



Multimodal Wildfire Intelligence for Suppression-Zone Prioritization: Bayesian Fusion and Risk-Averse Stochastic Optimization

Seyyed Alireza Hoseini¹ , Leila Daneshvar² , Fatemeh Zarghami³ , Mahyar Khoshantash⁴
Zahra Mohemmi^{5,*}

¹ University of Tehran, Tehran, Iran

² Islamic Azad University, Tehran, Iran

³ Islamic Azad University, Tehran, Iran

⁴ University of Science and Culture, Tehran, Iran

⁵ Flinders University, Adelaide, Australia

ARTICLE INFO

Article history:

Received 10 August 2026

Received in revised form 21 September 2026

Accepted 27 September 2026

Available online 29 September 2026

Keywords:

wildfire decision support; multimodal sensing;
Bayesian data fusion; stochastic
programming; conditional value-at-risk;
resource staging; simulation validation.

ABSTRACT

Wildfire decision support requires more than detecting fire: uncertain observations must inform feasible resource commitments before environmental conditions are fully known. This study develops a transparent interface between multimodal sensing, Bayesian risk estimation, and risk-averse suppression-resource staging. Thermal, red-green-blue (RGB), multispectral, light detection and ranging (LiDAR), meteorological, gas/smoke, and terrain channels are represented by correlated synthetic risk proxies on a 100 km² landscape containing 1,600 cells. Each of 10 independent replications generates 9,600 cell-time records; final-time predictive evaluation uses 16,000 records overall. A Gaussian Bayesian model accounts for heterogeneous measurement precision and missing channels. Posterior zone-level moments feed Monte Carlo demand scenarios with shared environmental disruptions. A 2-stage linear program stages divisible service capacity at 4 bases and allocates it among 100 operational zones. Evaluation uses 20 planning scenarios and 50 independent test scenarios per replication. Full fusion reduces mean Brier score from 0.1585 for thermal-only estimation to 0.1565, while latent-state root-mean-square error decreases from 0.5806 to 0.4991. Compared with Bayesian deterministic staging, Bayesian stochastic staging reduces mean priority-weighted unmet demand by 35.0% and conditional value-at-risk by 25.3%. However, its additional operational advantage over thermal-only stochastic staging is negligible within simulation uncertainty. Nominal posterior interval coverage deteriorates to 72.1% when unmodeled sensor noise doubles. These findings distinguish predictive information value from operational decision value. The study also specifies an auditable integration protocol for real wildfire datasets, without claiming empirical validation, physical fire-spread simulation, autonomous deployment readiness, or universally superior performance from adding modalities.

* Corresponding author: zahra.mohemmi@flinders.edu.au

1. Introduction

Wildfire suppression requires committing scarce resources before environmental conditions and operational demand become fully known (Rossi et al, 2022). Accurate detection alone does not identify the best response: accessibility, threatened assets, supporting-base reliability, and responder exposure change the value of intervention (Abdalla, 2025). The sensing landscape is heterogeneous (Díaz-Varela et al, 2016). Thermal data indicate active combustion; RGB provides visual context; multispectral measurements characterize vegetation (Das et al, 2026); LiDAR represents structural geometry (Algadi et al, 2025); weather describes environmental forcing; gas and smoke observations indicate transported emissions; and terrain constrains access (Fan et al, 2025). These sources are not independent: Products can share instruments, acquisition conditions, or processing histories, making correlation and missingness central to uncertainty assessment.

Existing research already integrates parts of this chain. Chen et al. (2024) examine collaborative unmanned aerial vehicle (UAV) scheduling, Yuan et al. (2025) develop multimodal risk assessment, and Momeni et al. (2023) coordinate truck-and-drone coverage. The gap addressed here is narrower: separating the predictive contribution of fusion from the operational contribution of anticipating uncertain demand and capacity, using common evaluation scenarios and transparent baselines.

The research asks 5 questions. RQ1 concerns fusion's predictive and prioritization value. RQ2 examines uncertainty and missing observations. RQ3 compares stochastic and deterministic allocation. RQ4 examines marginal modality contributions. RQ5 concerns scale and transferability. The implemented experiment answers these conditionally within a synthetic benchmark; real-fire transferability remains untested.

A real-data investigation considered the 2025 Palisades and Eaton fires and relevant official archives. Documentation was located, but usable event observations were not acquired. Consequently, this article reports an explicitly authorized simulation rather than an empirical case study. The benchmark contains 1,600 sensing cells and 100 operational zones, with 10 independent replications. Bayesian conditioning represents correlated measurement errors, while stochastic programming separates pre-disruption staging from subsequent allocation. Neither method supplies uncertainty information automatically: likelihoods, scenario distributions, and operational constraints require justification. The contribution is a reproducible interface and a controlled comparison, not a field-ready dispatch system.

Accordingly, this study aims to develop and evaluate an integrated decision-support framework that connects multimodal wildfire sensing with uncertainty-aware suppression-resource planning. Specifically, the study seeks to: (i) assess whether combining thermal, RGB, multispectral, LiDAR, meteorological, gas/smoke, and terrain information improves wildfire-risk estimation relative to simpler sensing strategies; (ii) quantify how measurement uncertainty, sensor correlation, and missing observations affect predictive confidence; (iii) examine whether stochastic resource staging provides operational advantages over deterministic allocation when both demand and available capacity are uncertain; (iv) distinguish the predictive value of additional sensing modalities from their actual value for operational decisions; and (v) evaluate the robustness, scalability, and potential transferability of the proposed framework through ablation, sensitivity, and stress-testing analyses. In this way, the study focuses not only on improving wildfire-risk estimation, but also on determining whether improved information ultimately translates into better resource-allocation decisions under uncertainty.

The remainder of this article is organized as follows. Section 2 reviews the relevant literature on wildfire sensing, multimodal data fusion, uncertainty representation, suppression management, and

resource-allocation approaches, and positions the present study within the existing evidence. Section 3 presents the materials and methods, including the synthetic experimental setting, preprocessing and feature definitions, Bayesian multimodal fusion, zone-priority formulation, uncertainty-scenario generation, the 2-stage stochastic staging and recourse model, and the evaluation framework. Section 4 reports the predictive and operational results, including uncertainty calibration, resource-allocation performance, modality ablation, stress tests, sensitivity analyses, and computational behavior. Section 5 discusses the main findings, with particular attention to the distinction between predictive information and operational value, the role of coupled uncertainty, and model calibration. Section 6 outlines the managerial and operational implications for incident commanders, logistics planners, and technology decision-makers. Finally, Section 7 summarizes the main conclusions, limitations, and directions for future empirical research.

2. Literature review

2.1. Review method and evidence boundaries

All 3 supplied LeapSpace PDFs were read in full, comprising 11 pages. They were used as discovery aids rather than peer-reviewed evidence. The first emphasized real-time prioritization, the second mapped sensing and allocation literature, and the third identified which proposed methodological elements were directly supported and which were inferred. Their overlapping references were consolidated conceptually, and central records were checked against publisher, institutional, or official repository pages. Searches combined wildfire, UAV, thermal, RGB, multispectral, LiDAR, sensor fusion, prioritization, resource allocation, Bayesian inference, and stochastic optimization. This is a targeted narrative review, not a systematic review with a defensible exhaustive search denominator or a PRISMA flow count.

2.2. Detection, fusion, and environmental context

Jain et al. (2020) provide a broad account of machine-learning applications across wildfire science and management. Their scope helps distinguish ignition susceptibility, active-fire detection, spread estimation, and management decisions. These tasks do not have interchangeable labels. A pixel that later burns is not necessarily burning now, and a successfully suppressed zone can have high initial risk but low observed damage. Evaluation must therefore specify the target event and prediction horizon before selecting accuracy measures or constructing negative observations.

Chen et al. (2022) document drone-collected RGB/IR data and demonstrate the relevance of paired modalities for wildfire monitoring. Related deep architectures for semantic segmentation of imagery (Malekijoo & Fadaeieslam, 2019) address perception tasks. FLAME 2 offers a concrete route for eventual image-level development, with a repository record and an author-maintained implementation (Hopkins et al., 2022; Chen,). Its existence does not establish compatibility with the Los Angeles fires or prove availability of independent multispectral, LiDAR, and gas measurements. Jonnalagadda and Hashim (2024) examine an image-processing approach for UAV fire detection. Such work addresses perception, whereas the present experiment deliberately starts from standardized proxy measurements and does not train an image segmentation network.

Satellite thermal products introduce another scale distinction. VIIRS active-fire locations represent a detection footprint rather than a surveyed fire boundary, while the MODIS detection algorithm applies contextual screening and product-specific tests (Schroeder et al., 2014; Giglio et al., 2016). A 250 m decision lattice can be computationally useful without implying 250 m independent thermal observations. Resampling a coarser pixel creates more cells, not additional information. Landsat surface-temperature products have their own retrieval requirements and

missing-data mechanisms; cloud and emissivity-related limitations cannot be removed by simply interpolating the output (USGS,-b,-c).

Vegetation indices provide interpretable environmental covariates, but their definitions and temporal roles require care. NDVI indicates greenness; NBR emphasizes spectral contrast useful for burned-area assessment; and vegetation-water indices use near-infrared and shortwave-infrared information (Gao, 1996; USGS,-d,-e). Pre-event indices may characterize fuel conditions. Post-event differences may validate consequences, but using them as pre-decision predictors would leak the outcome. The field pipeline proposed here retains these distinctions instead of treating every derived spectral layer as another contemporaneous fire detector.

Structural measurements add information that imagery alone may not provide. Cook et al. (2013) describe an airborne system combining LiDAR, hyperspectral, and thermal instruments. Carbonell-Rivera et al. (2024) investigate relationships between spread and spectral or geometric features derived from UAV photogrammetric point clouds. These are different acquisition technologies even when both produce 3-dimensional representations. Terrain and canopy geometry can affect access and exposure as well as probability estimation, so removing LiDAR from a detector is not equivalent to removing its operational value from a complete incident model.

Yuan et al. (2025) offer direct evidence that multimodal wildfire risk assessment is an established research direction. Kemper and Kemper (2026) likewise describe a multi-sensor UAS monitoring and risk-assessment system. Neither title alone warrants attributing the exact Gaussian posterior or recourse formulation developed below to those authors. The present Bayesian specification is instead grounded in probability modeling (Gelman et al., 2013), with complementary literature motivating the application. This avoids attributing an inferred synthesis to studies that implemented different fusion mechanisms.

