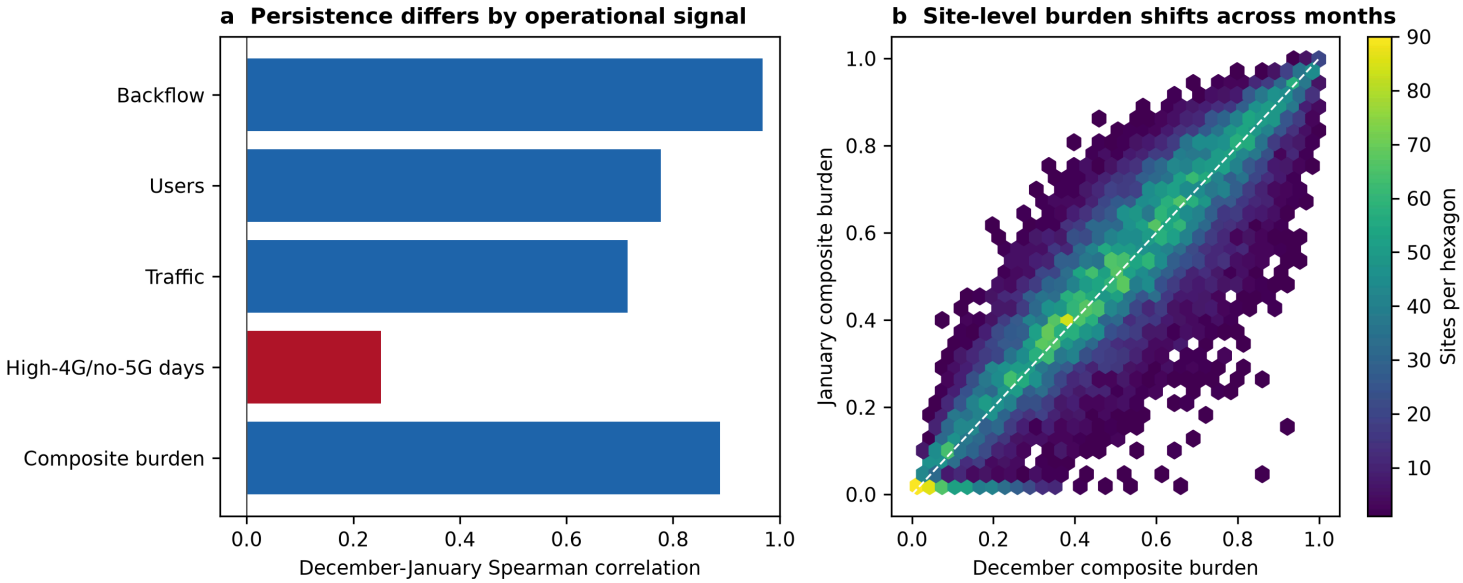


Figure 2

Temporal stability of operational signals. a, Spearman correlations between December and January site metrics. b, Joint density of December and January composite burden across 14,129 physical sites.

