PYROADAPT: ADAPTING WILDFIRE PREDICTION UNDER SPATIAL HETEROGENEITY AND TEMPORAL SHIFT

Anonymous authors
Paper under double-blind review

ABSTRACT

Prediction of wildfire occurrence is a rare-event problem compounded by spatial heterogeneity and temporal distribution shift, as fire occurrences are vastly outnumbered by non-occurrences, and predictor–fire relationship varies across space and time. Models trained on historical fire data may perform poorly under new conditions and require adaptation to the target distribution before operational use. We propose PYROADAPT, a pretrain–retrieve–rank framework that adapts a pretrained model to target conditions by retrieving historical locations with similar conditions and fine-tuning on the retrievals through risk ranking. For spatial adaptation, we condition risk on terrain, ecoregion embeddings, and fire rates, accounting for spatial context in the retrieval, and learn risk ordering from same-day fire–nonfire cell pairs. We compare direct ranking, residual pairwise DPO (RDPO), and selective ranking through a unified score-gap formulation that characterizes their gradient allocation. Over California (discretized into 666 $0.25^\circ \times 0.25^\circ$ grid cells), these objectives raise daily average precision from 21.62% for continued focal fine-tuning to 24.35–24.57%, and Top5% recall from 18.70% to 22.11–22.79%. Under a fixed daily detection budget of 34 cells (5% area), selective ranking captures 344 additional positive cell–days. For fires in the top 5%/10%/20% of dry matter consumption, selective ranking raises recall by 39.70/28.18/20.50 percentage points, respectively. Furthermore, rolling evaluations over Yosemite show that the gains from ranking persist under temporal distribution shift. Together, these results show that PYROADAPT prioritizes the most fire-prone locations under a daily budget constraint and detects more extreme fire events.

1 INTRODUCTION

Climate change is fueling more extreme wildfires (Abatzoglou et al., 2025), with devastating impacts on health (Qiu et al., 2025), ecosystems (Byrne et al., 2024), and the economy (Wang et al., 2020). Extreme wildfires are challenging to predict (Wang et al., 2022), as they emerge from the complex interplay of fire weather (Coen et al., 2018; Duane et al., 2021; Swain, 2021), terrain (Sharples et al., 2012), vegetation fuels (Rao et al., 2023; Swain, 2021), and human factors such as ignition and suppression (Keeley & Syphard, 2025; Kreider et al., 2024; Miller et al., 2020; Wu et al., 2023). Climate change further shifts the spatial pattern, seasonality, and distribution of fire weather (Jones et al., 2022) and fuel conditions (Baltzer et al., 2021; Ellis et al., 2022; Halofsky et al., 2020), so the predictor–fire relationships a model learns from historical data may not hold in the year it is deployed.

Predicting wildfire occurrence can be posed as scoring each location on each day for its propensity for fire, conditioned on environmental covariates and spatial context (Jain et al., 2020; Kondylatos et al., 2022). Three properties of the data render this problem difficult. First, positive labels (fire occurrences) are rare, as most location–days record no fire, yielding severe class imbalance and few positive labels relative to the volume of data (He & Garcia, 2009; Chawla et al., 2002; Liu, 2009). Second, the distribution of weather and fuel conditions changes from year to year (Abatzoglou & Williams, 2016; Jones et al., 2022), so inputs at deployment may come from a shifted distribution. Third, the mapping from covariates to risk varies across space, as terrain, vegetation, and ecoregion mediate how a given set of conditions translates into fire risk (Parisien & Moritz, 2009; Hawbaker et al., 2013; Syphard et al., 2024). These factors compound: a model must learn predictor–fire relationships across heterogeneous landscapes from relatively few positive labels while generalizing to conditions not represented in the training record. Consequently, risk scores learned from past data may be poorly matched to the deployment year.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

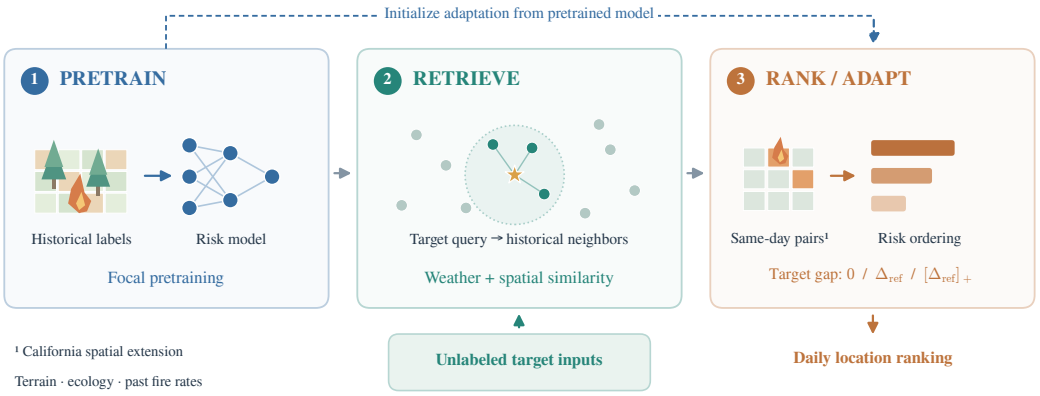

Figure 1: PYROADAPT overview. Given unlabeled target covariates, we retrieve their labeled historical neighbors (Step 2) and use the corresponding labels to adapt the pretrained model (Step 3). Target gaps (Step 3) correspond to direct ranking, RDPO, and selective ranking, respectively. For the California case, we add spatial covariates and same-day pairs for daily location ranking. Dashed arrows denote parameter initialization.

We introduce PYROADAPT, a unified *Pretrain–Retrieve–Rank* framework for batch-transductive adaptation using complete target-period covariates without target labels. Figure 1 summarizes the data flow. Historical focal pretraining initializes a risk model. Unlabeled target inputs retrieve k nearest labeled historical examples in standardized weather and spatial covariate space (Cover & Hart, 1967; Khandelwal et al., 2020). Their labels define positive–negative pairs for fine-tuning without focal supervision, using a ranking loss that penalizes scoring positive examples below negative examples (Burges et al., 2005; Joachims, 2002).

We analyze direct ranking, residual pairwise DPO (RDPO) (Rafailov et al., 2023), and selective ranking through a unified score-gap formulation. Direct ranking uses a zero target for adapted positive–negative score gaps; RDPO uses frozen reference gaps, and selective ranking retains pretrained targets only on correctly ordered pairs. Differentiating each loss with respect to the adapted gap reveals how these targets allocate gradients across correctly ordered and misordered pairs. For the binary-label RDPO applied to Yosemite, we derive its reduction and conditional preference-risk minimizer, distinguishing label-based classification supervision from cross-input ranking.

To address spatial heterogeneity over the California domain, we condition fire risk on terrain, learned ecoregion embeddings, and past fire rates. These inputs represent spatially varying predictor–fire relationships and provide context for historical retrieval. We restrict pairwise ranking to same-day pairs, because correct ranking across days may result from day-to-day variation rather than correctly resolved spatial differences. We evaluate domain-wide prioritization under a fixed daily budget and, separately, recall when dry matter (DM) consumption exceeds certain thresholds. These diagnostics measure budgeted occurrence ranking and sensitivity to high observed activity.

Our contributions are:

- **A unified adaptation framework.** PYROADAPT combines historical risk pretraining, target-conditioned retrieval, and ranking adaptation without target labels.
- **A budgeted-ranking certificate and selective target.** We show that same-day pairwise losses with nonnegative targets, including direct and selective ranking, bound the number of avoidable misses under a fixed daily budget.
- **Daily prioritization over a heterogeneous grid.** Spatial context and same-day ranking support temporal adaptation across historically observed California locations.
- **Budgeted ranking and high-DM recall.** Ranking adaptation outperforms continued focal fine-tuning: at the same daily budget, selective ranking captures about 344 additional positive location–days (after neighborhood smoothing). All three ranking objectives also improve recall of high-DM fires across severity thresholds.

Experimental sections and appendices detail each evaluation protocol.

2 PROBLEM SETUP

Let x denote input features for a location–day and $y \in \{0, 1\}$ indicate whether the neighborhood-smoothed estimate of burned dry matter is positive. Historical labeled examples form $\mathcal{D}_{\text{train}} = \{(x_i, y_i)\}_{i=1}^{n}$. At adaptation time, the learner receives an unlabeled *transductive* batch $\mathcal{X}_{\text{target}} = \{x_j^t\}_{j=1}^{m}$ from a later year; it never trains on target-year labels.

A classifier produces a scalar logit $z_\theta(x)$ with $p_\theta(y = 1 \mid x) = \sigma(z_\theta(x))$. Historical risk pretraining produces parameters θ_{ref} and logit $r(x)$, which provide the starting point for adaptation. We call this model the pretrained model and, when its frozen scores define loss targets, the reference model. Every ranking variant initializes at $\theta_{\text{start}} = \theta_{\text{ref}}$ and fine-tunes the inherited representation on retrieved historical labels. We retain the subscript ref for the pretrained checkpoint; RDPO and selective ranking also use its frozen scores to define their loss targets. We evaluate **ranking** (AUROC, noninterpolated average precision (AP; reported as AUPRC in Yosemite)), **detection at a fixed rule** (recall, F1 using a historical F1-selected threshold), and **top-ranked event recall** (pooled top-20%/30% recall in Yosemite, daily Top5% recall in California).

Fire-prone location prioritization under a daily budget. For date t, let $\eta_{t,i} = \Pr(Y_{t,i} = 1 \mid x_{t,i})$ denote the fire propensity of candidate cell i. Under budget B, the policy maximizing the expected number of captured fire-positive cells is $S_t^* = \arg\max_{|S| \le B} \sum_{i \in S} \eta_{t,i} = \text{TopB}_i\, \eta_{t,i}$. The model substitutes its learned score for the unknown $\eta_{t,i}$. In California, $B = \max\{1, \lceil 0.05|\mathcal{C}_t| \rceil\} = 34$ of 666 cells per day; daily AP, recall, and NDCG assess this within-date ordering on fire dates. Pairs are drawn across days in the Yosemite case and within the same day in the California case (Section 4.3).

3 RELATED WORK

Data-driven wildfire prediction. Data-driven wildfire prediction complements process-based models (Jain et al., 2020; Reichstein et al., 2019), using reanalysis and remote sensing inputs (Abatzoglou, 2013; van der Werf et al., 2025) to model fire risk at regional to continental scales (Huot et al., 2022; Kondylatos et al., 2022; Prapas et al., 2023). Our work studies how historically pretrained risk models can be adapted through retrieval and pairwise fine-tuning, with temporal evaluations of the Yosemite case and spatial prioritization across the domain of California.

Long-tailed and cost-sensitive learning. Standard remedies for severely imbalanced wildfire labels (He & Garcia, 2009) include oversampling (Chawla et al., 2002), focal loss and class-balanced reweighting (Lin et al., 2017; Cui et al., 2019), logit adjustment (Menon et al., 2021; Cao et al., 2019), and decoupled classifiers (Kang et al., 2020). Cost-sensitive learning formalizes asymmetric misclassification costs (Elkan, 2001). Proposition 4 connects the binary-label RDPO preference-risk minimizer to cost-sensitive logit adjustment; a weighted-focal control tests static reweighting.

Adaptation under distribution shift. Importance weighting (Shimodaira, 2000; Sugiyama & Kawanabe, 2012), domain-adaptation theory (Ben-David et al., 2010), adversarial domain alignment (Ganin et al., 2016), and distributionally robust optimization (Sagawa et al., 2020; Koh et al., 2021) address source–target mismatch. Test-time methods adapt with unlabeled target data (Sun et al., 2020; Wang et al., 2021). Our approach instead retrieves labeled historical neighbors (Cover & Hart, 1967; Khandelwal et al., 2020; Lewis et al., 2020) conditioned on the target batch, making it batch-transductive (Vapnik, 1998) rather than source-only.

Preference optimization and pairwise ranking. DPO reparameterizes preference optimization as a closed-form loss (Rafailov et al., 2023), with extensions to alternative divergences (Azar et al., 2024) and unpaired objectives (Ethayarajh et al., 2024). Pairwise learning to rank—RankNet (Burges et al., 2005), LambdaMART (Burges, 2010), listwise (Cao et al., 2007), and SVM-based (Joachims, 2002) approaches (see Liu 2009 for a survey)—optimizes inter-example order. RDPO compares reference-relative preferences over candidates. Its binary-label specialization, binary-label RDPO (Jiang & Sun, 2026), compares labels for one input and reduces to a reference-centered classification surrogate; cross-input RDPO instead compares positive and negative candidate inputs. Direct, RDPO, and selective ranking are alternative objectives for adapting the same pretrained risk model.

4 METHOD

Given a labeled historical archive $\mathcal{D}_{\text{train}}$ and an unlabeled target batch $\mathcal{X}_{\text{target}}$, PYROADAPT produces an adapted scoring model through *Pretrain–Retrieve–Rank*. The stages separate representation learning, target-conditioned sample selection, and decision-aligned adaptation, while keeping target labels withheld until evaluation.

4.1 PRETRAIN–RETRIEVE–RANK

Pretrain. Historical focal training learns a risk representation with parameters θ_{ref} and score $r(x)$ from labeled source years, before any target-year outcomes are observed. We copy this checkpoint and fine-tune the full representation rather than fitting a new model from the smaller retrieved set. The checkpoint initializes every adaptation policy, while its frozen, evaluation-mode scores also define the targets used by RDPO and selective ranking. Reference scores are detached, so gradients update only the adapted policy. Direct ranking shares the same initialization but does not use reference scores in its loss, giving all variants a controlled common starting point.

Retrieve. Continuous covariates are standardized using historical training statistics. Each unlabeled target input then retrieves its k nearest labeled historical inputs in covariate space. This uses target covariates to identify the relevant part of the historical distribution without using target outcomes. The local adaptation set is the union of the retrieved examples,

$$\mathcal{D}_{\text{local}} = \bigcup_{x^t \in \mathcal{X}_{\text{target}}} \left\{ (x_i, y_i) \in \mathcal{D}_{\text{train}} : x_i \in \text{KNN}_k(x^t; \mathcal{D}_{\text{train}}) \right\}. \tag{1}$$

Taking a union deduplicates historical rows retrieved by multiple target inputs, preventing target density from implicitly reweighting the loss. The resulting set is target-conditioned but historically supervised. Let $\mathcal{D}_+ = \{(x, y) \in \mathcal{D}_{\text{local}} : y = 1\}$ and $\pi_+ = |\mathcal{D}_+|/|\mathcal{D}_{\text{local}}|$. We fix $k = 5$ throughout repeated and rolling evaluations. Retrieval uses weather covariates in the Yosemite setting and also incorporates spatial context in the California setting (Section 4.3). This batch-transductive step uses the full set of target covariates only to construct $\mathcal{D}_{\text{local}}$.

Retrieval performs sample selection rather than label propagation: a target input does not inherit a neighbor's label. The target batch instead identifies a local historical slice for adaptation under covariate shift. In rolling evaluation, the archive, standardization statistics, and causal history features use only information available before the target year. The method is therefore transductive in target covariates but label-free for the target period.

This design assumes useful overlap between historical and target covariates. KNN always returns neighbors even when their absolute distances are large, so retrieval cannot reconstruct a target regime absent from the archive. Starting from the global pretrained representation reduces reliance on the smaller local subset, while held-out rolling years test whether the locality assumption transfers. We keep k and the distance representation fixed across evaluation years rather than tuning them on target outcomes.

Rank. Historical labels convert $\mathcal{D}_{\text{local}}$ into positive–negative pairs (x^+, x^-). The comparisons penalize event–non-event inversions and are invariant to a common score shift, aligning training with AP and budgeted recall. Pair construction follows the downstream decision: Yosemite's direct and selective variants compare inputs across the retrieved set, whereas California pairs cells within a date to prioritize the most fire-prone locations under a daily monitoring budget. Write $s_\theta = z_\theta$ for the adapted logit and define the adapted and frozen pretrained score gaps

$$\Delta_\theta = s_\theta(x^+) - s_\theta(x^-), \qquad \Delta_{\text{ref}} = r(x^+) - r(x^-). \tag{2}$$

For a detached target τ, the per-pair loss is $\ell_\tau(\Delta_\theta) = \text{softplus}[-\beta_{\text{rank}}(\Delta_\theta - \tau)]/\beta_{\text{rank}}$, with $\beta_{\text{rank}} > 0$. Direct ranking sets $\tau = 0$ and rewards any increase of the positive score relative to the negative (Burges et al., 2005); RDPO and selective ranking instead derive τ from the pretrained gap (Section 4.2). Minibatches sample admissible pairs while adaptation updates the full predictor, including learned embeddings. Afterward, the retrieved set and frozen reference are discarded; inference uses one forward pass through the adapted model, with no nearest-neighbor query or second deployed model.