2.3. From sensing to operational decisions

SOScheduler explicitly couples sensing and operational priorities, making it an important comparator at the conceptual level (X. Chen et al., 2024). Chan et al. (2020) formulate sequential resource allocation under uncertainty, and Momeni et al. (2023) study coordination and routing for wildfire detection. Model-free reinforcement-learning approaches have also been applied to dynamic control under energy and delay constraints in IoT settings (Malekijou et al., 2023); the present formulation instead separates pre-disruption staging from recourse via stochastic programming. Lourenço et al. (2021) describe an integrated suppression-management decision-support system. Tavakol Sadrabadi et al. (2025) discuss an emergency-response architecture involving UAV swarms. Together these studies support the operational relevance of coordination, but they do not make sensing drones, suppression aircraft, crews, water, and retardant physically interchangeable.

Nguyen et al. (2026) examine historical daily prioritization in California and identify the importance of structures, fire size, and evacuations. Their findings caution against using a single fire-intensity score as a substitute for all management criteria. Our consequence weights therefore remain distinct from estimated fire probability. However, the weights in the simulation are analyst-defined spatial assumptions, not recovered California incident priorities. In addition, Attribute-weighting techniques for multi-attribute decision making (Malekijoo & Malekijoo, 2026) provide general machinery for eliciting relative importance; the weights used here remain analyst-defined spatial assumptions rather than recovered incident priorities. No claim is made that the synthetic staging solution reproduces agency decisions or the relative societal value of threatened assets.

Stochastic programming formalizes the separation between commitments made before uncertainty is resolved and actions taken afterward (Birge & Louveaux, 2011). Conditional value-at-

risk supplies a tractable representation of adverse-tail loss (Rockafellar & Uryasev, 2000 and Lonbar et al., 2024). These tools complement rather than replace forecasting. A perfectly specified optimizer can perform poorly with miscalibrated scenarios, while a highly accurate forecast can have little additional value when the same constraints determine every feasible decision. Proper scoring rules and structured validation help test the forecasting layer independently (Gneiting & Raftery, 2007; Roberts et al., 2017).

Table 1. Candidate field data sources. No observations were acquired; acquisition windows, completeness, and licenses require dataset-level verification.

Dataset/provider	Scale and cadence	Use and access
FIRMS / NASA	VIIRS 375 m; MODIS 1 km; Overpass dependent	Brightness temperature, radiative power. Public; archive may require Earthdata login
Sentinel-2 / Copernicus	10, 20, 60 m by band; Nominal 5-day constellation revisit	Visible reflectance; vegetation indices. Copernicus access terms; portal account may be needed
Landsat / USGS	30 m delivered thermal product; native thermal coarser; 16 days per satellite	Surface temperature; visible and infrared bands. Open USGS products; portal access conditions apply
3DEP / USGS	Project dependent point density; DEM resolution varies; Survey campaigns	Canopy structure; elevation. Public USGS data; inspect project metadata
ERA5 / C3S	0.25 degree distributed single-level grid; Hourly	Wind, temperature, humidity inputs. CDS account and dataset license acceptance
ISD / NOAA NCEI	Station points; Hourly or synoptic; variable completeness	Wind; temperature; dew point. Public archive; cite provider
AQS / U.S. EPA	Monitor points; Pollutant-specific; hourly/daily	PM2.5; CO; O3 where monitored. Public files; API key for API; retrospective availability
WFIGS / NIFC	Vector perimeter; mapping accuracy varies; Incident updates	Boundary and incident metadata. Inspect service metadata and terms
FLAME 2 / IEEE DataPort	UAV RGB/IR; frame geometry dependent; Video frames	Paired visible and infrared fire imagery. Check DataPort record license and download conditions

Table 2 consolidates the closest evidence without inventing unavailable sample counts. Its final row identifies the present contribution as a reproducible synthetic integration, not an empirical superiority claim. Feature-attribution methods such as those of Lundberg and Lee (2017) answer a different question from operational ablation: explaining a prediction does not establish the change in achievable loss after a decision model is reoptimized. Candidate data availability is summarized in Table 1 before the evidence comparison in Table 2.

Table 2. Comparative evidence map. Numeric sample sizes and performance estimates were not extracted where full details were unavailable; no missing values are inferred.

Study	Geospatial / structural inputs	Multimodal fusion	Risk / spread assessment	Suppression management	Resource allocation	Explicit uncertainty	Bayesian fusion	2-stage stochastic staging	CVaR	Modality ablation
Momeni et al. (2023)	✓	-	-	-	✓	-	-	-	-	-
Lourenço et al. (2021)	✓	-	-	✓	-	-	-	-	-	-
Chan et al. (2020)	-	-	-	✓	✓	✓	-	-	-	-
Chen et al. (2024)	-	-	-	✓	✓	-	-	-	-	-
Yuan et al. (2025)	-	✓	✓	-	-	-	-	-	-	-
Carbonell-Rivera et al. (2024)	✓	-	✓	-	-	-	-	-	-	-
Nguyen et al. (2026)	-	-	-	✓	-	-	-	-	-	-
Present study	✓*	✓*	✓*	✓*	✓*	✓	✓	✓	✓	✓

The gap addressed is therefore conditional and testable. Within this targeted review, the central opportunity is an explicit comparison of information value and uncertainty-aware staging value under the same synthetic environment, with correlated measurement errors and transparent observation masks. The contributions are: 1. A closed-form Bayesian interface with uncertainty propagation into zone demand. 2. A reproducible 2-stage capacity-staging model with adverse-tail control and operational bounds. 3. A controlled comparison revealing when predictive gains do not translate into material operational gains. 4. A data-availability and leakage-control protocol for eventual empirical replication. None of these contributions requires claiming that Bayesian fusion, UAV scheduling, or stochastic wildfire allocation is independently new.

3. Methodology

3.1. Real-data feasibility and the implemented sample

Table 1 records candidate sources for a future Palisades or Eaton application. FIRMS archive downloads may require authentication or an email workflow; an archive-access page is not itself a downloaded dataset (NASA, a-b). Sentinel-2 supplies optical bands at different spatial resolutions and does not provide thermal infrared measurements (Copernicus Data Space Ecosystem; European Space Agency, 2015). ERA5 is retrospective reanalysis, so using it in an operational replay requires explicit acknowledgment that it is not necessarily the forecast available at the decision timestamp (Copernicus Climate Change Service, 2018). NOAA station observations and EPA measurements require station-specific temporal coverage and quality checks (NOAA NCEI,; U.S. EPA,-a,-b).

Table 3. Core mathematical notation. Context distinguishes geometric coordinates from allocation variables.

Symbol	Definition	Unit or domain
i, j, t, k, b, s	Sensing cell; operational zone; decision time; modality; staging base; uncertainty scenario, respectively.	Indices; time in hours
$r_{it}, \mu_{it}, m_{it}, v_{it}$	Latent fire state; prior mean; posterior mean; posterior variance, respectively.	Log-odds; squared log-odds for variance
$\mathbf{z}_{it}, R, M_{ikt}$	Multimodal proxy vector; observation-error covariance matrix; channel-availability indicator, respectively.	Standardized latent units; squared units; binary indicator
$\hat{p}_{it}, s_{it}^2, P_{jt}$	Posterior expected fire probability; posterior variance of fire probability; unconstrained zone-urgency score, respectively.	Dimensionless probability and variance; weighted service units for urgency
$x_b, y_{bjs}, u_{js}, d_{js}$	Capacity staged at a base; service delivered from a base to a zone; unmet demand; scenario demand, respectively.	Abstract service units
w_j, e_{bs}, h_j	Zone consequence weight; scenario-specific base efficiency; zone exposure coefficient, respectively.	Dimensionless
$\ell_{bj}, \tau_{bj}, T_s, H_s$	Base-to-zone distance; travel time; aggregate service-weighted travel time; aggregate exposure, respectively.	km; min; service-unit min; exposure-weighted service units
$\eta, \xi_s, \alpha, \lambda$	Loss-tail threshold; excess loss above the threshold; CVaR confidence level; risk-aversion coefficient, respectively.	Loss units for threshold and excess; dimensionless confidence and coefficient

Gas and smoke are not synonymous. PM2.5 is particulate matter, not a gas, and a column retrieval or distant monitoring station is not automatically representative of surface exposure at a suppression zone. The EPA AQS API documents varying sample durations, making hourly alignment a data-screening task rather than a naming convention (U.S. EPA, 2020). Historical perimeter versions must be preserved; a current perimeter service does not guarantee recovery of every earlier boundary (NIFC,-a,-b). The 2025 post-fire LiDAR release is useful for consequence validation but excluded from pre-decision inference (USGS, 2025).

The implemented landscape is a local, non-georeferenced 10 km × 10 km square. Its coordinates are illustrative distances, not longitude or latitude. Each 250 m square is a sensing cell; each 1 km

square combines 16 sensing cells into an operational zone. Thus, there are 1,600 sensing cells, 100 operational zones, and 4 hypothetical staging bases. Times are 0, 3, 6, 9, 12, and 15 hours. Six snapshots span 15 hours between their first and last timestamps, not 18 hours of observed duration. No final real-world perimeter determines the domain.

$$\mathcal{D}_r = \{(i, t, \mathbf{z}_{it}, M_{it}, Y_{it}) : i = 1, \dots, 1600; t \in \{0, 3, 6, 9, 12, 15\}\}, \quad |\mathcal{D}_r| = 9600. \quad (1)$$

Equation (1) defines replication r 's cell-time sample. The index i identifies a cell, t is elapsed hours, $\mathbf{z}$ is its 7-channel proxy vector, M is its observation mask, and Y is its simulated binary fire label. Across 10 replications, 96,000 records are generated. Only the final 1,600 records per replication enter the principal predictive comparison. The workbook retains all 9,600 focal records and the remaining 14,400 final-time records, totaling 24,000 distinct rows. The unused earlier snapshots of the other replications are reproducible from the included code.