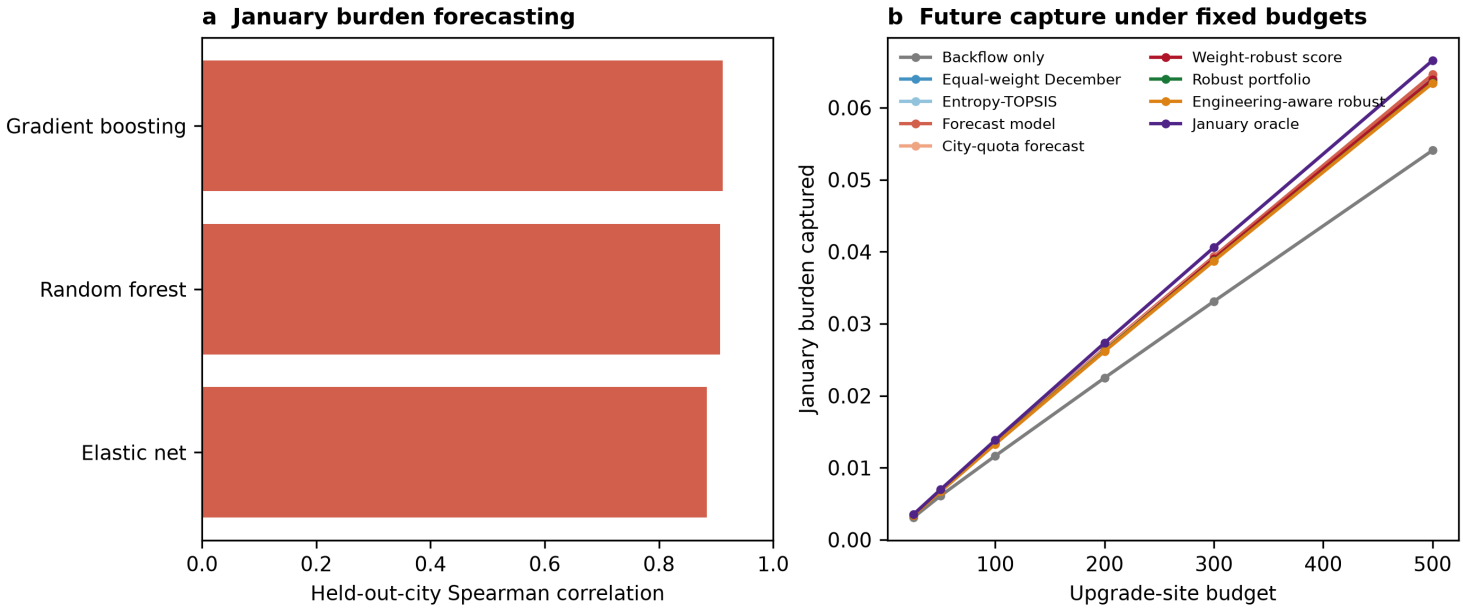


Figure 3

Future-burden prediction and cumulative portfolio capture. a, Cross-city held-out Spearman correlation for three forecasting models. b, January composite burden captured as the site budget increases. The January oracle uses future information and is shown only as an upper benchmark.

