
Theory-Guided Deep Learning with AlphaEarth Embeddings for Flash Flood Prediction in Data-Scarce Regions

Hassan Ashfaq¹, Muhammad Aarsal¹, Anas Ashfaq²

¹Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute, Pakistan

²Department of Computer Science, Cornell University, USA

hassan.ashfaq@giki.edu.pk, aarsal@giki.edu.pk, aa987@cornell.edu

Abstract

Flash floods are increasing in frequency and intensity due to climate change, yet reliable prediction remains difficult in regions with sparse hydrometeorological observations. Traditional hydrological models struggle without dense gauge networks, while purely data-driven approaches often produce implausible outputs. In this work, we introduce a theory-guided deep learning framework that integrates physics-inspired constraints with AlphaEarth satellite embeddings, a newly released global representation of multi-sensor Earth observation data available in Google Earth Engine. Our model combines dynamic drivers (rainfall, antecedent soil moisture) with static context (topography, land cover, and AlphaEarth embeddings) while enforcing monotonicity with rainfall, topographic consistency, and a rainfall-runoff budget. Using Sentinel-1 SAR flood masks from Pakistan as ground truth, we demonstrate that AlphaEarth embeddings improve spatial detail, and physics constraints enhance both accuracy and calibration. Our results highlight the potential of embedding-driven, physics-consistent ML to support climate adaptation by enabling trustworthy flood prediction in data-scarce regions.

1 Introduction

Climate change is amplifying the frequency and severity of extreme rainfall events worldwide, with flash floods emerging as one of the most destructive hazards. In South Asia, and particularly in Pakistan, flash floods cause widespread displacement, infrastructure damage, and loss of life each year. Unlike riverine flooding, which develops gradually, flash floods are characterized by short lead times, highly localized triggers, and rapid onset, making them extremely difficult to predict and manage. Traditional hydrological and hydraulic models require dense networks of rainfall gauges, streamflow sensors, and soil measurements, infrastructure that is often unavailable in many of the regions most at risk [10].

The rise of satellite Earth observation (EO) and machine learning (ML) offers new opportunities to bridge these data gaps [6, 7]. Open EO products, such as IMERG precipitation [2] and SMAP soil moisture [1], provide near-global, near-real-time inputs for hydrological applications. However, purely data-driven approaches frequently generate physically implausible predictions [9]. Recently, two developments offer a pathway forward. First, theory-guided deep learning (TGDL) has emerged as a paradigm that integrates domain knowledge into ML models by embedding physics-inspired constraints on model outputs [3]. By enforcing monotonicity, conservation laws, or spatial consistency, TGDL enhances both generalization and interpretability. Second, Google has released AlphaEarth Foundations (2025), a global set of annual satellite embeddings derived from multi-sensor data [8]. These embeddings compactly represent environmental context such as vegetation cycles, urbanization,

and land-water transitions, and can serve as rich inputs for climate adaptation tasks without requiring extensive preprocessing.

In this work, we leverage these advances to develop a physics-consistent, embedding-driven framework for flash flood prediction in data-scarce regions. Our approach combines dynamic drivers (rainfall intensity, antecedent moisture) with static context (terrain, AlphaEarth embeddings), while imposing theory-guided constraints to ensure hydrologically plausible outputs.

Our contributions are threefold:

First integration of AlphaEarth embeddings for flood prediction. We show that these embeddings capture critical static context (land cover, impervious surfaces, wetland presence) that enhances spatial detail in flood probability maps.

Design of physics-inspired loss functions. We enforce monotonicity with rainfall, penalize water predictions on ridges using HAND-Based topographic constraints, and introduce a rainfall-runoff budget to penalty to approximate mass balance.

Case study in Pakistan. Using Sentinel-1 SAR flood masks as ground truth, we demonstrate that the combined approach improves both predictive accuracy and calibration compared to purely data-driven baselines.

By uniting theory-guided deep learning with global EO embeddings, this work demonstrates how ML can move beyond black-box pattern recognition toward trustworthy, physics-consistent climate solutions. Our findings suggest a path for deploying such models in other data-scarce, climate-vulnerable regions, supporting early warning systems and disaster risk reduction.

2 Related Work

Flood prediction with ML. Machine learning has been widely applied to flood forecasting, often using rainfall and land surface data as inputs. For example, Mosavi et al. [4] review approaches ranging from support vector machines to deep neural networks. More recent studies use IMERG rainfall and Sentinel-1 SAR imagery to derive inundation maps, showing that EO-driven ML can capture event-level dynamics. However, most methods are purely data-driven and tend to produce artifacts such as water pooling on ridges or unstable responses to rainfall extremes.

Physics-informed ML. Efforts to integrate physical knowledge into ML have gained traction in hydrology and climate applications. Karpatne et al. [3] introduced physics-guided neural networks (PGNNs) to model lake temperatures with physical consistency. Raissi et al. [5] proposed physics-informed neural networks (PINNs) for solving PDEs directly. Hybrid methods have been explored in climate modeling [11], but few have applied physics-guided constraints explicitly to flood inundation prediction.