4.2 THREE ADAPTATION TARGETS AND GRADIENT ANALYSIS

The three objectives share initialization and pair supervision, with different target gaps. Setting $\tau = \alpha\Delta_{\text{ref}}$ for $\alpha \in [0, 1]$ gives

$$\mathcal{L}_\alpha = -\frac{1}{\beta_{\text{rank}}}\mathbb{E}_{(x^+,x^-)} \log \sigma[\beta_{\text{rank}}(\Delta_\theta - \alpha\Delta_{\text{ref}})]\,. \tag{3}$$

At $\alpha = 0$, direct ranking optimizes the adapted gap; at $\alpha = 1$, RDPO optimizes its change from the pretrained gap (Rafailov et al., 2023). Under $\pi_\theta(i \mid \mathcal{C}) \propto \exp s_\theta(x_i)$, the chosen–rejected log-ratio difference is $\Delta_\theta - \Delta_{\text{ref}}$. Intermediate α retains part of the reference margin. Yosemite applies the same construction to binary labels; Appendix A.1 gives its reduction and conditional optimum.

Selective target. Selective ranking assigns a zero target to pretrained misordered pairs and retains the pretrained gap for correctly ordered pairs. Setting $\alpha(x^+, x^-) = \mathbf{1}\{\Delta_{\text{ref}} > 0\}$ yields

$$\mathcal{L}_{\text{sel}} = -\frac{1}{\beta_{\text{rank}}}\mathbb{E} \log \sigma\big[\beta_{\text{rank}}\big(\Delta_\theta - [\Delta_{\text{ref}}]_+\big)\big]\,, \quad [u]_+ = \max(0, u). \tag{4}$$

It equals direct ranking on pretrained errors and RDPO on correctly ordered pairs. The reference is detached for both reference-dependent targets. These soft losses favor increasing the gap; shared parameters, finite training, and regularization do not guarantee correction or retention of every pair. Proposition 5 gives both direct and selective ranking a finite-sample certificate for avoidable Top-B misses. Selective ranking additionally provides a positive-reference-margin certificate. California evaluates these objectives at the same daily budget, where all three improve mean recall over continued focal (Section 6).

4.2.1 GRADIENT ALLOCATION ACROSS PAIR ORDERINGS

Differentiating the per-pair loss with respect to the adapted gap reveals how each target allocates training pressure to different pair orderings.

Proposition 1 (Pairwise target gradient allocation)**.** *For a fixed pair and detached reference, the magnitude of the negative derivative of the per-pair loss with respect to Δ_θ is*

$$g_\alpha = \sigma[-\beta_{\text{rank}}(\Delta_\theta - \alpha\Delta_{\text{ref}})]\,. \tag{5}$$

If initial policy and reference scores coincide, then

$$g_\alpha^{(0)} = \sigma[-\beta_{\text{rank}}(1-\alpha)\Delta_{\text{ref}}]\,, \qquad \frac{\partial g_\alpha^{(0)}}{\partial \alpha} = \beta_{\text{rank}}\Delta_{\text{ref}} g_\alpha^{(0)}(1 - g_\alpha^{(0)}). \tag{6}$$

Thus decreasing α increases the initial gap-gradient magnitude on reference-wrong pairs ($\Delta_{\text{ref}} < 0$), while increasing α increases it on reference-correct pairs ($\Delta_{\text{ref}} > 0$). At $\alpha = 1$ it equals $1/2$ for every pair.

The initialization identity assumes equal scores. Policy training-mode and pretrained-model evaluation-mode batch normalization can break this equality; Appendix A.11 checks realized scalar weights. The proposition concerns gap gradients, with parameter-gradient scope detailed in Appendix A.

Under score-equal initialization, selective ranking has gap-gradient magnitude $\max\{\sigma(-\beta_{\text{rank}}\Delta_{\text{ref}}), 1/2\}$: it preserves direct ranking's larger weight on pretrained errors and RDPO's weight on correctly ordered pairs (Appendix Figure 4).

When the targets differ. The scaled reference gap controls how strongly the choice of target changes the scalar gradient weight.

Proposition 2 (Pairwise target sensitivity bound)**.** *For a fixed pair and detached reference, g_α is Lipschitz in α with*

$$|g_\alpha - g_{\alpha'}| \le \frac{\beta_{\text{rank}}\,|\Delta_{\text{ref}}|}{4}\,|\alpha - \alpha'|\,, \qquad \alpha, \alpha' \in [0, 1], \tag{7}$$

and at score-equal initialization $g_0^{(0)} - g_1^{(0)} = \sigma(-\beta_{\text{rank}}\Delta_{\text{ref}}) - \frac{1}{2} = -\beta_{\text{rank}}\Delta_{\text{ref}}/4 + O((\beta_{\text{rank}}\Delta_{\text{ref}})^3)$. For the selective target,

$$\mathbb{E}\big|g_{\text{sel}}^{(0)} - g_1^{(0)}\big| \le \frac{\beta_{\text{rank}}}{4}\,\mathbb{E}[|\Delta_{\text{ref}}|\,\mathbf{1}\{\Delta_{\text{ref}} < 0\}]\,. \tag{8}$$

Small $\beta_{\text{rank}}|\Delta_{\text{ref}}|$ makes the weights close; larger gaps allow stronger separation. Selective ranking departs from RDPO only on reference-wrong pairs. Exceeding a negative RDPO target can leave $\Delta_\theta < 0$, whereas direct and selective ranking use zero targets on those pairs. These are soft objectives, so task-level correction and retention require empirical evaluation. Appendix A.11 provides the proof and measured margin scales.

4.3 Daily risk prioritization under spatial heterogeneity

The gradient allocation analysis characterizes how each ranking target weights pairwise updates; applying these objectives to a spatially heterogeneous system also requires specifying how locations are represented and compared. In California, we adapt the Pretrain–Retrieve–Rank framework by incorporating location-dependent context into risk prediction and retrieval, and by constructing same-day spatial pairs that align adaptation with daily statewide prioritization.

4.3.1 Location-dependent risk representation

We apply these objectives to domain-wide spatial risk prioritization. Let i index a cell, t a date, $w_{t-1,i}$ its previous-day weather, a_i its coordinates and terrain, e_i its ecological region, and $\mathcal{H}_{<\text{yr}(t),i}$ its outcomes from strictly earlier calendar years. The conditional risk is

$$\eta_{t,i} = \Pr\big(Y_{t,i} = 1 \mid w_{t-1,i}, a_i, e_i, \mathcal{H}_{<\text{yr}(t),i}, d_t\big)\,, \tag{9}$$

where d_t encodes seasonality and $\text{yr}(t)$ denotes the calendar year. California pretrains a shared model on

$$u_{t,i} = \big[w_{t-1,i},\ a_i,\ h_{\text{yr}(t),i},\ d_t,\ e_i\big]\,, \tag{10}$$

where h summarizes cell fire rates from strictly earlier years and e_i is the raw EPA Level III category (U.S. Environmental Protection Agency, n.d.; Omernik & Griffith, 2014). With $c_{t,i} = [w_{t-1,i}, a_i, h_{\text{yr}(t),i}, d_t]$, the pretrained and adapted models use learned four-dimensional embeddings:

$$r(u_{t,i}) = f_{\text{ref}}([c_{t,i}, E_{\text{ref}}(e_i)]), \qquad s_\theta(u_{t,i}) = f_\theta([c_{t,i}, E_\theta(e_i)]). \tag{11}$$

This shared representation permits location-dependent baseline risks and weather–landscape interactions. The adapted embedding starts from its pretrained counterpart and is fine-tuned with the full model. The 26 raw fields include weather, terrain, cell history, seasonality, and one region category. History summaries use full historical label grids from strictly earlier years; target-year outcomes never enter them.

Retrieval with weather and spatial context. California retrieval uses standardized continuous covariates, excluding history length, plus a fixed ecoregion one-hot vector in Euclidean distance. This representation is separate from the predictor's learned embedding and allows borrowing across regions without restricting neighbors to the same region. Equation 1 forms the deduplicated adaptation set with $k = 5$. Appendix D defines the inputs, causal-history shrinkage, and exact distance.

4.3.2 Same-day spatial comparisons

We construct positive–negative pairs from the same historical day because cross-date ranking can reward separating high-risk from low-risk dates without resolving which locations are riskier within a date. We therefore construct. To see the cancellation, write $s_\theta(u_{t,i}) = b_\theta(t) + q_\theta(t,i)$, where $b_\theta(t)$ is any additive date-specific offset. Then

$$s_\theta(u_{t,i^+}) - s_\theta(u_{t,i^-}) = q_\theta(t,i^+) - q_\theta(t,i^-). \tag{12}$$

The ranking term therefore cannot improve by raising all scores on a high-risk date: it must distinguish positive from negative cells within that date. Both within-region and cross-region comparisons are allowed.

For a minibatch $B \subset \mathcal{D}_{\text{local}}$, let $\mathcal{T}_B$ contain its dates with both labels. For each $t \in \mathcal{T}_B$, sample one positive cell i_t^+ and one negative cell i_t^- uniformly within their classes. Set $\Delta_{\theta,t} = s_\theta(u_{t,i_t^+}) - s_\theta(u_{t,i_t^-})$, with $\Delta_{\text{ref},t}$ defined analogously. The pairwise term is

$$\mathcal{L}_{\text{pair}}(B;\tau) = \frac{1}{|\mathcal{T}_B|}\sum_{t\in\mathcal{T}_B} \frac{\text{softplus}[-\beta_{\text{rank}}(\Delta_{\theta,t} - \tau_t)]}{\beta_{\text{rank}}}, \tag{13}$$

with zero contribution when $\mathcal{T}_B$ is empty. The targets analyzed in Section 4.2 are

$$\tau_t^{\text{direct}} = 0, \qquad \tau_t^{\text{RDPO}} = \Delta_{\text{ref},t}, \qquad \tau_t^{\text{selective}} = [\Delta_{\text{ref},t}]_+, \tag{14}$$

where $[u]_+ = \max(0, u)$. RDPO ranks the implicit residual $h_\theta^{\text{res}}(u) = s_\theta(u) - r(u)$. California fine-tunes with the pairwise objective, without focal supervision:

$$\mathcal{L}_{\text{CA}} = \mathbb{E}_{B,\,\text{pair draws}}[\lambda_{\text{rank}}\mathcal{L}_{\text{pair}}(B;\tau) + \lambda_s M_{\text{loss}}(B)]\,, \tag{15}$$

where M_{loss} is the TabNet mask term. We use $\beta_{\text{rank}} = \lambda_{\text{rank}} = 0.1$ and $\lambda_s = 0.05$; the single all-feature attention group makes the mask term zero. Focal loss is detached and monitored only during ranking adaptation. Dates without both classes contribute no preference pair. No positive-only preference term or margin cap is used. Both California and Yosemite exclude focal supervision from ranking adaptation; California uses same-day spatial pairs, whereas Yosemite's direct and selective variants use cross-date pairs at $\beta_{\text{rank}} = 1$. Appendix D details optimization and evaluation, and Appendix E describes broader-area extensions.

5 YOSEMITE: TEMPORAL SHIFT

Setup. For the Yosemite case, we evaluate *temporal shift* on a fixed 3×3 grid. The same nine cells appear in historical and target years, holding the candidate geography fixed while assessing year-to-year adaptation. We use daily GridMET meteorology (Abatzoglou, 2013) and GFED-derived dry matter (van der Werf et al., 2025) with twelve weather inputs. The main split contains 62,460 training grid–days through 2020 (7,613 positives) and 3,285 target grid–days in 2021 (363 positives). The historically pretrained TabNet (Arik & Pfister, 2021) is adapted for 100 epochs with AdamW at 10^{-4}; retrieval uses $k = 5$. The adapted local policies share the same retrieved rows, frozen pretrained model, initialization, and seeded batches. The main seeds are $\{11, 17, 23, 29, 37\}$; target labels are used only for evaluation. Full architecture and optimization details appear in Appendix B. We additionally train **joint focal + direct ranking** from random initialization on all historical rows, using the same 50-epoch Adam protocol as the pretrained focal model and a direct-ranking coefficient of one. This from-scratch baseline does not use target-dependent retrieval; its full configuration is in Appendix B.2.

The six local comparison policies are: the **frozen pretrained model** (no local adaptation), **continued focal** (local focal fine-tuning), **weighted focal** (Equation 23), **direct pairwise ranking** (Equation 3 at $\alpha = 0$), **selective pairwise ranking** (Equation 4), and **binary-label RDPO** (Equation 17). All three ranking adaptations fix $\beta = 1$ and preference coefficient one, with no focal supervision or extra positive-label term ($\lambda_2 = 0$). Direct and selective ranking compare positive–negative inputs across dates; binary-label RDPO compares true and complementary labels for each input. These configurations share training settings while differing in candidate definitions. Bold table entries indicate the best mean per column. Expanding historical windows end in 2018, 2019, and 2020 for rolling tests in 2019, 2020, and 2021, with seeds $\{11, 17, 23\}$. Uncertainty intervals, metric definitions, and evaluation scope are in Appendix B.

Results. Pairwise fine-tuning improves pooled ranking on the 2021 holdout (Table 1). Direct ranking raises AUROC from 70.32% for the pretrained model to 71.23%, and AUPRC from 21.23% to 22.82%. Selective ranking attains AUROC of 71.51% and AUPRC of 22.94%. The pointwise binary-label RDPO baseline also improves over the pretrained model, reaching AUROC of 71.00% and AUPRC of 22.59%. These results show that the inherited risk representation supports several subsequent adaptation objectives.

Rolling evaluation extends the comparison across 2019–2021. Figure 2 reports paired gains over each pretrained checkpoint across nine year–seed cases, and Appendix Table 11 gives absolute scores. Direct ranking attains mean AUROC of 74.19% and AUPRC of 24.30%, compared with 73.12% and 22.85% for the pretrained model and 73.08% and 22.08% for continued focal. Relative to binary-label RDPO, direct ranking gains 2.06 AUROC, 2.25 AUPRC, 3.69 R@20, and 5.97 R@30 percentage points (differences computed before rounding). This supports explicit pairwise fine-tuning as an effective adaptation option under temporal shift. The binary-label RDPO configuration evaluated here, with no focal supervision and a fixed training schedule, shows unstable performance across test years; this observation is specific to the evaluated configuration and does not establish a general limitation of RDPO.

The from-scratch joint focal and direct-ranking baseline attains rolling AUROC of 71.73% and AUPRC of 21.43%, compared with 74.19% and 24.30% for local direct adaptation. The paired

Table 1: Results for Yosemite in 2021, showing mean ± SD (%) over five seeds, with the highest mean in each column in bold. Ranking adaptation fixes $\beta = 1$ without focal supervision. R@20 and R@30 pool all target location–days.