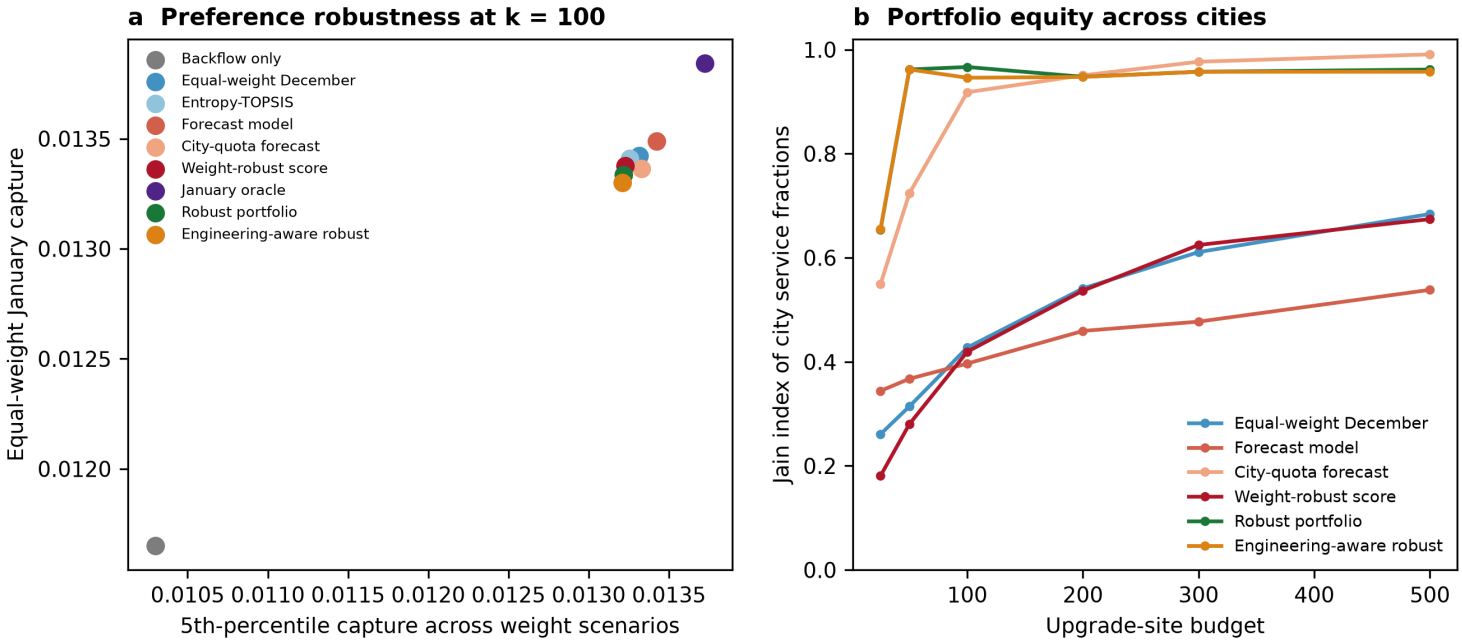


Figure 4

Preference robustness and city equity. a, Fifth-percentile preference-scenario capture versus equal-weight January capture at k = 100. b, Jain index of city service fractions across site budgets.

