Control-Aware Nonlinear Modeling and LLM-Enhanced Decision Support for Coupled Flood–Epidemic Response under Multi-Criteria Trade-offs

Raymond Houé Ngouna[1], Daniel Cobbinah[2]
John Addor Awuah[2], Gabriel Kolaye Guilsou[3] and Bernard Archimede[1]

Abstract— **In this study, we present an interpretable, control-aware framework for convergent urban flood–epidemic response, based on a coupled system of nonlinear ordinary differential equations (ODEs). The model links reduced-order hydrology, infrastructure degradation, displacement, contamination, epidemic dynamics, and operational capacity with bounded state domains that preserve physical meaning and enable robust simulation. On top of this nonlinear ODE core, we integrated transparent baselines, a one-step model predictive controller (MPC), a constraint-aware optimizer with explicit budget and safety criteria, a proximal policy optimization (PPO)-style reinforcement-learning (RL) policy, and hybrid/Pareto policy variants. To improve usability for non-technical stakeholders, we added a read-only, ground-truth-first LLM explanation layer: Python pre-computes per-phase causal premises (triggers, switch drivers, threshold exceedances, and policy mismatches) from the trajectory, and the LLM rewrites only the per-phase control narrative into human-readable prose. A four-backend evaluation — `llama3.2:3b` (raw-payload ablation), `GPT-4o-mini` and `gemma4:e2b` (premise layer), and `llama3-ft` (fine-tuned on premise–prose pairs) — identifies `gemma4:e2b` (7.2 GB, free/local) as the best inference-time trade-off and `llama3-ft` (6 GB, bf16) as a fine-tuned alternative achieving near-complete faithfulness without prompt engineering. Multi-seed stochastic evaluations confirm criterion-dependent non-dominance across four adaptive policies: no single policy is simultaneously optimal for mortality, contamination, shelter safety, reward, and budget, establishing the framework as a multi-criteria decision-support tool rather than a single-winner ranker.**

I. Introduction

Urban flood events increasingly co-occur with public health emergencies, creating convergent crises in which hydrology, infrastructure damage, displacement, contamination, and epidemic spread interact through nonlinear feedback. Decision-makers must trade off competing objectives (e.g., evacuation vs. healthcare mobilization) under tight budget and safety constraints; however, many existing models are either too complex for rapid policy exploration or too opaque to support accountable decision-making. This study addresses this gap by proposing an interpretable, control-aware nonlinear ODE framework for flood–epidemic response, paired with transparent baseline controllers and adaptive/composite policies including PPO, MPC, hybrid-adaptive, and Pareto-dominant control. We further provide a ground-truth-first LLM explanation layer in which Python pre-computes all per-phase causal premises from the trajectory (triggers, switch drivers, threshold exceedances, and policy mismatches), and the LLM rewrites only the per-phase control narrative into human-readable prose under strict faithfulness constraints, keeping the explanation layer fully separated from the simulator and decision logic.

A. Research Questions

- **RQ1**: How can a reduced-order nonlinear ODE model capture the dominant flood–health feedbacks while remaining interpretable and well-posed?
- **RQ2**: How do adaptive and composite control policies (MPC, RL, hybrid, Pareto) rank across competing criteria — mortality, contamination, budget, and safety violations — when the simulator is used as a multi-criteria decision-support tool?
- **RQ3**: Which, if any, outcomes are improved by PPO-style adaptive control in this nonlinear setting when action trajectories remain explainable?
- **RQ4**: How can a ground-truth-first LLM explanation layer eliminate causal attribution errors and improve the factual fidelity of per-phase policy narratives for non-technical stakeholders?

B. Contributions

The main contributions are:

- A control-aware, interpretable nonlinear ODE model with bounded state domains for convergent flood–health dynamics.
- An integrated policy stack comprising transparent/interpretable baselines (no-intervention, rule-based, constraint-aware) and adaptive/composite policies (MPC, PPO, hybrid-adaptive, Pareto-dominant), evaluated against explicit budget and safety criteria under multi-seed stochastic scenarios.
- A multi-seed evaluation framework supporting stochastic and deterministic rainfall scenarios, with policy comparisons reported as criterion-dependent trade-offs rather than single-winner rankings.
- A ground-truth-first LLM explanation layer: Python pre-computes phase triggers, switch drivers, threshold exceedances, and policy mismatches; the LLM rewrites only the prose narrative under faithfulness constraints. A four-backend evaluation

[1]Université de Technologie de Tarbes Occitanie Pyrénées, Laboratoire Génie de Production, 47 avenue d'Azereix 65000 Tarbes, France, raymond.houe-ngouna@uttop.fr, bernard.archimede@uttop.fr

[2]Department of Mathematics, Statistics and Actuarial Science, Takoradi Technical University, Ghana, daniel.cobbinah@ttu.edu.gh, john.addor@ttu.edu.gh

[3]Department of Mathematics and Computer Science, Faculty of Science, University of Maroua, P.O. Box 814, Maroua, Cameroon, kolayegg@gmail.com

uses `llama3.2:3b` (raw payload) as an architecture ablation, `GPT-4o-mini`/`gemma4:e2b` on the premise layer as a controlled model comparison, and `llama3-ft` (fine-tuned Llama 3.2-3B, 6 GB) as a demonstrated no-prompt-engineering alternative; `gemma4:e2b` (7.2 GB, temperature = 0) achieves 100% factual fidelity and `llama3-ft` near-complete faithfulness without prompt engineering.

Section II positions this work relative to flood, infrastructure, WASH (water, sanitation & hygiene) and epidemic modeling, and control-oriented decision-support approaches. Section III presents the interpretable controlled ODE system and intervention levers, along with the policy-layer abstraction used for comparisons. Section IV details the scenario design, RL training, and evaluation metrics. Section V presents a multi-seed, multi-policy evaluation under a coupled crisis stress test and consistency checks via LLM explanations, followed by a discussion of feasibility and interpretability trade-offs. Section VI concludes.