Method	Recall	F1	AUROC	AUPRC	R@20	R@30
Pretrained	57.80 ± 5.70	30.61 ± 0.75	70.32 ± 0.51	21.23 ± 1.27	42.31 ± 1.15	57.08 ± 2.54
Joint focal + direct	52.73 ± 6.77	29.39 ± 1.77	69.40 ± 1.00	20.27 ± 0.77	40.39 ± 2.85	54.82 ± 3.19
Continued focal	57.02 ± 6.80	29.93 ± 1.91	70.16 ± 1.11	21.07 ± 1.62	41.60 ± 2.71	56.31 ± 3.63
Weighted focal	56.09 ± 6.07	29.94 ± 1.56	70.02 ± 1.13	21.00 ± 1.53	41.10 ± 2.89	56.03 ± 3.72
Direct ranking	60.50 ± 2.37	**31.31** ± 1.11	71.23 ± 0.62	22.82 ± 1.10	44.90 ± 1.66	**58.68** ± 2.15
binary-label RDPO	61.27 ± 5.86	30.79 ± 1.41	71.00 ± 0.81	22.59 ± 1.70	43.86 ± 2.62	57.36 ± 3.25
Selective ranking	**61.32** ± 3.50	31.25 ± 0.86	**71.51** ± 0.44	**22.94** ± 1.09	**45.56** ± 1.37	**58.68** ± 1.87

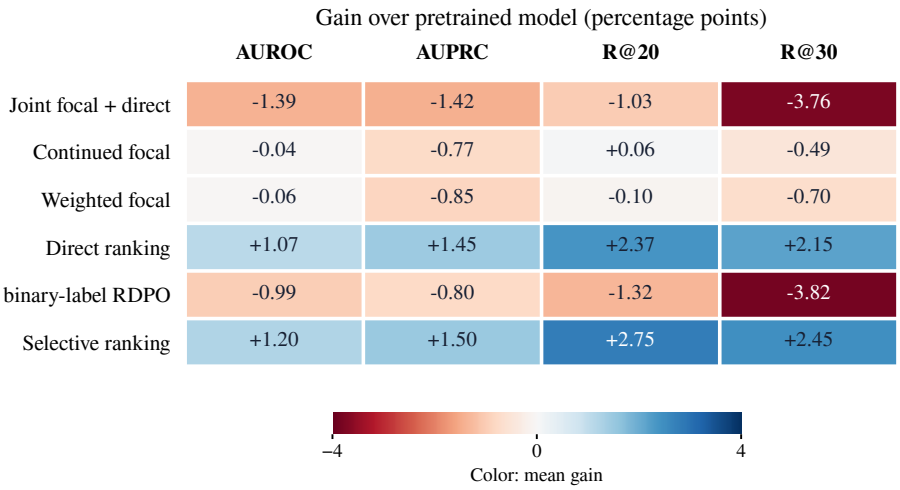

Figure 2: Paired Yosemite ranking gains over the pretrained model, 2019–2021, in percentage points. Cells and color show mean gains over three per-seed averages across years. Joint focal + direct starts from random initialization on all historical rows; local adaptation starts from the pretrained focal model. Binary-label RDPO is the pointwise preference baseline, with $\beta = 1$ and no focal term. See Appendix Table 11 for absolute scores.

local-minus-joint AUROC gain is 2.46 points; the AUPRC gain is 2.86 points. This comparison evaluates the complete local-adaptation configuration, with differences in data access and training schedules documented in Appendix B.2.

On Yosemite's nine-cell grid, the daily Top5% rule selects one cell, so this diagnostic operates at a different decision scale from statewide monitoring. Detailed direct–selective comparisons, annual results, operating points, and same-day pair diagnostics appear in Appendix C. California evaluates adaptation with same-day training pairs on the complete statewide candidate grid.

6 California: Statewide Risk Prioritization

Complete-grid evaluation. California has diverse terrain, ecology, and fire regimes. This domain is discretized into 666 grid cells (0.25°). Historical risk pretraining uses 2002–2021 data, and adaptation targets 2022. The fixed 20% historical sample contains 972,892 cell–days. Batch-transductive $k = 5$ retrieval queries all 243,090 unlabeled target covariates, yielding 363,643 unique historical examples for fine-tuning. Evaluation covers the complete daily grid, with 14,760 positives across 351 fire dates. The daily Top5% rule selects 34 cells per date, or 12,410 alarms in the 2022 evaluation. Within each of seeds {11, 22, 33}, adapted models share the pretrained checkpoint, scaler, and retrieved examples; pairwise variants use $\lambda_{\text{rank}} = \beta_{\text{rank}} = 0.1$ and the same pair-sampling rule, with focal supervision excluded from all three ranking objectives. See Appendix D for the full protocol.

Table 2: Results for the entire California domain in 2022, showing Mean ± SD (%) over three seeds. Ranking adaptation excludes focal supervision. Daily metrics average fire dates; global metrics pool cell–days.

Method	Daily AP	Daily R@5%	Global AP	AUROC
XGBoost	14.24 ± 0.07	11.64 ± 0.23	11.02 ± 0.02	66.27 ± 0.04
Pretrained spatial focal	21.10 ± 1.24	17.98 ± 1.36	19.30 ± 0.63	78.34 ± 0.37
Continued focal	21.62 ± 0.33	18.70 ± 0.74	20.07 ± 0.34	78.60 ± 0.21
Direct ranking	24.35 ± 0.69	22.25 ± 1.39	20.80 ± 1.53	79.46 ± 0.50
RDPO	24.47 ± 0.42	22.11 ± 1.92	**20.89** ± 1.39	**79.47** ± 0.47
Selective ranking	**24.57** ± 0.52	**22.79** ± 2.07	20.71 ± 1.13	79.38 ± 0.33

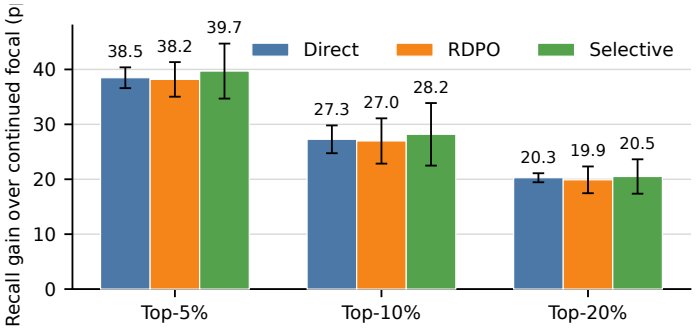

Figure 3: California recall gain over continued focal within retrospective high-DM strata. Mean and sample SD of paired seed differences; each policy uses its historical F1-selected threshold, with unmatched alarm rates.

Location prioritization under a tight daily budget. Without focal supervision during adaptation, direct, RDPO, and selective ranking attain daily AP of 24.35%, 24.47%, and 24.57%, respectively, compared with 21.62% for continued focal fine-tuning (Table 2). Their macro Top5% recall is 22.25%, 22.11%, and 22.79%, versus 18.70% for continued focal. These gains show the value of same-day ranking supervision for daily location prioritization. Proposition 5 connects the nonnegative-target losses used by direct and selective ranking to a finite-sample certificate for avoidable Top-B misses.

At 34 cells per day, selective captures a mean of 3,210.3 positive cell–days versus 2,866.0 for continued focal: 344 additional captures, or approximately 12%, at the same number of alarms (Appendix Table 19). Direct and RDPO add 311 and 301, respectively. These are neighborhood-smoothed positive cell–days, not independent incidents.

Thresholded recall within high-DM strata. Among the 2,193 raw-DM-positive test examples, we evaluate training-threshold recall within the highest-DM 5%, 10%, and 20%. All ranking variants identify a much larger fraction than continued focal: direct gains 38.48/27.27/20.27 points, RDPO 38.18/26.97/19.89, and selective 39.70/28.18/20.50 (Figure 3). These gains describe sensitivity under each policy's historical F1-selected threshold. Target-year DM defines retrospective strata only and supplies no severity supervision. Alarm rates are not matched here, so this diagnostic does not establish severity forecasting or equal-budget severe-event discrimination.

Isolating retrieval. A separate three-seed continued-focal ablation compares spatial KNN, a size-matched random subset, and the full historical sample at exactly 6,800 updates. KNN improves mean daily AP by 0.66 points over random and 0.36 over full-history adaptation, with positive paired differences in every seed. This isolates source selection under a common loss and compute budget; gains in Top5% recall remain less conclusive (Appendix Table 18).

7 Conclusion

PyroAdapt combines historical risk pretraining, target-conditioned retrieval, and ranking adaptation without focal supervision. Over the entire domain of California, all three ranking objectives improve mean daily AP and Top5% recall over continued focal. At a daily budget of 34 locations, the selective configuration captures approximately 344 additional neighborhood-smoothed positive cell–days. Ranking also increases recall within retrospective high-DM strata at training-selected thresholds.

AI USE STATEMENT

In this work, we used generative AI tools to help develop the theoretical framework, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, provide feedback on experimental design, implement methods, clean and reformat dataset. We did *not* use generative AI tools to generate synthetic data sets, propose or refine hypotheses, assist with translation, support qualitative and thematic data analysis, or interpret results. Additionally, we used generative AI tools to help design an overview of the framework (Figure 1), create Figures 2 and 3 from experimental results, create and edit software code, summarize and analyze existing literature on wildfire research, draft the paper based on results, edit the paper to improve readability, and format references. The authors have reviewed all AI-assisted work, including the theoretical claims, experimental design and implementation, software code, and results. The authors take full responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work involves no human subjects, no personally identifying information, and no new dataset release; all inputs are public environmental products (Abatzoglou, 2013; van der Werf et al., 2025; U.S. Environmental Protection Agency, n.d.) whose licenses permit research use. Because risk ranking directs limited monitoring capacity, both error modes are costly—a missed positive can leave a developing fire unobserved, while a false alarm consumes capacity another location needed—so we report ranking, fixed-threshold detection, and top-ranked event recall separately (Section 2) rather than a single headline number. Two risks deserve mention: location prioritization can disadvantage regions with sparse historical support, since our history features shrink such cells toward a statewide prior, so deployments should audit performance by region rather than only in aggregate; and our label is a neighborhood-smoothed occurrence indicator rather than a verified ignition, so predictions should not be read as ignition forecasts. We therefore claim no operational readiness, and regard premature operational use—without the deployment-matched evaluation of Appendix E—as the principal misuse risk of this work.

REPRODUCIBILITY STATEMENT

Theoretical claims are stated in Section 4 and proved in Appendix A, which also gives the expected-batch coefficient qualification and conditions for order preservation. Yosemite data, architecture, optimizer settings, retrieval, policy definitions, seeds, and the rolling protocol appear in Section 5 and Appendix B, including the fixed-configuration table and initial gradient-norm checks. The California protocol is in Appendix D: label smoothing, the fixed 20% stratified sample and its data seed, causal history features with shrinkage and cold-start rules, retrieval distance, optimizer settings, stopping rule, XGBoost configuration, and training-only threshold selection; metric definitions, including pooled R@q versus daily Top5% selection and noninterpolated average precision, accompany the results. The protocol records pretraining epoch caps, fixed loss settings, and the exclusion of focal supervision from ranking adaptation. All datasets are public, and our code and configuration files will be released.

REFERENCES

John T. Abatzoglou. Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(1):121–131, 2013.

John T. Abatzoglou and A. Park Williams. Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, 113(42):11770–11775, 2016.

John T. Abatzoglou, Crystal A. Kolden, Alison C. Cullen, Mojtaba Sadegh, Emily L. Williams, Marco Turco, and Matthew W. Jones. Climate change has increased the odds of extreme regional forest fire years globally. *Nature Communications*, 16(1), 7 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-61608-1. URL http://dx.doi.org/10.1038/s41467-025-61608-1.

Sercan Ö. Arik and Tomas Pfister. Tabnet: Attentive interpretable tabular learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 6679–6687, 2021.

Mohammad Gheshlaghi Azar, Mark Rowland, Bilal Piot, Daniel Guo, Daniele Calandriello, Michal Valko, and Rémi Munos. A general theoretical paradigm to understand learning from human feedback. In *International Conference on Artificial Intelligence and Statistics*, pp. 4447–4455, 2024.

Jennifer L. Baltzer, Nicola J. Day, Xanthe J. Walker, David Greene, Michelle C. Mack, Heather D. Alexander, Dominique Arseneault, Jennifer Barnes, Yves Bergeron, Yan Boucher, Laura Bourgeau-Chavez, Carissa D. Brown, Suzanne Carrière, Brian K. Howard, Sylvie Gauthier, Marc-André Parisien, Kirsten A. Reid, Brendan M. Rogers, Carl Roland, Luc Sirois, Sarah Stehn, Dan K. Thompson, Merritt R. Turetsky, Sander Veraverbeke, Ellen Whitman, Jian Yang, and Jill F. Johnstone. Increasing fire and the decline of fire adapted black spruce in the boreal forest. *Proceedings of the National Academy of Sciences*, 118(45):e2024872118, 2021. doi: 10.1073/pnas.2024872118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2024872118.

Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine Learning*, 79(1–2):151–175, 2010.

Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *Proceedings of the 22nd International Conference on Machine Learning*, pp. 89–96, 2005. doi: 10.1145/1102351.1102363. URL https://www.microsoft.com/en-us/research/publication/learning-to-rank-using-gradient-descent/.

Christopher J. C. Burges. From RankNet to LambdaRank to LambdaMART: An overview. Technical Report MSR-TR-2010-82, Microsoft Research, 2010.

Brendan Byrne, Junjie Liu, Kevin W. Bowman, Madeleine Pascolini-Campbell, Abhishek Chatterjee, Sudhanshu Pandey, Kazuyuki Miyazaki, Guido R. van der Werf, Debra Wunch, Paul O. Wennberg, Coleen M. Roehl, and Saptarshi Sinha. Carbon emissions from the 2023 Canadian wildfires. *Nature*, 633(8031):835–839, 8 2024. ISSN 1476-4687. doi: 10.1038/s41586-024-07878-z. URL http://dx.doi.org/10.1038/s41586-024-07878-z.

Kaidi Cao, Colin Wei, Adrien Gaidon, Nikos Arechiga, and Tengyu Ma. Learning imbalanced datasets with label-distribution-aware margin loss. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. Learning to rank: From pairwise approach to listwise approach. In *Proceedings of the 24th International Conference on Machine Learning*, pp. 129–136, 2007.

Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16: 321–357, 2002.

Janice L. Coen, E. Natasha Stavros, and Josephine A. Fites-Kaufman. Deconstructing the King megafire. *Ecological Applications*, 28(6):1565–1580, 2018. doi: 10.1002/eap.1752. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/eap.1752.

Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9268–9277, 2019.

Andrea Duane, Marc Castellnou, and Lluís Brotons. Towards a comprehensive look at global drivers of novel extreme wildfire events. *Climatic Change*, 165(3–4), 4 2021. ISSN 1573-1480. doi: 10.1007/s10584-021-03066-4. URL http://dx.doi.org/10.1007/s10584-021-03066-4.

Charles Elkan. The foundations of cost-sensitive learning. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, pp. 973–978, 2001.

Todd M. Ellis, David M. J. S. Bowman, Piyush Jain, Mike D. Flannigan, and Grant J. Williamson. Global increase in wildfire risk due to climate-driven declines in fuel moisture. *Global Change Biology*, 28(4):1544–1559, 2022. doi: 10.1111/gcb.16006. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.16006.

Kawin Ethayarajh, Winnie Xu, Niklas Muennighoff, Dan Jurafsky, and Douwe Kiela. Model alignment as prospect theoretic optimization. In *International Conference on Machine Learning*, 2024.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On calibration of modern neural networks. In *International Conference on Machine Learning*, pp. 1321–1330, 2017.

Jessica E. Halofsky, David L. Peterson, and Brian J. Harvey. Changing wildfire, changing forests: the effects of climate change on fire regimes and vegetation in the Pacific Northwest, USA. *Fire Ecology*, 16(1), 1 2020. ISSN 1933-9747. doi: 10.1186/s42408-019-0062-8. URL http://dx.doi.org/10.1186/s42408-019-0062-8.

Todd J. Hawbaker, Volker C. Radeloff, Susan I. Stewart, Roger B. Hammer, Nicholas S. Keuler, and Murray K. Clayton. Human and biophysical influences on fire occurrence in the united states. *Ecological Applications*, 23(3):565–582, 2013. doi: 10.1890/12-1816.1. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/12-1816.1.

Haibo He and Edwardo A. Garcia. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9):1263–1284, 2009.

Fantine Huot, R. Lily Hu, Nita Gober, Matthias Foël, Luke Schwindling, Kira Touçe, Yinuo Wang, Anahita Banerjee, Mark Erickson, Sara Mehr, et al. Next day wildfire spread: A machine learning dataset to predict wildfire spreading from remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13, 2022.

Piyush Jain, Sean C.P. Coogan, Sriram Ganapathi Subramanian, Mark Crowley, Steve Taylor, and Mike D. Flannigan. A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 28(4):478–505, 2020. doi: 10.1139/er-2020-0019. URL https://doi.org/10.1139/er-2020-0019.

Enyi Jiang and Wu Sun. Environment-adaptive preference optimization for wildfire prediction. *arXiv preprint arXiv:2605.12435*, 2026.

Thorsten Joachims. Optimizing search engines using clickthrough data. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 133–142, 2002.

