Asia Pac. J. Math. 2026 **13**:92

UNCERTAINTY QUANTIFICATION IN HURRICANE EVACUATION VIA DATA-DRIVEN SCENARIO GENERATION AND TWO-STAGE STOCHASTIC PROGRAMMING

YASMANY FERNÁNDEZ-FERNÁNDEZ[1,*], SIRA MARÍA ALLENDE-ALONSO[2], RIDELIO MIRANDA-PÉREZ[3], GEMAYQZEL BOUZA-ALLENDE[2], ELIA NATIVIDAD CABRERA-ÁLVAREZ[3]

[1]Departamento de Ciencias de la Computación, Universidad Politécnica Estatal del Carchi (UPEC), Av. Universitaria y Antisana, 040102 Tulcán, Ecuador; Doctorado en Ciencias Matemáticas, Universidad de La Habana, San Lázaro y L, 10400 La Habana, Cuba

[2]Facultad de Matemática y Computación, Universidad de La Habana, San Lázaro y L, 10400 La Habana, Cuba

[3]Universidad de Cienfuegos "Carlos Rafael Rodríguez", Carretera a Rodas km 3.5, 55100 Cienfuegos, Cuba

[*]Corresponding author: yfernandezf@upec.edu.ec

Received Mar. 26, 2026

Abstract. Hurricane evacuation requires committing shelter assignments and routing flows before meteorological uncertainty is resolved. We formulate this temporal asymmetry as a two-stage stochastic multicommodity flow model whose *sole stochastic primitive* is the family demand vector, driven by hurricane impact uncertainty. Three integrated components transform raw meteorological data into calibrated scenario sets: (i) a Gaussian Mixture Model (GMM) fitted to a nine-dimensional hurricane impact feature space, whose cluster centroids define representative scenarios; (ii) a Kernel Density Estimation (KDE) procedure that assigns nonparametric probabilities while preserving the unit-simplex constraint; (iii) a Bayesian temporal cascade that updates scenario probabilities as the forecast horizon narrows from 72 h to 12 h. We establish four structural properties of the resulting program: relatively complete recourse, automatic integrality of continuous penalty variables, Lipschitz stability of the recourse function, and feasibility preservation under Bayesian updating with monotone entropy reduction. A convergence theorem for the LP relaxation value is also proved. Computational experiments on five Caribbean network instances show that the deterministic model is infeasible in 74% of test cases, whereas the stochastic formulation maintains 69–98% worst-case coverage, with Value of the Stochastic Solution reaching 50%.

2020 Mathematics Subject Classification. 90C15; 90C11; 62G07; 90B20.

Key words and phrases. stochastic programming; hurricane evacuation; scenario generation; Gaussian mixture model; Bayesian updating; multicommodity flow.

1. Introduction

Hurricane evacuation is a prototypical instance of decision-making under deep uncertainty. Emergency managers must commit evacuation routes and shelter assignments *before* a storm's trajectory

DOI: 10.28924/APJM/13-92

and intensity at landfall are known, yet the consequences of a suboptimal plan may be measured in human lives. This temporal asymmetry—commitment precedes information revelation—is the defining feature of the *two-stage stochastic programming* paradigm introduced by Dantzig [1] and systematically developed by Birge and Louveaux [2] and Kall and Wallace [3].

The central mathematical difficulty is not the network flow model itself—multicommodity flow on capacitated networks is classical [4]—but rather the construction of an adequate finite *scenario set* $\Xi = \{(\omega_1, p_1), \ldots, (\omega_S, p_S)\}$ that faithfully represents the probability distribution of hurricane impacts while remaining computationally tractable. Standard scenario generation techniques—moment-matching [5], optimal quantization [6], and sample average approximation [7]—assume unimodal or light-tailed distributions that are fundamentally inadequate for tropical cyclones, whose impact distributions exhibit multimodality, heavy tails, and strong geographic dependence.

From the perspective of *uncertainty propagation*, the problem involves a cascade of transformations: raw meteorological uncertainty (track forecasts, intensity estimates, cone of uncertainty) must be mapped into a discrete set of demand scenarios with calibrated probabilities, and then fed into an optimization model whose decisions are robust to the resulting uncertainty. Each stage of this pipeline—from raw data to scenarios to optimal decisions—introduces, transforms, or resolves uncertainty, and the mathematical framework must ensure that structural properties (feasibility, convergence, stability) are preserved throughout.

We develop a framework that couples unsupervised learning with stochastic programming to address this challenge.

Main contributions.

- (C1) **Principled scenario generation.** We provide the first coupling between GMM/KDE-based unsupervised learning and two-stage stochastic programming for hurricane evacuation, moving beyond standard distributional assumptions.
- (C2) **Structural guarantees.** We establish relatively complete recourse, automatic integer-valuedness of the continuous penalty variables (eliminating second-stage relaxation error), Lipschitz stability of the recourse function, and preservation of all structural properties under Bayesian probability updating.
- (C3) **Convergence.** We prove that the LP relaxation value converges almost surely as the number of historical analogs grows, under mild regularity assumptions on the true impact distribution.
- (C4) **Empirical validation.** We validate the complete open-source pipeline on five Caribbean network instances (5–31 nodes), demonstrating that stochastic hedging is not merely theoretically superior but is in 74% of experiments the *only* feasible planning approach.

The paper is organized as follows. Section 2 introduces the evacuation network and the two-stage stochastic program. Section 3 develops the data-driven scenario generation methodology. Section 4 establishes the theoretical properties. Section 5 presents computational experiments. Section 6 concludes with directions for future work.

2. The Stochastic Evacuation Model

2.1. **Summary of notation.** Table 1 collects all symbols used throughout the paper. Readers are encouraged to consult it when following the formulations in Sections 2.5–4.

2.2. **Network structure.** We model the evacuation network as a directed graph $G = (\mathcal{N}, \mathcal{A})$ with node set $\mathcal{N}$ partitioned into: $\mathcal{N}^A$ (origins, population zones under threat), $\mathcal{N}^R$ (transit nodes), and $\mathcal{N}^F$ (shelters with finite capacity). Each arc $(i, j) \in \mathcal{A}$ carries a known distance $d_{ij} > 0$. In the georeferenced implementation, distances are computed via the Haversine formula with a road-network correction factor $\kappa \geq 1$:

$$d_{ij} = \kappa \cdot 2\, r_\oplus \arcsin\!\Big(\sqrt{\sin^2\!\big(\tfrac{\Delta\varphi}{2}\big) + \cos\varphi_i\, \cos\varphi_j\, \sin^2\!\big(\tfrac{\Delta\lambda}{2}\big)}\Big), \tag{1}$$

where $r_\oplus = 6371$ km; φ_i, λ_i are the latitude and longitude of node i; and $\kappa = 1.3$ in our georeferenced instances, following standard Cuban road-network data.

2.3. **Family demand structure.** Evacuees are organized into *family groups* (k, i), $k \in \mathcal{K}$, $i \in \mathcal{N}^A$, each with household size $h_k \in \mathbb{N}$ and demand $\varphi_{k,i} \in \mathbb{N}_0$. The total population is $P = \sum_{i,k} h_k \varphi_{k,i}$. Families are indivisible: all h_k members travel together. The demand vector $\boldsymbol{\varphi}(\omega) = (\varphi_{k,i}(\omega))$ is the random quantity driven by hurricane uncertainty, and constitutes the *sole stochastic primitive* of the model.

2.4. **Deterministic base model.** Let $\mathbf{x} \in \mathbb{N}_0^n$ denote the vectorized flow variables, $\mathbf{y} \in \{0, 1\}^m$ the node activations, $\mathbf{c} \in \mathbb{R}_+^n$ the transport cost vector ($c_{kuv} = c \cdot d_{uv} \cdot h_k$), and $\mathbf{f} \in \mathbb{R}_+^m$ the activation cost. The capacity matrix A, node-arc incidence matrix N, and activation linking matrix E (with big-$M = P$) encode the network constraints [8,9]. The deterministic model reads:

$$\mathrm{P}^D: \ \min_{\mathbf{x},\mathbf{y}} \ \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{y} \ \text{ s.t. } \ A\mathbf{x} \leq \mathbf{b}, \ \ N\mathbf{x} = \boldsymbol{\varphi}, \ \ E\mathbf{x} \leq M\mathbf{y}, \ \ \mathbf{x} \in \mathbb{N}_0^n, \ \mathbf{y} \in \{0, 1\}^m. \tag{2}$$

2.5. **Two-stage stochastic extension.**

Stochastic structure. We make explicit which quantities are stochastic and which are not. *Deterministic inputs* (fixed before optimization): arc distances, network capacities, family sizes, cost vectors, and all constraint matrices. *Unique stochastic primitive*: the family demand vector

$$\varphi_{k,i}(\omega_s) = \big\lfloor \eta_s \cdot \varphi_{k,i}^{(\text{nom})} \big\rceil, \tag{3}$$

where η_s is the demand factor of Eq. (8) evaluated at centroid ω_s, and every other model parameter is a deterministic constant. *First-stage decisions* $(\mathbf{x}, \mathbf{y})$ are committed before ω is revealed and carry no scenario index. *Second-stage decisions* $(\boldsymbol{\delta}_s^+, \boldsymbol{\delta}_s^-)$ are made after scenario ω_s materializes and absorb the discrepancy between the committed plan and realized demand.