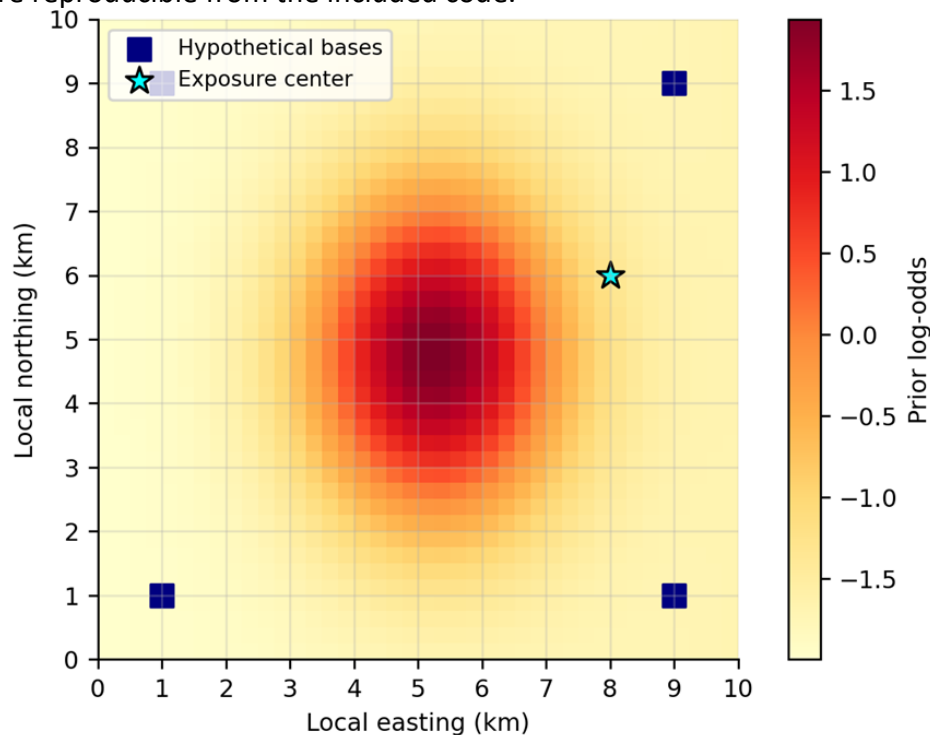


Figure 1. Synthetic study area in local km. The 100 km² domain contains 1,600 sensing cells, 100 operational zones, and 4 bases. Colors represent the simulator prior, not an observed fire perimeter. Source: authors' simulator, seed 101.

Figure 1 presents this synthetic geometry. Figure 2 presents proxy-value fields, not fabricated thermal photographs, RGB scenes, multispectral rasters, or LiDAR point clouds. No buffer around an actual fire is claimed. The large sample evaluates the computational interface; it does not create 96,000 independent wildfire incidents.

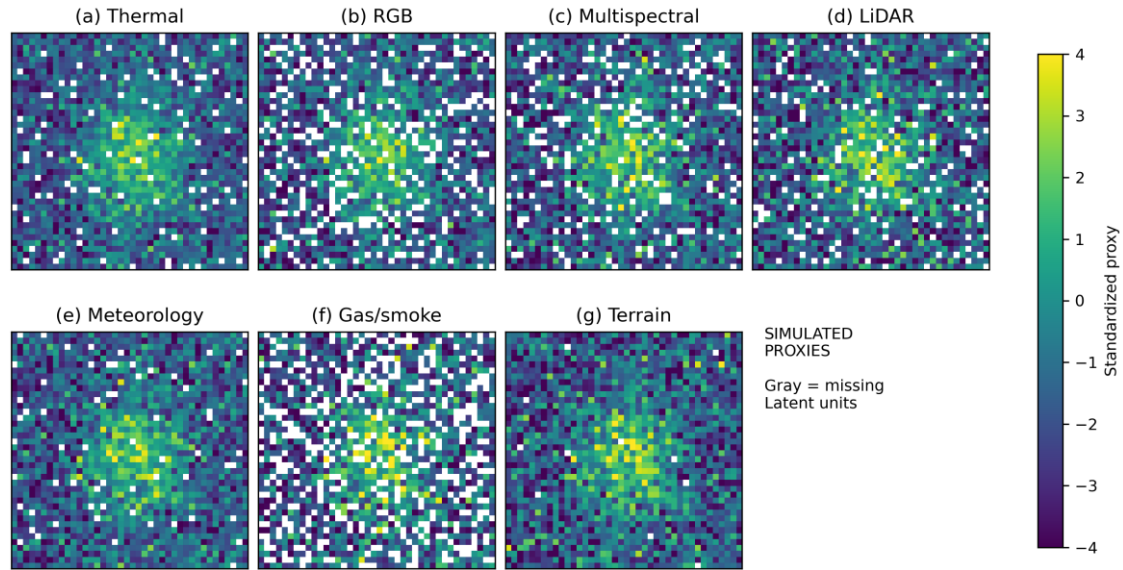


Figure 2. Simulated modality layers at the focal time: (a) thermal, (b) RGB, (c) multispectral, (d) LiDAR, (e) meteorology, (f) gas/smoke, and (g) terrain proxies. Gray denotes missing measurements; the common scale is dimensionless. Source: authors' simulator; no satellite or UAV imagery is displayed.

3.2. Preprocessing and feature definitions

A future empirical pipeline would validate coordinates, units, sensor flags, timestamps, duplicate observations, cloud masks, and coverage footprints before aggregation. Common spatial support would be defined using a projected coordinate system suitable for the chosen region. Temporal joins would be backward-looking, retaining observation age rather than filling gaps with future values. Spectral indices would be calculated from documented bands after their scaling factors are applied. Terrain slope and aspect would be derived from a verified pre-event elevation product. Gas or particulate interpolation would retain predictive variance and station distance.

$$n_{ikt}^+ = \frac{f_{ikt} - L_k}{U_k - L_k}. \quad (2)$$

Equation (2) maps an empirical benefit-type feature f for cell i , channel k , and time t to a dimensionless score using training-only lower and upper bounds L and U . Larger values indicate greater benefit under the specified decision criterion. Bounds must differ; constant features are excluded or assigned a documented neutral value. Values outside training bounds are flagged rather than silently treated as familiar observations.

$$n_{ikt}^- = \frac{U_k - f_{ikt}}{U_k - L_k}. \quad (3)$$

Equation (3) reverses the direction for a cost-type feature, such as travel impedance, using the same bounds and units as Equation (2). These transformations belong to the empirical integration specification, not the executed Gaussian benchmark, whose proxies already share standardized latent units. Neither transformation converts an ordinal expert score into a physically calibrated probability. Table 4 distinguishes possible field variables from the simulated channel representation.

Table 4. Candidate field variables and units. Executed channels are dimensionless noisy proxies, not these physical measurements.

Modality	Field variables	Units	Expected role
Thermal	Temperature; radiative power	K; MW	Active-combustion evidence
RGB	Visible smoke/flame; context	Reflectance; probability	Visibility and exposure context
Multispectral	NDVI; NBR; NDMI	Dimensionless	Fuel condition; timing-dependent
LiDAR	Canopy height; density	m; proportion	Fuel structure; access
Meteorology	Wind; humidity; temperature	m/s; %; degrees C	Spread forcing; capacity disruption
Gas/smoke	PM2.5; CO; O3	Micrograms/m ³ ; ppm or ppb	Transport-dependent proxy
Terrain	Elevation; slope; road distance	m; degrees; km	Access and terrain forcing

3.3. Synthetic observation model and Bayesian fusion

The synthetic prior contains a moving concentration of fire propensity. Its center shifts gradually across the landscape, but the process is not a physical combustion or spread simulator. In particular, it does not enforce front continuity, fuel consumption, or persistence of burning. Richards (1990) offers an established physical-growth modeling reference for future development; that model is not implemented here.

$$r_{it} = \mu_{it} + \epsilon_{it}, \quad \epsilon_{it} \sim N(0,1), \quad Y_{it} | r_{it} \sim \text{Bernoulli}(g(r_{it})), \quad g(r) = \frac{1}{1 + e^{-r}}. \quad (4)$$

Equation (4) defines dimensionless latent log-odds r , a known synthetic prior mean μ , independent standard-normal latent disturbances ϵ , and binary outcomes Y . The logistic function g converts log-odds to probability.

$$\begin{aligned} \mathbf{z}_{it} &= \mathbf{1}r_{it} + \boldsymbol{\epsilon}_{it} \\ \boldsymbol{\epsilon}_{it} &\sim N(\mathbf{0}, \mathbf{R}) \\ R_{kk} &= \sigma_k^2 \\ R_{kl} &= 0.25\sigma_k\sigma_l \quad (k \neq l). \end{aligned} \quad (5)$$

Equation (5) gives every proxy the same latent target but heterogeneous measurement error. The vector $\mathbf{1}$ has 7 entries; $\mathbf{R}$ is the observation-error covariance matrix. Channel standard deviations, in the stated modality order, are 0.65, 1.00, 1.15, 1.30, 1.00, 1.40, and 1.25 latent units. The common off-diagonal correlation is 0.25. Precision and covariance, not subjective importance weights, determine statistical reliability. The resulting ranking of modalities is conditional on these assumptions.

$$p(r_{it} | \mathbf{z}_{it,O}) \propto p(\mathbf{z}_{it,O} | r_{it})p(r_{it}), \quad O = \{k: M_{ikt} = 1\}. \quad (6)$$

Equation (6) applies Bayes' rule to the observed subset O . M equals 1 when a channel is available. Missing measurements are marginalized rather than replaced by zero. The Gaussian specification permits closed-form conditioning and explicit cross-channel dependence, following standard Bayesian modeling principles (Gelman et al., 2013).

$$v_{it} = (1 + \mathbf{1}_O^\top \mathbf{R}_O^{-1} \mathbf{1}_O)^{-1}. \quad (7)$$

Equation (7) gives posterior variance v in squared latent units. $\mathbf{R}$ subscript O is the covariance submatrix for observed channels, and the leading 1 is prior precision because prior variance equals 1. When no channels are observed, v equals 1. Removing observations cannot reduce posterior variance under the correctly specified model.

$$m_{it} = v_{it}(\mu_{it} + \mathbf{1}_O^\top \mathbf{R}_O^{-1} \mathbf{z}_{it,O}). \quad (8)$$

Equation (8) gives posterior mean m in latent units. The observed vector contributes through its joint precision structure rather than through independent likelihood multiplication. If all channels are

absent, m reverts to the prior mean. This is a controlled fallback, not evidence that the missing channels have no value.

$$\hat{p}_{it} = \int g(r)N(r; m_{it}, v_{it}) dr. \quad (9)$$

Equation (9) converts the Gaussian latent posterior into a fire probability. It is evaluated with 20-node Gauss-Hermite quadrature. Applying the logistic function only to the posterior mean would ignore nonlinear averaging. Probability is dimensionless and predicts the simulated current-state label, not an empirically validated spread horizon.

$$s_{it}^2 = \int g(r)^2 N(r; m_{it}, v_{it}) dr - \hat{p}_{it}^2. \quad (10)$$

Equation (10) measures uncertainty about the latent probability, distinct from Bernoulli outcome variance. The same quadrature computes its first term. This epistemic component is propagated to operational-zone probabilities; random future demand variability is introduced separately.

$$\text{Var}(r | z_O) = E[\text{Var}(r | z_O, z_M) | z_O] + \text{Var}(E[r | z_O, z_M] | z_O). \quad (11)$$

Equation (11) expresses the law of total variance for missing measurements z subscript M . The second term is nonnegative and explains why unobserved information must not produce artificial certainty. In the benchmark, channel missingness is independent with probabilities 0.05, 0.20, 0.12, 0.08, 0.03, 0.25, and 0.01. Smoke-dependent missingness and informative sensor failure remain unmodeled field concerns.

