Methodology: Probabilistic Optimization Framework for Red-Light Camera Site Selection in British Columbia

This section outlines a rigorous, implementation-ready methodology for optimally allocating Intersection Safety Cameras (ISC), commonly referred to as Red-Light Cameras (RLC), across British Columbia (BC). The framework maximizes the **Net Societal Benefit (NSB)** through probabilistic optimization that explicitly accounts for statistical uncertainty in crash predictions and treatment effects, while incorporating spatial deterrence benefits (Halo and Spillover Effects) from both existing and candidate camera locations.

1. Goal and Analytical Prerequisites

The primary objective is to maximize the aggregate expected NSB generated by the RLC program by selecting a portfolio of new sites $i \in I_{new}$ to supplement existing installations $m \in M_{existing}$, subject to implementation constraints. The methodology employs a **Full Bayesian (FB) hierarchical spatial model** to accurately forecast pre-treatment crash frequency with quantified uncertainty, thereby establishing the **Potential for Improvement (PFI)** at each candidate site with confidence intervals.

1.1 Foundational Data Requirements

All analyses require comprehensive historical crash data categorized by intersection, crash type, and severity, along with site-specific covariates:

Crash Data Structure: - **Crash Types** (t): Angle (RA), Rear-End (RE), Other (Oth) - **Severity Levels** (s): Fatal (K), Injury (I), Property Damage Only (O) - **Format**: Historical annual crash frequency $\text{History}_{i,t,s}$ for each site i

Detailed Crash Type Categories (Raw Data): The following collision types from crash records are aggregated into the three analytical categories:

Analytical Category	Includes Raw Categories
Angle (RA)	SIDE IMPACT, CONFLICTED
Rear-End (RE)	REAR END, REAR TO REAR
Other (Oth)	HEAD ON, SIDE SWIPE - OPPOSITE DIRECTION, SIDE SWIPE - SAME DIRECTION, OVERTAKING, MULTIPLE IMPACTS
Excluded	SINGLE VEHICLE, UNDETERMINED

Site Covariates: - Annual Average Daily Traffic (AADT) for major and minor approaches - Number of lanes (major and minor approaches) - Road classification (major and minor approaches) - Municipality jurisdiction - Geographic coordinates (for spatial structure)

Spatial Structure: - Network topology with all signalized intersections - Direct (Euclidean) distance matrix $d_{i,j}$ between all intersection pairs - Existing RLC locations $M_{existing}$ with installation dates

1.2 Full Bayesian Hierarchical Spatial Model

The crash prediction model is specified as a multivariate Poisson-lognormal regression with Conditional Autoregressive (CAR) spatial structure:

Model Specification:

$$\text{Crash}_{i,t,s} \sim \text{Poisson}(\lambda_{i,t,s})$$

$$\log(\lambda_{i,t,s}) = \alpha_{t,s} + \beta_{t,s}^T X_i + u_i + \epsilon_i$$

Where: - X_i = vector of site covariates (AADT, lanes, road class, municipality) - u_i = spatially structured random effect (CAR prior) - ϵ_i = unstructured random effect (independent) - $\alpha_{t,s}, \beta_{t,s}$ = crash type and severity-specific intercepts and coefficients

Prior Distributions: - Non-informative priors on all regression coefficients - CAR spatial structure for u_i based on direct distance adjacency - Gamma priors on precision parameters

Estimation: Full Bayesian inference via MCMC (Markov Chain Monte Carlo) produces posterior distributions for all parameters, enabling uncertainty quantification in crash predictions.