Matthew W. Jones, John T. Abatzoglou, Sander Veraverbeke, Niels Andela, Gitta Lasslop, Matthias Forkel, Adam J. P. Smith, Chantelle Burton, Richard A. Betts, Guido R. van der Werf, Stephen Sitch, Josep G. Canadell, Cristina Santín, Crystal Kolden, Stefan H. Doerr, and Corinne Le Quéré. Global and regional trends and drivers of fire under climate change. *Reviews of Geophysics*, 60(3):e2020RG000726, 2022. doi: 10.1029/2020RG000726. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020RG000726. e2020RG000726 2020RG000726.

Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, and Yannis Kalantidis. Decoupling representation and classifier for long-tailed recognition. In *International Conference on Learning Representations*, 2020.

Jon E. Keeley and Alexandra D. Syphard. Causal analysis of fire regime drivers in California. *International Journal of Wildland Fire*, 34(12):WF25166, 12 2025. ISSN 1049-8001. doi: 10.1071/WF25166. URL https://doi.org/10.1071/WF25166.

Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. Generalization through memorization: Nearest neighbor language models. In *International Conference on Learning Representations*, 2020.

Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay Balsubramani, Weihua Hu, Michihiro Yasunaga, Richard Lanas Phillips, Irena Gao, et al. WILDS: A benchmark of in-the-wild distribution shifts. In *International Conference on Machine Learning*, pp. 5637–5664, 2021.

Spyros Kondylatos, Ioannis Prapas, Michele Ronco, Ioannis Papoutsis, Gustau Camps-Valls, María Piles, Miguel-Ángel Fernández-Torres, and Nuno Carvalhais. Wildfire danger prediction and understanding with deep learning. *Geophysical Research Letters*, 49(17):e2022GL099368, 2022. doi: 10.1029/2022GL099368. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GL099368. e2022GL099368 2022GL099368.

Mark R. Kreider, Philip E. Higuera, Sean A. Parks, William L. Rice, Nadia White, and Andrew J. Larson. Fire suppression makes wildfires more severe and accentuates impacts of climate change and fuel accumulation. *Nature Communications*, 15(1), 3 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-46702-0. URL http://dx.doi.org/10.1038/s41467-024-46702-0.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*, volume 33, pp. 9459–9474, 2020.

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988, 2017.

Tie-Yan Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 3(3):225–331, 2009.

Aditya Krishna Menon, Sadeep Jayasumana, Ankit Singh Rawat, Himanshu Jain, Andreas Veit, and Sanjiv Kumar. Long-tail learning via logit adjustment. In *International Conference on Learning Representations*, 2021.

Rebecca K. Miller, Christopher B. Field, and Katharine J. Mach. Barriers and enablers for prescribed burns for wildfire management in California. *Nature Sustainability*, 3(2):101–109, 1 2020. ISSN 2398-9629. doi: 10.1038/s41893-019-0451-7. URL http://dx.doi.org/10.1038/s41893-019-0451-7.

James M. Omernik and Glenn E. Griffith. Ecoregions of the conterminous united states: Evolution of a hierarchical spatial framework. *Environmental Management*, 54(6):1249–1266, 2014. doi: 10.1007/s00267-014-0364-1. URL https://pubs.usgs.gov/publication/70125309.

Marc-André Parisien and Max A. Moritz. Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecological Monographs*, 79(1):127–154, 2009. doi: 10.1890/07-1289.1. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/07-1289.1.

Ioannis Prapas, Spyros Kondylatos, Ioannis Papoutsis, Gustau Camps-Valls, Michele Ronco, Miguel-Ángel Fernández-Torres, Marta P. Guillemé, and Nuno Carvalhais. Deep learning methods for daily wildfire danger forecasting. *arXiv preprint arXiv:2111.02736*, 2023.

Minghao Qiu, Jessica Li, Carlos F. Gould, Renzhi Jing, Makoto Kelp, Marissa L. Childs, Jeff Wen, Yuanyu Xie, Meiyun Lin, Mathew V. Kiang, Sam Heft-Neal, Noah S. Diffenbaugh, and Marshall Burke. Wildfire smoke exposure and mortality burden in the USA under climate change. *Nature*, 647(8091):935–943, 9 2025. ISSN 1476-4687. doi: 10.1038/s41586-025-09611-w. URL http://dx.doi.org/10.1038/s41586-025-09611-w.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D. Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, volume 36, pp. 53728–53741, 2023.

Krishna Rao, A. Park Williams, Noah S. Diffenbaugh, Marta Yebra, Colleen Bryant, and Alexandra G. Konings. Dry live fuels increase the likelihood of lightning-caused fires. *Geophysical Research Letters*, 50(15):e2022GL100975, 2023. doi: https://doi.org/10.1029/2022GL100975. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GL100975. e2022GL100975 2022GL100975.

Markus Reichstein, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, and Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743):195–204, 2019.

D. R. Roberts, V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8):913–929, 2017. doi: 10.1111/ecog.02881. URL https://nsojournals.onlinelibrary.wiley.com/doi/10.1111/ecog.02881.

Shiori Sagawa, Pang Wei Koh, Tatsunori B. Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. In *International Conference on Learning Representations*, 2020.

Jason J. Sharples, Richard H. D. McRae, and Stephen R. Wilkes. Wind–terrain effects on the propagation of wildfires in rugged terrain: fire channelling. *International Journal of Wildland Fire*, 21(3):282–296, 01 2012. ISSN 1049-8001. doi: 10.1071/WF10055. URL https://doi.org/10.1071/WF10055.

Hidetoshi Shimodaira. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, 90(2):227–244, 2000.

Masashi Sugiyama and Motoaki Kawanabe. *Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation*. MIT Press, 2012.

Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei A. Efros, and Moritz Hardt. Test-time training with self-supervision for generalization under distribution shifts. In *International Conference on Machine Learning*, pp. 9229–9248, 2020.

Daniel L. Swain. A shorter, sharper rainy season amplifies California wildfire risk. *Geophysical Research Letters*, 48(5):e2021GL092843, 2021. doi: https://doi.org/10.1029/2021GL092843. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021GL092843. e2021GL092843 2021GL092843.

Alexandra D. Syphard, Santiago José Elías Velazco, Miranda Brooke Rose, Janet Franklin, and Helen M. Regan. The importance of geography in forecasting future fire patterns under climate change. *Proceedings of the National Academy of Sciences*, 121(32):e2310076121, 2024. doi: 10.1073/pnas.2310076121. URL https://www.pnas.org/doi/abs/10.1073/pnas.2310076121.

U.S. Environmental Protection Agency. Ecoregions. https://www.epa.gov/eco-research/ecoregions, n.d. Accessed September 17, 2026.

Guido R. van der Werf, James T. Randerson, Dave van Wees, Yang Chen, Louis Giglio, Joanne Hall, Roland Vernooij, Mingquan Mu, Samiha Binte Shahid, Kelley C. Barsanti, et al. Landscape fire emissions from the 5th version of the global fire emissions database (GFED5). *Scientific Data*, 12 (1):1870, 2025.

Vladimir N. Vapnik. *Statistical Learning Theory*. Wiley, 1998.

Daoping Wang, Dabo Guan, Shupeng Zhu, Michael Mac Kinnon, Guannan Geng, Qiang Zhang, Heran Zheng, Tianyang Lei, Shuai Shao, Peng Gong, and Steven J. Davis. Economic footprint of California wildfires in 2018. *Nature Sustainability*, 4(3):252–260, 12 2020. ISSN 2398-9629. doi: 10.1038/s41893-020-00646-7. URL http://dx.doi.org/10.1038/s41893-020-00646-7.

Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully test-time adaptation by entropy minimization. In *International Conference on Learning Representations*, 2021.

S. S.-C. Wang, Y. Qian, L. R. Leung, and Y. Zhang. Interpreting machine learning prediction of fire emissions and comparison with firemip process-based models. *Atmospheric Chemistry and Physics*, 22(5):3445–3468, 2022. doi: 10.5194/acp-22-3445-2022. URL https://acp.copernicus.org/articles/22/3445/2022/.

Xiao Wu, Erik Sverdrup, Michael D. Mastrandrea, Michael W. Wara, and Stefan Wager. Low-intensity fires mitigate the risk of high-intensity wildfires in California's forests. *Science Advances*, 9(45): eadi4123, 2023. doi: 10.1126/sciadv.adi4123. URL https://www.science.org/doi/abs/10.1126/sciadv.adi4123.

A Derivations and ordering conditions

A.1 Binary-label RDPO: objective and analytical results

Yosemite applies binary-label RDPO to the label pair $(y, 1-y)$ for one input. With $s_y = 2y - 1 \in \{-1, +1\}$, its preference loss is

$$\ell_{\text{pref}}(z, r, y) = -\frac{1}{\beta} \log \sigma\left(\beta \left[\log \frac{p_\theta(y \mid x)}{p_{\text{ref}}(y \mid x)} - \log \frac{p_\theta(1-y \mid x)}{p_{\text{ref}}(1-y \mid x)}\right]\right). \tag{16}$$

The $1/\beta$ scale leaves the isolated preference-risk minimizer unchanged.

With uniform averages over retrieved rows, the evaluated baseline optimizes

$$\mathcal{L}_{\text{binary-RDPO}} = \mathbb{E}_{(x,y)\sim\mathcal{D}_{\text{local}}}[\ell_{\text{pref}}(z_\theta(x), r(x), y)] + \lambda_s M_{\text{loss}}. \tag{17}$$

We fix $\beta = 1$ and the preference coefficient to one, with no focal term and no extra positive-label preference term ($\lambda_2 = 0$). The mask coefficient is $\lambda_s = 0.05$; its realized term is zero.

A.1.1 Binary reduction and conditional minimizer

Proposition 3 (Binary reduction)**.** *Assume $\beta > 0$, $y \in \{0, 1\}$, and let $z, r \in \mathbb{R}$ be Bernoulli policy and reference logits, respectively. Then Equation 16 equals*

$$\ell_{\text{pref}}(z, r, y) = -\frac{1}{\beta} \log \sigma(\beta s_y (z - r)). \tag{18}$$

Its gradient and curvature with respect to z are

$$\frac{\partial \ell_{\text{pref}}}{\partial z} = -s_y \, \sigma(-\beta s_y (z - r)), \tag{19}$$

$$\frac{\partial^2 \ell_{\text{pref}}}{\partial z^2} = \beta \, \sigma(\beta s_y (z - r)) \, \sigma(-\beta s_y (z - r)) > 0. \tag{20}$$

With label-policy gaps $\Delta_\theta = s_y z_\theta(x)$ and $\Delta_{\text{ref}} = s_y r(x)$, this is Equation 3 at $\alpha = 1$. Uniform averaging and the absence of an extra positive-label term give

$$a_0 = a_1 = 1. \tag{21}$$

Allowing general fixed class coefficients makes the reference's effect on the conditional optimum explicit.

Proposition 4 (DPO-only population minimizer)**.** *Fix $a_0, a_1 > 0$ and $\beta > 0$. At a point x with local positive probability $0 < \eta(x) < 1$ and fixed reference $r(x)$, the conditional preference risk $R(z \mid x) = \eta a_1 \ell_{\text{pref}}(z, r, 1) + (1 - \eta) a_0 \ell_{\text{pref}}(z, r, 0)$ is strictly convex in z with unique minimizer*

$$z^*_{\text{pref}}(x) = r(x) + \frac{1}{\beta}\left[\operatorname{logit} \eta(x) + \log \frac{a_1}{a_0}\right]. \tag{22}$$

For the evaluated setting, $\beta = 1$ and $a_0 = a_1 = 1$, so $z^*_{\text{pref}}(x) = r(x) + \operatorname{logit} \eta(x)$. The objective is therefore a pointwise binary-classification preference loss rather than the positive–negative input-ranking loss used for California. It uses no focal term or additional positive-label weight. The reference contributes an input-dependent offset: a reference nondecreasing in η preserves ordering, whereas a sufficiently misordered reference can reverse it.

A.1.2 Weighted focal control

We retain a static weighted-focal control for cost-sensitive adaptation. It multiplies the aligned focal loss by fixed class weights

$$c_0 = 2, \qquad c_1 = 2 + 0.1/\pi_+. \tag{23}$$

These weights are fixed using retrieved-set prevalence and are independent of minibatch composition. This is a separate class-reweighting baseline, rather than a coefficient match to the pure binary-label RDPO objective (Appendix B).

A.2 Proof of Proposition 3

For a Bernoulli model with logit z,

$$\log \frac{p_\theta(1 \mid x)}{p_\theta(0 \mid x)} = z, \qquad \log \frac{p_\theta(0 \mid x)}{p_\theta(1 \mid x)} = -z. \tag{24}$$

The same identities hold for the reference logit r. If $y = 1$, the expression inside the square brackets of Equation 16 is $z - r$; if $y = 0$, it is $-(z - r)$. Both cases equal $s_y(z - r)$, which yields Equation 18. Let $u = \beta s_y(z - r)$. Since $d[-\log \sigma(u)]/du = -\sigma(-u)$, the chain rule gives

$$\frac{\partial \ell_{\text{pref}}}{\partial z} = -s_y \sigma(-u). \tag{25}$$

Differentiating again gives $\beta\sigma(u)\sigma(-u) > 0$, where strict positivity follows from $\beta > 0$ and finite z, r.

A.3 Proof of Proposition 4

Fix $a_0, a_1 > 0$ and $\beta > 0$. At a fixed x, write $d = z - r$ and $0 < \eta = \Pr_{Q_{\text{local}}}(Y = 1 \mid X = x) < 1$. The conditional preference risk is

$$R(d) = -\frac{1}{\beta} \left[\eta a_1 \log \sigma(\beta d) + (1 - \eta) a_0 \log \sigma(-\beta d)\right]. \tag{26}$$

Differentiating twice gives

$$R'(d) = -\eta a_1 \sigma(-\beta d) + (1 - \eta) a_0 \sigma(\beta d), \tag{27}$$

$$R''(d) = \beta \left[\eta a_1 + (1 - \eta) a_0\right] \sigma(\beta d)\sigma(-\beta d) > 0. \tag{28}$$

Hence R is strictly convex. Its unique stationary point, and therefore its unique global minimizer, satisfies $R'(d) = 0$. Set $q = \sigma(\beta d)$. Then $(1 - \eta)a_0 q = \eta a_1(1 - q)$, so

$$q = \frac{\eta a_1}{\eta a_1 + (1 - \eta) a_0}. \tag{29}$$

Taking odds and solving for d yields

$$d^* = \frac{1}{\beta} \left[\operatorname{logit} \eta + \log(a_1/a_0)\right], \tag{30}$$

which proves Equation 22.

A.4 Conditions for order preservation

For two inputs x_i, x_j with $0 < \eta(x_j) < \eta(x_i) < 1$ and the same $a_0, a_1 > 0$ and $\beta > 0$,

$$z^*_{\text{pref}}(x_i) - z^*_{\text{pref}}(x_j) = r(x_i) - r(x_j) + \frac{1}{\beta} \left[\operatorname{logit} \eta(x_i) - \operatorname{logit} \eta(x_j)\right]. \tag{31}$$

If $\eta(x_i) > \eta(x_j)$ and $r(x_i) \geq r(x_j)$, the reference difference is nonnegative and the posterior-log-odds difference is strictly positive, so the order is preserved. If the reference is misordered, a reversal occurs whenever

$$r(x_j) - r(x_i) > \frac{1}{\beta} \left[\operatorname{logit} \eta(x_i) - \operatorname{logit} \eta(x_j)\right]. \tag{32}$$

Equality gives a tie; strict inequality gives a reversal. The condition identifies how the frozen pretrained offset affects the ordering of the pointwise preference-risk minimizer.

A.5 Full-set versus minibatch normalization

The pure binary-label RDPO baseline averages its preference term over all rows in each minibatch; both label coefficients remain one. For fixed pointwise scores, an i.i.d. minibatch mean is an unbiased estimate of the retrieved-distribution preference risk. Training-mode batch normalization makes policy scores depend on the batch, so this identity does not establish the exact neural-network training risk. The weighted-focal control instead uses the fixed retrieved-set weights in Equation 23.

A.6 Scope of the conditional minimizer

Proposition 4 analyzes the preference component with fixed coefficients and unrestricted deterministic logits. The implemented baseline uses this preference loss without focal supervision or extra positive-label weighting; the TabNet mask term is zero. Shared parameters and batch normalization couple the scores, so the neural-network optimizer trajectory is outside the unrestricted pointwise result.