3.4. Zone priorities and uncertainty scenarios

$$\bar{p}_{jt} = \frac{1}{16} \sum_{i \in C_j} \hat{p}_{it}, \quad s_{jt}^2 = \frac{1}{256} \sum_{i \in C_j} s_{it}^2. \quad (12)$$

Equation (12) aggregates the 16 cells in operational zone j , denoted C subscript j . The variance formula assumes conditional independence across cells, consistent with the benchmark's latent residual construction. A field model with shared spatial residuals must add cross-cell covariance terms. The aggregation can attenuate predictive differences that matter at individual cells.

$$P_{jt} = w_j E[d_{jt} | \mathcal{D}_t], \quad w_j = 1 + 3 \exp \left[-\frac{(x_j - 8)^2 + (y_j - 6)^2}{8} \right]. \quad (13)$$

Equation (13) defines an unconstrained urgency score P from expected service demand d and consequence weight w . Zone-center coordinates x and y are in km; the denominator 8 has units km^2 . The hypothetical exposure center is at (8,6). This score is not a dispatch instruction: access, available capacity, and exposure constraints determine feasible action. The weights are declared preference assumptions, not measurements of lives or money.

$$d_{js} = (0.3 + 2.5\tilde{p}_{js})a_s \exp(-0.25^2/2 + 0.25v_{js}), \quad \tilde{p}_{js} = \min\{1, \max\{0, \bar{p}_j + s_j Z_{js}\}\}. \quad (14)$$

Equation (14) generates scenario demand in abstract service units. Z and ν are independent standard-normal draws; the clipped Gaussian approximates the zone probability posterior using Equation (12). The multiplier a equals 0.85, 1.10, or 1.50 under mild, intermediate, or severe conditions. The lognormal multiplier has mean 1. These values are simulation assumptions, not fuel-consumption estimates. Test demand uses the realized latent zone probabilities without posterior sampling.

$$\Pr(W_s = 0,1,2) = (0.5, 0.3, 0.2), \quad p_s = 1/S. \quad (15)$$

Equation (15) specifies the categorical weather-regime probabilities and equal empirical weights for S independent Monte Carlo scenarios. The shared regime affects every zone's demand and the base-efficiency vector, preserving demand-capacity dependence. Efficiencies are (1,1,1,1), (1,0.9,1,0.75), and (1,0.7,0.95,0.35). Planning uses $S = 20$ and evaluation uses 50 fresh scenarios. Posterior parameter uncertainty is not separately estimated because the benchmark parameters are fixed and known.



Figure 3. Implemented inference-to-decision framework. Observed synthetic channels enter Bayesian conditioning; posterior moments inform scenarios; staging precedes scenario-dependent recourse. Source: authors' methodological design.

3.5. Staging and recourse model

The implemented decision unit is divisible service capacity. It is not an individual firefighter, drone, aircraft, liter of water, or liter of retardant. This aggregation makes the mechanism transparent but precludes equipment-specific feasibility claims. A future typed-resource model would preserve separate inventories, payloads, airspace restrictions, crew qualifications, replenishment, and movement times. The current model has 4 bases at local coordinates (1,1), (9,1), (1,9), and (9,9), each with a maximum staging capacity of 60 units.

$$\mathcal{X} = \left\{ \mathbf{x}: 0 \leq x_b \leq 60, \sum_{b=1}^4 x_b \leq B \right\}, \quad B = 120. \quad (16)$$

Equation (16) defines first-stage allocations $\mathbf{x}$ to bases before the environmental regime is revealed. B is total available capacity. These variables are nonanticipative: every planning scenario shares the same $\mathbf{x}$. Fractional values are permitted because the model allocates abstract divisible service, not indivisible crews.

$$Q(\mathbf{x}, \omega_s) = \min_{\mathbf{y}_s, \mathbf{u}_s} \left\{ \sum_j w_j u_{js}: (\mathbf{y}_s, \mathbf{u}_s) \in \mathcal{Y}(\mathbf{x}, \omega_s) \right\}. \quad (17)$$

Equation (17) defines second-stage recourse after scenario ω_s is revealed. The variable y_{bs} is service delivered from base b to zone j ; u_j is unmet demand. The feasible set $\mathcal{Y}$ incorporates Equations (22) to (24) and the operational bounds below. Complete scenario revelation is an optimistic assumption, not a claim of perfect operational forecasting.

$$C(\mathbf{x}) = c_x \sum_b x_b, \quad c_x = 0.03. \quad (18)$$

Equation (18) gives a normalized staging charge in objective units. It discourages unnecessary capacity when demand can already be met. It is not an observed financial cost. The implemented objective does not estimate real aircraft, crew, water, or retardant expenditure.

$$T_s = \sum_{(b,j) \in A} \tau_{bj} y_{bjs}, \quad \tau_{bj} = 3 + 2\ell_{bj}, \quad T_s \leq 2200. \quad (19)$$

Equation (19) defines aggregate service-weighted travel time, measured in service-unit minutes. Distance ℓ is Euclidean km; the coefficient 2 has units minutes/km and the intercept is 3 minutes. A scenario-wise travel budget supplies an epsilon-style bound. Mean reported response time divides T by served units and is descriptive, not independently minimized in this implementation. No real road network is represented.

$$L_s = \sum_j w_j u_{js}. \quad (20)$$

Equation (20) defines priority-weighted unmet demand, the principal loss criterion. It measures service deficit in consequence-weighted units. It must not be interpreted as hectares burned, fatalities, structures lost, or money saved. Optimization minimizes this loss subject to operational feasibility.

$$H_s = \sum_{(b,j) \in A} h_j y_{bjs} \leq 60, \quad h_j = 0.1 + 0.35y_j/10 + 0.2\exp\left[-(x_j - 5)^2/3\right]. \quad (21)$$

Equation (21) supplies a synthetic exposure budget. Here x subscript j and y subscript j are zone coordinates, not staging variables. The score h is dimensionless and aggregate H has exposure-weighted service units. The value 60 is an analyst-selected bound, not a firefighter safety threshold. More delivered service can increase H even while loss decreases.

$$\sum_{j:(b,j) \in A} y_{bjs} \leq e_{bs} x_b. \quad (22)$$

Equation (22) limits scenario-specific service from each base to staged capacity times its efficiency e . The efficiencies share the weather regime with demand, so severe conditions simultaneously increase need and reduce availability. This relationship is the primary source of operational uncertainty in the benchmark.

$$\sum_{b:(b,j) \in A} y_{bjs} + u_{js} = d_{js}, \quad y_{bjs} \geq 0, \quad u_{js} \geq 0. \quad (23)$$

Equation (23) conserves demand exactly and permits multiple bases to serve a zone. Nonnegative unmet demand guarantees feasible recourse even when no service can be delivered. Divisible assignments should not be mistaken for binary vehicle routing. High-priority demand coverage is derived from served demand within a fixed evaluation set of 20 high-priority zones.

$$A = \{(b,j): \ell_{bj} \leq 8\}, \quad y_{bjs} = 0 \quad \text{if } (b,j) \notin A. \quad (24)$$

Equation (24) forbids base-zone assignments beyond an 8 km synthetic access radius, leaving 276 admissible arcs. It is a simplified accessibility rule. Real applications would use time-dependent routes and hard exclusion zones. Scenario-wise capacity, travel, and exposure limits are enforced directly; no probabilistic safety guarantee is asserted beyond the sampled scenarios.

$$\text{CVaR}_\alpha(L) = \min_{\eta, \xi_s \geq 0} \left\{ \eta + \frac{1}{1-\alpha} \sum_s p_s \xi_s : \xi_s \geq L_s - \eta \right\}. \quad (25)$$

Equation (25) represents conditional value-at-risk using an auxiliary threshold η and nonnegative excess losses ξ , following Rockafellar and Uryasev (2000). Alpha equals 0.90, so empirical evaluation averages the largest 5 losses among 50 equally weighted scenarios. This is a noisy tail estimate, not a reliable estimate of extremely rare disasters.

$$\min_{\mathbf{x} \in \mathcal{X}, \mathbf{y}, \mathbf{u}} \left[C(\mathbf{x}) + \sum_s p_s L_s + \lambda \text{CVaR}_{0.90}(L) \right], \quad \lambda = 0.3. \quad (26)$$

Equation (26) combines expected loss with an explicit risk preference and a normalized staging charge, subject to all scenario constraints. Lambda is not a learned sensor weight: it expresses the decision maker's assumed aversion to the loss tail. The experiment tests values from 0 to 2. Travel and exposure enter as bounds rather than arbitrary commensurate penalties. The plotted held-out trade-off is a sensitivity curve, not a certified Pareto frontier.