II. Related Work

Prior research has typically modeled the components of flood-health crises in isolation. Hydrologic studies focus on rainfall–runoff and inundation dynamics using high-fidelity hydraulic or reduced-order storage–outflow representations [1]–[3]. Infrastructure damage and service failure are often treated as separate fragility or loss models linked to hazard intensity and exposure [4], [5]. Displacement and shelter pressure are commonly studied within humanitarian or social vulnerability frameworks, emphasizing baseline vulnerability, hazard-triggered mobility, and return dynamics [6]–[8]. Environmental contamination and waterborne disease processes are modeled in epidemiology and WASH literature, capturing pathogen shedding, flood-driven mobilization, and sanitation-mediated decay [9]–[12].

Epidemic dynamics and outcomes are typically studied using compartmental models that emphasize infection, recovery, and mortality in relation to healthcare capacity [13], [14]. Economic loss is frequently modeled as a cumulative index derived from direct flood damage and indirect disruptions [15]–[17]. While each line of work provides mechanistic insight, the resulting models rarely capture the cross-sector feedback that arises when flood impacts, displacement, contamination, and healthcare degradation co-evolve. This motivates an integrated, control-aware formulation that retains interpretability while enabling policy evaluation across the interacting subsystems.

III. Model and Control Formulation

This section formalizes the nonlinear ODE model, defines the control inputs, and positions the optimization and policy layers used in this study.

A. Problem Statement

We sought a reduced-order, interpretable nonlinear dynamical model that captures the essential couplings between flood hazards, infrastructure damage, population displacement, environmental contamination, epidemic spread, and response capacity. The model must (i) preserve physical and operational meaning, (ii) remain computationally efficient for simulation-based policy evaluation, and (iii) admit well-posed trajectories over admissible domains to support learning and optimization of the policy. These requirements motivate the use of a controlled ODE formulation with bounded state variables and explicit intervention levers.

B. Controlled Nonlinear ODE Model

We model the coupled flood–health crisis dynamics as a controlled ODE system:

$$\dot{X}(t) = F\big(t, X(t), u(t)\big), \quad X(0) = X^0 \in \mathcal{D}, \tag{1}$$

where $X(t)$ collects all state variables, $u(t)$ collects bounded intervention controls, and $\mathcal{D}$ denotes the admissible state domain (forward-invariant). The vector field F is constructed from bounded monotone kernels so that the physical meaning and domain constraints are preserved.

The state vector includes hydrologic and infrastructure variables (flood depth, flow, damage, and service failure), displacement and contamination variables, epidemic states (vulnerable, infected, recovered, and deaths), and operational capacities (healthcare, sanitation, personnel, and stress). This structure captures key feedback pathways: flooding and infrastructure degradation amplify displacement and contamination, contamination and crowding jointly drive infection pressure, and operational capacities modulate recovery, mortality, and infrastructure resilience.

C. State Variables

We collected the state variables as:

$$X(t) = \big[X_1, \dots, X_{13}, X_{\text{san}}, X_{\text{team}}, X_{\text{str}}\big]^\top, \tag{2}$$

where: X_1: areal-average flood depth, X_2: signed flood velocity, X_3: physical infrastructure damage (normalized), X_4: infrastructure service failures (normalized), X_5: displaced people (stock), X_6: shelter pressure / overcrowding, X_7: environmental contamination intensity, X_8: vulnerable (at-risk, uninfected) population (stock), X_9: infected people (prevalence stock), X_{10}: recovered people (stock), X_{11}: deaths (cumulative stock), X_{12}: healthcare capacity (normalized), X_{13}: economic loss (normalized), X_{san}: sanitation / water-treatment capability (normalized), X_{team}: available emergency personnel (normalized), X_{str}: operational stress index (normalized).

Economic loss is modeled as an endogenous consequence variable rather than a directly controllable state.

We assume the admissible domains: X_1, X_5, X_6, X_7, X_8, X_9, X_{10}, $X_{11} \in \mathbb{R}_+$, $X_2 \in \mathbb{R}$, X_3, X_4, X_{12}, X_{13}, X_{san}, X_{team}, $X_{\text{str}} \in [0, 1]$. Control inputs are bounded: u_3, u_4, u_6, u_{12}, u_{san}, $u_{\text{team}} \in [0, 1]$.

D. Full Coupled System

a) Construction Hypotheses: The equations of the coupled nonlinear system reflect the dominant flood–health mechanisms—threshold effects, saturation, contamination-driven transmission, and capacity-limited recovery— based on the established environmental and health literature.

b) Detailed Model: The full 16-state controlled ODE system is given below. Equations (3)–(5) govern flood depth and flow velocity with epidemic-feedback terms ϕ_1–ϕ_2; Equations (6)–(7) govern infrastructure damage and service failures with feedback terms ϕ_3–ϕ_4; Equations (8)–(9) govern displacement and shelter pressure; Equations (10)–(14) govern contamination, the SIR epidemic sub-model, and cumulative mortality; and Equations (15)–(18) govern healthcare, economic loss, sanitation, and personnel capacities.

$$\dot{X}_1 = \alpha_R R(t) - k_1 X_1 - \beta_{12} X_1 |X_2| + \beta_{23} |X_2| X_3 + \phi_1 (1 - X_{12}), \tag{3}$$

where $R(t)$ is the stochastic rainfall input modelled as a compound Poisson process:

$$R(t) = \mu_r + \varepsilon(t) + \mathbf{1}[P(t) = 1] \cdot \xi(t),\ \xi \sim \mathcal{N}^+(\mu_p, \sigma_p^2), \tag{4}$$

with $P(t) \sim \text{Bernoulli}(p)$ an independent pulse indicator and ξ a positive-truncated normal pulse. The baseline μ_r captures mean precipitation; $\varepsilon(t)$ is a small zero-mean noise term; and the pulse term adds sporadic heavy-rainfall events. Default values: $\mu_r = 0.60$, $p = 0.30$, $\mu_p = 2.50$, $\sigma_p = 0.50$.