A.7 Proof of Proposition 1

For one candidate pair, write $u = \beta_{\mathrm{rank}}(\Delta_\theta - \alpha\Delta_{\mathrm{ref}})$ and hold the reference fixed. The chain rule gives

$$\frac{\partial \ell_\alpha}{\partial \Delta_\theta} = -\sigma(-u), \qquad \frac{\partial^2 \ell_\alpha}{\partial \Delta_\theta^2} = \beta_{\mathrm{rank}}\sigma(u)\sigma(-u) > 0. \tag{33}$$

Substituting score-equal initialization, $\Delta_\theta = \Delta_{\mathrm{ref}}$, gives Equation 6. Differentiating with respect to α gives the displayed derivative, whose sign equals the reference gap. Therefore the initial direct-ranking gap gradient is greater than $1/2$ on wrong pairs, less than $1/2$ on correct pairs, and equal to $1/2$ on ties; RDPO gives $1/2$ for all pairs. The per-pair parameter gradient is

$$\nabla_\theta \ell_\alpha = -g_\alpha\big[\nabla_\theta s_\theta(x^+) - \nabla_\theta s_\theta(x^-)\big]. \tag{34}$$

Different pair Jacobians and cancellation in their average prevent treating the scalar factor as a universal parameter-gradient norm comparison. Likewise, adaptive optimizers need not take proportionally larger steps.

A.8 Selective target and gradient allocation

For $\Delta_{\mathrm{ref}} \le 0$, $[\Delta_{\mathrm{ref}}]_+ = 0$ and selective equals direct ranking. For $\Delta_{\mathrm{ref}} > 0$, selective equals RDPO. The target has no policy derivative because the reference is detached. Under score-equal initialization its gradient factor is

$$g_{\mathrm{sel}}^{(0)} = \begin{cases} \sigma(-\beta_{\mathrm{rank}}\Delta_{\mathrm{ref}}), & \Delta_{\mathrm{ref}} \le 0, \\ 1/2, & \Delta_{\mathrm{ref}} > 0. \end{cases} \tag{35}$$

Figure 4 illustrates this analytic allocation. For any of these one-sided per-pair losses, the derivative remains negative at every finite gap. The isolated surrogate favors increasing the gap rather than imposing a hard condition $\Delta_\theta > \tau$. In a shared neural network, finite training, competing pairs, and regularization can prevent crossing zero or retaining a previously correct order. This distinction is visible in the correction/retention diagnostics.

A.9 Fire-prone location prioritization under a daily budget

For date t, let P_t and N_t be the numbers of positives and negatives, let $T_{\theta,t}^B$ be the positives among the B highest-scored cells, and define the avoidable misses relative to a label oracle as $q_t^B = \min(B, P_t) - T_{\theta,t}^B$. Let $\mathcal{P}_t = \{(i,j) : y_{t,i} = 1,\ y_{t,j} = 0\}$ contain all same-day positive–negative pairs. For pair-dependent targets τ^{ij}, write $\ell_\tau^{ij} = \ell_{\tau^{ij}}(\Delta_\theta^{ij})$. Direct and selective ranking use the nonnegative targets 0 and $[\Delta_{\mathrm{ref}}^{ij}]_+$, respectively.

Proposition 5 (Nonnegative-target daily-budget certificate)**.** *For $P_t > 0$ and any targets $\tau^{ij} \ge 0$,*

$$\frac{q_t^B}{P_t} \le \frac{\beta_{\mathrm{rank}}}{P_t \log 2} \sum_{(i,j)\in\mathcal{P}_t} \ell_\tau^{ij}. \tag{36}$$

Thus the bound applies to both direct and selective ranking. Selective ranking additionally satisfies

$$\ell_{\mathrm{sel}}^{ij} \le \frac{\log 2}{\beta_{\mathrm{rank}}} \quad \Longrightarrow \quad \Delta_\theta^{ij} \ge [\Delta_{\mathrm{ref}}^{ij}]_+. \tag{37}$$

Each avoidable miss can be matched to a selected negative that outranks an unselected positive, giving $\Delta_\theta^{ij} \le 0$. For any $\tau^{ij} \ge 0$, every such inversion has $\ell_\tau^{ij} \ge \log 2/\beta_{\mathrm{rank}}$; summing over the matched pairs proves Equation 36. The second statement follows from monotonicity of softplus.

The Top-B certificate therefore applies to both direct and selective ranking. Selective's additional certificate says that sufficiently low loss removes a negative ordering on a pretrained inversion and retains the positive reference margin on a correctly ordered pair. RDPO can meet a negative target without correcting an inversion, whereas direct ranking does not retain a positive reference margin.

Proposition 6 (Weighting induced by the minibatch sampler)**.** *Fix batch-independent per-example scores and suppose minibatches are exchangeable among examples sharing a date and label. For dates with $P_t N_t > 0$, define the full-pair mean*

$$\overline{L}_{\tau,t} = \frac{1}{P_t N_t} \sum_{(i,j)\in\mathcal{P}_t} \ell_\tau^{ij} \tag{38}$$

and let

$$w_t = \mathbb{E}_{\mathcal{B}}\left[\frac{\mathbf{1}\{t \in \mathcal{T}_{\mathcal{B}}\}}{|\mathcal{T}_{\mathcal{B}}|}\right], \tag{39}$$

where the ratio is zero when $\mathcal{T}_{\mathcal{B}}$ is empty. Uniform within-class pair draws in Equation 13 satisfy

$$\mathbb{E}_{\mathcal{B},\,\text{pair draws}}[\mathcal{L}_{\text{pair}}(\mathcal{B};\tau)] = \sum_t w_t \overline{L}_{\tau,t}. \tag{40}$$

For any target date weights π_t with $w_t > 0$ whenever $\pi_t > 0$, direct and selective ranking consequently satisfy

$$\sum_t \pi_t \frac{q_t^B}{P_t} \le \frac{\beta_{\text{rank}}}{\log 2}\left(\max_{t:\pi_t>0} \frac{\pi_t N_t}{w_t}\right) \mathbb{E}_{\mathcal{B},\,\text{pair draws}}[\mathcal{L}_{\text{pair}}(\mathcal{B};\tau)]\,. \tag{41}$$

Exchangeability makes every full-date positive–negative pair receive the same expected coefficient, while the probability that a date is eligible and the normalization by $|\mathcal{T}_{\mathcal{B}}|$ produce w_t. Equation 41 exposes the mismatch with uniformly weighted macro daily recall: the constant depends on both date eligibility and N_t and can be large. It is a finite-sample statement for the training dates, not a population guarantee for an unseen year. Training-mode batch normalization also makes a score depend on the other examples in its realized batch; in that setting Equation 40 characterizes the sampler's induced weights but does not identify SGD with optimization of one fixed full-pair empirical loss.

A.10 Sensitivity to the pairwise target

All members share the same pair supervision and differ only through the target $\alpha\Delta_{\text{ref}}$, so they separate only insofar as that target is appreciable on the scale set by β_{rank}.

Proposition 2 limits differences in scalar gap-gradient weights when $\beta_{\text{rank}}|\Delta_{\text{ref}}|$ is small. The selective attenuation bound additionally restricts its departure from RDPO to reference-wrong pairs. If there are no such pairs, its departure is zero; otherwise the truncated moment equals their probability times the conditional mean absolute gap. These statements concern loss anchoring, not the pretrained initialization shared by all variants. They do not establish equivalent parameter updates, stopping epochs, or ranking performance. Negative RDPO targets can be exceeded while the final pair remains misordered; all these logistic losses nevertheless continue to favor larger gaps.

Proposition 2 controls the magnitude of scalar-weight differences, while Proposition 1 describes their allocation across pair strata. Larger scaled pretrained margins allow larger differences between objectives. Task-level benefits are evaluated separately through daily AP and top-5% recall; the bounds do not predict performance.

A.11 Proof of Proposition 2 and the measured margin scale

From Proposition 1, $\partial g_\alpha/\partial\alpha = \beta_{\text{rank}}\Delta_{\text{ref}} g_\alpha(1-g_\alpha)$ for a detached reference. Since $g(1-g) \le 1/4$ for $g \in (0,1)$, the mean value theorem gives Equation 7. Thus the scalar gradient factors lie in a band of width $\beta_{\text{rank}}|\Delta_{\text{ref}}|/4$ around any one member. This bound does not by itself constrain final ranking gains. At score-equal initialization $g_\alpha^{(0)} = \sigma[-\beta_{\text{rank}}(1-\alpha)\Delta_{\text{ref}}]$, so $g_0^{(0)} - g_1^{(0)} = \sigma(-\beta_{\text{rank}}\Delta_{\text{ref}}) - 1/2$; expanding $\sigma(-u) = 1/2 - u/4 + u^3/48 + O(u^5)$ gives the stated cubic

Table 3: Frozen pretrained score-gap scale, measured on held-out positive–negative pairs (cross-date for Yosemite, same-date on the fixed 20% diagnostic sample for California) and averaged over seeds. $g_\alpha^{(0)}$ is the score-equal-initialization gradient weight of Equation 6; RDPO fixes it at $1/2$ for every pair, and rel. expresses the direct–RDPO separation as a fraction of that $1/2$. Proposition 2 bounds each mean separation by $\beta_{\text{rank}}\mathbb{E}|\Delta_{\text{ref}}|/4$. These are scalar gradient weights at a single model state: they bound how far the targets can differ, not whether any difference improves ranking. Table 4 repeats the diagnostic on sampled-query diagnostic batches without assuming score equality.

Setting	β_{rank}	$\mathbb{E}\lvert\Delta_{\text{ref}}\rvert$	$\beta_{\text{rank}}\mathbb{E}\lvert\Delta_{\text{ref}}\rvert$	Ref. pair acc. (%)	$\mathbb{E}\lvert g_0^{(0)}-g_1^{(0)}\rvert$	rel. (%)	$\mathbb{E}\lvert g_{\text{sel}}^{(0)}-g_1^{(0)}\rvert$
Yosemite	1	0.77	0.77	70.3	0.1568	31.4	0.0320
California	0.1	1.01	0.10	76.8	0.0253	5.1	0.0026

remainder. For the selective target $g_{\text{sel}}^{(0)} - g_1^{(0)} = [\sigma(-\beta_{\text{rank}}\Delta_{\text{ref}}) - 1/2]_+$, which vanishes whenever $\Delta_{\text{ref}} \geq 0$; bounding the surviving term by $\beta_{\text{rank}}|\Delta_{\text{ref}}|/4$ and taking expectations over pairs yields Equation 8. The reference-wrong event restricts selective's departure from RDPO; when its probability is zero, the indicator-moment bound is zero without an undefined conditional expectation.

The asymmetry between the two endpoints is also structural. For $\Delta_{\text{ref}} \ll 0$ the RDPO condition $\Delta_\theta > \Delta_{\text{ref}}$ is satisfied at values with $\Delta_\theta < 0$, so exceeding the RDPO target alone does not certify a corrected ordering; the logistic loss still favors further increases. Direct ranking retains the pretrained parameters and representation but omits the frozen pretrained model margin from its loss, including on pairs the reference already orders correctly. The selective target applies a zero target to reference-misranked pairs and retains positive reference margins elsewhere. These targets specify training emphasis; held-out pair accuracy and top-ranked event recall assess the resulting decisions.

Table 3 evaluates these quantities on the frozen pretrained model using held-out positive–negative pairs: all $1{,}060{,}686$ cross-date pairs per seed for Yosemite, matching its cross-date training pairs, and all $355{,}227$ same-date pairs per seed from the fixed 20% California diagnostic sample. This frozen-reference diagnostic is independent of the adaptation loss; all adapted-policy results use the complete grid. Yosemite logits are recovered from the stored probabilities. These are evaluation pairs, so they measure the reference's margin *scale* rather than the exact training-pair distribution drawn from $\mathcal{D}_{\text{local}}$; no quantity here is fitted. California's measured $\mathbb{E}|g_0^{(0)} - g_1^{(0)}| = 0.0253$ is close to the Equation 7 bound $\beta_{\text{rank}}\mathbb{E}|\Delta_{\text{ref}}|/4 = 0.0253$ at the reported precision, consistent with a near-linear regime; at nonzero margins the Lipschitz bound is generally strict, so the displayed equality reflects rounding. Yosemite's 0.157 falls further below its bound of 0.193 as the sigmoid begins to saturate. Sweeping β_{rank} on the same references, the relative separation $\mathbb{E}|g_0^{(0)} - g_1^{(0)}|/(1/2)$ rises from 5.1% at $\beta_{\text{rank}} = 0.1$ to 12.4% at 0.25, 23.2% at 0.5, 39.3% at 1 and 78.4% at 5. All percentages in this subsection are seed means, matching the table. A β_{rank} sweep would have to span roughly this range for the variants to differ appreciably, but changing β_{rank} rescales both losses rather than isolating the reference term, so direct ranking and RDPO require the same tuning opportunity for an interpretable comparison.

Real training batches without the score-equality assumption. Table 3 assumes policy and reference scores coincide at initialization. They do not: the policy runs training-mode and the frozen pretrained model evaluation-mode batch normalization, so $\Delta_{\theta,0} \neq \Delta_{\text{ref}}$ even though the parameters are copied. Table 4 therefore recomputes $g_{\text{direct}} = \sigma(-\beta_{\text{rank}}\Delta_{\theta,0})$, $g_{\text{RDPO}} = \sigma[-\beta_{\text{rank}}(\Delta_{\theta,0} - \Delta_{\text{ref}})]$ and the selective analogue on California diagnostic batches from the original sampled-query retrieval set and pair draws at initialization, pooling 25 batch draws ($3{,}751$ pairs per seed) with the parameters held at initialization. The realized departure is substantial, $\mathbb{E}|\Delta_{\theta,0} - \Delta_{\text{ref}}| = 0.28$, yet at the trained $\beta_{\text{rank}} = 0.1$ the realized mean separations, 0.0237 for direct–RDPO and 0.0013 for selective–RDPO, match their score-equal counterparts to the reported precision. The score-equal approximation thus matches the realized mean separations at the trained setting. The 99th percentile of $|g_{\text{direct}} - g_{\text{RDPO}}|$ is 0.097, showing larger differences on some pairs. Pretrained pair accuracy is 83.2% on retrieved training pairs and 76.8% on held-out pairs, which helps explain the smaller selective–RDPO separation in training. These measurements describe pairwise initialization-level scalar weights, with no focal gradient or parameter update; they do not report trained-policy performance for either retrieval setting.

Table 4: Scalar pairwise gradient weights on California diagnostic batches from sampled-query retrieval at initialization; only pairwise gradients are measured. The adapted model uses training-mode batch normalization and the pretrained model evaluation mode. Measurements pool 25 batch draws (3,751 pairs per seed) and average model seeds 11, 22, 33; no parameters are updated. Although $\mathbb{E}|\Delta_{\theta,0} - \Delta_{\text{ref}}| = 0.28$, realized mean separations match the score-equal approximation at the reported precision. † marks the trained β_{rank}. The sweep describes scalar-weight differences; changing β_{rank} reshapes both losses.

	$\lvert g_{\text{direct}} - g_{\text{RDPO}}\rvert$		$\lvert g_{\text{sel}} - g_{\text{RDPO}}\rvert$		Score-equal	Bound
β_{rank}	mean	p99	mean	p99	mean	$\beta_{\text{rank}}\mathbb{E}\lvert\Delta_{\text{ref}}\rvert/4$
0.1†	0.0237	0.0974	0.0013	0.0203	0.0237	0.0238
0.25	0.0580	0.2297	0.0033	0.0503	0.0580	0.0594
0.5	0.1095	0.3968	0.0065	0.0992	0.1091	0.1188
1	0.1881	0.5982	0.0126	0.1880	0.1862	0.2377
2	0.2829	0.8180	0.0226	0.3297	0.2782	0.4754
5	0.3928	0.9846	0.0388	0.5872	0.3874	1.1884

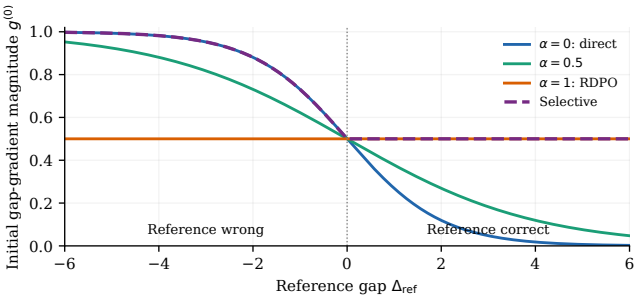

Figure 4: Analytic initial gap-gradient magnitudes for Equation 3 and selective ranking, assuming equal initial scores and $\beta_{\text{rank}} = 1$. Curves characterize the loss under the stated initialization assumption.