Model Outputs: The FB model generates posterior distributions (mean and 95% credible intervals) for aggregate crash counts:

1. **Total Predicted Count ($\Lambda_{i,total}$):** Total expected crashes at site i (all types and severities combined), including site-specific spatial random effect u_i
2. **Total Expected Count ($\mu_{i,total}$):** Total expected crashes at site i based on systemic covariates only, excluding site-specific random effect (set $u_i = 0$)

Disaggregation by Crash Type and Severity: To obtain type- and severity-specific predictions, the aggregate predicted count is disaggregated using observed crash proportions at each site:

$$\Lambda_{i,t,s} = \Lambda_{i,total} \cdot \frac{\text{Observed}_{i,t,s}}{\sum_{t,s} \text{Observed}_{i,t,s}}$$

$$\mu_{i,t,s} = \mu_{i,total} \cdot \frac{\text{Observed}_{i,t,s}}{\sum_{t,s} \text{Observed}_{i,t,s}}$$

Where $\text{Observed}_{i,t,s}$ is the historical crash count at site i for crash type t and severity s.

Rationale: This approach preserves the site-specific crash type and severity distribution observed in historical data while applying the FB model's RTM-corrected total crash prediction and uncertainty quantification.

1.3 Potential for Improvement (PFI) Quantification

The PFI identifies and ranks locations where safety interventions are most likely to yield substantial crash reductions. PFI is calculated from the FB model's posterior distributions using aggregate totals:

PFI (Absolute Difference):

$$\text{PFI}_{\text{Diff},i} = \Lambda_{i,total} - \mu_{i,total}$$

PFI (Relative Risk Ratio):

$$\text{PFI}_{\text{Ratio},i} = \frac{\Lambda_{i,total}}{\mu_{i,total}}$$

Where: - $\Lambda_{i,total} = \sum_{t,s} \Lambda_{i,t,s}$ = Total predicted crashes (all types/severities) - $\mu_{i,total} = \sum_{t,s} \mu_{i,t,s}$ = Total expected crashes based on covariates only

Both metrics are available as: - **Mean values** (point estimates) - **95% Credible Intervals** (2.5th and 97.5th percentiles)

The disaggregated Predicted Counts $\Lambda_{i,t,s}$ serve as the RTM-corrected baseline for all benefit calculations.

2. Optimization Framework and Objective Function

2.1 Decision Variables

- I_{new}: Set of all candidate intersections for new RLC deployment (approximately 200 sites)
- $M_{existing}$: Set of existing RLC installations (approximately 140 cameras)
- X_i: Binary decision variable. $X_i = 1$ if an RLC is installed at new site $i \in I_{new}$, 0 otherwise
- J: Complete set of all signalized intersections in the jurisdiction (approximately 1,500 sites)

2.2 Objective Function: Maximize Expected Net Societal Benefit (NSB)

The objective function maximizes the expected NSB across the portfolio, accounting for uncertainty in crash predictions and CMF effectiveness:

$$\text{Maximize} \quad E[\text{NSB}] = \sum_{i \in I_{new}} X_i \cdot E[B_{i,\text{Total}} - C_{i,\text{Total}}]$$

Where expectations are taken over the posterior distributions of crash predictions and CMF distributions.

2.3 Policy and Budget Constraints

1. **Budget Constraint:**

$$\sum_{i \in I_{new}} X_i \cdot K \leq \text{Budget}_{max}$$

Where $K = \$150{,}000$ is the constant capital cost per site. This constraint can be equivalently expressed as:

$$\sum_{i \in I_{new}} X_i \leq \left\lfloor \frac{\text{Budget}_{max}}{150{,}000} \right\rfloor$$

2. **Maximum Sites Constraint:**

$$\sum_{i \in I_{new}} X_i \leq R_{max}$$

3. **Minimum Performance Threshold:**

$$E[\text{PFI}_{\text{Ratio},i}] \geq 1.0 \quad \text{AND} \quad E[\Lambda_{i,total}] \geq 4 \text{ crashes/year} \quad \forall i \in I_{new}$$

Note: If the budget constraint is more restrictive than the sites constraint (i.e., $\text{Budget}_{max}/K < R_{max}$), then the budget constraint is binding; otherwise, the sites constraint is binding.