$$\dot{X}_2 = \alpha_{21} X_1 s_f - \beta_{23} X_3 |X_2| X_2 + \phi_2 X_9 X_2, \tag{5}$$

where $s_f \in \{-1, 1\}$ is the flow sign parameter (+1 for flow toward the population center, -1 for flow away from it), so velocity is driven by accumulated water depth while its direction is encoded by s_f,

$$\dot{X}_3 = \alpha_3 (1 - X_3) \frac{[X_1 - h_3]_+}{1 + [X_1 - h_3]_+} A_f(X_1) \frac{|X_2|}{1 + \eta_3 |X_2|} - \rho_3 u_3 X_3 + \phi_3 (1 - X_{12}), \tag{6}$$

$$\dot{X}_4 = \alpha_4 (1 - X_4) \frac{[X_1 - h_4]_+}{1 + [X_1 - h_4]_+} \frac{|X_2|}{1 + \eta_4 |X_2|} X_3^{p_4} - \rho_4 u_4 X_4 + \phi_4 X_9. \tag{7}$$

The four epidemic-feedback coefficients $\phi_1, \ldots, \phi_4$ introduce reverse coupling from the health subsystem back to the hydrological and infrastructure sub-models. $\phi_1(1 - X_{12})$ captures drainage maintenance failure when healthcare is overwhelmed; $\phi_2 X_9 X_2$ represents indirect flow-management disruption caused by workforce absenteeism and degraded operational coordination during epidemic peaks; $\phi_3(1 - X_{12})$ models accelerated infrastructure deterioration under reduced maintenance capacity; and $\phi_4 X_9$ accounts for service degradation through workforce absenteeism.

$$\dot{X}_5 = b_5 + \beta_5 [X_1 - h_5]_+ A_f(X_1) |X_2| X_3^{p_5} - k_5 X_5, \tag{8}$$

$$\dot{X}_6 = b_6 + \alpha_6 X_5 - \rho_6 u_6 X_6, \tag{9}$$

$$\dot{X}_7 = \sigma_9 \frac{X_9}{1 + X_9} + \sigma_F A_f(X_1) \frac{|X_2|}{1 + \eta_7 |X_2|} - k_7 X_7 - \kappa_7 X_{\text{san}} X_7, \tag{10}$$

$$\dot{X}_8 = b_8 + \beta_8 [X_1 - h_8]_+ |X_2| \frac{X_6}{1 + X_6} \frac{X_7}{1 + X_7} - \lambda(t) X_8 - k_8 X_8, \tag{11}$$

$$\dot{X}_9 = \lambda(t) X_8 - \gamma(X_{12}) X_9 - \mu(X_{12}) X_9, \tag{12}$$

$$\dot{X}_{10} = \gamma(X_{12}) X_9, \tag{13}$$

$$\dot{X}_{11} = \mu(X_{12}) X_9 + \theta_F A_f(X_1) \frac{|X_2|}{1 + \eta_{11} |X_2|} + \theta_D \frac{X_5}{1 + X_5}, \tag{14}$$

$$\dot{X}_{12} = \rho_{12} u_{12} (1 - X_{12}) - X_{12} (\delta_0 + \delta_3 X_3 + \delta_4 X_4 + \delta_{13} X_{13}), \tag{15}$$

$$\dot{X}_{13} = (1 - X_{13}) \Big(\alpha_{13} A_f(X_1) \frac{|X_2|}{1 + \eta_{13} |X_2|} X_3^{p_{13}} + \chi_5 \frac{X_5}{1 + X_5} + \chi_9 \frac{X_9}{1 + X_9} \Big), \tag{16}$$

$$\dot{X}_{\text{san}} = \rho_{\text{san}} u_{\text{san}} (1 - X_{\text{san}}) - \delta_{\text{san}} X_{\text{san}}, \tag{17}$$

$$\dot{X}_{\text{team}} = \rho_{\text{team}} u_{\text{team}} (1 - X_{\text{team}}) - X_{\text{team}} (\delta_0 + \delta_{13} X_{13} + \delta_{\text{str}} X_{\text{str}}). \tag{18}$$

The operational stress index is evaluated as the bounded diagnostic $X_{\text{str}} = (X_4 + X_5)/(1 + X_4 + X_5)$, coupling service failure and displacement pressure; it enters personnel dynamics through δ_{str}.

where

$$A_f(X_1) = \frac{[X_1 - h_A]_+}{1 + [X_1 - h_A]_+}, \tag{19}$$

$$\lambda(t) = \beta \frac{X_7}{1 + X_7} \frac{X_6}{1 + X_6}, \tag{20}$$

$$\gamma(X_{12}) = \gamma_0 + \gamma_1 X_{12}, \tag{21}$$

$$\mu(X_{12}) = \mu_0 + \mu_1 (1 - X_{12})^r, \quad r \geq 1. \tag{22}$$

and $[z]_+ = \max(z, 0)$ ensures nonnegative threshold activation. In Eq. (19), $A_f \in [0, 1)$ is a saturating flood-activation kernel: it is zero below the threshold h_A and approaches one asymptotically, ensuring bounded activation and preserving the physical admissibility of the dynamics. In Eq. (20), $\lambda(t)$ is the force of infection: the product of saturating contamination and crowding fractions scales the transmission rate β, so infection pressure grows jointly with waterborne hazard and shelter density. In Eq. (21), $\gamma(X_{12})$ is the healthcare-modulated recovery rate: a baseline γ_0 is augmented by $\gamma_1 X_{12}$, so recovery accelerates when healthcare capacity is available and collapses to γ_0 when $X_{12} \to 0$. In Eq. (22), $\mu(X_{12})$ is the healthcare-dependent mortality rate: the baseline μ_0 is raised by a convex term $\mu_1(1 - X_{12})^r$ that grows rapidly as healthcare is depleted, capturing excess mortality under overwhelmed health systems.

The coefficients in the full system equations are:

- **Rainfall process:** μ_r (baseline intensity), p (pulse probability), μ_p (pulse mean), σ_p (pulse std.).
- **Hydrology/infrastructure:** α_R, α_{21}, s_f, α_3, α_4, k_1, β_{12}, β_{23}, η_3, η_4, p_4, ρ_3, ρ_4, h_3, h_4, h_A, ϕ_1, ϕ_2, ϕ_3, ϕ_4.
- **Displacement/shelter:** $b_5, \beta_5, h_5, p_5, k_5, b_6, \alpha_6, \rho_6$.
- **Contamination:** $\sigma_9, \sigma_F, \eta_7, k_7, \kappa_7$.
- **Vulnerability/infection:** $b_8, \beta_8, h_8, k_8, \beta, \gamma_0, \gamma_1, \mu_0, \mu_1, r$.
- **Deaths:** $\theta_F, \eta_{11}, \theta_D$.
- **Healthcare:** $\rho_{12}, \delta_0, \delta_3, \delta_4, \delta_{13}$.
- **Economy:** $\alpha_{13}, \eta_{13}, p_{13}, \chi_5, \chi_9$.
- **Sanitation/personnel:** $\rho_{\text{san}}, \delta_{\text{san}}, \rho_{\text{team}}, \delta_{\text{str}}$.