A.12 Pretrained score structure and scale

The relevant pair distribution is the one induced by the candidate task: for daily prioritization, pairs should share a date. Reference accuracy can be summarized by $\Pr(\Delta_{\text{ref}} > 0) + \frac{1}{2}\Pr(\Delta_{\text{ref}} = 0)$, supplemented by performance near the inspection cutoff and margin diagnostics stratified by correct, wrong, and tied pairs. Aggregate AUROC does not fix those conditional margin distributions. Replacing r by $ar + b$ for $a > 0$ leaves its ordering unchanged but replaces $\beta_{\text{rank}}\Delta_{\text{ref}}$ by $a\beta_{\text{rank}}\Delta_{\text{ref}}$. Matching AUROC therefore does not control the reference-dependent objective or its gradients. With score-equal initialization, changing the common initial policy/reference score scale also changes $g_\alpha^{(0)}$ for $\alpha < 1$.

B Experimental configuration and metric verification

B.1 Yosemite data and architecture details

The full Yosemite file has 70,491 rows from 1997–2022 and 9,219 positives. The nominal protocol starts in 2001, but no 2001 rows exist, so effective training begins in 2002. The twelve inputs are maximum/minimum temperature and relative humidity, precipitation, wind speed/direction, vapor-pressure deficit, energy-release component, burning index, and 100-hour/1000-hour fuel moisture. Labels satisfy the supplied smoothed $\text{DM} > 0$; coordinates, time, dry matter, and carbon are excluded from inputs. Standardization fits only historical training rows. Retrieval uses the complete unlabeled target batch.

The pretrained TabNet has widths 8, four steps, $\gamma = 1.3$, sparsemax, two independent/shared gated linear units, and virtual batch size 128. It uses 50 epochs, Adam at 5×10^{-3}, and batches of 4096. Historical F1 selects each model's classification threshold.

B.2 FROM-SCRATCH JOINT FOCAL AND DIRECT RANKING

The joint baseline initializes a fresh TabNet with the same architecture and seed as the pretrained focal model and trains on the complete historical window. Its objective is

$$\mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{focal}} + \mathbb{E}_{(x^+,x^-)} \operatorname{softplus}(-\Delta_\theta) + 0.05 M_{\text{loss}}. \tag{42}$$

Focal supervision uses $\gamma = 2$ and positive-class weight $1 - \pi_+$, where π_+ is historical training prevalence. Direct ranking uses coefficient one and $\beta_{\text{rank}} = 1$, with every positive–negative pair in each shuffled minibatch, including cross-date pairs. We retain the pretraining Adam optimizer at 5×10^{-3}, batch size 4096, 50 epochs, and final checkpoint. The mask term remains zero. No reference scores, pretrained policy weights, or target-dependent retrieval enter joint optimization. Historical training F1 chooses its classification threshold; target labels enter only final evaluation.

This tests one global joint-training configuration against subsequent local adaptation. Training data, optimizer, initialization, and update counts differ from the local policies, so the comparison does not isolate initialization alone or match total computation. The from-scratch joint baseline uses the same training schedule as the pretrained focal model; the local methods additionally perform 100 adaptation epochs. The initial-gradient analysis assumes pretrained initialization and therefore does not describe the joint baseline's random starting point. The rolling 2021 results reuse the corresponding joint executions after verifying identical data windows and reference comparison checkpoints.

B.3 PAIRWISE POLICY DETAILS

All three local ranking configurations use $\beta = 1$, preference coefficient one, and no focal supervision during adaptation. Direct and selective ranking use every positive–negative input pair within a shuffled minibatch, including cross-date pairs. Binary-label RDPO compares the true and complementary labels for each input, with no extra positive-label term ($\lambda_2 = 0$). Each retains the same $+0.05 M_{\text{loss}}$ term, whose realized value is zero. The configurations share pretrained checkpoints, retrieved rows, seeded minibatches, AdamW settings, and 100 adaptation epochs. Their candidate definitions differ, so this comparison does not isolate only the cross-input reference target. No interpolation sweep is evaluated.

The weighted-focal control retains fixed weights $c_0 = 2$ and $c_1 = 2 + 0.1/\pi_+$ on the aligned focal loss. It is a separate static reweighting baseline; it does not match the pure binary-label RDPO coefficients.

Gradient matching is an auxiliary scaling control. The actual TabNet mask term is zero in these experiments, and AdamW approximately cancels uniform loss scaling; near-identical fixed/matched results should therefore not be overinterpreted as a broad robustness result.

B.4 ROLLING EVALUATION AND METRIC COMPUTATION

The rolling aggregate and daily-ranking tables report means $\pm$ 95% Student-t seed-CI half-widths. For rolling, average the three test years separately within each seed, then compute $t_{0.975,n-1} s/\sqrt{n}$ across $n = 3$ seed averages; the 2021 daily table uses $n = 5$ seed scores.

In 2021, 57 dates contain positives, of which 42 have both labels. With nine candidates per day, $\lceil 0.05 \times 9 \rceil = 1$: the nominal daily Top5% metric is Top1 at an effective selection fraction of 11.11%. Mixed-day AP excludes all-positive dates. We report seed variability and paired directions without row-level significance tests.

B.5 FIXED CONFIGURATIONS

The 2021 local set has 11,187 unique historical rows and 1,582 positives ($\pi_+ = 0.1414$), identical across policies and seeds. The TabNet group-attention matrix contains one all-ones group, yielding a zero mask-loss term throughout training. The sparsity term enters the objective with coefficient $+0.05$.

Table 5: Shared Yosemite configuration and objective-specific differences.

Component	Value
Features / pretrained model	12 weather variables / TabNet, $n_d = n_a = 8$, 4 steps
Pretraining objective / optimizer	focal, $\gamma = 2$ / Adam, lr 5×10^{-3}
Pretraining epochs	50
Retrieval	training-standardized Euclidean KNN, $k = 5$
Local optimizer / epochs / batch	AdamW, lr 10^{-4} / 100 / 4096
binary-label RDPO $\beta, \lambda_1, \lambda_2$	$1.0, 1.0, 0.0$
Pairwise ranking β_{rank}	1.0
Ranking adaptation focal term	absent for direct, selective, and binary-label RDPO
Sparsity term	$+0.05M_{\mathrm{loss}}$ (zero in these runs)
Main / rolling seeds	$\{11, 17, 23, 29, 37\}$ / $\{11, 17, 23\}$
Checkpoint / threshold	final epoch / historical-training F1 optimum
Joint initialization / training set	random / all historical rows
Joint optimizer / epochs / coefficients	Adam, lr 5×10^{-3} / 50 / focal 1, direct ranking 1
Direct / fixed selective coefficient	1.0 / 1.0
Selective target	$[\Delta_{\mathrm{ref}}]_+$, detached, no cap

B.6 Pair identity and gradient matching

The direct and selective runs use exactly the same retrieved historical examples and seeded minibatch orders. Every minibatch contributes the full Cartesian product of its positive and negative rows; empty-class batches would be skipped. Both methods have the same per-epoch pair counts and no skipped batches. Selective reference logits are cached in evaluation mode and detached. Policy training mode retains the original batch-normalization behavior.

For each seed, an independent copy reproduces the first actual training batch without advancing the training shuffle generator. Let G_{direct} and G_{sel} be the gradients of the unweighted pairwise terms with respect to all trainable parameters. The matched coefficient is

$$\lambda_{\mathrm{match}} = \frac{\|G_{\mathrm{direct}}\|_2}{\|G_{\mathrm{sel}}\|_2}. \tag{43}$$

The two norms are finite and nonzero, and the direct coefficient is one. The initial weighted norm then agrees exactly, but later gradients and optimization trajectories need not agree. Table 6 reports measured norms. Because the other loss term is zero, this control largely scales the entire objective; AdamW's normalization makes its predictions almost identical to fixed-coefficient selective. The control therefore addresses initial scaling rather than broader objective sensitivity.

B.7 Evaluation and reuse

Each classification threshold maximizes F1 over all historical training predictions. For global recall at fraction q, the model selects $\lceil qm \rceil$ target grid–days and divides captured positives by all target positives. Daily diagnostics instead select $\max\{1, \lceil 0.05|\mathcal{C}_t| \rceil\}$ cells on each date; stable score ties follow original candidate-row order, without labels. Daily AP uses exact score-tie handling and excludes zero-positive dates; mixed-day AP further excludes all-positive dates. Reference-state diagnostics use strict positive/negative reference gaps; tied pairs are counted separately in the diagnostics.

The reused rolling 2021 cases have identical frozen pretrained models, adapted model states, retrieved examples, data partitions, and hyperparameters to their corresponding main executions. The five-seed and rolling summaries therefore contain overlapping 2021 cases. Classification, daily-ranking, and pair-stratum metrics are independently recomputed from final predictions. All training losses and prediction scores are finite; no margin cap is used.

Table 6: Initial ranking parameter-gradient norms on the first actual minibatch. Matched coefficient: $\|G_{\text{direct}}\|_2/\|G_{\text{selective}}\|_2$. Rolling 2021 reuses the identical main executions.

Year	Seed	$\|G_{\text{direct}}\|_2$	$\|G_{\text{selective}}\|_2$	Coefficient	Pairs
2021	11	0.1702	0.5082	0.3350	2000943
2021	17	0.1896	0.4687	0.4045	2155120
2021	23	0.1409	0.4714	0.2989	2012775
2021	29	0.2067	0.5316	0.3888	1938300
2021	37	0.1939	0.6093	0.3181	2045148
2019	11	0.2927	1.3228	0.2213	1688415
2019	17	0.2834	1.1193	0.2532	1763823
2019	23	0.2955	1.1031	0.2679	1865628
2020	11	0.1364	0.4049	0.3369	1997980
2020	17	0.1273	0.3217	0.3957	2051008
2020	23	0.1383	0.3267	0.4235	1868679

Table 7: Daily ranking: mean $\pm$ 95% Student-t seed-CI half-width. Use five seeds for 2021; for rolling, average three years within each of three seeds. CIs condition on these test years. AP excludes zero-fire days; mixed AP also excludes all-positive days. Top1 is one of nine cells (11.11% budget). All values are percentages.

Experiment	Method	Daily AP	Mixed AP	R@Top1	NDCG@Top1
2021	Pretrained	77.24 ± 1.54	69.11 ± 2.09	9.24 ± 1.24	59.65 ± 6.35
	Joint focal + direct	77.40 ± 0.71	69.33 ± 0.97	9.04 ± 0.68	58.95 ± 3.96
	Direct ranking	77.38 ± 1.37	69.30 ± 1.86	9.24 ± 1.09	60.00 ± 6.43
	binary-label RDPO	**78.12** ± 0.27	**70.31** ± 0.37	**10.03** ± 0.89	**64.21** ± 5.68
	Selective ranking	77.78 ± 0.93	69.85 ± 1.26	9.39 ± 0.91	60.70 ± 5.47
Rolling	Pretrained	81.68 ± 3.58	73.10 ± 4.55	11.68 ± 2.19	72.20 ± 9.73
	Joint focal + direct	80.80 ± 3.44	72.23 ± 2.97	11.50 ± 1.07	71.37 ± 5.87
	Direct ranking	**82.56** ± 3.49	72.96 ± 3.69	11.95 ± 0.58	74.09 ± 2.35
	binary-label RDPO	81.97 ± 5.69	**73.41** ± 5.91	**11.99** ± 1.83	**74.52** ± 8.90
	Selective ranking	82.52 ± 5.59	72.97 ± 5.30	11.68 ± 1.10	72.97 ± 5.59

C Additional Yosemite results

The Yosemite result tables use a fixed 9pt font. Bold marks the best displayed mean in each column, including ties, within each experiment group (and pair scope where applicable). Larger is better except for Miss, where smaller is better.

C.1 Daily risk-ranking diagnostics

Table 7 reports daily candidate-ordering metrics with 95% Student-t confidence intervals over seeds. For 2021 there are five seeds; rolling intervals use the three per-seed averages across the fixed 2019–2021 test years. These intervals describe training randomness conditional on the evaluated data, rather than uncertainty over future years. Daily AP and Top1 recall complement pooled ranking by assessing the within-date candidate task.

Table 8: 2021 classification results, five seeds. Mean ± SD; percentages. Miss is one minus recall.

Method	Accuracy	Recall	F1	Miss
Pretrained	71.09 ± 2.15	57.80 ± 5.70	30.61 ± 0.75	42.20 ± 5.70
Joint focal + direct	$\mathbf{72.12} \pm 1.75$	52.73 ± 6.77	29.39 ± 1.77	47.27 ± 6.77
Continued focal	70.59 ± 1.63	57.02 ± 6.80	29.93 ± 1.91	42.98 ± 6.80
Weighted focal	71.09 ± 1.35	56.09 ± 6.07	29.94 ± 1.56	43.91 ± 6.07
Direct ranking	70.67 ± 0.83	60.50 ± 2.37	$\mathbf{31.31} \pm 1.11$	39.50 ± 2.37
binary-label RDPO	69.60 ± 2.15	61.27 ± 5.86	30.79 ± 1.41	38.73 ± 5.86
Selective ranking	70.19 ± 1.40	$\mathbf{61.32} \pm 3.50$	31.25 ± 0.86	$\mathbf{38.68} \pm 3.50$

Table 9: Rolling classification results. Means over nine year–seed cases; percentages.

Method	Accuracy	Recall	F1	Miss
Pretrained	74.63	56.01	29.29	43.99
Joint focal + direct	**75.36**	50.97	28.16	49.03
Continued focal	73.50	59.05	29.36	40.95
Weighted focal	73.63	58.57	29.35	41.43
Direct ranking	72.52	64.61	**30.60**	35.39
binary-label RDPO	70.83	60.24	28.17	39.76
Selective ranking	72.19	**65.29**	30.55	**34.71**

C.2 Complete classification metrics

Tables 8 and 9 supplement the global ranking tables with accuracy, recall, F1, and miss rates. Each model uses its own historically selected F1 operating point. These threshold metrics complement the ranking comparisons at the same selected fraction.

C.3 Complete rolling summaries

Table 10 reports annual recall for the seven policies as mean ± SD over three seeds per year. Table 11 preserves the numerical comparisons shown in Figure 2, including 95% seed confidence intervals for all six metrics.

C.4 Year-specific ranking and paired differences

Direct and selective ranking have similar rolling pooled means: AUROC of 74.19% and 74.32%, AUPRC of 24.30% and 24.35%, and daily AP of 82.56% and 82.52%, respectively. Tables 12–15 give annual AUROC, AUPRC, R@20, and R@30. Each cell reports the three-seed mean ± SD. Target sets contain 92 positives in 2019, 540 in 2020, and 363 in 2021 among approximately 3,285 grid–days. Changing prevalence strongly affects AUPRC, so compare methods within a year. The overall rolling means weight each year–seed case equally. Selective improves AUROC over direct ranking in 7/9 cases and AUPRC in 6/9; the tables retain all annual comparisons.

C.5 Reference-correct and reference-wrong pairs

Table 17 conditions on reference gap sign, with equal-weight averages across cases. Pooled pairs match the cross-date training interpretation; same-day pairs assess the daily candidate task. More wrong-pair corrections can coexist with lower correct-pair retention. Margin expansion does not ensure crossing zero on a wrong pair, nor does aggregate correction ensure improved daily top-one recall.

Table 10: Recall by test year, three seeds per year. Mean ± SD; percentages. The 2021 column uses three of the five main seeds.

Method	2019	2020	2021
Pretrained	38.41 ± 9.37	74.07 ± 9.65	55.56 ± 6.75
Joint focal + direct	37.32 ± 7.86	61.79 ± 9.84	53.81 ± 5.82
Continued focal	46.01 ± 4.39	76.23 ± 9.24	54.91 ± 7.99
Weighted focal	45.29 ± 5.14	76.23 ± 9.56	54.18 ± 7.50
Direct ranking	51.09 ± 9.48	82.96 ± 2.77	59.78 ± 2.90
binary-label RDPO	48.91 ± 8.90	74.51 ± 4.39	57.30 ± 2.98
Selective ranking	**51.81** ± 12.69	**83.83** ± 3.80	**60.24** ± 4.43

Table 11: Rolling evaluation, nine year–seed cases. Mean ± 95% seed-CI half-width, conditional on these test years: average the three years within each seed, then use a Student-t interval over three seeds. Percentages; bold marks the largest displayed mean.