3. Crash Modification Factors and Cost Parameters

3.1 CMF Distributions by Crash Type and Severity

CMFs are modeled as random variables following Beta distributions with specified means and 95% credible intervals for probabilistic optimization:

Angle (Side-Impact) Crashes:

Severity	Mean CMF	95% CI	Beta(α, β)	Crash Reduction
Fatal (K)	0.70	[0.60, 0.85]	Beta(33.6, 14.4)	30%
Injury (I)	0.75	[0.65, 0.85]	Beta(37.5, 12.5)	25%
PDO (O)	0.82	[0.70, 0.95]	Beta(29.5, 6.5)	18%

Rear-End Crashes:

Severity	Mean CMF	95% CI	Beta(α, β)	Change
Fatal (K)	1.02	[0.95, 1.10]	Beta(399.0, 391.2)	+2%
Injury (I)	1.05	[1.00, 1.12]	Beta(175.0, 166.3)	+5%

Severity	Mean CMF	95% CI	Beta(α, β)	Change
PDO (O)	1.11	[1.05, 1.20]	Beta(89.1, 80.2)	+11%

Other Crash Types:

For "Other" crash types (including HEAD ON, SIDE SWIPE, OVERTAKING, MULTIPLE IMPACTS), the CMF is assumed to be 1.0 (no effect) based on literature indicating RLCs do not significantly affect these collision types. Therefore, the "Other" category contributes zero to the net safety benefit calculation and is omitted from the benefit formula.

Beta Distribution Parameterization: Beta distribution parameters (α, β) are derived from the specified mean μ and 95% credible interval [L, U] using quantile matching: - Mean constraint: $\mu = \alpha/(\alpha + \beta)$ - Quantile constraints: $P(X \leq L) = 0.025$ and $P(X \leq U) = 0.975$

Parameters are solved numerically to satisfy all three constraints simultaneously using nonlinear equation solvers (e.g., `scipy.optimize` in Python or `optim` in R).

3.2 Crash Cost Valuations (2025 C$)

Societal crash costs:

Severity	Mean Cost	95% Confidence Range
Fatal (K)	$315,000	[$200,000, $500,000]
Injury (I)	$65,000	[$50,000, $80,000]
PDO (O)	$7,050	[$5,000, $10,000]

3.3 Implementation Cost Parameters

Parameter	Value	Note
Capital Cost per Site (K)	$120,000	Installation Cost
	$37,000	Annual O&M Cost
Capital Recovery Factor (CRF)	**0.136**	$\frac{0.06(1.06)^{10}}{(1.06)^{10} - 1} = 0.1359$

4. Benefit and Cost Calculations

4.1 Annual Safety Benefit Calculation

The safety benefit is the net monetized crash reduction, accounting for angle crash reductions and rear-end crash increases. The calculation is performed within each Monte Carlo iteration using sampled crash predictions and CMFs, then averaged.

Monte Carlo Implementation: For each simulation iteration $n = 1, \ldots, N_{sim}$:

1. Sample crash predictions: $\Lambda_{i,t,s}^{(n)}$ from posterior distributions

2. Sample CMFs: $\text{CMF}_{t,s}^{(n)}$ from Beta distributions
3. Calculate safety benefit for iteration n:

$$B_{i,\text{Safety}}^{(n)} = \underbrace{\sum_{s\in\{K,I,O\}} \Lambda_{i,\text{RA},s}^{(n)} \cdot C_s \cdot (1 - \text{CMF}_{\text{RA},s}^{(n)})}_{\text{Angle crash reduction benefit (positive)}}$$

$$- \underbrace{\sum_{s\in\{K,I,O\}} \Lambda_{i,\text{RE},s}^{(n)} \cdot C_s \cdot (\text{CMF}_{\text{RE},s}^{(n)} - 1)}_{\text{Rear-end crash increase cost (negative)}}$$

4. Average across iterations:

$$E[B_{i,\text{Safety}}] = \frac{1}{N_{sim}} \sum_{n=1}^{N_{sim}} B_{i,\text{Safety}}^{(n)}$$

Note: The "Other" crash type term is omitted as $\text{CMF}_{\text{Oth}} = 1.0$, resulting in zero contribution: $(1 - 1.0) = 0$.