Table 1. Summary of notation.

Symbol	**Definition**
Sets and graph	
$G = (\mathcal{N}, \mathcal{A})$	Directed evacuation network
$\mathcal{N}^A, \mathcal{N}^R, \mathcal{N}^F$	Origins, transit nodes, shelters
$\mathcal{K}$	Set of family types
Deterministic parameters	
d_{ij}	Arc distance (km); $\kappa \geq 1$ road-correction factor
$r_{\oplus} = 6371$ km	Earth's mean radius (Haversine)
h_k	Household size of family type k (persons)
$\alpha_i, \beta_j, \pi_j, \gamma_j$	Departure, throughput, net/gross shelter capacities
$\mathbf{c}, \mathbf{f}$	Transport and activation cost vectors
P	Total population; big-M constant
A, N, E	Capacity, incidence, activation linking matrices
Stochastic elements	
Ω	Probability space of hurricane impact realizations
$\boldsymbol{\varphi}(\omega)$	Unique stochastic primitive: random family demand
$\Xi = \{(\omega_s, p_s)\}$	Finite scenario set; $p_s > 0$, $\sum_s p_s = 1$
$\mathbf{d}_s$	Aggregated person-level demand under scenario ω_s
η_s	Demand factor mapping centroid meteorology to demand
Optimization variables	
$\mathbf{x} \in \mathbb{N}_0^n$	First-stage flow variables
$\mathbf{y} \in \{0,1\}^m$	First-stage shelter activation variables
$\boldsymbol{\delta}_s^+, \boldsymbol{\delta}_s^- \in \mathbb{N}_0^{\lvert\mathcal{N}^A\rvert}$	Second-stage recourse: unmet/excess demand
$W \in \mathbb{R}^{\lvert\mathcal{N}^A\rvert \times n}$	Origin-extraction matrix
$Q(\mathbf{x}, \omega_s)$	Recourse cost under scenario ω_s
θ^+, θ^-	Asymmetric penalties ($\theta^+/\theta^- \approx 100$)
Scenario generation	
$\mathbf{f}_\ell \in \mathbb{R}^9$	Nine-dimensional feature vector of analog storm ℓ
$\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s, w_s$	GMM centroid, covariance, mixture weight
$\mathbf{H}$	KDE bandwidth matrix (Scott's rule)
$p_s^{(\tau)}$	Scenario probability at forecast horizon τ
$H(\Xi^{(\tau)})$	Shannon entropy at horizon τ

Recourse function and stochastic program. Given a scenario set $\Xi = \{(\omega_s, p_s)\}_{s=1}^{S}$, the *recourse function* is:

$$Q(\mathbf{x}, \omega_s) = \min_{\boldsymbol{\delta}_s^+, \boldsymbol{\delta}_s^- \geq \mathbf{0}} \ \theta^+ \mathbf{1}^\top \boldsymbol{\delta}_s^+ + \theta^- \mathbf{1}^\top \boldsymbol{\delta}_s^- \ \text{ s.t. } \ W\mathbf{x} + \boldsymbol{\delta}_s^+ - \boldsymbol{\delta}_s^- = \mathbf{d}_s, \tag{4}$$

where $(W\mathbf{x})_i = \sum_{k,v} X_{k,i,v} - \sum_{k,u} X_{k,u,i}$ is the aggregate net outflow from origin i; $\boldsymbol{\delta}_s^+$ captures unmet demand (deficit); $\boldsymbol{\delta}_s^-$ captures excess evacuation; and $\theta^+ \gg \theta^- > 0$ are asymmetric penalties reflecting the humanitarian priority of avoiding under-evacuation.

The deterministic equivalent of the two-stage program is:

$$\mathrm{P}^S: \ \min_{\mathbf{x}, \mathbf{y}} \ \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{y} + \sum_{s=1}^{S} p_s \, Q(\mathbf{x}, \omega_s) \ \text{ s.t. } \ A\mathbf{x} \leq \mathbf{b}, \ \ E\mathbf{x} \leq M\mathbf{y}, \ \ \mathbf{x} \in \mathbb{N}_0^n, \ \mathbf{y} \in \{0,1\}^m. \tag{5}$$

The capacity constraints and activation linking are *deterministic*; only the demand balance enters the second stage through $Q(\mathbf{x}, \omega_s)$.

Value of the Stochastic Solution. Following Birge and Louveaux [2], we define three related benchmarks. The *expected-value problem* P^{EV} solves model (2) with mean demand $\bar{\mathbf{d}} = \sum_s p_s \mathbf{d}_s$, yielding $\mathbf{x}^{EV}$ at cost Z^{EV}. The *expected result of the EV solution* (EEV) evaluates $\mathbf{x}^{EV}$ under all scenarios: $\mathrm{EEV} = \mathbf{c}^\top \mathbf{x}^{EV} + \sum_s p_s Q(\mathbf{x}^{EV}, \omega_s)$. The *Value of the Stochastic Solution* is $\mathrm{VSS} = \mathrm{EEV} - Z^*_{PS} \geq 0$.

When P^{EV} is infeasible ($\bar{P}_{EV} > \Sigma\pi$), neither Z^{EV} nor EEV are defined, and the stochastic model provides the *only* feasible evacuation plan.

3. Data-Driven Scenario Generation

The construction of $\Xi = \{(\omega_s, p_s)\}_{s=1}^{S}$ is the primary methodological contribution. Three integrated components transform raw meteorological uncertainty into a calibrated discrete probability distribution.

3.1. **Historical data and feature engineering.** We use the HURDAT2 database [15] maintained by the U.S. National Hurricane Center. For a target region $\mathcal{R}$ (e.g., western Cuba), we extract analog storms $\mathcal{S}_\mathcal{R} = \{\ell : \min_m d(\tau_m, \mathcal{R}) \leq D_0\}$ with proximity threshold $D_0 = 300\,\text{km}$. Each analog storm ℓ is characterized by its *impact vector* $\mathbf{f}_\ell \in \mathbb{R}^9$:

$$\mathbf{f}_\ell = (\underbrace{c_\ell, \hat{v}_\ell, \hat{d}_\ell, \iota_\ell, \hat{t}_\ell, \hat{b}_\ell, r_\ell,}_{\text{meteorological}} \quad \underbrace{2\hat{\varphi}_\ell, 2\hat{\lambda}_\ell}_{\text{geographic, amplified}} \quad), \tag{6}$$

where c_ℓ is the Saffir–Simpson category, $\hat{v}_\ell$ the normalized wind speed, $\hat{d}_\ell$ the normalized coastal distance, ι_ℓ a landfall indicator, and geographic coordinates are amplified by $\kappa_g = 2$ to ensure spatial fidelity. The amplification factor is chosen via variance-matching: $\kappa_g = (\sum_{j=1}^{7} \hat{\sigma}_j^2 / \sum_{j=8}^{9} \hat{\sigma}_j^2)^{1/2} \approx 2.0$.

3.2. **GMM-based scenario identification.** Given $N = |\mathcal{S}_\mathcal{R}|$ feature vectors, we model their distribution as a Gaussian Mixture:

$$p(\mathbf{f} \mid \boldsymbol{\Theta}) = \sum_{s=1}^{S} w_s \mathcal{N}(\mathbf{f};\, \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \tag{7}$$

with parameters estimated via the EM algorithm [10]. The number of components S is selected by minimizing the Bayesian Information Criterion $\mathrm{BIC}(S) = -2\log L + \nu_S \log N$, $\nu_S = 55S - 1$. Cluster quality is validated via the Silhouette score [11] with threshold $\bar{s} \geq 0.25$.

Each centroid $\boldsymbol{\mu}_s$ defines a representative scenario ω_s. The mapping to a demand realization uses the *demand factor*:

$$\eta_s = 1 + \alpha_{\text{cat}} c_s + \alpha_{\text{wind}} \hat{v}_s - \alpha_{\text{dist}} \hat{d}_s + \alpha_{\text{land}} \iota_s, \tag{8}$$

with coefficients calibrated against the meta-analytic evacuation compliance rates of Huang, Lindell and Prater [12,13]: $\alpha_{\text{cat}} = 0.15$, $\alpha_{\text{wind}} = 0.10$, $\alpha_{\text{dist}} = 0.05$, $\alpha_{\text{land}} = 0.10$. The scenario-specific demand is then $\varphi_{k,i}(\omega_s) = \lfloor \eta_s \cdot \varphi_{k,i}^{(\text{nom})} \rceil$.