Method	Recall	F1	AUROC	AUPRC	R@20	R@30
Pretrained	56.01 ± 19.65	29.29 ± 4.71	73.12 ± 4.16	22.85 ± 6.82	43.03 ± 9.38	60.81 ± 13.20
Joint focal + direct	50.97 ± 9.48	28.16 ± 1.70	71.73 ± 0.97	21.43 ± 2.19	42.00 ± 5.91	57.06 ± 1.82
Continued focal	59.05 ± 16.11	29.36 ± 3.96	73.08 ± 3.97	22.08 ± 4.57	43.09 ± 7.28	60.32 ± 11.63
Weighted focal	58.57 ± 16.40	29.35 ± 3.31	73.06 ± 3.92	22.00 ± 4.55	42.93 ± 7.76	60.11 ± 11.45
Direct ranking	64.61 ± 12.17	**30.60** ± 2.32	74.19 ± 2.21	24.30 ± 2.60	45.41 ± 0.46	62.96 ± 6.63
binary-label RDPO	60.24 ± 12.70	28.17 ± 2.30	72.13 ± 3.19	22.05 ± 6.00	41.72 ± 9.17	57.00 ± 8.07
Selective ranking	**65.29** ± 16.92	30.55 ± 2.47	**74.32** ± 1.96	**24.35** ± 2.89	**45.79** ± 4.08	**63.26** ± 6.53

D California data and experimental protocol

D.1 Spatially smoothed occurrence and fixed subsampling

The pilot uses 666 GFED cells whose centers lie within California. Previous-day GridMET weather predicts the next day's label. Let $D_{t,j}$ denote the raw daily dry-matter estimate and d_{ij} Euclidean distance between cell centers in latitude/longitude degrees. Following the supplied Yosemite smoothing convention,

$$W_{ij} = \frac{\exp[-d_{ij}^2/(2\sigma_g^2)]\mathbf{1}\{d_{ij} \le c_g\}}{\sum_{j'} \exp[-d_{ij'}^2/(2\sigma_g^2)]\mathbf{1}\{d_{ij'} \le c_g\}}, \qquad Y_{t,i} = \mathbf{1}\left\{\sum_j W_{ij} D_{t,j} > 0\right\}, \tag{44}$$

with $\sigma_g = 0.25°$ and $c_g = 0.4°$. The support and normalization are restricted to the California grid. For nonnegative dry matter, the binary rule is exactly $Y_{t,i} = \mathbf{1}\{\exists j : d_{ij} \le c_g,\ D_{t,j} > 0\}$ in real arithmetic. Thus the cutoff determines binary support; the Gaussian bandwidth changes smoothed amounts but not this zero-threshold occurrence support. A positive label can arise from a neighboring positive cell within the cutoff; this target represents neighborhood-smoothed occurrence, not an independently measured ignition or an extreme-fire category. On the full 2022 grid, smoothing increases prevalence from 0.902% to 6.072%. Distances in degrees do not define a constant physical length across a continental area; a broader-area implementation should reassess the smoothing scale in an appropriate geographic metric before comparing labels across latitudes.

The final training period is 2002–2021 and testing is 2022. Uniform sampling without replacement within each historical year–label stratum retains 20% of training rows. This fixed seed-11 sample contains 972,892 rows, including 76,785 positives. Retrieval queries all 243,090 unlabeled target grid–days and evaluation covers all 666 cells on each of 365 dates. The 2022 evaluation contains 14,760 positives across 351 positive dates. Target outcomes are read for evaluation after the predictor-only

Table 12: AUROC by rolling test year, three seeds per year. Mean ± SD; percentages.

Method	2019	2020	2021
Pretrained	67.91 ± 2.00	81.33 ± 2.71	70.12 ± 0.60
Joint focal + direct	66.95 ± 2.12	78.88 ± 2.64	69.38 ± 1.41
Continued focal	68.32 ± 2.04	81.23 ± 1.58	69.69 ± 1.22
Weighted focal	68.37 ± 2.06	81.27 ± 1.62	69.54 ± 1.18
Direct ranking	69.09 ± 1.87	**82.63** ± 0.62	70.85 ± 0.46
binary-label RDPO	66.95 ± 1.73	78.90 ± 1.89	70.53 ± 0.41
Selective ranking	**69.13** ± 1.86	82.60 ± 0.77	**71.23** ± 0.32

Table 13: AUPRC by rolling test year, three seeds per year. Mean ± SD; percentages.

Method	2019	2020	2021
Pretrained	4.35 ± 0.30	43.69 ± 7.45	20.50 ± 0.65
Joint focal + direct	4.34 ± 0.50	39.98 ± 3.54	19.98 ± 0.92
Continued focal	4.43 ± 0.31	41.71 ± 4.07	20.09 ± 1.18
Weighted focal	4.44 ± 0.32	41.50 ± 4.19	20.05 ± 1.09
Direct ranking	4.52 ± 0.42	**46.08** ± 3.07	22.29 ± 0.31
binary-label RDPO	4.15 ± 0.24	40.53 ± 6.87	21.47 ± 0.49
Selective ranking	**4.56** ± 0.47	45.90 ± 3.78	**22.57** ± 0.60

candidate covariates are constructed; no target labels determine candidate membership or supervise adaptation.

D.2 Causal cell-history features

For prediction year Y, let $\mathcal{I}_{i,A}$ be the full available grid–days for cell i whose years belong to a past-year set A. Define $S_{i,A} = \sum_{t \in \mathcal{I}_{i,A}} Y_{t,i}$ and $N_{i,A} = |\mathcal{I}_{i,A}|$. With $A_Y = \{2002, \dots, Y-1\}$ (empty for $Y = 2002$) and $\kappa_A > 0$, the statewide past prevalence and shrunk rate are

$$\overline{\pi}_{<Y} = \begin{cases} \dfrac{\sum_i S_{i,A_Y}}{\sum_i N_{i,A_Y}}, & \sum_i N_{i,A_Y} > 0, \\ 0.01, & \sum_i N_{i,A_Y} = 0, \end{cases} \qquad \widehat{\pi}_{i,A} = \frac{S_{i,A} + \kappa_A \overline{\pi}_{<Y}}{N_{i,A} + \kappa_A}. \tag{45}$$

The all-history feature uses $A = \{2002, \dots, Y-1\}$ and the three-year feature uses its last three years, both with $\kappa_A = 365$ pseudo-days. For calendar month m, the monthly feature restricts the all-history counts to that month and uses 30 pseudo-days toward the same statewide all-month past prevalence. In 2002 the cold-start prior is 0.01. We also supply the history length $Y - 2002$ and seasonality $[\sin(2\pi\,\mathrm{DOY}(t)/365.25), \cos(2\pi\,\mathrm{DOY}(t)/365.25)]$.

These counts are constructed from full, spatially smoothed historical labels before subsampling, using only years strictly before each input's prediction year. They do not update with target-year outcomes.

D.3 Pretraining, retrieval, and adaptation

Terrain consists of elevation mean, standard deviation, minimum, maximum, and mean slope, aggregated from the approximately 4-km GridMET terrain grid to GFED cells. Cell centers receive EPA Level III categories by polygon membership, with nearest-polygon fallback when necessary. The category index is embedded in four dimensions inside the TabNet predictor. All other inputs are standardized using the appropriate historical pretraining period.

Table 14: R@20 by rolling test year, three seeds per year. Mean ± SD; percentages.

Method	2019	2020	2021
Pretrained	34.42 ± 5.47	52.90 ± 7.44	41.78 ± 1.24
Joint focal + direct	35.14 ± 8.79	50.19 ± 5.30	40.68 ± 3.13
Continued focal	35.87 ± 3.92	53.09 ± 2.65	40.31 ± 2.56
Weighted focal	36.23 ± 5.47	53.09 ± 3.24	39.49 ± 2.03
Direct ranking	**36.96** ± 1.09	55.37 ± 0.32	43.89 ± 0.89
binary-label RDPO	33.70 ± 3.92	48.40 ± 5.27	43.07 ± 2.14
Selective ranking	36.59 ± 3.82	**55.86** ± 1.54	**44.90** ± 1.46

Table 15: R@30 by rolling test year, three seeds per year. Mean ± SD; percentages.

Method	2019	2020	2021
Pretrained	55.43 ± 7.61	70.99 ± 6.88	56.01 ± 2.92
Joint focal + direct	50.36 ± 4.90	65.80 ± 5.87	55.00 ± 3.05
Continued focal	55.07 ± 6.64	71.17 ± 3.85	54.73 ± 3.98
Weighted focal	54.71 ± 6.99	71.36 ± 3.44	54.27 ± 3.86
Direct ranking	56.16 ± 6.99	75.43 ± 1.13	57.30 ± 1.46
binary-label RDPO	49.64 ± 3.14	65.25 ± 5.21	56.11 ± 1.57
Selective ranking	**56.52** ± 5.75	**75.49** ± 2.33	**57.76** ± 1.88

Let J contain continuous feature indices other than history length. The model-independent retrieval representation is

$$v(u) = \left[(u_J - \mu_J)/s_J^{\text{scale}},\ \text{onehot}(e)\right], \qquad d(u,u') = \|v(u) - v(u')\|_2. \tag{46}$$

The division is elementwise. Here μ_J and s_J^{scale} are the fitted training scaler's mean and positive scale values: a population-standard-deviation estimate (ddof zero), with scale one for constant or numerically constant fields. With 25 continuous fields and a four-dimensional categorical embedding, the predictor expands 26 raw fields to 29 internal inputs; retrieval omits history length and uses 24 continuous fields plus 13 one-hot categories (37 dimensions). Its one-hot contribution to squared distance is $2\mathbf{1}\{e \neq e'\}$. There is no learned embedding distance, regional restriction, distance threshold, or retrieval-weight tuning. The primary union of exact $k = 5$ neighbors over all 243,090 unlabeled target rows gives 363,643 unique historical rows with prevalence 0.075907. Within each of model seeds 11, 22, and 33, all ranking policies and continued focal use the same retrieved historical examples and frozen pretrained model. Sampled positive–negative pair identities agree across the three ranking methods over common epochs; different stopping epochs still imply different total update counts.

Pretrained models use the same TabNet widths/steps as Yosemite, Adam at 0.005, zero weight decay, and a 120-epoch spatial cap. The focal loss for a set with prevalence π is

$$\ell_{\text{focal}}(s,y) = -w_y(1-p_y)^2 \log p_y, \qquad p_y = \sigma((2y-1)s), \quad w_1 = 1-\pi,\ w_0 = \pi. \tag{47}$$

The continued-focal control recomputes these weights from the retrieved set. For direct, RDPO, and selective ranking, focal loss is detached and logged only; its coefficient is zero and it supplies no gradient. All local policies use AdamW at 10^{-4}, weight decay 0.01, and batches of 4096. The epoch cap is twice the actual pretraining epoch count, with a minimum 15 epochs. The stopping rule compares adjacent five-epoch means of training loss; the absolute relative change must be below 0.001 for two consecutive checks. It monitors pretraining focal loss or the local total objective. For ranking adaptation this is the pairwise objective plus the zero-valued mask term; the monitored focal loss does not select stopping epochs.

The final spatial pretrained models for seeds 11 and 33 reach the 120-epoch cap without a recorded plateau pass; seed 22 passes the rule. Direct ranking stops after 77/158/76 epochs for seeds 11/22/33,

Table 16: Paired selective-minus-direct differences (percentage points), calculated before rounding.

Experiment	Variant	Δ AUC	Δ AP	Δ R@20	Δ R@30	Δ Recall	Δ F1
2021	Selective ranking	**+0.27**	**+0.13**	**+0.66**	**+0.00**	**+0.83**	**-0.06**
Rolling	Selective ranking	**+0.13**	**+0.05**	**+0.38**	**+0.29**	**+0.68**	**-0.06**

Table 17: Held-out pair diagnostics, averaged over five seeds (2021) or nine year–seed cases (rolling). Corrected/retained means policy gap > 0; expanded means policy gap exceeds the reference gap on correct pairs. Reference ties are excluded. Percentages.

Experiment	Scope	Method	Wrong corrected	Correct retained	Correct expanded
2021	Pooled	Direct ranking	12.02	96.23	91.80
		Selective ranking	**14.45**	95.59	**92.28**
2021	Same day	Direct ranking	8.87	90.77	**81.42**
		Selective ranking	**10.59**	89.78	81.39
Rolling	Pooled	Direct ranking	16.71	95.67	90.34
		Selective ranking	**18.54**	95.13	**91.14**
Rolling	Same day	Direct ranking	18.65	88.38	72.89
		Selective ranking	**22.04**	87.07	**74.49**

RDPO after 126/158/64, and selective after 126/94/64. All nine ranking adaptations pass the local training-loss stopping rule. The single all-ones group-attention matrix makes the mask-loss term zero, so these experiments use the full input group rather than learning sparse spatial feature selection. Policy batch normalization is in training mode and pretrained-model batch normalization in evaluation mode, so matching initial weights need not produce identical scores.

The XGBoost baseline uses the same fixed sampled training rows. Its settings are 1,000 rounds, depth 6, learning rate 0.03, row and column sampling 0.8, minimum child weight 5, and L2 penalty 1. Its F1 threshold is selected on all sampled 2002–2021 training rows scored by the final model.

D.4 Retrieval and source-selection ablation

We first isolate retrieval under continued focal adaptation across model seeds 11, 22, and 33 on the complete 666-cell daily grid. Each policy starts from its seed-matched pretrained checkpoint and receives exactly 6,800 AdamW updates with batch size 4,096, learning rate 10^{-4}, and weight decay 0.01. The spatial-KNN and uniformly sampled random sets both contain 363,643 historical rows; full-archive adaptation uses all 972,892 rows. Fixing both selection size for the matched comparison and optimizer updates avoids confounding target-conditioned selection with data volume or compute. The frozen seed-matched sources receive no adaptation.

Spatial KNN improves daily AP over the matched random subset in every seed. The paired mean gain is 0.66 percentage points (95% t-interval $[0.29, 1.04]$); relative to the full archive it is 0.36 points ($[0.01, 0.71]$). Mean Top5% recall, global AP, and AUROC also favor spatial KNN over random, but their three-seed paired intervals include zero. Spatial adaptation does not uniformly improve on the frozen source: its mean daily-AP gain is 0.45 points, but seed 11 decreases by 0.53 points and the paired interval includes zero. Thus target-conditioned spatial selection has reproducible daily-AP value relative to matched random and full-history focal adaptation under this fixed budget, while the evidence does not establish a uniform improvement over no adaptation or across all metrics.

Table 18: California retrieval ablation under continued focal adaptation on 2022. Values are mean ± sample SD over seeds 11, 22, and 33. Each adapted policy starts from its seed-matched pretrained checkpoint and receives exactly 6,800 optimizer updates; the frozen source receives none. Random is sampled uniformly without replacement and exactly matches the spatial-KNN set size. Metrics use the complete daily grid and are percentages. Bold marks the best displayed value.

Historical data	**Rows**	**Daily AP**	**R@5%**	**Global AP**	**AUROC**
Frozen pretrained source	–	21.10±1.24	17.98±1.36	19.30±0.63	78.34±0.37
Full historical archive	972,892	21.19±0.40	18.21±0.96	19.43±0.41	78.42±0.33
Random matched subset	363,643	20.88±0.53	17.90±1.19	19.21±0.48	78.33±0.33
Spatial one-hot KNN, $k = 5$	363,643	**21.55±0.42**	**18.44±0.88**	**19.99±0.36**	**78.57±0.31**

Table 19: Complete-grid Top5% selection: 34 cells per date and 12,410 alarms per policy. Counts are neighborhood-smoothed positive cell–days, not independent incidents. Means over three seeds. Ranking adaptation has no focal term; paired gains use same-seed continued focal.

Method	FN	Captured	Paired gain
Continued focal	11894.00	2866.00	0.00
Direct ranking	11583.33	3176.67	310.67
RDPO	11593.33	3166.67	300.67
Selective ranking	**11549.67**	**3210.33**	**344.33**

D.5 Thresholds and independent metric checks

Final policy thresholds maximize F1 on all sampled 2002–2021 training rows scored by that policy. They are applied unchanged to the complete 2022 grid. No test outcomes select a threshold, checkpoint, or calibration map; no probability calibration is fitted. The saved sigmoid probabilities determine operating-point counts, while actual checkpoint logits determine all ranking metrics to avoid saturation ties.