Key Clarification: - When CMF < 1 (crash reduction), the term $(1 - \text{CMF})$ is positive, contributing to benefits. - When CMF > 1 (crash increase), the term $(\text{CMF} - 1)$ is positive, but is subtracted (negative contribution to net benefit).

4.2 Total Annual Benefit

$$E[B_{i,\text{Total}}] = E[B_{i,\text{Safety}}]$$

4.4 Total Annualized Cost

$$C_{i,\text{Total}} = (K \cdot \text{CRF}) + O = (\$120{,}000 \times 0.136) + \$37{,}000 = \$53{,}320$$

4.5 Performance Metrics with Uncertainty

For each candidate site, calculate probabilistic performance metrics from Monte Carlo simulations:

Net Societal Benefit (NSB): - Mean: $E[\text{NSB}_i] = E[B_{i,\text{Total}}] - C_{i,\text{Total}}$ - 95% CI: $[\text{NSB}_{i,0.025}, \text{NSB}_{i,0.975}]$ - Probability positive: $P(\text{NSB}_i > 0) = \frac{\#\{\text{NSB}_i^{(n)} > 0\}}{N_{sim}}$

Benefit-Cost Ratio (B/C): - Mean: $E[\text{B/C}_i] = E[B_{i,\text{Total}}]/C_{i,\text{Total}}$ - 95% CI: $[\text{B/C}_{i,0.025}, \text{B/C}_{i,0.975}]$

Return on Investment (ROI): - Mean: $E[\text{ROI}_i] = E[\text{NSB}_i]/C_{i,\text{Total}}$ - 95% CI: $[\text{ROI}_{i,0.025}, \text{ROI}_{i,0.975}]$

5. Spatial Effects: Halo and Spillover Benefits

5.1 Spatial Benefit Formulation

Spatial benefits from RLC installations (both existing and new) are modeled using an exponential distance-decay function applied to direct (Euclidean) distance:

$$\text{Spatial Influence}_{i,k} = 0.06 \cdot e^{-1.10 \cdot d_{i,k}}$$

Where: - Maximum benefit: 6% crash reduction at distance = 0 km - Decay constant: $\lambda = 1.10$ (from literature) - Distance metric: Direct (Euclidean) distance $d_{i,k}$ in kilometers

Rationale: Exponential decay is consistent with spatial deterrence literature, where the effect of enforcement visibility diminishes rapidly with distance.

5.2 Cumulative Spatial Benefit at Each Site

For any site k in the network, the total spatial benefit received from all RLC installations is:

$$\text{Total Spatial Benefit}_k = \sum_{m \in M_{existing}} \text{Spatial Influence}_{m,k} + \sum_{i \in I_{new}} X_i \cdot \text{Spatial Influence}_{i,k}$$

This formulation captures: 1. **Existing camera influence**: Benefits already accruing from installed RLCs 2. **New camera influence**: Additional benefits from candidate sites if selected

6. Optimization Scenarios

Scenario A: Direct Benefits Only (No Spatial Effects)

This conservative baseline scenario considers only the direct treatment effect at the installation site.

Objective Function:

$$\text{Maximize} \quad E[\text{NSB}_A] = \sum_{i \in I_{new}} X_i \cdot (E[B_{i,\text{Total}}] - C_{i,\text{Total}})$$

Characteristics: - No spatial spillover effects - Benefits strictly localized to treated intersection - Simplest computational approach - Provides lower-bound estimate of program benefits

---

Scenario B: Proximity-Based Halo Effect (Local Spillover)

This scenario incorporates localized safety improvements at nearby untreated intersections within the effective range.