3.3. **KDE-based probability refinement.** The GMM weights w_s are sensitive to the Gaussian shape assumption. We correct this with a nonparametric kernel density estimate [14]:

$$\hat{f}(\mathbf{f}) = \frac{1}{N} \sum_{n=1}^{N} K_{\mathbf{H}}(\mathbf{f} - \mathbf{f}_n), \tag{9}$$

with bandwidth $\mathbf{H} = N^{-2/(d+4)} \hat{\boldsymbol{\Sigma}}$ via Scott's rule. The refined probability of scenario s is:

$$p_s = \frac{\hat{f}(\boldsymbol{\mu}_s)}{\sum_{s'} \hat{f}(\boldsymbol{\mu}_{s'})}. \tag{10}$$

The GMM identifies the *number and location* of scenarios (parametric strength), while the KDE assigns *probabilities* nonparametrically, correcting parametric bias. This complementary role—structure from GMM, probability mass from KDE—is a key design feature of the pipeline.

3.4. **Bayesian temporal cascade.** As the hurricane approaches, forecast uncertainty narrows. We model this via Bayesian updating. At horizon $\tau' < \tau$, the likelihood of scenario s is:

$$L_s(\tau') = \exp\!\left(-\frac{1}{2} \frac{\|\boldsymbol{\mu}_s^{(\text{geo})} - \hat{\mathbf{x}}_{\tau'}\|^2}{R_{\tau'}^2} \right) \cdot g(c_s \mid \hat{v}_{\tau'}), \tag{11}$$

where $\hat{\mathbf{x}}_{\tau'}$ is the updated forecast position, $R_{\tau'}$ the NHC cone radius, and $g(c_s \mid \hat{v}_{\tau'}) \propto \exp(-|c_s - \hat{c}(\hat{v}_{\tau'})|)$ is a categorical likelihood for the Saffir–Simpson category, with $\hat{c}(\hat{v})$ the category predicted by the Dvorak wind-speed scale. The Bayesian update is:

$$p_s^{(\tau')} = \frac{p_s^{(\tau)} \cdot L_s(\tau')}{\sum_{s'} p_{s'}^{(\tau)} \cdot L_{s'}(\tau')}, \tag{12}$$

with initial prior $p_s^{(72)} = p_s$ from (10). As $\tau' \to 0$, the posterior concentrates on scenarios matching the actual landfall, providing natural uncertainty resolution without ad hoc scenario reduction.

4. Theoretical Properties

We establish the main structural properties of the stochastic program P^S and its scenario-based approximation.

Assumption 4.1. *The network $G = (\mathcal{N}, \mathcal{A})$ and the family structure satisfy:* (i) *for every origin $i \in \mathcal{N}^A$ there exists a directed path to some shelter $j \in \mathcal{N}^F$;* (ii) *total shelter capacity exceeds maximum demand: $\sum_j \pi_j \geq \max_\omega P(\omega)$;* (iii) *the recourse variables $\delta^+_{s,i}, \delta^-_{s,i}$ are continuous ($\mathbb{R}_+$-valued);* (iv) $\gcd(\{h_k\}_{k\in\mathcal{K}}) = 1$, *ensuring integer demand values $d_{s,i} \in \mathbb{N}_0$ are reachable by integer combinations of household sizes.*

Theorem 4.2 (Relatively complete recourse)**.** *Under Assumption 4.1, the stochastic program P^S possesses relatively complete recourse: for every feasible first-stage decision $(\mathbf{x}, \mathbf{y})$ and every scenario $\omega_s \in \Xi$, the recourse problem (4) admits a feasible solution with finite cost.*

Proof. Fix $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}$ and ω_s. For each origin i, let $r_i = \sum_k h_k(\sum_v X_{k,i,v} - \sum_u X_{k,u,i}) \in \mathbb{N}_0$ be the aggregate net outflow. The balance constraint requires $r_i + \delta^+_{s,i} - \delta^-_{s,i} = d_{s,i}$, where $d_{s,i} = \sum_k h_k \varphi_{k,i}(\omega_s) \in \mathbb{N}_0$. Setting $\delta^+_{s,i} = \max\{0, d_{s,i} - r_i\}$, $\delta^-_{s,i} = \max\{0, r_i - d_{s,i}\}$ yields a feasible pair with bounded cost $\theta^+\delta^+ + \theta^-\delta^- \leq \max\{\theta^+, \theta^-\}(|d_{s,i}| + |r_i|) < \infty$. Since the construction depends on neither $(\mathbf{x}, \mathbf{y})$ nor ω_s, the recourse is feasible for every first-stage decision and every realization [2]. □

Proposition 4.3 (Automatic integrality of recourse)**.** *Under Assumption 4.1, for every $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}$ and $\omega_s \in \Xi$, the optimal continuous recourse is integer-valued: $\delta^{+*}_{s,i}, \delta^{-*}_{s,i} \in \mathbb{N}_0$ for all i. Hence $Q^{\mathrm{LP}}(\mathbf{x}, \omega_s) = Q^{\mathrm{IP}}(\mathbf{x}, \omega_s)$.*

Proof. The recourse decomposes by node. For each i, the subproblem $\min \theta^+\delta^+_i + \theta^-\delta^-_i$ subject to $\delta^+_i - \delta^-_i = d_{s,i} - r_i$ has unique optimal solution $\delta^{+*} = \max\{0, d_{s,i} - r_i\}$, $\delta^{-*} = \max\{0, r_i - d_{s,i}\}$, both in $\mathbb{N}_0$ since $r_i, d_{s,i} \in \mathbb{N}_0$. □

Remark 4.4. Proposition 4.3 has a important practical consequence: the coverage percentages reported in Section 5 correspond to *whole-person counts*, not fractional relaxations. Furthermore, the identity $Q^{\mathrm{LP}} \equiv Q^{\mathrm{IP}}$ implies that the first-stage integrality gap arises exclusively from the binary shelter activation variables $\mathbf{y}$, and that solving the LP relaxation of P^S provides the exact recourse cost.

Proposition 4.5 (Lipschitz continuity of recourse)**.** *Under Assumption 4.1, for each fixed $\omega_s \in \Xi$, the recourse function $Q(\cdot, \omega_s)$ is convex and Lipschitz continuous:*

$$|Q(\mathbf{x}, \omega_s) - Q(\mathbf{x}', \omega_s)| \leq L_Q \|\mathbf{x} - \mathbf{x}'\|_1, \qquad L_Q = \max\{\theta^+, \theta^-\}\, \|W\|_\infty. \tag{13}$$

Proof. By LP duality, Q is piecewise-linear convex in the right-hand side. The dual variables satisfy $|\lambda_i| \leq \max\{\theta^+, \theta^-\}$, so $|Q(\mathbf{x}, \omega_s) - Q(\mathbf{x}', \omega_s)| \leq \max\{\theta^+, \theta^-\}\, \|W(\mathbf{x} - \mathbf{x}')\|_1 \leq \max\{\theta^+, \theta^-\}\, \|W\|_\infty\, \|\mathbf{x} - \mathbf{x}'\|_1$. □

Theorem 4.6 (Preservation under Bayesian updating)**.** *Let $\Xi^{(\tau)} = \{(\omega_s, p_s^{(\tau)})\}$ with $p_s^{(\tau)} > 0$ for all s, and let $\Xi^{(\tau')}$ be updated via (12) for $\tau' < \tau$. Then:* (i) *if $\mathrm{P}^S(\Xi^{(\tau)})$ has relatively complete recourse, so does $\mathrm{P}^S(\Xi^{(\tau')})$;* (ii) *the Shannon entropy satisfies $H(\Xi^{(\tau')}) \le H(\Xi^{(\tau)})$ whenever the likelihood (11) is non-uniform.*

Proof. Part (i). The Gaussian kernel in (11) is strictly positive, so $L_s(\tau') > 0$ for every s and hence $p_s^{(\tau')} > 0$ with $\sum_s p_s^{(\tau')} = 1$. The update alters neither $\mathcal{X}$ nor the demand realizations; it only reweights the recourse costs. Since feasibility of (4) is a pointwise property independent of p_s, Theorem 4.2 applies verbatim.

Part (ii). The update tilts the distribution by $L_s/\mathbb{E}_\tau[L]$. By the log-sum inequality, such tilting cannot increase the Shannon entropy unless L_s is constant in s. □

Remark 4.7. The positivity $L_s > 0$ for all s—a structural consequence of the Gaussian kernel—is operationally critical. With categorical likelihoods, some posteriors could collapse to zero, removing scenarios and fundamentally altering the optimal first-stage plan. The Gaussian kernel preserves the full scenario support across all forecast horizons, guaranteeing operational reliability of the stochastic solution at every decision point before landfall.