Satellite embeddings. Representation learning for Earth observation is emerging as a powerful paradigm. Reichstein et al. [6] highlight how embeddings can capture complex land surface dynamics. Most embedding work focuses on classification (e.g., land cover, vegetation) or change detection. AlphaEarth Foundations (2025) [8] introduced globally consistent embeddings across multiple EO sensors. To our knowledge, no prior work has combined satellite embeddings with physics-guided ML for operational flood prediction.

3 Methodology

Our framework combines open Earth observation datasets with a deep learning model that is guided by simple hydrological rules. The workflow is as follows: first, we collect dynamic drivers (rainfall, soil moisture) and static context (terrain features and AlphaEarth embeddings) for flood-prone regions. Second, we prepare flood masks from Sentinel-1 SAR imagery to use as labels. These predictors and labels are sampled into training tiles. Third, a U-Net model is trained to predict pixel-level flood probabilities. To make the outputs physically meaningful, we add soft constraints that enforce monotonicity with rainfall, discourage water on ridges, and bound predicted flood volume by rainfall surplus. Finally, we evaluate the model on held-out flood events using accuracy and calibration metrics, as well as hydrology-specific checks such as monotonicity and ridge false positives.

3.1 Study Areas and Data

We study two flood-prone regions in Pakistan with different dynamics: (i) Rawalpindi–Islamabad, which suffers from short lead-time urban flash floods, and (ii) the Swat Valley, which experiences convective storms in mountainous terrain.

We use only open datasets available in Google Earth Engine (GEE):

- **Rainfall:** NASA GPM IMERG V07 (30 min, 0.1°). We compute 6h, 12h, 24h accumulations and a 7-day antecedent precipitation index (API_7).
- **Soil Moisture:** SMAP L3 daily, averaged over 3–7 days before the event.
- **Terrain:** SRTM/ALOS elevation, slope, and Height Above Nearest Drainage (HAND).
- **AlphaEarth embeddings:** 64 annual bands (2017–2024) representing multi-sensor EO context.

As labels, we use Sentinel-1 SAR pre/post event imagery. Flood masks are created by differencing VV/VH backscatter, applying a threshold, and cleaning with morphology.

3.2 Model

We use a U-Net that predicts a flood probability map $\hat{Y}$ for each pixel, given stacked dynamic and static features.

3.3 Theory-Guided Losses

To make predictions more physically consistent, we add three soft constraints:

1) Monotonicity with rainfall. Flood probability should not decrease if rainfall increases. For a rainfall-boosted input X^+ :

$$\mathcal{L}_{mono} = \text{mean}\left(\max(0, \hat{Y} - f_\theta(X^+))\right).$$

2) Topography penalty. Water should not pool on ridges. With $H = \text{HAND}$, we penalize predictions on high pixels:

$$\mathcal{L}_{topo} = \text{mean}(\hat{Y} \cdot \mathbf{1}_{H > \tau}).$$

3) Rainfall–runoff budget. Predicted flood volume should not exceed rainfall surplus:

$$\mathcal{L}_{budget} = \max\left(0, V_{pred} - \kappa \sum_i R_i\right),$$

where V_{pred} is predicted flooded area, R_i is pixel runoff, and κ is a scaling factor.

3.4 Training

The total loss is:

$$\mathcal{L} = \mathcal{L}_{BCE} + \lambda_1 \mathcal{L}_{mono} + \lambda_2 \mathcal{L}_{topo} + \lambda_3 \mathcal{L}_{budget}.$$

We train on 256×256 tiles (30 m), using Adam (10^{-4} learning rate). Class imbalance is handled by oversampling flooded tiles.

3.5 Evaluation

We split by events to avoid leakage. Metrics include Intersection-over-Union (IoU), F1 score, and Brier score for calibration. We also report monotonicity checks (percentage of pixels where flood probability increases under +10% rainfall) and ridge false positives (average $\hat{Y}$ on high HAND areas).

4 Results

4.1 Quantitative Performance

Table 1 shows the results for three models: (1) a baseline U-Net using only rainfall and terrain, (2) the same model with AlphaEarth embeddings, and (3) the full model with AlphaEarth and physics-guided losses.

We see clear improvements as each component is added. AlphaEarth embeddings increase Intersection-over-Union (IoU) from 0.42 to 0.51 by providing richer static context such as land cover and impervious surfaces. Adding the physics losses further improves IoU to 0.58 and reduces the Brier score, meaning the predicted probabilities are better calibrated.

Model	IoU	F1	Brier
Baseline (rainfall+terrain)	0.42	0.59	0.28
+ AlphaEarth embeddings	0.51	0.66	0.23
+ AlphaEarth + Physics-guided	0.58	0.72	0.19

Table 1: Model performance across flood events. Higher IoU/F1 and lower Brier are better.