Effective Range: $D_{Halo} = 1.0$ km (direct distance)

B.1 Halo Benefit Calculation

For each site j in the candidate set I_{new}, calculate the halo benefit received from nearby RLC installations:

Halo benefit from existing cameras:

$$B_{j,\text{Halo,existing}} = \sum_{\substack{m \in M_{existing} \\ d_{m,j} \leq 1.0}} 0.06 \cdot e^{-1.10 \cdot d_{m,j}} \cdot E[\Lambda_{j,total}] \cdot \bar{C}_j$$

Halo benefit from new cameras:

$$B_{j,\text{Halo,new}} = \sum_{\substack{i \in I_{new}, i \neq j \\ d_{i,j} \leq 1.0}} X_i \cdot 0.06 \cdot e^{-1.10 \cdot d_{i,j}} \cdot E[\Lambda_{j,total}] \cdot \bar{C}_j$$

Where: - $E[\Lambda_{j,total}]$ = Expected total crashes at site j (from FB model) - $\bar{C}_j = \sum_s \frac{E[\Lambda_{j,s}]}{E[\Lambda_{j,total}]} \cdot C_s$ = Weighted average crash cost at site j

Total halo benefit for site j:

$$B_{j,\text{Halo,total}} = B_{j,\text{Halo,existing}} + B_{j,\text{Halo,new}}$$

Critical Note on Double-Counting Prevention: The summation for new cameras explicitly excludes $i = j$ (i.e., $i \neq j$), ensuring that a site does not receive halo benefits from itself. This prevents double-counting of the direct treatment effect. If site j is selected for its own camera ($X_j = 1$), it receives: - Its own direct benefit: $E[B_{j,\text{Total}}] - C_{j,\text{Total}}$ - Halo benefits from OTHER cameras (both existing and other newly selected) - But NOT halo benefit from itself (enforced by $i \neq j$ constraint)

B.2 Objective Function (Scenario B)

$$\text{Maximize} \quad E[\text{NSB}_B] = \sum_{i \in I_{new}} X_i \cdot (E[B_{i,\text{Total}}] - C_{i,\text{Total}}) + \sum_{j \in I_{new}} B_{j,\text{Halo,total}}$$

Key Features: 1. Accounts for existing camera spatial influence on all candidate sites 2. New camera placements consider areas already receiving halo benefits vs. coverage gaps 3. Synergistic effects where multiple cameras reinforce coverage are captured 4. Mathematical structure prevents double-counting through explicit $i \neq j$ exclusion

---

Scenario C: Jurisdictional Spillover Effect (Network-Wide Coverage)

This scenario optimizes for system-wide deterrence and equitable network coverage by prioritizing cameras that fill gaps in the spatial distribution.

Effective Range: $D_{Spillover} = 2.0$ km (direct distance)

C.1 Site Spillover Index (SSI) Calculation

The SSI measures how effectively a candidate camera would improve network-wide coverage, accounting for existing camera influence.

Step 1: Calculate existing coverage at all sites

For every site k in the complete network J:

$$\text{Coverage}_{k,existing} = \sum_{m \in M_{existing}} 0.06 \cdot e^{-1.10 \cdot d_{m,k}}$$

Step 2: Calculate marginal coverage benefit from candidate site i

For each candidate site $i \in I_{new}$, calculate how much additional coverage it provides to all sites k within the spillover range:

$$\text{SSI}_i = \sum_{\substack{k \in J \\ d_{i,k} \leq 2.0}} \frac{0.06 \cdot e^{-1.10 \cdot d_{i,k}}}{1 + \text{Coverage}_{k,existing}}$$

Interpretation: - **High SSI**: Candidate camera fills a coverage gap (sites with low existing coverage in vicinity) - **Low SSI**: Area already well-covered by existing cameras - **Denominator weighting**: $(1 + \text{Coverage}_{k,existing})$ prioritizes underserved areas - **Distance cutoff**: Only sites within 2.0 km contribute to SSI (beyond this, influence < 0.5%)

Step 3: Monetize SSI

Convert SSI to dollar benefits using site-specific crash costs and predicted crashes:

$$\text{Monetized SSI}_i = \sum_{\substack{k \in J \\ d_{i,k} \leq 2.0}} \frac{0.06 \cdot e^{-1.10 \cdot d_{i,k}}}{1 + \text{Coverage}_{k,existing}} \cdot E[\Lambda_{k,total}] \cdot \bar{C}_k$$

Where: - $E[\Lambda_{k,total}]$ = Expected total crashes at site k (from FB model) - $\bar{C}_k = \sum_s \frac{E[\Lambda_{k,s}]}{E[\Lambda_{k,total}]} \cdot C_s$ = Weighted average crash cost at site k

Interpretation: This formula calculates the expected dollar benefit of improved coverage at site i by summing across all sites k within range, weighted by: 1. Distance-decayed influence from i to k 2. How underserved site k currently is (inverse of existing coverage) 3. The actual crash risk and cost at site k

Computational Note: With approximately 1,500 signalized intersections in J and a 2.0 km distance cutoff, the SSI calculation for each candidate involves summing over roughly 10-50 nearby sites (depending on urban density), making this computationally tractable.