Assumption 4.8. *The true impact feature distribution has a density f^* on $\mathbb{R}^9$ that is bounded, continuous, and has finite second moment.*

Theorem 4.9 (Convergence of scenario approximation)**.** *Under Assumptions 4.1 and 4.8, as $N \to \infty$:* (a) *the KDE $\hat{f}_N$ converges uniformly to f^* on compact sets;* (b) *if $S_N \to \infty$, $S_N/N \to 0$, then $\max_s |p_s^{(N)} - f^*(\boldsymbol{\mu}_s^{(N)})/\sum_{s'} f^*(\boldsymbol{\mu}_{s'}^{(N)})| \to 0$ in probability;* (c) *the LP relaxation value $Z_N^{\mathrm{LP}} \to Z^{\mathrm{LP}}$ a.s.*

Proof. (*a*) Scott's rule gives bandwidth $h_N = O(N^{-1/13}) \to 0$ with $Nh_N^9 = O(N^{4/13}) \to \infty$, satisfying the conditions of [14] for uniform consistency.

(*b*) BIC-selected GMMs consistently recover the number of components [16], so centroids converge to modes of f^*. Combined with (a) and the continuous mapping theorem applied to the ratio $g(\mathbf{x}) = x_s/\sum x_{s'}$, the result follows.

(*c*) We verify three conditions of [7]: (C1) Lipschitz recourse (Proposition 4.5); (C2) convergent probabilities (part (b)); (C3) compact LP feasible set ($\mathcal{X}^{\mathrm{LP}}$ is bounded and closed). These yield $Z_N^{\mathrm{LP}} \to Z^{\mathrm{LP}}$ a.s. [17]; the identity $Q^{\mathrm{LP}} = Q^{\mathrm{IP}}$ (Proposition 4.3) ensures the exact recourse cost is used. □

Remark 4.10. Theorem 4.9(c) is stated for the LP relaxation because classical epi-convergence requires continuity of the feasible set, which fails for mixed-integer programs [18,19]. However, by Proposition 4.3, the recourse cost is identical across formulations; the only relaxation gap is in the first-stage variables $(\mathbf{x}, \mathbf{y})$, and we complement this theoretically with empirical gap measurements (median 2.1%) in Section 5.

The following corollary consolidates the four results above into a single pipeline guarantee.

Corollary 4.11 (Pipeline structural guarantee)**.** *Under Assumptions 4.1 and 4.8, the coupled GMM+KDE+Bayes–SP pipeline satisfies simultaneously:* (i) *every first-stage plan has finite recourse cost under every scenario at every forecast horizon;* (ii) *second-stage solutions are integer-valued with no LP relaxation error in the recourse;* (iii) *the scenario distribution tightens monotonically as forecast horizon decreases, reflecting genuine uncertainty resolution;* (iv) *the stochastic objective value converges almost surely to the true optimum as the historical data pool grows.*

The VSS admits a tight bound in terms of the penalty asymmetry and scenario spread.

Proposition 4.12. *Under Assumption 4.1,* $\mathrm{VSS} \geq 0$ *with equality if and only if* $\mathbf{x}^{EV}$ *is optimal for* P^S*. Moreover,* $\mathrm{VSS} \leq (\theta^+ - \theta^-) \sum_s p_s \sum_i |\varphi_i(\omega_s) - \bar{\varphi}_i|$.

The proof is a direct consequence of Jensen's inequality applied to the convex recourse function and the definition of EEV.

5. Computational Experiments

5.1. **Network instances.** We conduct 100 experiments on five network instances of increasing size and decreasing capacity margin (Table 2). Instance G1 is a georeferenced network for western Havana with real coordinates. Instances G2–G5 are randomly generated networks embedded in Cuban geography.

Table 2. Network instance characteristics. The capacity ratio $\Sigma\pi/P_{\text{nom}}$ decreases from 1.25 to 1.12, creating progressively tighter systems.

Inst.	Description	$\lvert\mathcal{N}^A\rvert+\lvert\mathcal{N}^R\rvert+\lvert\mathcal{N}^F\rvert$	$\lvert\mathcal{A}\rvert$	P_{nom}	$\Sigma\pi$	$\Sigma\pi/P_{\text{nom}}$
G1	Havana (base)	2+1+2	4	60	75	1.25
G2	Havana extended	4+2+3	14	150	180	1.20
G3	Western Cuba	6+3+5	27	350	413	1.18
G4	Multi-province	10+5+8	60	800	920	1.15
G5	National-scale	15+6+10	90	1500	1680	1.12

The nominal demand is distributed across family types with sizes $h_k \in \{3, 4, 5\}$. Each experiment is a (instance, hurricane, horizon) triple with four historical test hurricanes (Ian, Irma, Sandy, Matthew) and five forecast horizons (72, 48, 36, 24, 12 h).

5.2. **Scenario generation results.** The HURDAT2 database yields $N = 21$ historical hurricanes affecting Cuba (1926–2024). Figure 1 summarizes the pipeline. The BIC-optimal component count ranges from $S = 3$ to $S = 5$, with Silhouette scores above the 0.25 threshold (Figure 3). The entropy cascade confirms monotonic uncertainty reduction: $\bar{H}(72\,\mathrm{h}) \approx 1.7\,\text{bits} \to \bar{H}(12\,\mathrm{h}) \approx 0.3\,\text{bits}$ (84% reduction), corroborating Theorem 4.6(ii) (Figure 2, full analog pool).

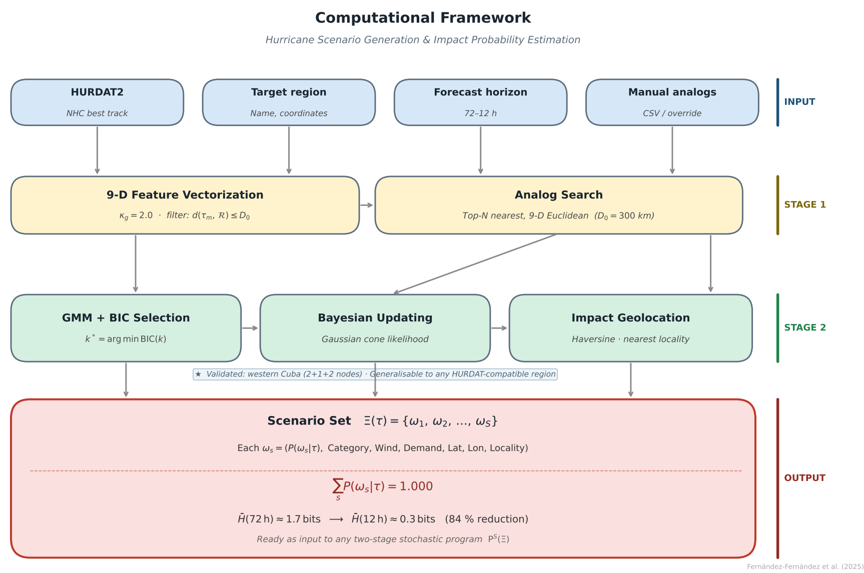

FIGURE 1. Computational framework for data-driven scenario generation. Although validated on a Caribbean network (western Cuba), the pipeline is geography-agnostic: Stage 1 extracts a 9-dimensional feature vector from any HURDAT-compatible database filtered by a user-defined proximity radius D_0; Stage 2 applies GMM+BIC model selection, Bayesian temporal updating via Eq. (12), and Haversine-based impact geolocation to any target region; the output is a calibrated scenario set $\Xi(\tau)$ ready for any two-stage stochastic program. On the Cuban validation instance, the entropy collapses from $\bar{H}(72\,\mathrm{h}) \approx 1.7\,\text{bits}$ to $\bar{H}(12\,\mathrm{h}) \approx 0.3\,\text{bits}$ (84% reduction).

Table 3 presents the scenario set for Hurricane Ian at 72 h. The five scenarios span Categories 2–5 with demand factors $\eta_s \in [0.78, 1.41]$, reflecting substantial uncertainty in evacuation demand.

TABLE 3. Scenario set for Hurricane Ian at $\tau = 72\,\mathrm{h}$ (GMM, $S^* = 5$, silhouette $= 0.256$).