All are assumed nonnegative unless stated otherwise; thresholds $h_\cdot \in \mathbb{R}$ and exponents satisfy $p_4, p_5, p_{13} \geq 1$, $r \geq 1$. All 60 ODE parameter values used in Section V are listed in Table III (Appendix I); entries with two values denote Exp. I / Exp. II.

c) Cascading Amplification: The feedback structure of the coupled model suggests the possibility of cascading amplification mechanisms under strong nonlinear coupling [18]. In particular, flood shocks can propagate through displacement, contamination, epidemic pressure, healthcare-capacity degradation, infrastructure deterioration, and operational-capacity stress, thereby indirectly aggravating flood impacts and service failures. The model therefore captures reinforcing cross-domain interactions relevant to policy design under convergent crisis conditions.

E. Control Inputs

We define bounded intervention controls

$$u(t) = \big[u_3, u_4, u_6, u_{12}, u_{\text{san}}, u_{\text{team}}\big], \tag{23}$$

representing building protection/repair, service restoration, evacuation and shelter expansion, healthcare mobilization, sanitation/water treatment, and personnel mobilization. Each control is constrained to $[0, 1]$ and acts either by accelerating the recovery and mitigation processes or by increasing the operational capacities.

F. Policy Layers and Constraints

On top of the nonlinear ODE core, we evaluated a no-intervention reference and multiple policy layers: a rule-based controller (thresholded responses), a constraint-aware controller using explicit budget and safety criteria, one-step model-predictive control (short-horizon optimization), PPO-style adaptive control, and hybrid/Pareto composite policies. All policies share the same state and control definitions; within each experiment, they are evaluated on matched stochastic seeds and rainfall realizations so that differences reflect the decision layer rather than scenario changes.

IV. Experimental Setup

This section describes the scenario generation, policy baselines, training configuration, and evaluation metrics used to compare the decision policies under convergent flood–health dynamics.

A. Scenarios and Simulation Horizon

We simulated T discrete decision steps of the controlled ODE system. The reported policy comparisons use stochastic rainfall pulses: rainfall is generated from a baseline intensity plus random shocks, and each seed defines one rainfall realization. For fairness, all policies within the same experiment and seed use the same rainfall realization and initial conditions. The simulator also supports deterministic exponential-decay rainfall profiles, but the multi-seed results in Table I use the stochastic pulse settings specified in Section V.

B. Policies Compared

In this study, we evaluated two policy families built on the same controlled ODE simulator. The first family compares transparent baselines: no intervention; a threshold-triggered *rule-based* controller that fires fixed action bundles when state variables cross predefined safety thresholds; and a *constraint-aware* controller that gates those same bundles against an explicit cumulative budget cap and per-step safety-criterion filters. The second family compares adaptive or composite policies: *one-step MPC*, which selects the greedy action maximising the immediate reward at each step without multi-step planning; the *trained PPO* policy (Section IV-C); a *hybrid-adaptive* policy that routes to constraint-aware, heuristic, or MPC sub-policies according to per-step flood/epidemic hazard scores; and a *Pareto-dominant* composite that augments the constraint-aware bundle with rule-based overrides chosen to avoid simultaneous domination across all safety-cap criteria. All policies share the same state, action space, and scenario realizations, enabling controlled comparisons across decision paradigms.

We define the decision problem as a finite-horizon Markov decision process (MDP).

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, P, r, \gamma, T), \tag{24}$$

where the state $s_t \in \mathcal{S}$ corresponds to the ODE state vector $X(t)$ (or an observation thereof), the action $a_t \in \mathcal{A}$ corresponds to the bounded intervention vector $u(t) = \big[u_3, u_4, u_6, u_{12}, u_{\text{san}}, u_{\text{team}}\big]$, and the transition kernel $P(s_{t+1} \mid s_t, a_t)$ is induced by nonlinear controlled dynamics and scenario drivers (e.g., rainfall).

The objective is to learn a policy $\pi_\theta(a \mid s)$ that maximizes the expected discounted return as follows:

$$\pi_\theta^\star = \arg\max_{\pi_\theta} \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{T-1} \gamma^t\, r(s_t, a_t)\right], \tag{25}$$

where the reward r encodes multi-objective trade-offs between health outcomes, service disruption, economic loss, and intervention costs, and $\gamma \in (0,1)$ is the discount factor. Explicitly, the step reward is:

$$r(s_t, s_{t+1}, u_t) = -\big(w_I\, X_9^{t+1} + w_C\, X_7^{t+1} + w_D\, [\Delta X_{11}]_+ \\ + w_E\, [\Delta X_{13}]_+ + w_u\, \|u_t\|_1\big), \tag{26}$$

where $[\Delta X_i]_+ = \max(X_i^{t+1} - X_i^t, 0)$ denotes the one-sided increment (the agent is penalized only when deaths or economic loss worsen). The default weights are $w_I = 1.0$ (infection level), $w_C = 0.6$ (contamination level), $w_D = 2.0$ (mortality increment), $w_E = 0.4$ (economic loss increment), and $w_u = 0.1$ (L_1 action cost). In the advanced-policy stress test we also report a resource-sensitive variant with $w_u = 0.55$ to make the intervention-cost trade-off explicit. Safety and budget limits are represented through explicit violation counts, constrained reward terms, or post hoc action projection.

We then evaluate each controller by inspecting outcome metrics, constraint violations, action trajectories, and LLM-based explanations of policy priorities.