C.2 Objective Function (Scenario C)

$$\text{Maximize} \quad E[\text{NSB}_C] = \sum_{i \in I_{new}} X_i \cdot (E[B_{i,\text{Total}}] - C_{i,\text{Total}}) + \omega_S \sum_{i \in I_{new}} X_i \cdot \text{Monetized SSI}_i$$

Where ω_S is a policy weight (typically 0.3 to 0.5) reflecting the relative importance of network coverage equity versus direct site benefits.

Key Features: 1. New cameras complement rather than duplicate existing coverage 2. Underserved areas receive priority through inverse coverage weighting 3. Network-wide deterrence effect is maximized 4. Site-specific crash risk ensures coverage prioritizes high-consequence areas 5. Distance cutoff ensures computational efficiency

---

7. Probabilistic Optimization Methodology

7.1 Monte Carlo Simulation Framework

To propagate uncertainty from crash predictions and CMFs through to final metrics:

Step 1: Generate Monte Carlo Samples - Draw $N_{sim} = 10{,}000$ samples from posterior distributions of $\Lambda_{i,t,s}$ (from FB model MCMC chains) - Draw $N_{sim} = 10{,}000$ samples from CMF distributions (Beta distributions) - For each sample $n = 1, \ldots, N_{sim}$: - Calculate $B_{i,\text{Safety}}^{(n)}$ using sampled crashes and CMFs (as specified in Section 4.1) - Calculate $\text{NSB}_i^{(n)} = B_{i,\text{Total}}^{(n)} - C_{i,\text{Total}}$

Step 2: Summary Statistics - Mean: $E[\text{NSB}_i] = \frac{1}{N_{sim}} \sum_{n=1}^{N_{sim}} \text{NSB}_i^{(n)}$ - 95% CI: $[\text{quantile}_{0.025}(\{\text{NSB}_i^{(n)}\}), \text{quantile}_{0.975}(\{\text{NSB}_i^{(n)}\})]$ - Probability of positive NSB: $P(\text{NSB}_i > 0) = \frac{\#\{\text{NSB}_i^{(n)} > 0\}}{N_{sim}}$

7.2 Linearization for Computational Tractability

The optimization problem with spatial effects (Scenarios B and C) contains nonlinear terms that must be linearized for efficient solution.

Scenario B Linearization

Nonlinear Term:

$$\sum_{j \in I_{new}} \sum_{\substack{i \in I_{new}, i \neq j \\ d_{i,j} \leq 1.0}} X_i \cdot H_{ij} \cdot \Phi_j$$

Where: - $H_{ij} = 0.06 \cdot e^{-1.10 \cdot d_{i,j}}$ (pre-computed halo influence) - $\Phi_j = E[\Lambda_{j,total}] \cdot \bar{C}_j$ (expected benefit potential at site j)

Linearization Strategy:

Pre-computation Phase: 1. Calculate and store H_{ij} for all pairs (i, j) where $i \neq j$ and $d_{i,j} \leq$ 1.0 km 2. This creates a sparse influence matrix with approximately $O(k \cdot n)$ non-zero entries, where k is the average number of neighbors within 1.0 km (typically 5-15) and $n = |I_{new}| \approx 200$