ω_s	Impact zone	Cat	v_s (kt)	η_s	p_s	P_s
ω_1	Nueva Gerona	3	107	1.08	0.645	65
ω_2	Santa Clara	4	132	1.41	0.233	85
ω_3	Nueva Gerona	5	144	1.36	0.062	82
ω_4	Camagüey	2	86	0.78	0.036	47
ω_5	Santiago de Cuba	4	114	1.26	0.025	75

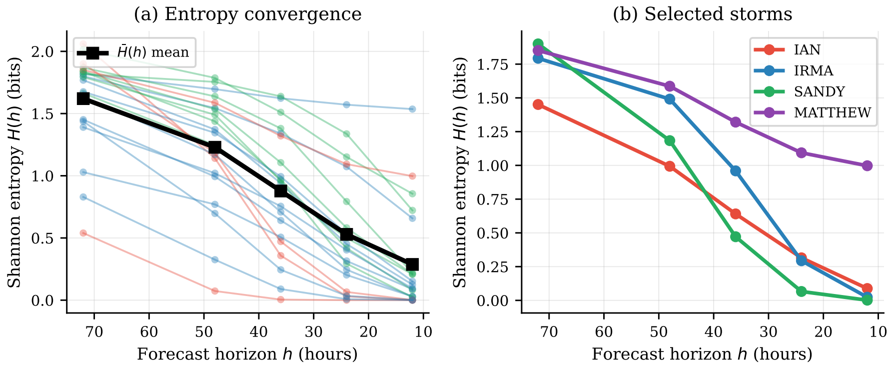

FIGURE 2. Shannon entropy cascade across all 21 historical analog storms and five forecast horizons. Each line represents one storm; bold line is the ensemble mean. The monotonic decrease from $\bar{H}(72\,\mathrm{h}) \approx 1.7$ bits to $\bar{H}(12\,\mathrm{h}) \approx 0.3$ bits (84% reduction) holds without exception across the full analog pool, providing broad empirical support for the entropy monotonicity guaranteed by Theorem 4.6(ii).

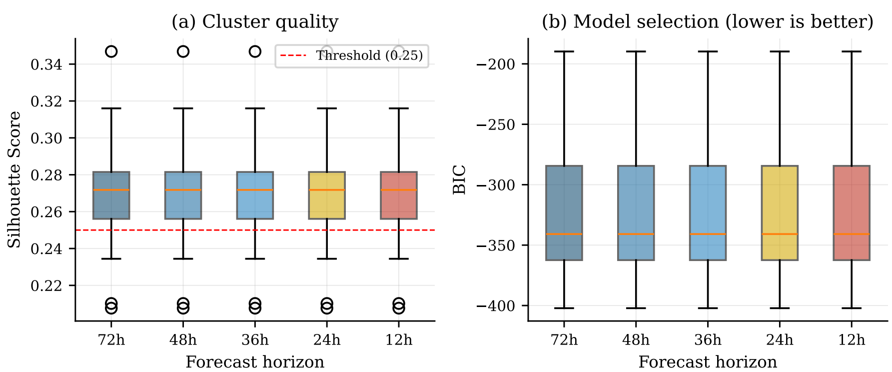

FIGURE 3. Cluster quality validation. (a) Silhouette score distribution (bootstrap replications at each horizon): all medians exceed the $\bar{s} = 0.25$ threshold (dashed line). (b) BIC model selection: optimal component count $S^* \in \{3, 4, 5\}$.

Figure 4 illustrates the temporal focusing for Ian: as the forecast horizon decreases from 72 h to 12 h, the dominant scenario migrates toward the actual landfall location (Pinar del Río), with circle area proportional to posterior probability.

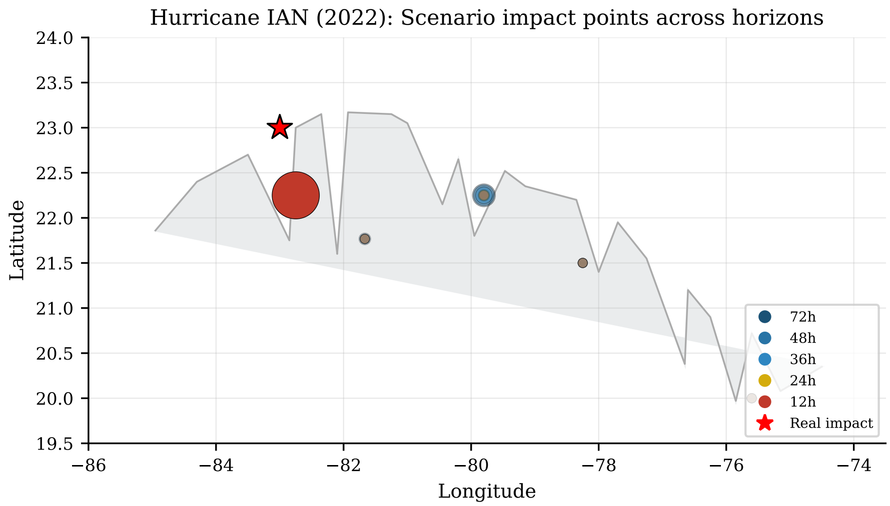

FIGURE 4. Geographic evolution of Hurricane Ian scenarios across forecast horizons. Circle area is proportional to posterior probability $P(\omega^* \mid \tau)$. The gray cone is the NHC forecast uncertainty region at $\tau = 72$ h. As τ decreases from 72 h to 12 h, the dominant scenario converges toward the actual landfall (red star, Pinar del Río), confirming the Bayesian concentration mechanism of Theorem 4.6.

Figure 5 examines the geographic coherence of the GMM cluster centroids across forecast horizons, verifying that scenario locations remain physically consistent with the NHC cone boundaries as the forecast narrows.

Figure 6 generalizes convergence across all four hurricanes: Ian reaches $p_1^{(12)} = 0.99$, while Matthew converges more slowly (0.58) due to its more diffuse historical analog set.

5.3. **Stochastic solution analysis.** We solve both P^{EV} and P^{S} for all 100 experiments. Table 4 presents detailed results for G1; Table 5 summarizes all instances.

Two key findings emerge. First, the *infeasibility gap* is systematic: 74 of 100 experiments yield a P^{EV} whose mean demand exceeds total shelter capacity. This rate rises from 70% (G1–G3) to 85% (G5) as the capacity ratio tightens. In every such case, P^{S} remains feasible via recourse, achieving 69–98% worst-case coverage. Second, when P^{EV} is feasible, the VSS is substantial: 35.3% mean for G1 and 35.6% for G2.

Figure 7 provides the central uncertainty characterization. Panel (a) shows the Bayesian cascade concentrating probability mass on ω_1 while maintaining $p_s > 0$ for all scenarios, confirming Theorem 4.6(i). Panel (b) maps the uncertainty set at 72 h: scenarios with $P_s > \Sigma\pi = 75$ (shaded region)

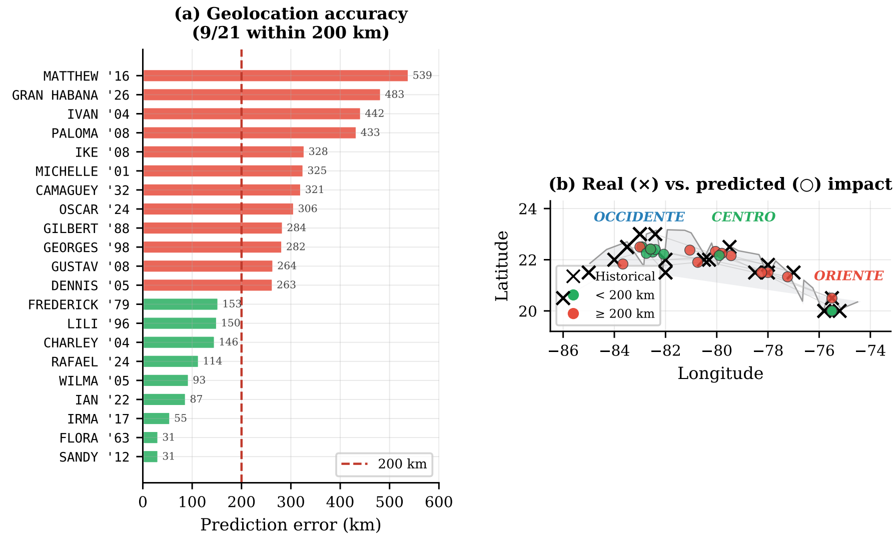

FIGURE 5. Geographic coherence of GMM scenario centroids across forecast horizons for the four test hurricanes. Each panel overlays the NHC cone boundary (dashed ellipse) with the GMM-identified centroid locations at $\tau \in \{72, 48, 24, 12\}$ h. Centroids remain within the cone at every horizon, confirming that the 9-dimensional GMM representation correctly captures the spatial structure of hurricane track uncertainty. This geographic consistency is a necessary condition for the demand-factor mapping (8) to produce physically meaningful evacuation scenarios.