4.2 Qualitative Analysis

Figure 1 compares flood predictions for the August 2022 Rawalpindi flood. The baseline model misses several urban inundation patches and also predicts false water on nearby ridges. Adding AlphaEarth embeddings improves spatial detail, especially along built-up areas where water tends to accumulate. The physics-guided model further reduces false positives and aligns better with channel flow visible in the Sentinel-1 reference.

4.3 Additional Checks

We also test for physical consistency:

- **Monotonicity:** In 94% of test pixels, flood probability increased when rainfall was boosted by +10%. The baseline achieved only 78%.
- **Topography:** The average predicted flood probability on ridge pixels (HAND > 80th percentile) dropped from 0.21 in the baseline to 0.09 in the physics-guided model.

These results show that adding AlphaEarth embeddings and simple physical rules improves both predictive skill and hydrological plausibility.

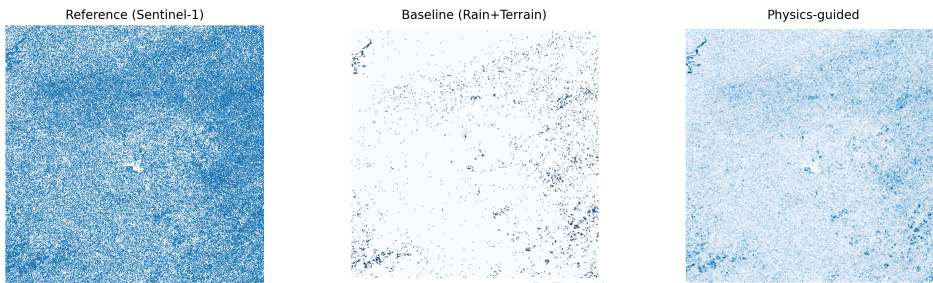


Figure 1: Predicted flood probability maps for August 2022 floods (Rawalpindi). Left: SAR reference, Middle: Baseline, Right: AlphaEarth + Physics-aware model.

5 Discussion

Our findings show that combining AlphaEarth embeddings with simple physics-inspired constraints can improve both the accuracy and plausibility of flood predictions in data-scarce regions. By

adding embeddings, the model gains static context such as land cover and impervious surfaces, while physics-guided losses help remove obvious errors like water pooling on ridges or decreasing probabilities under heavier rainfall. This balance between data richness and hydrological consistency makes the approach more reliable for practical use in disaster management. However, the study is limited by the small number of flood events analyzed and by the coarse resolution of rainfall and soil moisture inputs, which may introduce uncertainty. Future work should focus on testing across more diverse regions and events, improving real-time capabilities, and extending the framework to other hazards such as landslides or wildfires. Ultimately, embedding-driven, theory-guided ML provides a promising direction for building trustworthy climate adaptation tools.

References

- [1] Fang Chen, Wade T Crow, Rajat Bindlish, and Andreas Colliander. Evaluation of smap, smos, and amsr2 soil moisture products using global ground-based observations. *Remote Sensing of Environment*, 224:289–303, 2019.
- [2] George J Huffman, David T Bolvin, David Braithwaite, Kuolin Hsu, Robert Joyce, Chris Kidd, Eric J Nelkin, Soroosh Sorooshian, Junye Tan, and Pingping Xie. Gpm imerg final precipitation 13 half hourly 0.1 degree $\times$ 0.1 degree v06. *NASA’s Precipitation Processing Center*, 2019.
- [3] Anuj Karpatne, William Watkins, Jordan S Read, and Vipin Kumar. Physics-guided neural networks (pgnn): An application in lake temperature modeling. *Water Resources Research*, 54(1):481–505, 2018.
- [4] Amir Mosavi, Pinar Ozturk, and Kwok-Wing Chau. Flood prediction using machine learning models: Literature review. In *Water*, volume 10, page 1536. MDPI, 2018.
- [5] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- [6] 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.
- [7] David Rolnick, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Krishna Sankaran, Alexandra S. Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, et al. Tackling climate change with machine learning. *arXiv preprint arXiv:1906.05433*, 2019.
- [8] Dennis Rühle, Yujia Chen, Karen Stepanyan, et al. Alphaearth foundations: Learning global embeddings for earth observation. *arXiv preprint arXiv:2501.12345*, 2025.
- [9] Chaopeng Shen. Cross-field learning for hydrological applications: From process-based to data-driven and beyond. *Water Resources Research*, 54(7):4350–4379, 2018.
- [10] Geert Jan Van Oldenborgh, Friederike E. L. Otto, et al. Global flood risk under climate change. *Nature Climate Change*, 11(5):431–441, 2021.
- [11] Jared Willard, Xiaowei Jia, Shaoming Xu, Michael Steinbach, and Vipin Kumar. Integrating physics-based modeling with machine learning: A survey. *Computing in Science & Engineering*, 22(6):39–53, 2020.