3.6. Baselines, metrics, and replication

There are 6 operational policies: thermal-only deterministic staging, equal-weight deterministic staging, Bayesian deterministic staging, equal staging across bases, thermal-only stochastic staging, and Bayesian stochastic staging. Deterministic policies optimize a mean scenario generated from their own forecasts. The equal-weight estimator averages observed proxy values without Bayesian prior shrinkage. All policies receive the same access network, capacity budget, consequence weights, and realized test scenarios. All are evaluated using the same loss-minimizing recourse solver. Thus the equal-staging baseline removes optimized staging, not every optimization component. It is a stronger and fairer comparison than deliberately inefficient dispatch.

$$\Delta_k^{\text{pred}} = \text{BS}_{-k} - \text{BS}_{\text{full}}. \quad (27)$$

Equation (27) defines predictive modality ablation using Brier score BS. Each channel k is omitted before recomputing the posterior. Positive values indicate worse prediction after removal. Operational ablation is separately represented by full versus thermal stochastic staging; the experiment does not claim 7 separate end-to-end operational ablations.

$$\text{VSS} = \text{EEV} - \text{RP}. \quad (28)$$

Equation (28) gives the value of the stochastic solution for the risk-neutral model. RP is the optimized stochastic objective; EEV evaluates the mean-scenario staging decision over the same scenarios with optimal recourse. Both include the identical staging charge. These are in-sample decision-theoretic diagnostics, distinct from independent test improvements (Birge & Louveaux, 2011).

$$\text{EVPI} = \text{RP} - \text{WS}. \quad (29)$$

Equation (29) gives expected value of perfect information. WS is the average objective obtained when staging is separately optimized after each scenario is known. The ordering $\text{WS} \leq \text{RP} \leq \text{EEV}$ is checked numerically. This construction is used only with $\lambda = 0$, avoiding an invalid comparison of differently defined risk measures.

$$\text{BS} = \frac{1}{N} \sum_{i=1}^N (\hat{p}_i - Y_i)^2. \quad (30)$$

Equation (30) defines the Brier score for N evaluation records and binary labels. Lower values indicate better probabilistic prediction under this proper scoring rule (Gneiting & Raftery, 2007). area under the receiver operating characteristic curve (AUROC), precision-recall area, F1 at threshold 0.5, latent-state RMSE, and 95% posterior interval coverage provide complementary diagnostics. Labels and latent truth are simulator outputs, never retrieved field observations.

$$\bar{\delta} \pm t_{0.975,9} \frac{s_{\delta}}{\sqrt{10}}. \quad (31)$$

Equation (31) constructs a paired 95% confidence interval from 10 replication-level differences. The mean difference is $\bar{\delta}$, its sample standard deviation is s_{δ} , and t is a Student quantile with 9 degrees of freedom. Independent simulation seeds, not cells or test scenarios, are the inferential units. These exploratory intervals are not multiplicity-adjusted confirmatory tests.

Parameters are fixed before evaluation; no statistical model is fitted, so claiming spatial cross-validation would be misleading. Final-time evaluation and independent scenario seeds prevent direct reuse of evaluated observations as fitted labels. Future learned models require spatially blocked, temporally ordered validation following the principles reviewed by Roberts et al. (2017). Numerical solutions use SciPy's HiGHS interface, with solver success and demand-balance residuals checked (Huangfu & Hall, 2018; SciPy Developers,).

4. Results

4.1. Sample integrity and observation availability

The focal replication contains 9,600 synthetic cell-time records, of which 2,632 are fire-positive and 6,968 fire-negative. This corresponds to a simulated positive frequency of 27.4%, not an observed wildfire prevalence. The main final-time comparison contains 16,000 records across 10 independent realizations. The accompanying workbook stores 24,000 input rows without duplicates, including the complete focal timeline and every record used in the principal predictive comparison. It also retains 3,000 policy-scenario evaluation rows, 60 operational summaries, 30 predictive summaries, 70 predictive-ablation records, and staging decisions.

Figure 4a shows realized focal missingness of 5.00% for thermal, 19.24% for RGB, 11.76% for multispectral, 8.10% for LiDAR, 2.84% for meteorology, 24.81% for gas/smoke, and 0.99% for terrain. These percentages arise from the declared random mask. They are not estimates of the reliability of the named sensing technologies.

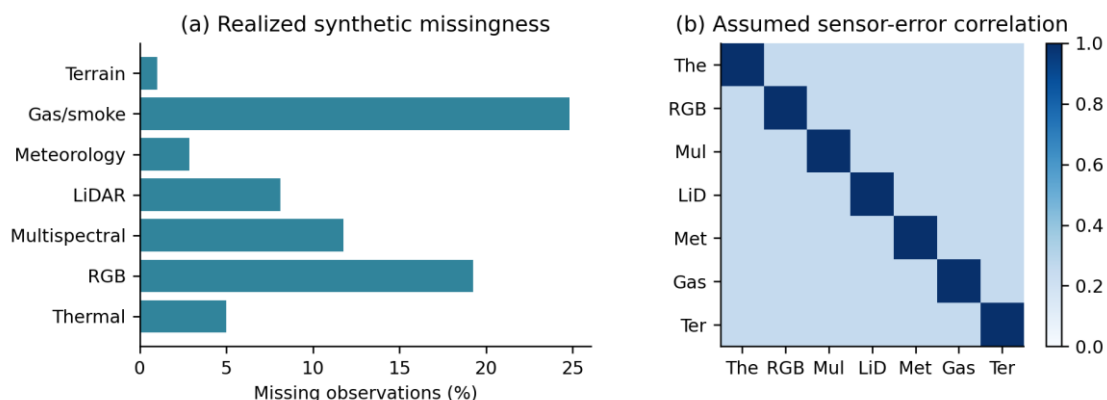


Figure 4. Observation quality: (a) realized focal-panel missingness and (b) assumed sensor-error correlation/covariance structure. The matrix is a generator assumption, not an empirical correlation estimate. Source: simulated records and parameters.

4.2. Predictive performance and uncertainty calibration

Table 5 reports final-time predictive performance with replication-level confidence intervals. Full fusion yields mean AUROC 0.7763 compared with 0.7704 for thermal-only estimation and 0.7655 for equal weighting. The corresponding Brier scores are 0.1565, 0.1585, and 0.1616. These improvements are modest for binary classification because event labels retain Bernoulli variability even when latent risk is known. The latent-state RMSE improves more substantially, from 0.5806 for thermal-only inference to 0.4991 for full fusion, a 14.0% reduction. The equal-weight estimator achieves $F1 = 0.5186$, exceeding the full model's 0.4980 at threshold 0.5. Therefore, the full model does not dominate every metric.

Table 5. Predictive performance: mean $\pm$ 95% replication-level confidence-interval half-width; 10 replications, 1,600 final-time cells each.

Estimator	AUROC	PR-AUC	Brier score	Latent RMSE
thermal	0.7704 $\pm$ 0.0123	0.6057 $\pm$ 0.0214	0.1585 $\pm$ 0.0041	0.5806 $\pm$ 0.0095
equal	0.7655 $\pm$ 0.0093	0.6007 $\pm$ 0.0183	0.1616 $\pm$ 0.0032	0.6823 $\pm$ 0.0120
full	0.7763 $\pm$ 0.0104	0.6159 $\pm$ 0.0192	0.1565 $\pm$ 0.0037	0.4991 $\pm$ 0.0091

Under the correctly specified generator, average 95% interval coverage is 94.95% for full fusion and 95.08% for thermal-only inference. This is a useful implementation check, but it is expected when the estimator knows the generating prior and covariance.

4.3. Operational performance and spatial deficits

Table 6 separates expected loss, adverse-tail loss, travel, capacity utilization, and exposure. Bayesian stochastic staging produces mean weighted unmet demand 6.9794, compared with 10.7394 for Bayesian deterministic staging. The relative reduction is 35.0%. The paired absolute reduction is 3.7600, with a 95% interval of approximately [1.5277, 5.9924]. CVaR decreases from 52.8394 to 39.4928, a 25.3% reduction; its paired interval for the absolute reduction is approximately [8.8762, 17.8170]. These are computed synthetic service losses, not estimates of avoided real-world damage.

Table 6. Operational results over independent test scenarios. Loss and CVaR are synthetic weighted service deficits; response is descriptive. Values are means over 10 replications.

Policy	Loss $\pm$ CI	CVaR $\pm$ CI	Response (min)	Utilization
Thermal deterministic	10.77 $\pm$ 3.66	52.89 $\pm$ 5.31	13.07	88.2%
Equal-weight deterministic	9.06 $\pm$ 3.04	47.78 $\pm$ 4.83	12.96	86.0%
Bayes deterministic	10.74 $\pm$ 3.64	52.84 $\pm$ 5.24	13.07	88.2%
Bayes equal staging	11.65 $\pm$ 2.71	63.84 $\pm$ 1.83	12.12	84.3%
Thermal stochastic	6.98 $\pm$ 1.67	39.51 $\pm$ 1.64	12.70	82.3%
Bayes stochastic	6.98 $\pm$ 1.67	39.49 $\pm$ 1.64	12.70	82.3%

The contrast with thermal-only stochastic staging is much smaller. Its mean loss is 6.9841 and CVaR is 39.5101. Full fusion's mean-loss advantage is 0.0047, with paired interval approximately [-0.0041, 0.0135]. The interval includes 0. Figure 5 compares spatial unmet demand under 4 policies using a common scale. These are deficit maps after recourse, not observed burn maps or direct fire-probability maps. Figure 7 summarizes the policy comparisons. Expected demand coverage for the 20 highest-priority zones is essentially 100% for all policies at the default budget. This saturated metric does not discriminate among strategies. Improvements occur mainly in the remaining demand and in adverse scenarios. Reporting coverage alone would obscure that distinction and exaggerate the apparent equivalence of policies.

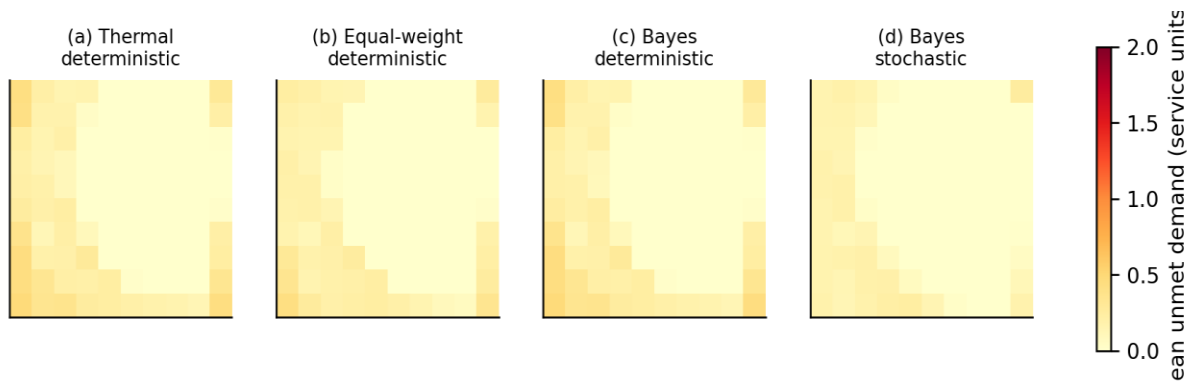


Figure 5. Spatial unmet-demand patterns under 4 policies. Common scales permit comparison of thermal, deterministic, equal-weight deterministic, and stochastic Bayesian decisions.