C. RL Training Configuration

The RL controller uses the same MDP state, reward in (26), and bounded action vector as the other policies. The `ppo_trained` checkpoint was trained for 80,000 environment steps using an MLP actor–critic with $L = 3$ hidden layers and $H = 128$ units per layer; all hyperparameters — including discount factor $\gamma = 0.981$, learning rate 8.8×10^{-5}, entropy coefficient 6.2×10^{-5}, clip range 0.21, and rollout length $n_{\text{steps}} = 128$ — were selected jointly via an Optuna hyperparameter search over 40 trials (objective: mean undiscounted episode return). The checkpoint is evaluated on the same seeds and rainfall realizations as the other controllers; PPO claims are interpreted through observed rollout metrics, not as a guarantee of constraint satisfaction.

D. Evaluation Metrics

For each policy and seed, we tracked end-of-horizon deaths $X_{11}(T)$, peak infection $\max_t X_9(t)$, peak contamination $\max_t X_7(t)$, peak operational stress $\max_t X_{\text{str}}(t)$, safety-violation counts $\#\{t : X_7(t) > 0.65\}$ and $\#\{t : X_6(t) > 2.0\}$, cumulative reward $\sum_{t=0}^{T-1} r_t$, and total intervention budget $\sum_t \|u_t\|_1$. These quantities cover health, environmental exposure, operational stress, feasibility, and resource use. We report mean±standard deviation over seeds and inspect time-varying action profiles to assess intervention priority and timing.

E. Explainability Layer

The explainability layer uses a ground-truth-first architecture: all causal premises are pre-computed by Python before any LLM is invoked, eliminating attribution errors by construction.

a) Phase detection: Consecutive steps sharing the same dominant action are grouped into phases $\mathcal{P}_i = [s_i, e_i]$, where $d_t = \arg\max_k |u_k(t)|$ ($d_t = \emptyset$ if $\max_k |u_k(t)| \le \epsilon$).

b) Primary trigger: For each phase, the phase-mean $\bar{x}_k^i = |\mathcal{P}_i|^{-1} \sum_{t \in \mathcal{P}_i} x_k(t)$ is computed over threshold-monitored variables. The primary trigger is

$$k^* = \arg\max_{k:\, \bar{x}_k^i \ge 0.6\theta_k}\ \rho_k^i, \qquad \rho_k^i = \tfrac{\bar{x}_k^i - \theta_k}{\theta_k}, \tag{27}$$

where θ_k is the safety threshold for indicator k and the filter $\bar{x}_k^i \ge 0.6\theta_k$ restricts to indicators near or above their threshold.

c) Switch driver: Between phases $\mathcal{P}_{i-1}$ and $\mathcal{P}_i$, the threshold-relative delta

$$\Delta_k^i = \tfrac{\bar{x}_k^i - \bar{x}_k^{i-1}}{\theta_k} \tag{28}$$

measures cross-phase change in scale-fair units (preventing bias between bounded variables such as $X_4 \in [0,1]$ and unbounded ones such as X_1). The switch driver k_{sw} is selected by a three-level fallback: (i) the action-aligned indicator that worsened most ($\Delta_k^i > \epsilon_{\min} = 0.005$); (ii) a saturated aligned indicator ($\bar{x}_k^i \ge \theta_k$, no further worsening); (iii) the indicator with the largest positive Δ_k^i.

d) Mismatch flag: A mismatch is flagged when $\mathcal{A}^{-1}(k^*) \ne d_i$ — the primary trigger's natural target (e.g., $X_7 \mapsto u_{\text{san}}$) differs from the chosen action d_i — a signature of PPO multi-objective trade-offs.

The LLM rewrites the action-choice explanation into prose preserving all numbers, listing all also-elevated indicators ($\bar{x}_k^i \ge 0.6\theta_k$, $k \ne k^*$), and including the mismatch explanation and saturation context when flagged. The remaining report sections are rendered directly from Python without LLM involvement.

V. Results

We evaluate four adaptive policies under a coupled-crisis stress test (T=80, baseline rainfall 0.70, pulse probability 0.35, pulse mean 3.00 (std 0.70), $X_1(0)$=0.8, $X_7(0)$=0.05, w_u=0.55, safety caps X_7>0.65 and X_6>2.0). Values are mean±std over seeds 0–29. LLM explanations use `gemma4:e2b` (7.2 GB, temperature = 0) as described in Section IV-E.

A. Results I — Advanced Policy Comparison

We compare one-step MPC, the trained PPO policy, hybrid-adaptive control, and a Pareto-dominant composite policy. The one-step MPC is the least expensive policy (5.0 ± 0.5 cumulative budget units) but leaves high deaths and violations, showing that low spending alone is not a safety criterion. The trained PPO policy achieves the lowest peak contamination (1.17 ± 0.04) and the lowest peak infection (0.48 ± 0.04); in the illustrative seed-0 and mean rollouts (Fig. 1), the policy briefly activates sanitation in the early steps before shifting to structural flood repair (u_3) as the dominant action across most of the horizon, as confirmed by the LLM explanation layer analysis in Section V-C. Hybrid-adaptive control achieves the highest mean reward among the evaluated adaptive policies (-198.7 ± 5.8), reflecting the explicit action-cost penalty. The Pareto-dominant composite obtains the strongest remaining safety outcomes: lowest deaths (18.9 ± 1.2), lowest peak stress (0.841 ± 0.022), and fewest shelter-cap violations ($37.6 \pm$

14.9), but it does so at very high cumulative budget cost (305.1 ± 4.3). This is the clearest expression of the multi-perspective result: better health/safety outcomes require more aggressive resource use, whereas reward- or budget-oriented criteria select different policies.

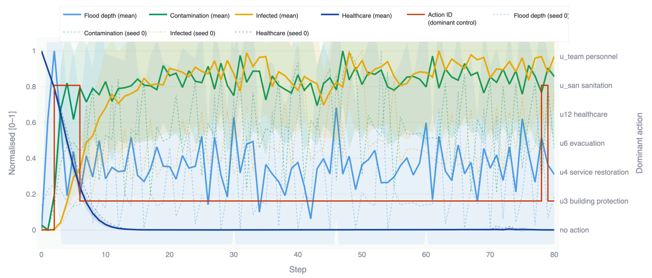

Fig. 1. PPO crisis trajectory under the coupled stress-test setting (mean + seed 0). The red step line gives the dominant active control, selected as the active control with the highest action value; equal-value ties are resolved by action ordering $[u_3, u_4, u_6, u_{12}, u_{\text{san}}, u_{\text{team}}]$; co-active controls are omitted from the action-ID overlay.