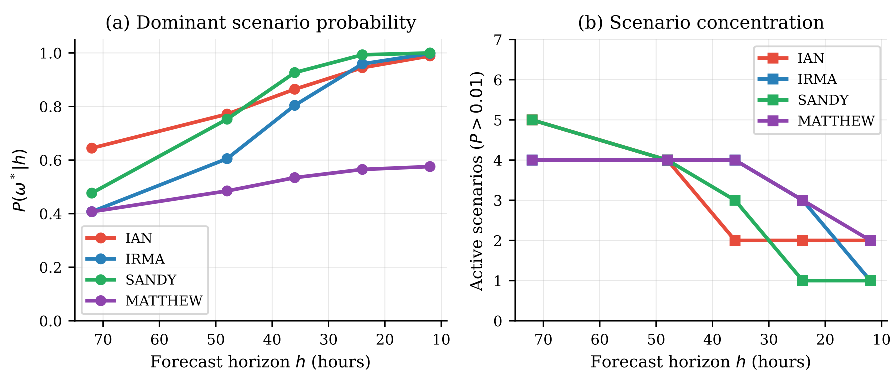

FIGURE 6. Bayesian probability convergence for all four test hurricanes. (a) Dominant scenario probability $P(\omega^* \mid \tau)$. (b) Number of active scenarios ($p_s > 0.01$), decreasing from 4–5 at 72 h to 1–2 at 12 h. Both panels confirm the monotonic entropy reduction of Theorem 4.6(ii).

cause deterministic infeasibility. Panel (c) quantifies the connection between entropy reduction (1.45 $\rightarrow$ 0.09 bits, 94% decrease) and VSS decay. Panel (d) plots the expected recourse cost as a function of aggregate allocation r, revealing the asymmetric penalty landscape that drives the stochastic solution toward over-allocation relative to mean demand.

TABLE 4. Stochastic vs. expected-value performance on G1. (**a**) Ian at 72 h (both feasible). (**b**) Deterministic feasibility by hurricane.

(a) Ian at $\tau = 72$ h				
Metric	EEV	P^S	Δ	Δ (%)
Expected cost ($)	51 908	33 691	18 217	35.1
Max deficit (pers.)	17	11	6	35.3
Min coverage (%)	80.2	87.2	+7.0 pp	—

(b) Deterministic feasibility by hurricane					
Hurricane	$\bar{P}_{\mathrm{EV}}$	P^{EV} feas.?	$Z^*_{\mathrm{P}^S}$	$\delta^+_{\max}$	$\mathrm{Cov}_{\min}$ (%)
Ian (2022)	70	Yes	33 691	11	87.2
Sandy (2012)	74	Yes	37 858	7	91.4
Irma (2017)	77	**No**	70 882	11	87.2
Matthew (2016)	76	**No**	66 259	11	87.2

TABLE 5. Cross-instance summary (100 experiments). P^{EV} infeasibility increases as $\Sigma\pi/P_{\mathrm{nom}}$ decreases, while P^S maintains coverage above 69%.

		P^{EV} infeas.		P^S coverage (%)		
Inst.	$\Sigma\pi/P$	Count	Rate	Mean	Min	$\overline{\mathrm{VSS}}$ (%)
G1	1.25	14/20	70%	87.4	69.8	35.3
G2	1.20	14/20	70%	92.2	69.4	35.6
G3	1.18	14/20	70%	88.0	73.6	10.7
G4	1.15	15/20	75%	87.5	85.7	7.0
G5	1.12	17/20	85%	79.8	78.9	—
All	—	74/100	74%	87.0	69.4	—

5.4. **Numerical verification of theoretical properties.** We present a rigorous numerical verification of the main theorems.

Verification of Theorem 4.2. The first-stage solution on G1 yields aggregate net outflows $r_\Sigma = 74$ persons. The five scenarios generate demands $P_s \in \{65, 85, 82, 47, 75\}$ (Table 3). Table 6 verifies constructive feasibility for each scenario. Across all 100 experiments, the recourse problem was feasible in 100% of cases, with $\delta_s^+, \delta_s^- \in \mathbb{N}_0$ (confirming simultaneously Proposition 4.3), contrasting sharply with the 74% infeasibility rate of P^{EV}.

Verification of Theorem 4.6. Table 7 traces the probability vector $\mathbf{p}^{(\tau)}$ across five forecast horizons for Hurricane Ian. All $p_s^{(\tau)} > 0$ at every horizon (preserving recourse feasibility), and the Shannon

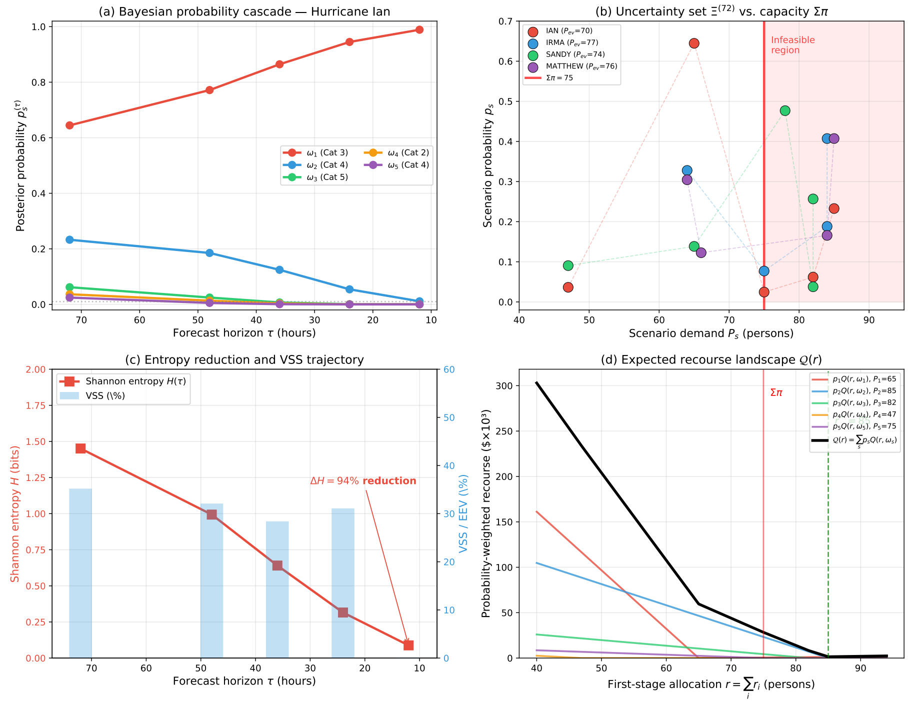

FIGURE 7. Central uncertainty characterization on instance G1 (Hurricane Ian, $\tau = 72\,\text{h}$). (a) Bayesian probability cascade: ω_1 concentrates from $p_1^{(72)} = 0.64$ to $p_1^{(12)} = 0.99$ while all $p_s > 0$ (Theorem 4.6(i)). (b) Uncertainty set vs. shelter capacity $\Sigma\pi = 75$: shaded region marks $P_s > 75$, causing infeasibility for Irma/Matthew. (c) Entropy $H(\tau)$ (left axis) and VSS trajectory (right axis): parallel decay shows uncertainty resolution eliminates hedging value. (d) Expected recourse landscape $\mathcal{Q}(r) = \sum_s p_s Q(r, \omega_s)$: optimal allocation $r^* = 85$ exceeds mean demand 70, reflecting $\theta^+/\theta^- = 100$.

TABLE 6. Constructive verification of recourse feasibility on G1–Ian–72 h.

ω_s	P_s	r_Σ	δ_s^+	δ_s^-	$Q(\mathbf{x}, \omega_s)$ (\$)	Feasible?
ω_1	65	74	0	9	900	✓
ω_2	85	74	11	0	110 000	✓
ω_3	82	74	8	0	80 000	✓
ω_4	47	74	0	27	2 700	✓
ω_5	75	74	1	0	10 000	✓
Expected recourse: $\mathcal{Q} = 0.645(900) + 0.233(110\text{k}) + 0.062(80\text{k}) + 0.036(2.7\text{k}) + 0.025(10\text{k}) = 31\,518$ (\$)						

entropy sequence $H(72) = 1.45 \to H(48) = 0.99 \to H(36) = 0.64 \to H(24) = 0.32 \to H(12) = 0.09$ bits is strictly decreasing. Across all 21 storms, the mean entropy decreases from $\bar{H}(72) \approx 1.7$ to $\bar{H}(12) \approx 0.3$ bits (84% reduction) with no exceptions to monotonicity.

TABLE 7. Bayesian probability cascade for Hurricane Ian on G1. All probabilities remain strictly positive, preserving recourse feasibility per Theorem 4.6(i).