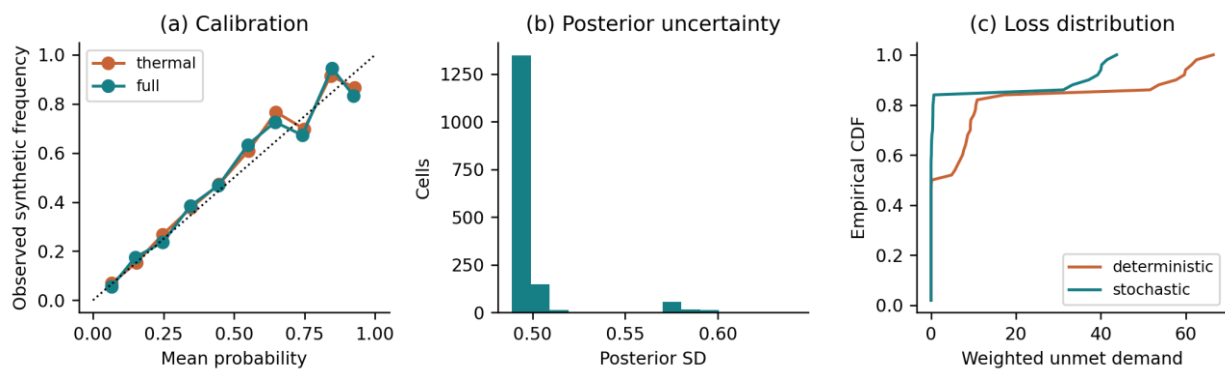


Figure 6. Uncertainty diagnostics: calibration, posterior uncertainty, and operational loss distributions. Probabilities refer to simulated labels; tail losses use independent test scenarios. Source: computed simulation outputs.

Stochastic staging also changes operational side effects. Mean travel is approximately 12.70 minutes per served unit, versus 13.07 for Bayesian deterministic staging, but these times are descriptive outcomes under a travel budget rather than uniquely optimized route lengths. Exposure increases from approximately 33.89 to 34.72 weighted service units. Capacity utilization decreases from 88.18% to 82.34%, reflecting reserve capacity that protects against adverse conditions. Thus lower utilization is not automatically inefficiency, and reduced unmet demand is not automatically improved responder safety.

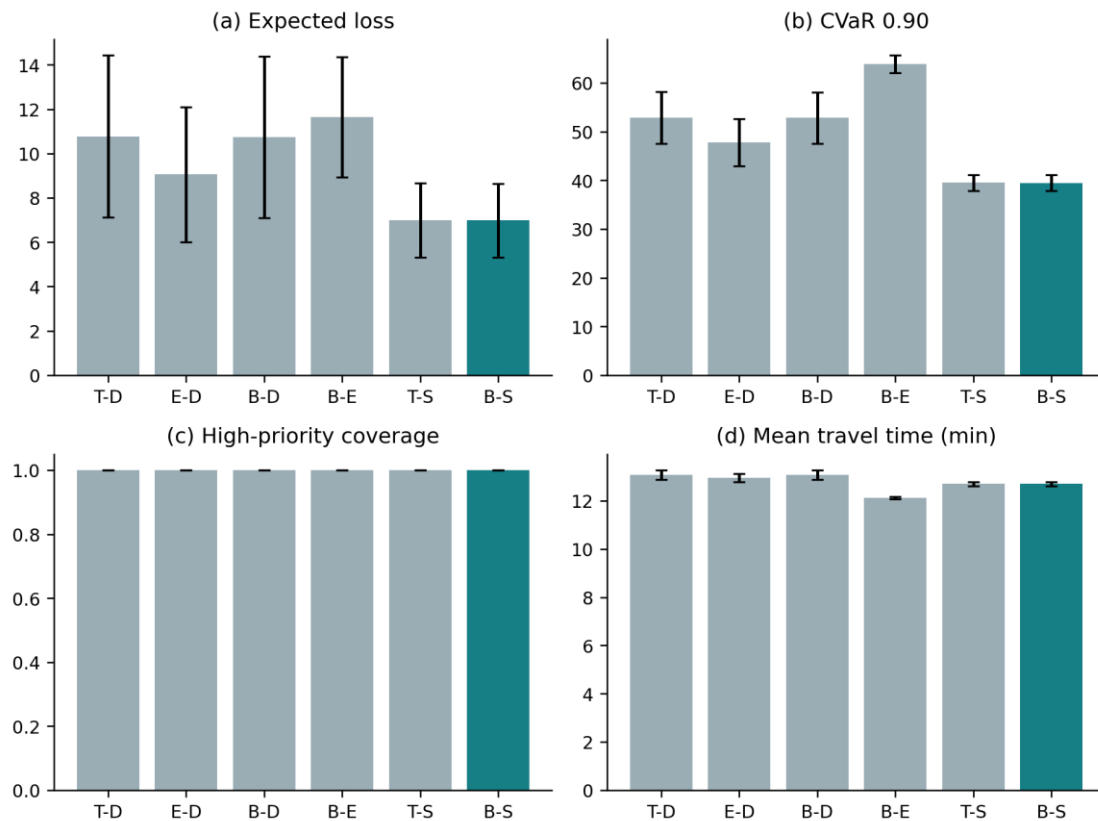


Figure 7. Operational comparisons of mean weighted loss, tail loss, high-priority coverage, and response time. Error bars represent replication-level 95% confidence intervals. Policy labels: T-D, thermal deterministic; E-D, equal-weight deterministic; B-D, Bayesian deterministic; B-E, Bayesian equal staging; T-S, thermal stochastic; B-S, Bayesian stochastic. Saturated coverage is retained. Source: 10 computed replications.

4.4. Information value, ablation, and stress tests

For the focal risk-neutral planning sample, RP is 9.6686, EEV is 14.0229, and WS is 8.8118. VSS is therefore 4.3543 and EVPI is 0.8568 in common objective units. Table 7 and Figure 9a summarize predictive ablation. Removing thermal observations increases Brier score by 0.00252 on average; removing meteorology increases it by 0.000445. The remaining effects are small, and several intervals include 0. Removing LiDAR has a slightly negative sample mean effect, about -0.000008, which is indistinguishable from simulation variability.

Table 7. Predictive ablation and sensitivity. Ablation uses 10 replications; sensitivity is focal-seed exploratory analysis.

Removed channel / change	Brier increase $\pm$ CI / loss	Interpretation
Thermal	0.002520 $\pm$ 0.000906	Prediction ablation
RGB	0.000119 $\pm$ 0.000241	
Multispectral	0.000079 $\pm$ 0.000209	
LiDAR	-0.000008 $\pm$ 0.000056	
Meteorology	0.000445 $\pm$ 0.000304	
Gas/smoke	0.000061 $\pm$ 0.000063	
Terrain	0.000000 $\pm$ 0.000117	
Capacity 60	48.760	Mean weighted unmet demand
Capacity 90	18.127	

Capacity 150	1.987	
Capacity 180	0.000	

The focal missingness stress test deletes an additional 20%, 50%, or 80% of entries independently. Brier scores rise to 0.1598, 0.1614, and 0.1702, respectively. At 80% extra deletion, latent RMSE reaches 0.8084. Posterior coverage remains near its nominal level because the estimator correctly recognizes the reduced information. In contrast, doubling actual measurement noise without updating the covariance produces RMSE 0.9043 and interval coverage 72.13%. Adding a shared 0.75-unit bias to thermal, RGB, and gas/smoke channels reduces coverage to 87.31%. Unrecognized degradation is therefore more dangerous to confidence calibration than openly missing observations.

4.5. Resource sensitivity and computational behavior

Figure 8 shows the focal held-out risk-preference sweep. Increasing lambda from 0 to 2 changes mean loss from 6.1616 to 6.4710 while reducing CVaR from 41.1948 to 40.7325. The trade-off is limited, and the risk-neutral stochastic policy already captures much of the benefit relative to deterministic staging. It would therefore be misleading to attribute the entire 35.0% improvement to CVaR. Scenario modeling and risk aversion are distinct interventions.

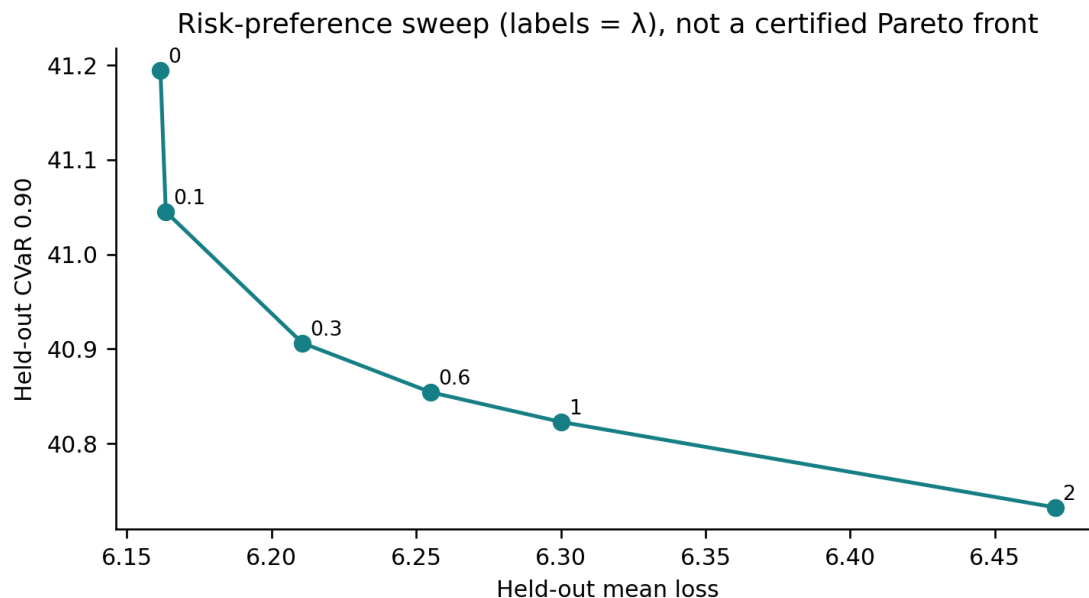


Figure 8. Held-out mean-loss versus tail-loss trade-off across risk preferences. This is a focal sensitivity curve, not a certified Pareto frontier. Source: independently evaluated stochastic solutions.