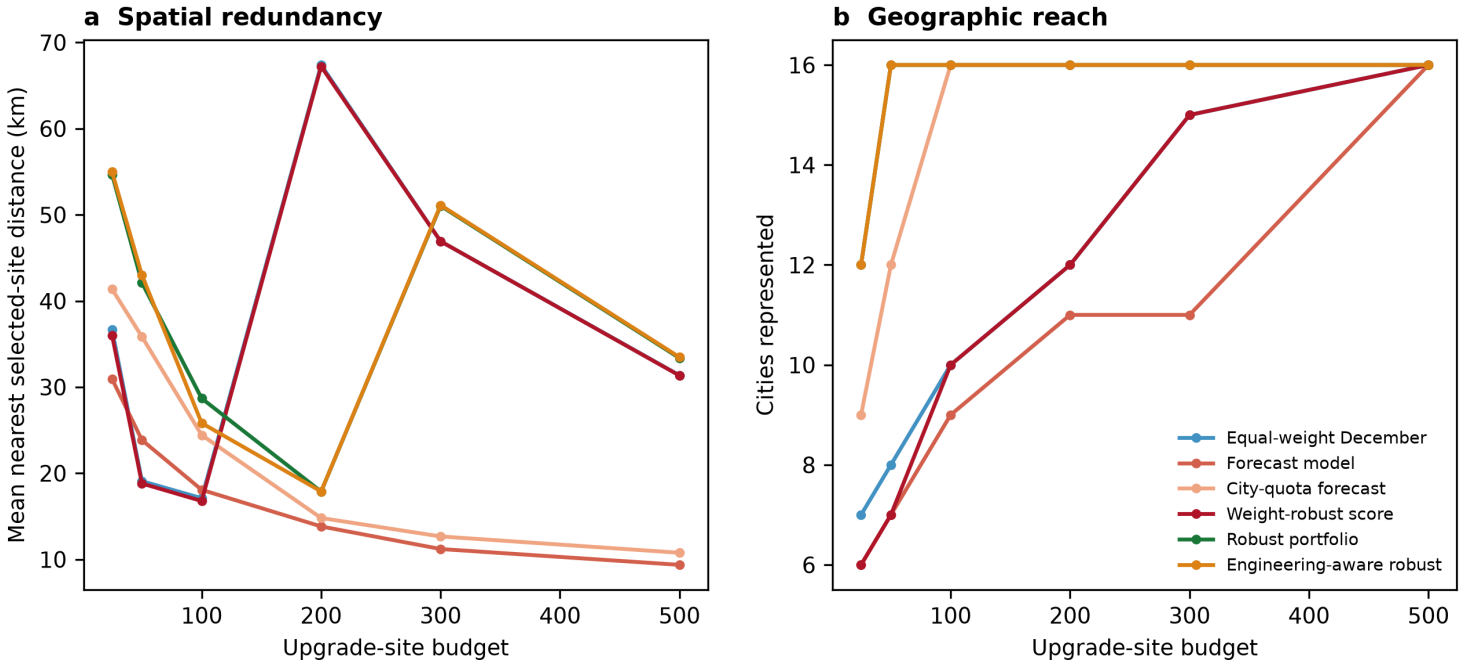


Figure 5

Spatial redundancy and geographic reach. a, Mean nearest-neighbour distance among selected sites; larger values indicate less local clustering. b, Number of cities represented in each portfolio.

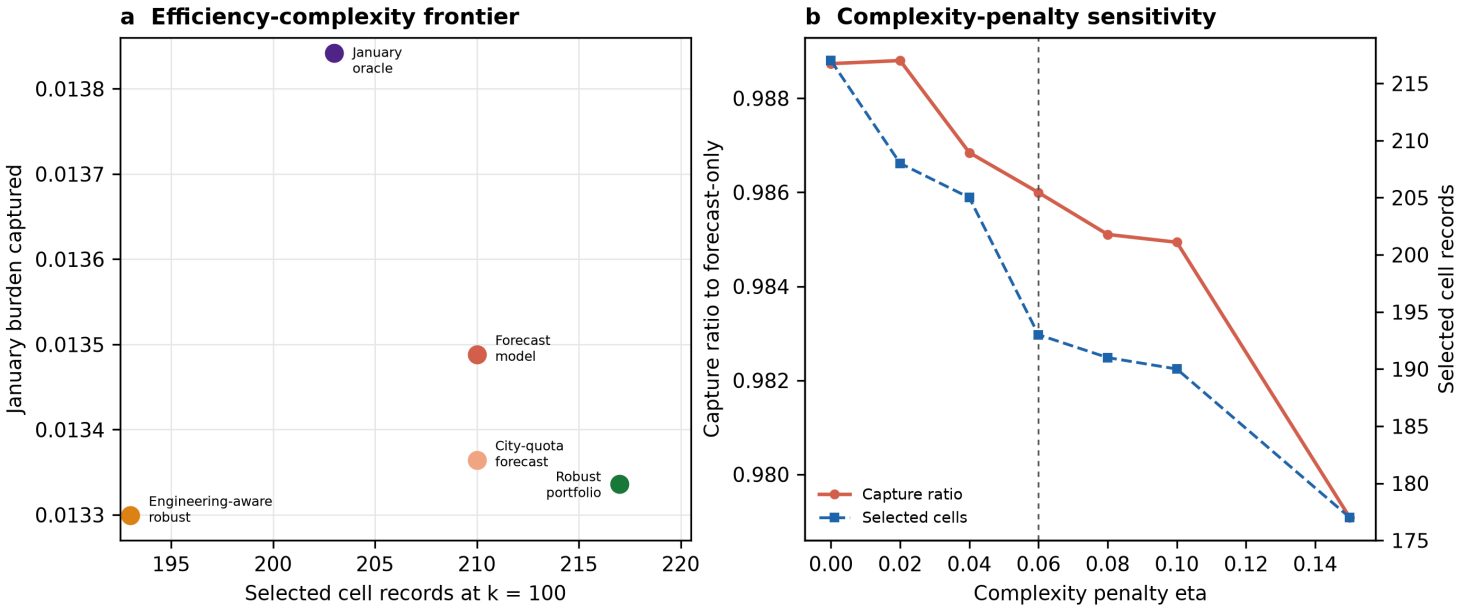


Figure 6

Engineering-complexity frontier. a, January burden capture at k = 100 against the selected cell-record proxy for the main deployable strategies. b, Sensitivity of the engineering-aware robust portfolio to the cell-count penalty eta; the default eta = 0.06 is marked by the vertical reference line.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- OnlineResource1.zip