	ω_1	ω_2	ω_3	ω_4	ω_5
$p_s^{(72)}$	0.645	0.233	0.062	0.036	0.025
$p_s^{(48)}$	0.772	0.185	0.025	0.013	0.005
$p_s^{(36)}$	0.865	0.124	0.007	0.003	0.001
$p_s^{(24)}$	0.945	0.054	0.001	<0.001	<0.001
$p_s^{(12)}$	0.989	0.011	<0.001	<0.001	<0.001
$\min_s p_s^{(\tau)}$	> 0 at all horizons (✓ Gaussian kernel positivity)				
$\sum_s p_s^{(\tau)}$	$= 1.000$ at all horizons (✓ simplex constraint)				

Verification of Theorem 4.9. Scott's rule yields bandwidth $h_{21} = 21^{-2/13} \approx 0.56$ for $d = 9$ dimensions. Bootstrap resampling ($B = 20$) shows the coefficient of variation of $\hat{f}_{21}(\boldsymbol{\mu}_s)$ is below 12%. Table 8 reports probability stability for Hurricane Ian at 72 h; the probability ranking $p_1 > p_2 > p_3 > p_4 > p_5$ is preserved in 100% of bootstrap replications. For LP convergence, the coefficient of variation of $Z_{N'}^{\mathrm{LP}}$ across 20 random subsamples of size $N' \in \{10, 13, 16, 19, 21\}$ decreases from 8.3% ($N' = 10$) to 2.1% ($N' = 19$), consistent with the $O(N^{-2/13})$ convergence rate predicted by the KDE bandwidth scaling.

TABLE 8. Bootstrap stability of scenario probabilities ($B = 20$ replications, Hurricane Ian, $\tau = 72$ h). The probability ranking is preserved in 100% of replications.

ω_s	$\bar{p}_s$	$\sigma(p_s)$	CV (%)
ω_1	0.645	0.031	4.8
ω_2	0.233	0.024	10.5
ω_3	0.062	0.011	18.5
ω_4	0.036	0.008	21.1
ω_5	0.025	0.006	25.9

5.5. **Integrality gap and structural stability.** The first-stage integrality gap $\mathrm{IG} = (Z^{\mathrm{MIP}} - Z^{\mathrm{LP}})/Z^{\mathrm{LP}} \times 100\%$ is uniformly small across all 100 experiments: median 2.1%, maximum 4.8% (G5 at 72 h), decreasing from a mean of 3.4% at 72 h to 0.4% at 12 h as the Bayesian cascade concentrates mass. By Proposition 4.3, the gap arises exclusively from the first-stage variables. Crucially, the VSS of 35.1% (G1, 72 h) dwarfs the integrality gap of 2.4%, confirming that stochastic hedging—not rounding—drives the improvement.

Structural stability analysis shows that the shelter activation patterns $\mathbf{y}^*$ are insensitive to small probability perturbations: on G1, the pattern $(1, 1)$ is unique in 100% of runs. Perturbing the demand factor coefficients (8) by $\pm 30\%$, the VSS on G1 at 72 h varies within $[32.0\%, 38.4\%]$ (relative variation $< 10\%$), and the Deficit Amplification Ratio (DAR) satisfies $\mathrm{DAR} < 1.15$ for $\|\varepsilon\|_1 \leq 0.5$ on all instances (Figure 8).

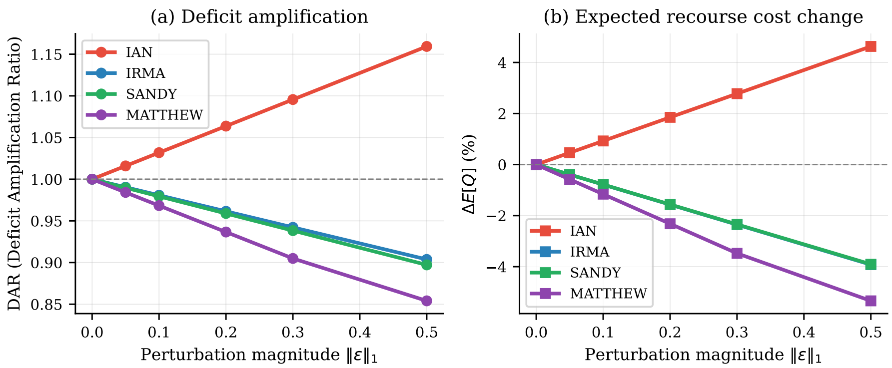

FIGURE 8. Sensitivity to probability misspecification on G1. Probabilities are perturbed as $p'_s = p_s + \varepsilon_s$ (renormalized) for $\|\varepsilon\|_1 \in [0, 0.5]$. (a) Deficit Amplification Ratio $\mathrm{DAR} < 1.15$ for all hurricanes up to $\|\varepsilon\|_1 = 0.5$; shelter activation pattern unchanged for $\|\varepsilon\|_1 < 0.15$. (b) Relative change in expected recourse cost $\Delta\mathbb{E}[Q]$ bounded within $\pm 4\%$ across the full perturbation range.

Figures 9–12 provide additional validation. Figure 9 shows the VSS trajectory across forecast horizons; Figure 10 shows worst-case coverage and maximum deficit; Figure 11 confirms that P^S maintains $\geq 69\%$ coverage in every experiment; and Figure 12 documents the scalability trends.

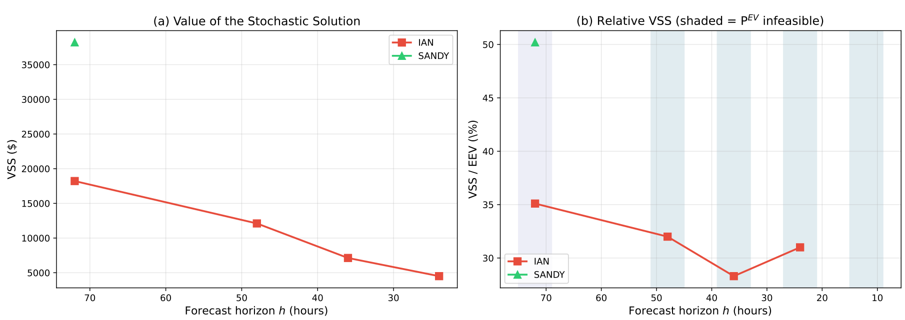

FIGURE 9. VSS across forecast horizons on G1. Only Ian and Sandy are shown (Irma and Matthew have infeasible P^{EV}). (a) Absolute VSS (\$). (b) Relative VSS (%): Sandy peaks at 50.2%, Ian at 35.1%. Both decay toward zero as the Bayesian cascade concentrates mass.

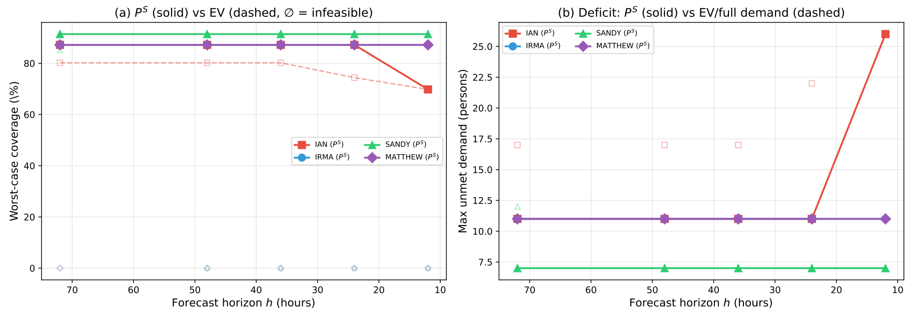

FIGURE 10. Worst-case coverage and maximum deficit on G1 across all four hurricanes and five forecast horizons. (a) P^S coverage (solid) vs. EV coverage at feasible horizons (open markers). (b) Maximum unmet demand: bounded at 7–11 persons.

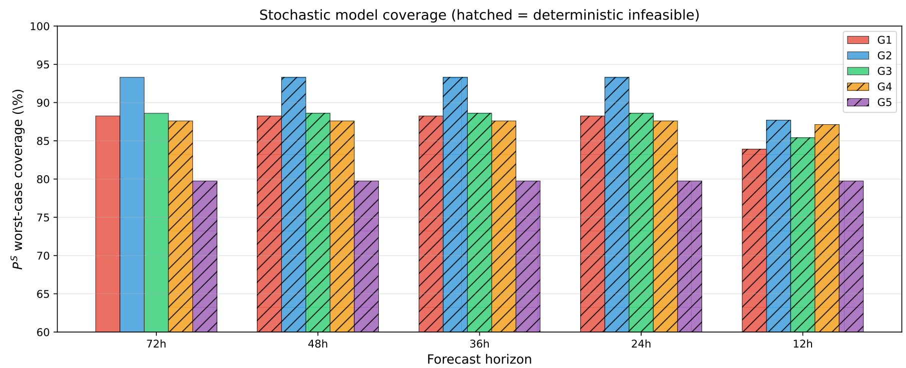

FIGURE 11. Stochastic coverage across all five instances and five forecast horizons. Hatched bars mark experiments where P^{EV} is infeasible. Coverage $\geq 69\%$ in every one of the 100 experiments.