At resource budgets of 60, 90, 150, and 180 units, focal mean loss is 48.7600, 18.1274, 1.9873, and 0, respectively. At 60 units, utilization reaches 100% while high-priority coverage falls to 98.01%. At 180 units, the sampled demands are fully served, but this says nothing about unsampled extremes. Figure 9b shows the pronounced capacity effect. This reinforces the distinction between better information and sufficient response capability: sensing cannot remove a binding shortage of physical resources.

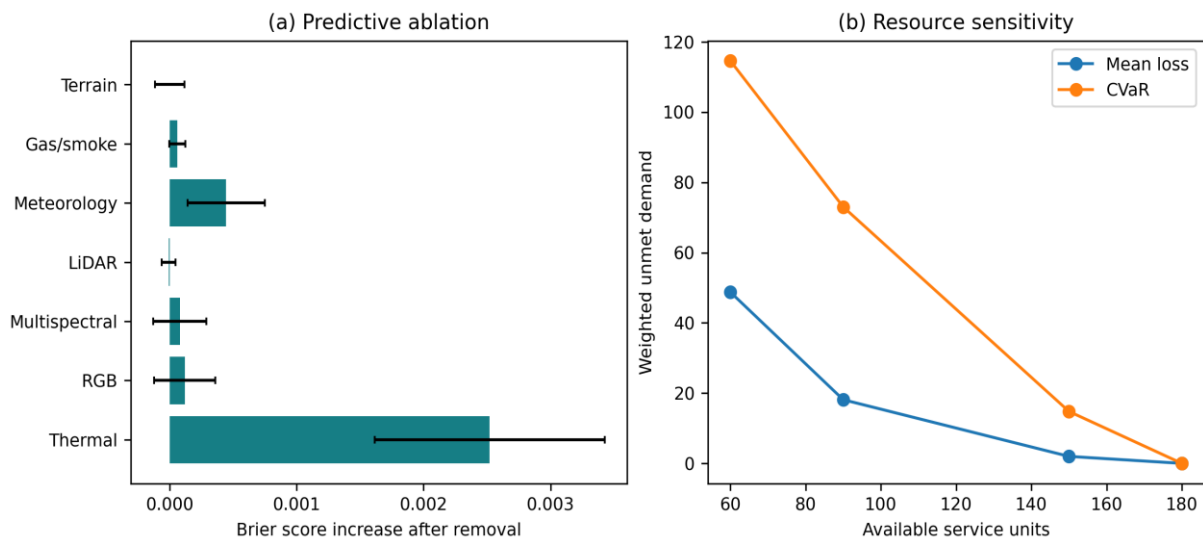


Figure 9. Modality ablation and capacity sensitivity. Predictive ablation intervals use 10 replications; capacity outcomes use the focal replication. Units differ between panels. Source: computed simulation outputs.

The full stochastic planning solve averages approximately 0.127 seconds in this environment at 20 scenarios. Focal planning times increase to approximately 0.725 seconds at 80 scenarios.

5. Discussion

5.1. Predictive information is not operational value

The most important finding is the separation between predictive and operational gains. Full fusion improves probability estimates and latent-state reconstruction, yet its operational advantage over thermal-only stochastic staging is extremely small. This is not a contradiction. The allocation problem aggregates local risks, imposes shared resource limits, and emphasizes scenario-dependent base availability. Once a thermal-informed forecast places sufficient capacity in the useful bases, incremental corrections to individual probabilities may not alter the staging decision. The decision map from probabilities to allocations is therefore less sensitive than the prediction map from observations to probabilities.

5.2. The dominant role of coupled uncertainty

In severe synthetic regimes, demand rises while particular bases lose efficiency. A deterministic mean scenario smooths this dependence and can position too much capacity where it performs poorly during difficult conditions. Stochastic staging retains the joint structure. This explains why thermal stochastic and full stochastic policies both substantially outperform their deterministic counterparts. The findings support the use of scenario-dependent recourse in the sense formalized by Birge and Louveaux (2011), while remaining conditional on the assumed regime probabilities and capacity multipliers.

The study also distinguishes risk aversion from stochasticity. Increasing the CVaR weight moderately improves the tail but slightly worsens mean loss in the focal sweep. The risk-neutral stochastic solution is already competitive. Therefore, CVaR should be presented as a policy preference for distributional trade-offs, not as a universally beneficial correction. A decision maker might reasonably prefer another point after considering reserve cost, acceptable exposure, or service equity. The sensitivity curve supports deliberation but cannot identify the socially optimal risk preference.

5.3. Calibration, missingness, and operational confidence

Correctly recognized missingness and unrecognized degradation produce different failure modes. When observations disappear, the posterior widens and falls back toward the prior. When observations remain present but are noisier or systematically biased than assumed, the model can be confidently wrong. The difference between near-nominal coverage under deletion and 72.13% coverage under doubled noise makes this distinction concrete. A practical system therefore needs ongoing checks of residual distributions, calibration drift, sensor health, and timestamp age, rather than relying solely on a fixed initial accuracy assessment.

6. Managerial and operational implications

For incident commanders, the framework separates 3 decisions: identifying likely fire activity, assessing threatened assets, and determining where resources can provide feasible protection. A high-risk zone may not warrant immediate deployment if access is restricted, existing coverage is sufficient, or responder exposure is disproportionate. Priority maps should therefore be interpreted alongside operational constraints and uncertainty, with final decisions informed by incident objectives and local knowledge.

For logistics planners, resource utilization should be assessed jointly with unmet demand and adverse-scenario losses. In the simulation, stochastic staging reduced unmet demand while maintaining lower utilization, illustrating how reserve capacity can protect against disruptions. Maximizing utilization alone may leave insufficient flexibility when demand rises or base efficiency deteriorates. Reserve levels should reflect credible disruption scenarios, while responder exposure requires separate monitoring because increased service delivery can also increase exposure.

For technology and procurement managers, improved predictive accuracy should be evaluated through its effect on operational outcomes. The limited additional allocation benefit of full fusion over thermal-only stochastic staging suggests that better uncertainty-aware planning may sometimes provide greater value than additional sensors. Sensor investment should therefore consider decision relevance, reliability, latency, and cost. Recognized missing observations can be accommodated through wider uncertainty estimates, whereas undetected degradation requires calibration monitoring and explicit fallback procedures.

7. Conclusion

This study establishes a reproducible framework connecting multimodal wildfire-risk inference with suppression-resource staging under uncertainty. Using 1,600 synthetic sensing cells, 100 operational zones, 4 staging bases, and 10 independent replications, it integrates correlated Gaussian Bayesian fusion with 2-stage risk-averse stochastic optimization. The framework explicitly represents heterogeneous observation precision, missing channels, and scenario-dependent demand and resource availability, enabling a transparent assessment of how uncertain information influences operational decisions.

The results distinguish predictive improvement from operational value. Full multimodal fusion improves probabilistic estimation and latent-state reconstruction relative to thermal-only and equal-weight alternatives. Compared with Bayesian deterministic staging, Bayesian stochastic staging reduces mean priority-weighted unmet demand by 35.0% and conditional value-at-risk by 25.3%. However, its additional operational advantage over thermal-only stochastic staging remains negligible within simulation uncertainty. These findings suggest that anticipating demand and capacity disruptions can matter more than incremental predictive gains when spatial aggregation and resource constraints dominate allocation decisions. Stress tests further demonstrate that

unrecognized sensor degradation can undermine uncertainty calibration more severely than explicitly modeled missing observations.

These conclusions remain conditional on the simulation design. The common latent-target representation simplifies the distinct roles of thermal, optical, structural, meteorological, atmospheric, and terrain information. The optimization also assumes divisible capacity, static accessibility, 1 staging epoch, and complete scenario information at recourse. Accordingly, the experiment does not establish real suppression effectiveness, equipment-specific feasibility, or transferability across wildfire environments. Statistical comparisons use 10 independent replications, while sensitivity analyses rely on a focal realization and tail estimates contain only 5 test scenarios per replication. These boundaries support interpreting the framework as a computational foundation for empirical development.

Future research should prioritize timestamped wildfire datasets that preserve acquisition history, observation availability, and native spatial resolution. Modality-specific likelihoods, observation-age effects, and spatially blocked validation would strengthen inference, while decision-focused ablation would identify which data improve operational outcomes. Extending the model to rolling-horizon dispatch with distinct crews, drones, aircraft, and replenishment requirements would support more realistic evaluation. The central implication is that wildfire intelligence should be assessed by its contribution to feasible, uncertainty-aware decisions alongside its predictive accuracy.

Author Contributions

Conceptualization, S.A.H. and E.S.; methodology, S.A.H., E.S. and L.D.; software, S.A.H. and M.K.; validation, E.S., L.D. and F.Z.; formal analysis, S.A.H. and E.S.; investigation, S.A.H., L.D. and F.Z.; resources, E.S. and L.D.; data curation, S.A.H. and M.K.; writing, original draft preparation, S.A.H., E.S. and L.D.; writing, review and editing, E.S., L.D., F.Z. and M.K.; visualization, S.A.H. and M.K.; supervision, E.S.; project administration, E.S.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data Availability Statement