B. Results II — Cross-Criterion Interpretation

Table I confirms criterion-dependent non-dominance across both experiments. In Exp. I, rule-based control achieves the best overall health outcomes and shelter safety, while constraint-aware control achieves fewer contamination-threshold violations and the lowest reward cost. In Exp. II, Pareto performs best on deaths, peak stress, and shelter violations; PPO for peak contamination and infection; hybrid-adaptive for reward; and MPC for budget only. Policy choice is therefore inseparable from the operative criterion.

C. Results III — LLM Explanation Layer (Faithfulness Evaluation)

Following the ground-truth-first architecture in Section IV-E, we evaluated the PPO seed-0 trajectory using `gemma4:e2b` (7.2 GB, Ollama on-device, temperature $= 0$). Python pre-computed five phases with, for each, a primary trigger ranked by relative exceedance, a complete also-elevated list, a threshold-relative switch driver ($\Delta = |\delta \bar{x}_k| / \theta_k \times 100\%$, preventing scale confusion between bounded variables such as $X_4 \in [0, 1]$ and unbounded ones such as X_1), and mismatch flags where the primary trigger's natural action differed from the chosen policy action. The LLM received these pre-computed phase bullets and rewrote only the action-choice explanation into prose.

The generated prose was evaluated against the Python premises element by element. All faithfulness criteria were satisfied across all phases: primary trigger with threshold percentage (4/4 action phases), complete also-elevated indicator list (3/3 phases), switch driver with Δ-notation (4/4), three-part mismatch explanation including target action, chosen action, and PPO multi-objective rationale (2/2 flagged phases), and saturation context where applicable. No hallucinated content or preamble was introduced. Fig. 2 shows the resulting action-choice explanation output.

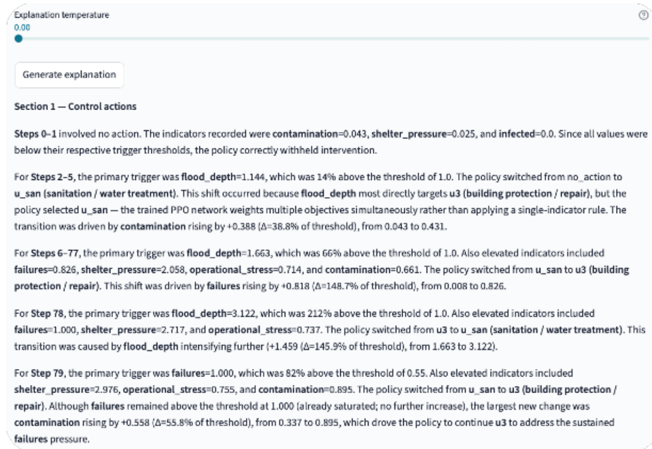

Fig. 2. Action-choice explanation prose generated by `gemma4:e2b` (7.2 GB, temperature $= 0$) from Python-computed premises for the trained PPO trajectory (seed 0, $T = 80$). Five phases are identified; all primary triggers, also-elevated indicators, threshold-relative switch drivers (Δ-notation), and mismatch flags are reproduced faithfully from the premise layer, achieving 100% factual fidelity.

a) Consistency check and model comparison: The generated action sequence — no action $\rightarrow u_{\text{san}} \rightarrow u_3$ dominant $\rightarrow$ brief u_{san} return $\rightarrow u_3$ resumption — is consistent with Table I and Fig. 1: the LLM introduces no actions absent from the simulator record. The auditability property holds: each stated reason can be verified against the Python premise layer. Table II uses `llama3.2:3b` (raw payload, no premise layer) as an *architecture ablation* and `GPT-4o-mini`/`gemma4:e2b` on the premise layer as a *controlled model comparison*.

`llama3.2:3b` (2 GB, free/local), without the premise layer, produced causal inversions, dropped switch drivers, and generated terse prose — confirming that the architecture, not model scale alone, is the primary driver of faithfulness failures.

`GPT-4o-mini` (hosted API), given the Python premise layer, achieved near-complete faithfulness and fluent prose suitable for non-expert readers, with one partial failure on also-elevated indicator lists. It requires a paid external API, making it unsuitable for offline deployment.

`gemma4:e2b` (7.2 GB, free/local), given the same premise layer at temperature $= 0$, achieves 100% faithfulness across all criteria and prose quality matching the API model — the best inference-time trade-off, combining full fidelity with no external dependency.

`llama3-ft` (fine-tuned Llama 3.2-3B, 6 GB bf16, free/local), trained on Python-verified premise–prose pairs without explicit prompt engineering, achieves near-complete faithfulness: all primary triggers, also-elevated lists, switch drivers with Δ-notation, and mismatch explanations are reproduced correctly; only the saturation context for one phase was absent.

D. Discussion

The LLM layer translates policy outcomes into structured narratives while remaining separated from the simulator

TABLE I

MULTI-SEED POLICY COMPARISON (N=30, MEAN±STD; LOWER BETTER EXCEPT $\sum r$). X_{11}: DEATHS; X_9/X_7/X_{str}: PEAK INFECTED/CONTAMINATION/STRESS; X_7>0.65/X_6>2.0: SAFETY-CAP VIOLATIONS; $\sum r$: REWARD; BUDGET: $\sum_t \|u_t\|_1$. EXP. I: DEFAULT (w_u=0.10); EXP. II: STRESS TEST (w_u=0.55); T=80 FOR BOTH.