The evaluation verifies common labels, dates, and coordinates, saved-probability agreement with recomputed logits, finite scores, and saved confusion counts. The reference checkpoint and retrieved-index hashes are checked for every ranking policy. Realized sampled preference-pair hashes match across the three objectives for the first 77, 94, and 64 epochs in seeds 11, 22, and 33, respectively. Different stopping epochs imply different total update counts. Tables report mean and sample SD over these three model seeds on the same test year; this variability concerns optimization randomness.

D.6 Pooled and daily evaluation metrics

Global AP and AUROC pool all 243,090 grid–days. Pooled R@q selects $\lceil qN \rceil$ highest raw scores across all dates and divides captured positives by 14,760, for $q = 0.20$ and 0.30, using the same partial-sort convention as Yosemite. Daily Top5%, Top20%, and Top30% instead select 34, 134, and 200 cells separately per date, with score ties ordered by latitude then longitude. Daily AP, recall, and NDCG average the 351 positive dates. Total FN counts uncaptured positive cell–days and is distinct from one minus macro daily recall. Threshold F1 is an auxiliary operating-point diagnostic.

For the raw-DM severity analysis in Figure 3, we restrict evaluation to the 2,193 cell–days with observed $D_{t,j} > 0$, order them by $D_{t,j}$, and take exactly $\lceil q\,2193 \rceil$ examples for $q \in \{0.05, 0.10, 0.20\}$. This gives 110, 220, and 439 cell–days, with boundary values 9.2034816×10^9, 3.885151744×10^9, and 2.177735936×10^9, respectively; none of these boundaries has a tie. Recall is the fraction whose model probability exceeds that policy's training-selected threshold. The plotted quantities are seed-wise recall differences from continued focal, summarized by their mean and sample SD. This diagnostic is retrospective: 2022 dry matter defines the severity strata but never enters model fitting, threshold selection, or deployment-time ranking.

Selective's Top5% advantage over RDPO is concentrated in seed 22, where the recall difference is 1.97 points; the other two differences are approximately zero and 0.07 points. This supports a small-budget prioritization result without a uniform advantage across seeds or budgets. The seed-22

Table 20: California 2022 complete-grid evaluation, 666 cells per date. Ranking adaptation excludes focal supervision. Mean ± sample SD over seeds 11, 22, 33; score metrics are percentages. R@20/R@30 pool raw scores across dates, matching Yosemite. Recall/F1 use each policy threshold selected on training data.

Method	AUPRC	AUROC	R@20	R@30	Recall	F1
XGBoost	11.02 ± 0.02	66.27 ± 0.04	38.38 ± 0.03	51.45 ± 0.06	31.54 ± 0.86	17.81 ± 0.07
Pretrained spatial focal	19.30 ± 0.63	78.34 ± 0.37	54.60 ± 0.61	68.55 ± 0.41	46.51 ± 1.58	26.34 ± 0.69
Continued focal	20.07 ± 0.34	78.60 ± 0.21	55.09 ± 0.64	68.70 ± 0.93	45.70 ± 1.77	$\mathbf{26.58} \pm 0.14$
Direct ranking	20.80 ± 1.53	79.46 ± 0.50	$\mathbf{56.41} \pm 0.92$	$\mathbf{70.51} \pm 0.91$	$\mathbf{54.80} \pm 5.07$	26.52 ± 0.31
RDPO	$\mathbf{20.89} \pm 1.39$	$\mathbf{79.47} \pm 0.47$	56.24 ± 0.91	70.42 ± 0.97	54.71 ± 5.26	26.42 ± 0.41
Selective ranking	20.71 ± 1.13	79.38 ± 0.33	56.16 ± 0.78	70.10 ± 0.42	54.20 ± 4.62	26.43 ± 0.43

Table 21: California 2022 complete-grid evaluation, 666 cells per date. Ranking adaptation excludes focal supervision. Mean ± sample SD over seeds 11, 22, 33; score metrics are percentages. Metrics average the 351 positive dates; daily selection uses 34 cells.

Method	Daily AP	R@Top5%	NDCG@Top5%
XGBoost	14.24 ± 0.07	11.64 ± 0.23	16.60 ± 0.26
Pretrained spatial focal	21.10 ± 1.24	17.98 ± 1.36	26.16 ± 2.08
Continued focal	21.62 ± 0.33	18.70 ± 0.74	27.06 ± 0.81
Direct ranking	24.35 ± 0.69	22.25 ± 1.39	31.05 ± 1.44
RDPO	24.47 ± 0.42	22.11 ± 1.92	31.15 ± 1.47
Selective ranking	$\mathbf{24.57} \pm 0.52$	$\mathbf{22.79} \pm 2.07$	$\mathbf{31.63} \pm 1.75$

selective and RDPO policies stop at 94 and 158 epochs, respectively; this comparison includes their different realized training durations.

D.7 Held-out pair strata and equal-budget misses

The pair audit enumerates all positive–negative pairs within each complete-grid 2022 date. The frozen seed-specific pretrained gap defines correct ($\Delta_{\mathrm{ref}} > 0$) and wrong ($\Delta_{\mathrm{ref}} < 0$) strata; ties are excluded from both. Correction and retention require $\Delta_\theta > 0$, and expansion requires $\Delta_\theta > \Delta_{\mathrm{ref}}$ on reference-correct pairs. These are evaluation diagnostics, not additional training supervision or guarantees on unseen pairs.

E Design for larger-area deployment: untested extensions

The California pilot combines spatial risk pretraining, target-conditioned retrieval, and ranking adaptation without focal supervision. The following untested design describes ways to adapt this pretrained foundation for larger-area deployment; the reported experiments use the architecture and protocol in Appendix D.

Partial pooling of spatial effects. A shared weather/landscape backbone can be complemented by a region residual,

$$r_{t,i} = f_0(w_{t-1,i}, a_i, h_{\mathrm{yr}(t),i}, d_t) + b_{e_i} + v_{e_i}^\top \psi(w_{t-1,i}, a_i), \tag{48}$$

where b_e represents regional baseline differences and v_e regional response differences. A penalty $\rho \sum_e (b_e^2 + \|v_e\|_2^2)$ with $\rho > 0$ partially pools poorly supported regions toward the shared model. This permits regularized regional intercept and response corrections without fitting independent predictors for every county. Their decomposition is generally nonunique when the shared model and regional effects are freely learned together. For an interpretable staged fit, hold f_0 and ψ fixed, center/normalize ψ on historical training rows, and impose $\sum_e n_e b_e = 0$ and $\sum_e n_e v_e = 0$ with fixed support weights $n_e > 0$. These conventions do not establish a causal interpretation. Neither this explicit head nor its penalty is implemented in the current TabNet embedding model. Unknown-region effects default to the shared component; they cannot be learned from a region's future labels.

Table 22: California 2022 complete-grid evaluation, 666 cells per date. Ranking adaptation excludes focal supervision. Mean ± sample SD over seeds 11, 22, 33; score metrics are percentages. Daily budgets select 34, 134, and 200 cells, respectively; recall averages positive dates.

Method	R@Top5%	R@Top20%	R@Top30%
Direct ranking	22.25 ± 1.39	**56.82** ± 1.47	**71.49** ± 0.66
RDPO	22.11 ± 1.92	56.41 ± 1.34	71.27 ± 0.74
Selective ranking	**22.79** ± 2.07	56.21 ± 1.36	71.22 ± 0.65

Table 23: California 2022 complete-grid evaluation, 666 cells per date. Ranking adaptation excludes focal supervision. Mean ± sample SD over seeds 11, 22, 33; score metrics are percentages. Each seed enumerates 8,865,724 same-date positive–negative pairs. Correction/retention require a positive policy gap; expansion requires a larger gap than the frozen reference on reference-correct pairs. Reference ties are excluded from both strata.

Method	Wrong corrected	Correct retained	Correct expanded
Direct ranking	33.14 ± 2.46	93.08 ± 0.99	92.62 ± 1.09
RDPO	32.93 ± 3.40	**93.17** ± 0.98	**92.72** ± 1.10
Selective ranking	**33.66** ± 3.32	92.90 ± 0.64	92.43 ± 0.72

Retrieval support and adaptive distance. A candidate distance family is

$$d_\omega^2(u,u') = \omega_w\|\widetilde{w}-\widetilde{w}'\|^2 + \omega_a\|\widetilde{a}_{\mathrm{land}}-\widetilde{a}'_{\mathrm{land}}\|^2 + \omega_h\|\widetilde{h}-\widetilde{h}'\|^2 + \omega_d\|d_t-d_{t'}\|^2 + \omega_e\mathbf{1}\{e\neq e'\} + \omega_{\mathrm{geo}}[d_{\mathrm{geo}}(i,i')/L]^2, \quad \omega_\cdot \geq 0. \tag{49}$$

Here i, i' are the cells associated with u, u'. The displayed squared quantity is a weighted dissimilarity. Its square root d_ω is a pseudometric on the represented inputs, and is a metric only if the positively weighted components distinguish them; zero weights can remove those distinctions. Here a_{land} excludes coordinates, d_{geo} is a geodesic or suitable projected distance, and $L > 0$ sets its scale. Select weights, support thresholds, and k on historical folds before the final test; compare against the pilot's unweighted standardized distance. Combining ecological similarity with weather similarity permits borrowing across regions while reducing unsupported neighbors. A fallback to the frozen global model when historical support is insufficient must be specified before evaluating its effect. The pilot's one-hot category term corresponds to a fixed categorical distance, not a learned metric or a support-aware fallback.

Within-region and cross-region preferences. A regional sampling study can use

$$\mathcal{L}_{\text{spatial pair}} = (1-\xi)\,\mathbb{E}_{q_{\mathrm{within}}}\ell(\Delta_\theta-\tau) + \xi\,\mathbb{E}_{q_{\mathrm{cross}}}\ell(\Delta_\theta-\tau), \quad \ell(m) = \mathrm{softplus}(-\beta_{\mathrm{rank}}m)/\beta_{\mathrm{rank}}, \tag{50}$$

with $0 \leq \xi \leq 1$, $\beta_{\mathrm{rank}} > 0$, and normalized same-date positive–negative pair distributions in both terms. Within-region pairs test local weather response; cross-region pairs connect score scales needed for area-wide prioritization. Regions or dates without both labels require an explicit skip-or-borrow rule. The pilot allows both pair types naturally but has no tuned mixture or regional reweighting. A fair loss comparison holds the sampled pairs, initialization, batches, and stopping rule fixed and selects each loss weight on historical folds; an equal fixed weight is only a diagnostic comparison.

Calibration with sparse-region shrinkage. If historical calibration diagnostics show persistent regional errors, a possible hierarchical map is

$$\widehat{p}_{t,i} = \sigma(a_{e_i}s_\theta(u_{t,i}) + b_{e_i}), \qquad a_e = \mathrm{softplus}(\alpha_0+\delta_e), \quad b_e = b_0+\gamma_e, \tag{51}$$

fitted to strictly forward historical scores with shrinkage on δ_e, γ_e. Obtain such scores through separate adaptation at each issue year using only information available then; pooling later-year histories can expose earlier calibration outcomes. Center the deviations with $\sum_e n_e\delta_e = \sum_e n_e\gamma_e = 0$ to fix the common-effect convention. Positive slopes preserve order *within* a region, but region-specific maps can change *cross-region* order. Therefore they require a new area-wide ranking evaluation, because the ordering invariance of a single global monotone map does not extend to region-specific maps. Use

Table 24: California 2022 complete-grid evaluation, 666 cells per date. Ranking adaptation excludes focal supervision. Mean ± sample SD over seeds 11, 22, 33; score metrics are percentages. Daily Top5% uses 12,410 alarms against 14,760 positives. Zero-hit dates have positives but no captured positive.

Method	FN	FP	Zero-hit dates
XGBoost	13028.00 ± 20.78	10678.00 ± 20.78	87.33 ± 1.53
Pretrained spatial focal	11985.00 ± 142.43	9635.00 ± 142.43	70.67 ± 8.50
Continued focal	11894.00 ± 60.26	9544.00 ± 60.26	67.00 ± 6.93
Direct ranking	11583.33 ± 101.48	9233.33 ± 101.48	40.33 ± 6.03
RDPO	11593.33 ± 131.46	9243.33 ± 131.46	42.67 ± 13.87
Selective ranking	**11549.67** ± 159.86	**9199.67** ± 159.86	**39.67** ± 15.89

one global monotone map when preserving the original statewide ranking is part of the deployment requirement. Keep the ranking loss in its score space unless a separately tested objective explicitly optimizes calibrated probabilities.

Staged, deployment-matched evaluation. First hold the shared backbone fixed and ablate coordinates, terrain, causal rates, and ecological embeddings, giving baselines the same historical information access. Then compare continued focal, direct, RDPO, and selective ranking on several chronological windows with historical-only selection of coefficients. Report pooled/daily AP, recall and FN at the same selected fraction, regional support, within/cross-region pair accuracy, and calibration both overall and by region. Separately test unseen locations through spatial blocks or held-out ecological regions (Roberts et al., 2017). Historical rates from a held-out region are legitimate only for a deployment scenario in which its past labels are available; a zero-shot geographic test must remove that information and provide an unknown-category fallback. Temporal and spatial uncertainty need date/spatial blocks, not independent row resampling.

Finally, replace annual-batch retrieval with rolling covariate windows for an online forecasting claim. At an issue time, all retrieval inputs, label histories, embedding vocabulary, and calibration fits must be available then. Evaluate the full daily candidate grid at the intended physical smoothing scale and selected fraction.

F Discussion and limitations

Pretraining as a foundation for adaptation. The evaluated framework uses a historically trained risk model as the starting point for subsequent fine-tuning. All local ranking variants inherit its parameters and representation. Direct ranking optimizes positive–negative score gaps; RDPO and selective use frozen pretrained gaps to define alternative loss targets. California compares these targets at fixed pairwise settings, while Yosemite compares binary-label and cross-input preferences. Both experiments exclude focal supervision from ranking adaptation; their pair contexts, temperatures, and stopping schedules remain dataset-specific. The experiments support ranking adaptation from the evaluated checkpoints. The fixed-update continued-focal ablation isolates source selection and supports a daily-AP benefit over matched random and full-history adaptation. It does not establish that retrieval is necessary, nor isolate its interaction with the ranking loss. Staging and pretrained-model quality are not separately controlled. The scalar-gradient analysis characterizes objective behavior under its stated assumptions, rather than establishing a generalization guarantee.

Deployment covariates and information access. Retrieval uses unlabeled deployment covariates, so the evaluation is batch-transductive. California's primary protocol uses covariates from the entire target year before adaptation and evaluates all 666 cells on every test date. Although each row uses previous-day weather, adaptation sees the future target covariate distribution. Daily budgets specify the evaluation decision, not an online information protocol; issue-time adaptation using only then-available covariates remains unevaluated. The spatial pretrained model receives 26 raw fields, including an ecological category represented by a learned embedding, and its causal history features draw on full historical label grids from prior years.

Evaluation scope. California evaluates one target year on locations present historically, so spatial heterogeneity here concerns temporal adjustment on a fixed grid; unseen-region generalization and spatial feature contributions are not isolated. The reported seed variation measures training randomness conditional on the evaluated years. The three ranking objectives use a common training-loss stopping rule but different realized update counts, which limits attribution of small differences solely to their targets. Fixed k and dataset-specific ranking temperatures are evaluated configurations, rather than demonstrated optima from a historical-forward sensitivity study. Yosemite's fixed joint-training comparison changes data access, optimizer, and update counts, so it does not isolate initialization alone.

Probability interpretation and calibration. Modern neural networks are often poorly calibrated (Guo et al., 2017). The no-focal ranking objectives optimize score gaps without constraining probabilities to be calibrated, so calibration is distinct from the ranking evaluation reported here. All reported California policies are evaluated without a fitted calibration map. Deployment requiring probability estimates should fit and evaluate calibration on strictly forward historical predictions, separately from ranking assessment.

High-activity diagnostic. High raw-DM strata are defined retrospectively for evaluation; training uses binary neighborhood-smoothed occurrence labels. The reported recall gains use policy-specific historical F1 thresholds and do not control alarm rates or demonstrate severity calibration. They support sensitivity to high observed activity under that threshold protocol. Positive cell–days may share the same underlying fire and should not be interpreted as counts of independent incidents.