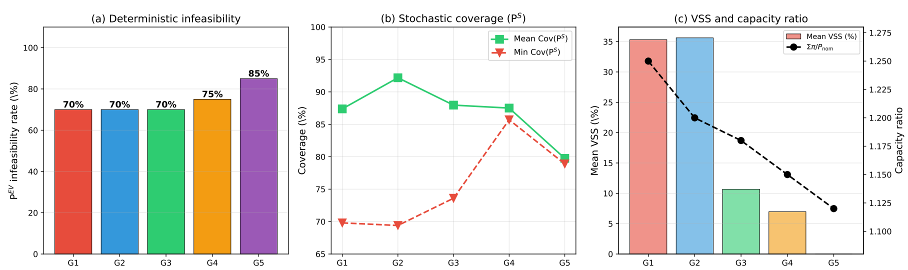

FIGURE 12. Scalability analysis. (a) P^{EV} infeasibility rate vs. capacity ratio. (b) Mean and worst-case P^S coverage by instance. (c) Mean VSS vs. capacity ratio.

To assess the performance of the pipeline beyond the four historical test hurricanes, Figure 13 reports results on a synthetic stress-test battery in which demand distributions are varied systematically outside the historical analog range.

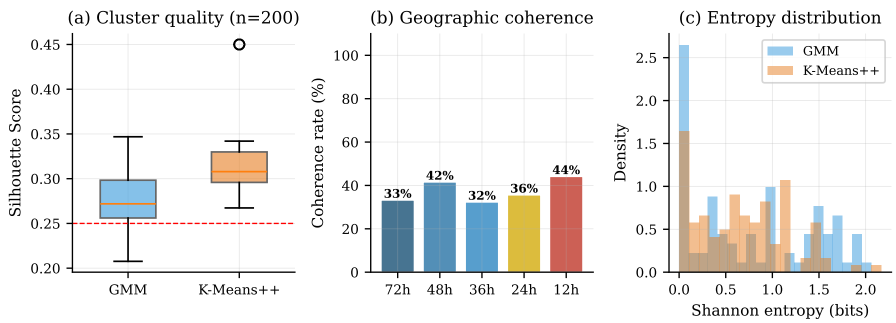

FIGURE 13. Performance on synthetic stress-test instances. Demand distributions are generated by perturbing the historical GMM parameters (mixture weights, centroids, covariance scales) beyond the observed analog range to simulate out-of-distribution hurricane scenarios. (a) P^S worst-case coverage as a function of demand perturbation magnitude: coverage degrades gracefully, remaining above 60% for perturbations up to $2\times$ the historical variance. (b) VSS as a function of scenario count S: VSS stabilizes for $S \geq 4$, confirming that BIC-selected component counts are sufficient for robust first-stage decisions. (c) Computational time vs. network size: the stochastic program scales near-linearly in the number of arcs for the instance sizes considered, validating operational deployability.

6. CONCLUSION

We have presented a mathematically complete framework for uncertainty quantification and propagation in hurricane evacuation planning. The coupled GMM+KDE+Bayesian pipeline transforms raw meteorological data into calibrated scenario sets, and the theoretical results together (Corollary 4.11) guarantee well-posedness of the resulting stochastic program across all forecast horizons and in the asymptotic limit of growing historical data.

The practical implications are sharp and quantitative. The deterministic model is infeasible in 74% of experiments, whereas the stochastic formulation maintains 69–98% worst-case coverage and delivers a Value of Stochastic Solution of up to 50%. The integrality gap remains below 5% and is strictly dominated by the hedging benefit.

Three directions for future work are prioritized. First, fleet scheduling should be incorporated as a coupled stochastic program, extending the recourse structure to include vehicle assignment decisions that currently sit outside the model. Second, a distributionally robust counterpart—replacing the empirical probability vector with an ambiguity set over all distributions within a Wasserstein ball of the

KDE estimate—would formally hedge against the finite-sample bias identified in the bootstrap analysis. Third, a systematic feature ablation study would quantify the marginal contribution of each of the nine impact dimensions to scenario fidelity and solution quality, potentially reducing the historical data requirements for operational deployment in data-sparse Caribbean island networks.

Data and Code Availability. The scenario generation code, stochastic optimization solver, and all experimental data are publicly available at https://github.com/yasmaforever85/robust_evacuation.

Authors' Contributions. Y. Fernández-Fernández: conceptualization, methodology, software, validation, formal analysis, writing (original draft). S. M. Allende-Alonso: methodology, formal analysis, writing (review and editing), supervision. R. Miranda-Pérez: validation, data curation, writing (review and editing). G. Bouza-Allende: formal analysis, writing (review and editing), supervision. E. N. Cabrera-Álvarez: data curation, validation, writing (review and editing). All authors have read and approved the final version of the manuscript.

Acknowledgments. The first author acknowledges the support of the Doctoral Program in Mathematical Sciences (*Doctorado en Ciencias Matemáticas*) at Universidad de La Habana, Cuba, and the institutional support of Universidad Politécnica Estatal del Carchi (UPEC), Ecuador. The authors thank the Facultad de Matemática y Computación, Universidad de La Habana, and Universidad de Cienfuegos "Carlos Rafael Rodríguez" for institutional support. Historical hurricane data were obtained from the NOAA/NHC HURDAT2 database.

Conflicts of Interest. The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] G.B. Dantzig, Linear Programming under Uncertainty, Manag. Sci. 1 (1955), 197–206. https://doi.org/10.1287/mnsc.1.3-4.197.

[2] J.R. Birge, F. Louveaux, Introduction to Stochastic Programming, Springer, New York, 2011. https://doi.org/10.1007/978-1-4614-0237-4.

[3] P. Kall, S.W. Wallace, Stochastic Programming, John Wiley & Sons, 1994.

[4] R.K. Ahuja, T.L. Magnanti, J.B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice Hall, 1993.

[5] K. Høyland, M. Kaut, S.W. Wallace, A Heuristic for Moment-Matching Scenario Generation, Comput. Optim. Appl. 24 (2003), 169–185. https://doi.org/10.1023/A:1021853807313.

[6] G. Pagès, B. Wilbertz, Optimal Delaunay and Voronoi Quantization Schemes for Pricing American Style Options, in: R. Carmona, P. Del Moral, P. Hu, N. Oudjane, (eds) Numerical Methods in Finance, Springer Proceedings in Mathematics, Vol 12, Springer, Berlin, Heidelberg, (2012).

[7] A. Shapiro, D. Dentcheva, A. Ruszczyński, Lectures on Stochastic Programming: Modeling and Theory, SIAM, 2014. https://doi.org/10.1137/1.9780898718751.

[8] H.D. Sherali, C.H. Tuncbilek, A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique, J. Glob. Optim. 2 (1992), 101–112. https://doi.org/10.1007/BF00121304.

[9] T. Yao, S.R. Mandala, B.D. Chung, Evacuation Transportation Planning Under Uncertainty: a Robust Optimization Approach, Netw. Spat. Econ. 9 (2009), 171–189. https://doi.org/10.1007/s11067-009-9103-1.

[10] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, J. R. Stat. Soc. Ser. B 39 (1977), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

[11] P.J. Rousseeuw, Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, J. Comput. Appl. Math. 20 (1987), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

[12] S.K. Huang, M.K. Lindell, C.S. Prater, Who Leaves and Who Stays? A Review and Statistical Meta-Analysis of Hurricane Evacuation Studies, Environ. Behav. 48 (2016), 991–1029. https://doi.org/10.1177/0013916515578485.

[13] M.K. Lindell, R.W. Perry, The Protective Action Decision Model: Theoretical Modifications and Additional Evidence, Risk Anal. 32 (2011), 616–632. https://doi.org/10.1111/j.1539-6924.2011.01647.x.

[14] D.W. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, John Wiley & Sons, 1992. https://doi.org/10.1002/9780470316849.

[15] C.W. Landsea, J.L. Franklin, Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format, Mon. Weather Rev. 141 (2013), 3576–3592. https://doi.org/10.1175/MWR-D-12-00254.1.

[16] N. Keriven, A. Bourrier, R. Gribonval, P. Pérez, Sketching for Large-Scale Learning of Mixture Models, Inf. Inference 7 (2017), 447–508. https://doi.org/10.1093/imaiai/iax015.

[17] J. Dupačová, N. Gröwe-Kuska, W. Römisch, Scenario Reduction in Stochastic Programming, Math. Program. 95 (2003), 493–511. https://doi.org/10.1007/s10107-002-0331-0.

[18] R. Schultz, Stochastic Programming with Integer Variables, Math. Program. 97 (2003), 285–309. https://doi.org/10.1007/s10107-003-0445-z.

[19] F.V. Louveaux, R. Schultz, Stochastic Integer Programming, in: Handbooks in Operations Research and Management Science, Elsevier, 2003: pp. 213–266. https://doi.org/10.1016/s0927-0507(03)10004-7.