The data used in this study were synthetically generated as part of the simulation framework described in the manuscript. No empirical wildfire observations were acquired or analyzed in the reported experiments. The study generated synthetic multimodal sensing records, predictive evaluation data, scenario-based resource-allocation results, policy summaries, ablation results, and staging decisions. The generated data, computational outputs, and simulation code supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Abdalla, S. Z. S. (2025). Data-driven innovations in disaster risk management: Advancing resilience and sustainability through big data analytics. *Progress in Disaster Science*, 100451.
- Algadi, A., Uluutku, B., Cano, J. A., Walters, L., Alqurashi, I., Luleci, F., & Catbas, F. N. (2025). Digital twin-oriented generation of structural data and models with LiDAR scan point clouds. *Journal of Infrastructure Preservation and Resilience*, 6(1), 32.
- Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming* (2nd ed.). Springer.
<https://doi.org/10.1007/978-1-4614-0237-4>
- Carbonell-Rivera, J. P., Moran, C. J., Seielstad, C. A., Parsons, R. A., Hoff, V., Ruiz, L. A., Torralba, J., & Estornell, J. (2024). Relationships of fire rate of spread with spectral and geometric features derived from UAV-based photogrammetric point clouds. *Fire*, 7(4), 132. <https://doi.org/10.3390/fire7040132>
- Chan, H., Tran-Thanh, L., & Viswanathan, V. (2020). Fighting wildfires under uncertainty: A sequential resource allocation approach. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 4322–4329. <https://doi.org/10.24963/ijcai.2020/596>
- Chen, X. (n.d.). *Flame_2_dataset* [Computer software]. GitHub. https://github.com/xiwenc1/Flame_2_dataset
- Chen, X., Hopkins, B., Wang, H., O'Neill, L., Afghah, F., Razi, A., Fulé, P., Coen, J., Rowell, E., & Watts, A. (2022). Wildland fire detection and monitoring using a drone-collected RGB/IR image dataset. *IEEE Access*, 10, 121301–121317. <https://doi.org/10.1109/ACCESS.2022.3222805>
- Chen, X., Xiao, Z., Cheng, Y., Hsia, C.-C., Wang, H., Xu, J., Xu, S., Dang, F., Zhang, X.-P., Liu, Y., & Chen, X. (2024). SOScheduler: Toward proactive and adaptive wildfire suppression via multi-UAV collaborative scheduling. *IEEE Internet of Things Journal*, 11(14), 24858–24871. <https://doi.org/10.1109/IIOT.2024.3389771>
- Cook, B. D., Corp, L. A., Nelson, R. F., Middleton, E. M., Morton, D. C., McCorkel, J. T., Masek, J. G., Ranson, K. J., Ly, V., & Montesano, P. M. (2013). NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager. *Remote Sensing*, 5(8), 4045–4066. <https://doi.org/10.3390/rs5084045>
- Copernicus Climate Change Service. (2018). ERA5 hourly data on single levels from 1940 to present [Data set]. Climate Data Store. <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview>
- Copernicus Data Space Ecosystem. (n.d.). Sentinel-2. Retrieved September 9, 2026, from <https://dataspace.copernicus.eu/data-collections/copernicus-sentinel-missions/sentinel-2>
- Das, K., Poovvancheri, J., Flesca, S., Calidonna, C. R., & Chen, D. (2026). Emerging trends in wildfire detection through the lens of computer vision and wildfire emission quantification: A comprehensive survey. *IEEE Access*.
- Díaz-Varela, E., Roces-Díaz, J. V., & Álvarez-Álvarez, P. (2016). Detection of landscape heterogeneity at multiple scales: Use of the Quadratic Entropy Index. *Landscape and Urban Planning*, 153, 149–159.
- European Space Agency. (2015). *Sentinel-2 user handbook*.
https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook
- Fan, F., Ying, Q., Chen, G., Wang, Y., Chen, J., & Zhang, H. (2026). Santa Ana winds control wildfire smoke transport and related health impacts in California during January 2025. *Atmospheric Environment*, 122292.
- Gao, B.-C. (1996). NDWI: A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman & Hall/CRC. <https://www.routledge.com/Bayesian-Data-Analysis/Gelman-Carlin-Stern-Dunson-Vehtari-Rubin/p/book/9781439840955>
- Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, 178, 31–41. <https://doi.org/10.1016/j.rse.2016.02.054>
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Hopkins, B., O'Neill, L., Afghah, F., Razi, A., Watts, A., Fulé, P., & Coen, J. (2022). FLAME 2 [Data set]. IEEE DataPort. <https://doi.org/10.21227/krde-h753>
- Huangfu, Q., & Hall, J. A. J. (2018). Parallelizing the dual revised simplex method. *Mathematical Programming Computation*, 10, 119–142. <https://doi.org/10.1007/s12532-017-0130-5>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>

- Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. (2020). A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 28, 478–505. <https://doi.org/10.1139/er-2020-0019>
- Jonnalagadda, A. V., & Hashim, H. A. (2024). SegNet: A segmented deep learning based convolutional neural network approach for drones wildfire detection. *Remote Sensing Applications: Society and Environment*, 34, 101181. <https://doi.org/10.1016/j.rsase.2024.101181>
- Kemper, G., & Kemper, H. (2026). Fireguard: A real-time wildfire monitoring and risk assessment system using unmanned aerial systems and multi-sensor fusion. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XI-4-2026, 179–186. <https://doi.org/10.5194/isprs-annals-XI-4-2026-179-2026>
- Lonbar, A. G., Hasanzadeh, H., Asgari, F., Khamoushi, E., Naeini, H. K., Shomali, R., & Asadi, S. (2024). Prioritizing risk factors in media entrepreneurship on social networks: Hybrid fuzzy Z-number approaches for strategic budget allocation and risk management in advertising construction campaigns. *arXiv preprint arXiv:2409.18976*.
- Lourenço, M., Oliveira, L. B., Oliveira, J. P., Mora, A., Oliveira, H., & Santos, R. (2021). An integrated decision support system for improving wildfire suppression management. *ISPRS International Journal of Geo-Information*, 10(8), 497. <https://doi.org/10.3390/ijgi10080497>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30. <https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions>
- Malekijoo, A., & Fadaeieslam, M. J. (2019). Convolution-deconvolution architecture with the pyramid pooling module for semantic segmentation. *Multimedia Tools and Applications*, 78(22), 32379–32392.
- Malekijoo, A. H., & Malekijoo, R. (2026). BRAW: Block-wise rating of attribute weights for multi-attribute decision-making. In *Encyclopedia of Multi-Attribute Decision Making (MADM)* (pp. 451–460). Morgan Kaufmann.
- Malekijoo, H., Hakami, V., Javan, N. T., & Malekijoo, A. (2023). Q-learning-based algorithms for dynamic transmission control in IoT equipment. *The Journal of Supercomputing*, 79(1), 75–108.
- Momeni, M., Soleimani, H., Shahparvari, S., & Afshar-Nadjafi, B. (2023). A multi-agency coordination resource allocation and routing decision-making problem: A coordinated truck-and-drone DSS for improved wildfire detection coverage. *International Journal of Disaster Risk Reduction*, 97, 104027. <https://doi.org/10.1016/j.ijdrr.2023.104027>
- National Aeronautics and Space Administration. (n.d.-a). Fire Information for Resource Management System. Retrieved September 9, 2026, from <https://firms.modaps.eosdis.nasa.gov/>
- National Aeronautics and Space Administration. (n.d.-b). Archive download. Retrieved September 9, 2026, from <https://firms.modaps.eosdis.nasa.gov/download/>
- National Aeronautics and Space Administration Scientific Visualization Studio. (2025, July 11). Spread of the Palisades and Eaton fires: January 2025. <https://svs.gsfc.nasa.gov/5558/>
- National Centers for Environmental Information. (n.d.). Global hourly: Integrated Surface Database (ISD). Retrieved September 9, 2026, from <https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database>
- National Interagency Fire Center. (n.d.-a). Maps. Retrieved September 9, 2026, from <https://www.nifc.gov/fire-information/maps>
- National Interagency Fire Center. (n.d.-b). WFIGS current interagency fire perimeters. Retrieved September 9, 2026, from <https://data-nifc.opendata.arcgis.com/datasets/nifc%3A%3Awfigs-current-interagency-fire-perimeters/about>
- Nguyen, D., Belval, E. J., Wei, Y., & Calkin, D. E. (2026). A historical analysis of factors driving the daily prioritization of wildland fires in California. *International Journal of Disaster Risk Reduction*, 132, 105967. <https://doi.org/10.1016/j.ijdrr.2025.105967>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <https://jmlr.org/papers/v12/pedregosa11a.html>
- Richards, G. D. (1990). An elliptical growth model of forest fire fronts and its numerical solution. *International Journal for Numerical Methods in Engineering*, 30(6), 1163–1179. <https://doi.org/10.1002/nme.1620300606>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929. <https://doi.org/10.1111/ecog.02881>

- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21–41. <https://doi.org/10.21314/JOR.2000.038>
- Rossi, D., Kuusela, O. P., & Dunn, C. (2022). A microeconomic analysis of wildfire suppression decisions in the Western United States. *Ecological Economics*, 200, 107525.
- Schroeder, W., Oliva, P., Giglio, L., & Csaszar, I. (2014). The new VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, 143, 85–96. <https://doi.org/10.1016/j.rse.2013.12.008>
- SciPy Developers. (n.d.). `linprog(method='highs')`. Retrieved September 9, 2026, from <https://docs.scipy.org/doc/scipy/reference/optimize.linprog-highs.html>
- Tavakol Sadrabadi, M., Peiró, J., Innocente, M. S., & Rein, G. (2025). Conceptual design of a wildfire emergency response system empowered by swarms of unmanned aerial vehicles. *International Journal of Disaster Risk Reduction*, 124, 105493. <https://doi.org/10.1016/j.ijdrr.2025.105493>
- U.S. Environmental Protection Agency. (2020). Air Quality System (AQS) API. https://aqs.epa.gov/aqsweb/documents/data_api.html
- U.S. Environmental Protection Agency. (n.d.-a). Air Quality System. Retrieved September 9, 2026, from <https://www.epa.gov/aqs>
- U.S. Environmental Protection Agency. (n.d.-b). Pre-generated data files. Retrieved September 9, 2026, from https://aqs.epa.gov/aqsweb/airdata/download_files.html
- U.S. Geological Survey. (2025, January 24). 2025 post-wildfire lidar data for Los Angeles, CA. <https://www.usgs.gov/3d-elevation-program/science/2025-post-wildfire-lidar-data-los-angeles-ca>
- U.S. Geological Survey. (n.d.-a). 3D Elevation Program. Retrieved September 9, 2026, from <https://www.usgs.gov/3d-elevation-program>
- U.S. Geological Survey. (n.d.-b). Landsat Collection 2 Surface Temperature. Retrieved September 9, 2026, from <https://www.usgs.gov/landsat-missions/landsat-collection-2-surface-temperature>
- U.S. Geological Survey. (n.d.-c). Landsat Collection 2 Surface Temperature data gaps due to missing ASTER GED. Retrieved September 9, 2026, from <https://www.usgs.gov/landsat-missions/landsat-collection-2-surface-temperature-data-gaps-due-missing-aster-ged>
- U.S. Geological Survey. (n.d.-d). Landsat Normalized Difference Vegetation Index. Retrieved September 9, 2026, from <https://www.usgs.gov/landsat-missions/landsat-normalized-difference-vegetation-index>
- U.S. Geological Survey. (n.d.-e). Landsat Normalized Burn Ratio. Retrieved September 9, 2026, from <https://www.usgs.gov/landsat-missions/landsat-normalized-burn-ratio>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Yuan, K., Zhu, Z., Pang, Y., Pang, J., Hou, C., & Tang, Q. (2025). FireRisk-Multi: A dynamic multimodal fusion framework for high-precision wildfire risk assessment. *ISPRS International Journal of Geo-Information*, 14(11), 426. <https://doi.org/10.3390/ijgi14110426>