Exp.	Policy	Deaths X_{11}	Peak X_9	Peak X_7	Peak X_{str}	Viol. $X_7 > 0.65$	Viol. $X_6 > 2.0$	Reward $\sum r$	Budget used
I	No policy	30.1 ± 0.8	0.647 ± 0.009	1.934 ± 0.075	0.949 ± 0.007	74.6 ± 1.5	71.6 ± 1.2	-175.3 ± 4.8	0.0 ± 0.0
I	Rule-based	$\mathbf{17.5\pm0.9}$	$\mathbf{0.293\pm0.021}$	$\mathbf{0.793\pm0.033}$	$\mathbf{0.809\pm0.030}$	20.0 ± 4.5	$\mathbf{31.6\pm14.4}$	-107.0 ± 4.5	293.1 ± 11.7
I	Constraint-aware	21.0 ± 1.1	0.362 ± 0.022	0.804 ± 0.027	0.842 ± 0.025	$\mathbf{17.1\pm6.3}$	67.5 ± 3.3	$\mathbf{-104.3\pm5.0}$	160.0 ± 0.0
II	MPC	36.7 ± 1.1	0.830 ± 0.011	2.496 ± 0.104	0.957 ± 0.005	77.4 ± 0.5	72.8 ± 1.0	-198.8 ± 6.0	$\mathbf{5.0\pm0.5}$
II	PPO	22.8 ± 1.4	$\mathbf{0.480\pm0.040}$	$\mathbf{1.170\pm0.040}$	0.860 ± 0.020	$\mathbf{40.5\pm5.5}$	52.1 ± 11.5	-219.6 ± 6.2	220.8 ± 3.6
II	Hybrid-adaptive	36.2 ± 1.1	0.832 ± 0.010	2.510 ± 0.094	0.957 ± 0.005	75.4 ± 1.4	72.3 ± 1.0	$\mathbf{-198.7\pm5.8}$	13.8 ± 1.5
II	Pareto-dominant	$\mathbf{18.9\pm1.2}$	0.550 ± 0.056	1.327 ± 0.052	$\mathbf{0.841\pm0.022}$	55.7 ± 3.1	$\mathbf{37.6\pm14.9}$	-265.6 ± 7.9	305.1 ± 4.3

TABLE II

FAITHFULNESS AND PROSE QUALITY ACROSS FOUR BACKENDS. ✓ = CORRECT; ○ = PARTIAL; × = FAILED/DROPPED.

Model	Premise	Trigger	Also-elev.	Switch	Mismatch	Prose	Access
`llama3.2:3b`	×	×	×	×	×	terse	free/local
`GPT-4o-mini`	✓	✓	○	✓	✓	fluent	paid API
`gemma4:e2b`	✓	✓	✓	✓	✓	fluent	free/local
`llama3-ft`	✓	✓	✓	✓	✓	fluent	free/local

and policy logic. The ground-truth-first architecture further ensures that all causal claims in the prose are pre-verified by Python: the LLM cannot invent trigger reasons or attribution — it can only paraphrase premises whose factual content has already been computed and checked.

Together, these results support the claim that nonlinear, control-aware modeling should be paired with explicit criteria and human-centered interpretability in crisis decision support.

VI. CONCLUSION

We introduced an interpretable control-aware nonlinear ODE framework for convergent flood–epidemic response and evaluated seven policies under multi-seed stochastic scenarios. Results confirm criterion-dependent non-dominance: rule-based control minimises deaths, peak infections, contamination, and shelter violations (transparent-baseline experiment); constraint-aware control gives the best reward at lower budget; among adaptive policies, PPO achieves the lowest peak contamination and infection; hybrid-adaptive gives the highest reward; MPC minimises spending but not risk; and Pareto-dominant control achieves the lowest deaths, peak stress, and fewest shelter violations at the highest cost. No policy simultaneously optimises all criteria: policy choice is inseparable from an explicit statement of the operative objective. A four-backend evaluation confirmed that premise grounding is the primary faithfulness driver: `gemma4:e2b` achieves 100% fidelity at zero API cost, and `llama3-ft`, fine-tuned on verified premise–prose pairs, achieves near-complete faithfulness without prompt engineering — demonstrating fine-tuned local models as a viable deployment path. Future work will address data calibration, richer constraints, extended training protocols, and multi-region extension; extending premise-based grounding to all explanation sections and embedding the layer within a human-in-the-loop interface are the primary interpretability directions.

REFERENCES

[1] V. T. Chow, D. R. Maidment, and L. W. Mays, *Applied Hydrology*. McGraw-Hill, 1988.

[2] P. D. Bates, M. S. Horritt, and T. J. Fewtrell, "A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling," *Journal of Hydrology*, vol. 387, no. 1–2, pp. 33–45, 2010.

[3] K. Beven, "Rainfall–runoff modelling: The primer," *Wiley*, 2012.

[4] M. Ouyang, "A review on modeling and simulation of interdependent critical infrastructure systems," *Reliability Engineering & System Safety*, 2014.

[5] S. Hallegatte, J. Rentschler, and J. Rozenberg, "Lifelines: The resilient infrastructure opportunity," World Bank Group, Washington, DC, Tech. Rep., 2019.

[6] I. D. M. Centre, "Global report on internal displacement," 2017.

[7] S. Association, "The sphere handbook: Humanitarian charter and minimum standards in humanitarian response," 2018.

[8] UNHCR, "Emergency handbook," 2014.

[9] P. A. Collender, O. C. Cooke, L. D. Bryant, T. R. Kjeldsen, and J. V. Remais, "Estimating the microbiological risks associated with inland flood events: Bridging theory and models of pathogen transport," *Critical Reviews in Environmental Science and Technology*, vol. 46, no. 23–24, pp. 1787–1833, 2016.

[10] S. C. Chapra, *Surface Water-Quality Modeling*. New York: McGraw-Hill, 1997.

[11] W. H. Organization, "Wash and health in emergencies," 2017.

[12] UNICEF, "Water, sanitation and hygiene in emergencies," 2019.

[13] R. M. Anderson and R. M. May, "Infectious diseases of humans: Dynamics and control," *Oxford University Press*, 1992.

[14] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, pp. 700–721, 1927.

[15] S. Hallegatte, "The indirect cost of natural disasters and an economic definition of macroeconomic resilience," World Bank, Policy Research Working Paper 7357, 2015.

[16] Centre for Research on the Epidemiology of Disasters (CRED) and United Nations International Strategy for Disaster Reduction (UNISDR), "The human cost of weather related disasters 1995–2015," UNISDR, Geneva, Tech. Rep., 2015.

[17] A. Delivorias and N. Scholz, "Economic impact of epidemics and pandemics," European Parliamentary Research Service, Briefing EPRS_BRI(2020)646195, 2020.

[18] S. H. Strogatz, *Nonlinear Dynamics and Chaos*. Reading, MA: Addison-Wesley, 1994.