Linear Reformulation: The halo benefit term is already linear in X_i since Φ_j and H_{ij} are constants: Halo Benefit $= \sum_{j \in I_{new}} \left[\sum_{\substack{i \in I_{new}, i \neq j \\ d_{i,j} \leq 1.0}} H_{ij} \cdot \Phi_j \cdot X_i \right] = \sum_{i \in I_{new}} X_i \left[\sum_{\substack{j \in I_{new}, j \neq i \\ d_{i,j} \leq 1.0}} H_{ij} \cdot \Phi_j \right]$

This can be compactly written as: Halo Benefit $= \sum_{i \in I_{new}} X_i \cdot \text{HPS}_i$

Where HPS_i (Halo Potential Score) is pre-computed: $\text{HPS}_i = \sum_{\substack{j \in I_{new}, j \neq i \\ d_{i,j} \leq 1.0}} H_{ij} \cdot \Phi_j$

Scenario B Optimization Problem (Linear Form): Maximize $\sum_{i \in I_{new}} X_i \cdot (E[B_{i,\text{Total}}] - C_{i,\text{Total}} + \text{HPS}_i) + \sum_{j \in I_{new}} B_{j,\text{Halo,existing}}$

Note: The existing halo benefit term $\sum_j B_{j,\text{Halo,existing}}$ is a constant (independent of X_i) and can be dropped from optimization.

Scenario C Linearization

Scenario C is already linear since SSI is pre-calculated: Objective $= \sum_{i \in I_{new}} X_i \cdot (E[B_{i,\text{Total}}] - C_{i,\text{Total}} + \omega_S \cdot \text{Monetized SSI}_i)$

All terms in parentheses are constants computed before optimization.

7.3 Solver Strategy: Hybrid Approach

Scenario A (Direct Benefits): - **Solver**: Mixed Integer Linear Programming (MILP). - **Library**: `scipy.optimize.milp` - **Method**: Exact deterministic optimization using expected NSB coefficients.

Scenarios B (Halo) & C (Spillover): - **Solver**: GPU-Accelerated Genetic Algorithm (Heuristic). - **Library**: cupy (custom implementation). - **Rationale**: The spatial interaction terms create a large-scale Quadratic Unconstrained Binary Optimization (QUBO) problem. A Genetic Algorithm is employed to find high-quality solutions efficiently within the computational constraints, leveraging the GPU for massive parallel evaluation of the objective function.

Genetic Algorithm Settings: - Population Size: 2,000 - Generations: 1,000 - Mutation Rate: 2% - Elitism: 10%

Solution Quality Check: - Verify optimality gap < 0.1% - Check constraint satisfaction (all constraints met) - Validate non-negativity of expected NSB for selected sites

7.4 Solution Approach for Probabilistic Optimization

Recommended Approach: Expectation-Based Optimization

1. **Pre-Optimization**: Run Monte Carlo simulation (10,000 iterations) to compute expected values $E[\text{NSB}_i]$ and $E[B_{i,\text{Total}}]$ for all candidate sites
2. **Preliminary Optimization**: Solve deterministic MILP using expected values as objective coefficients. The objective is to limit the search space for the Genetic Algorithm.
3. **Optimization**: Maximizing the overall net safety benefit, including the localized effects and spatial deterrence effect using GPU-accelerated Genetic Algorithm.
4. **Post-Optimization**: For the selected portfolio, calculate confidence intervals for portfolio-level metrics using the Monte Carlo samples

Rationale: This approach is computationally efficient (single MILP solve) and produces optimal solutions for the expected value objective. Alternative approaches (e.g., chance-constrained optimization) require multiple MILP solves and provide minimal improvement for this application.

Portfolio-Level Uncertainty Quantification: After obtaining optimal solution $X_i^* = \{0,1\}$, calculate portfolio metrics:

For each Monte Carlo sample n: Portfolio $\text{NSB}^{(n)} = \sum_{i:X_i^*=1} \text{NSB}_i^{(n)}$

Summary statistics: - Mean: $E[\text{Portfolio NSB}] = \frac{1}{N_{sim}} \sum_n \text{Portfolio NSB}^{(n)}$ - 95% CI: $[\text{quantile}_{0.025}, \text{quantile}_{0.975}]$ - Probability positive: $P(\text{Portfolio NSB} > 0)$