APPENDIX I
MODEL PARAMETERS

Table III lists all 60 ODE parameters. Single values are shared across scenarios; *default/stress* entries give the article-default and coupled-crisis stress-test values respectively.

TABLE III

ODE parameter values. Default/stress-test values where they differ; single values are shared.

Param.	Description	Val.	Param.	Description	Val.	Param.	Description	Val.	Param.	Description	Val.
Hydrology			*Infrastructure*			*Displacement / Shelter*			*Contamination*		
α_R	rainfall gain	0.80/1.10	α_3	damage growth	0.40	b_5	base displacement	0.10	σ_9	pathogen shedding	0.40/0.60
k_1	drainage rate	0.30	h_3	damage threshold	0.20	β_5	flood displ. rate	0.60	σ_F	flood mobilization	0.50/0.80
β_{12}	advective export	0.20/0.35	η_3	velocity sat.	0.50	h_5	displ. threshold	0.10	η_7	velocity saturation	0.50
β_{23}	ponding feedback	0.20	ρ_3	repair rate	0.30	p_5	displ. exponent	2.0	k_7	natural decay	0.30
α_{21}	depth-to-flow	0.50	α_4	failure growth	0.50	k_5	return rate	0.20	κ_7	sanit. efficacy	0.70
s_f	flow orientation	1.0	h_4	failure threshold	0.20	b_6	baseline shelter	0.05			
ϕ_1	drainage response fail.	0.10	η_4	velocity sat.	0.50	α_6	shelter growth	0.40			
ϕ_2	flow disruption	0.15	p_4	failure exponent	2.0	ρ_6	evacuation efficacy	0.40			
			ρ_4	restoration rate	0.30						
			ϕ_3	maint. cap. fail.	0.05						
			ϕ_4	svc. degradation	0.10						
Epidemic			*Healthcare*			*Economy / Mortality*			*Sanitation / Personnel*		
b_8	base recruitment	0.05	ρ_{12}	HC mobilization	0.60	θ_F	flood mortality	0.20	ρ_{san}	san. mobilization	0.60
β_8	exposure coeff.	0.40	δ_0	baseline decay	0.10	η_{11}	velocity sat.	0.50	δ_{san}	sanitation decay	0.20
h_8	exposure threshold	0.10	δ_3	damage strain	0.30	θ_D	displ. mortality	0.10	ρ_{team}	pers. mobilization	0.50
k_8	vuln. outflow	0.20	δ_4	failure strain	0.20	α_{13}	econ. loss rate	0.40	δ_{str}	stress strain	0.20
β	transmission rate	0.60/0.90	δ_{13}	economic strain	0.20	η_{13}	velocity sat.	0.50	h_A	flood activation	0.10
γ_0	base recovery	0.30				p_{13}	loss exponent	1.2			
γ_1	capacity recovery	0.30				χ_5	displ. loss coeff.	0.30			
μ_0	base mortality	0.10				χ_9	infect. loss coeff.	0.30			
μ_1	capacity mortality	0.20/0.35									
r	mortality exponent	2.0									

Appendix II
Well-Posedness of the Controlled Dynamics

a) Purpose: This appendix ensures that the simulator induces a unique trajectory for any admissible initial condition and bounded control input, and that the trajectory remains well-defined over the full simulation horizon.

b) Assumptions (mild regularity): (i) For each fixed x, $t \mapsto F(t, x, u(t))$ is measurable. (ii) For almost every t, $x \mapsto F(t, x, u(t))$ is locally Lipschitz on $\mathcal{D}$ (each component is built from sums/products of Lipschitz saturations and bounded kernels). (iii) On $\mathcal{D}$, F has at most linear growth in the non-normalized stock variables, while the normalized components remain bounded by construction.

c) Claim (existence, uniqueness, and global extension): For every $X^0 \in \mathcal{D}$, there exists a unique absolutely continuous solution $X : [0, \infty) \to \mathbb{R}^n$ satisfying the ODE for almost every $t \geq 0$. In particular, the solution is defined for a full simulation horizon (no finite-time blow-up).

d) Sketch of proof: By (i)–(ii), F satisfies the Carathéodory conditions; hence, there exists a unique local solution on $[0, T_{\max})$. To show global existence, note that normalized states (e.g., $X_3, X_4, X_{12}, X_{13}, X_{\mathrm{san}}, X_{\mathrm{team}}$) are bounded by design (see Appendix III for forward invariance of $\mathcal{D}$); because all normalized states remain bounded in $[0, 1]$, all cross terms are bounded by linear functions of the stock variables. For nonnegative stock variables, inflows are bounded by constants/bounded kernels and outflows are at most linear in the state, yielding inequalities of the form

$$\dot{X}_i(t) \leq c_{i,0} + c_{i,1} X_i(t)$$

for finite constants $c_{i,0}, c_{i,1}$. Grönwall's inequality implies boundedness on any finite interval, which rules out finite-time blow-up; thus, $T_{\max} = +\infty$.

Appendix III
Positive Invariance (Physical Admissibility)

For any admissible initial condition $X^0 \in \mathcal{D}$, the unique solution of the controlled system satisfies $X(t) \in \mathcal{D}$ for all $t \geq 0$. Indeed, for nonnegative stock variables $X_i \in \mathbb{R}_+$, each equation has the form "nonnegative inflows minus a nonnegative rate times X_i," implying $\dot{X}_i \geq 0$ on the boundary $X_i = 0$. For normalized capacity variables $Z \in [0, 1]$ (e.g., $X_3, X_4, X_{12}, X_{13}, X_{\mathrm{san}}, X_{\mathrm{team}}$), the dynamics satisfy $\dot{Z} \geq 0$ at $Z = 0$ and $\dot{Z} \leq 0$ at $Z = 1$, ensuring forward invariance of $[0, 1]$. The diagnostic stress index $X_{\mathrm{str}} = (X_4 + X_5)/(1 + X_4 + X_5)$ also remains in $[0, 1]$. The velocity variable $X_2 \in \mathbb{R}$ is unconstrained. The epidemic-to-flow feedback term $\phi_2 X_9 X_2$ is interpreted as an indirect operational disruption effect rather than a direct hydrodynamic interaction. Hence, the admissible domain $\mathcal{D}$ is positively invariant, and all trajectories remain physically meaningful.