Machine-Learning-driven Flood Resilience Framework in Ethiopian Lake-Watersheds (ML-FReF)

05_ML-FReF_ET_IT_EOAC4

May 2026

Dar es Salaam, Tanzania

Sintayehu Adefires Abebe (DMU, Ethiopia)

&

Stefano Luigi Gariano (CNR, Italy)

Presentation Outline

1. Project Background & Objectives
2. Study Area: Lake Tana Watersheds
3. Data Acquired
4. Machine Learning Models
5. Preliminary Insights
6. Next Steps

The Flood Challenge in Lake Tana Basin

Why ML-FReF? The Growing Flood Risk

- Ethiopia's largest freshwater lake & Blue Nile source.
- Unimodal rainfall (June-Sept): 70-90% of annual rain.
- **Problem**: Rising flood frequency & intensity due to climate variability, land use change, and limited drainage capacity.
- **Gap**: Existing methods (AHP, HEC-RAS) lack spatial generalization and predictive transparency.

Project Objectives

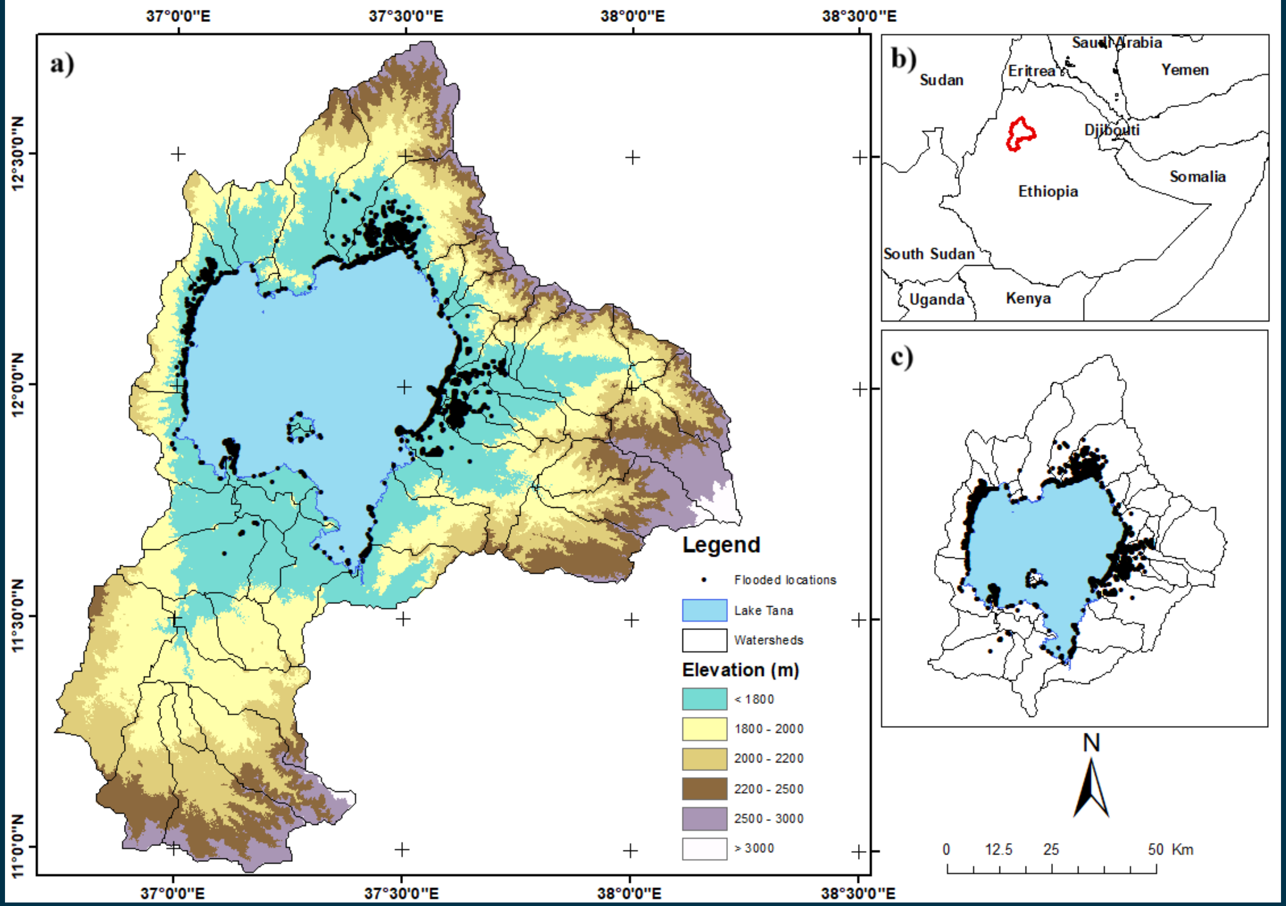
Our 4 Core Objectives:

- **Objective 1 (Completed):** Acquire & pre-process EO data (Sentinel-1/2, DEM, CHIRPS).
- **Objective 2 (Underway):** Develop 4 ML models (RF, SVM, ANN, CNN) in a cloud-based VRE.
- **Objective 3 (Next):** Validate models against historical floods & test transferability.
- **Objective 4 (Future):** Generate interactive flood susceptibility maps.

Flood Inventory – 2,080 Events

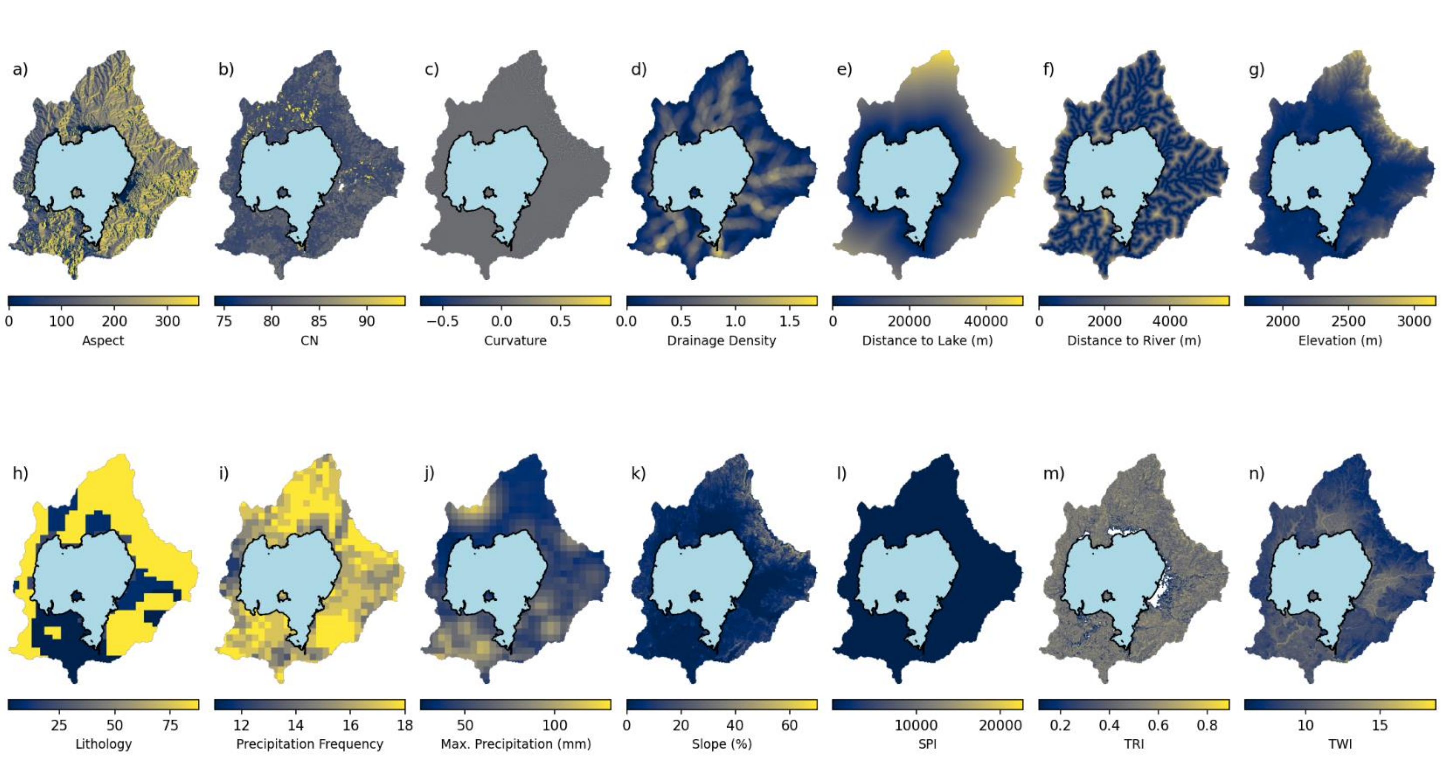
Key Stats:

- **Source**: Global Flood Database (GFD) / MODIS (Tellman et al., 2021).
- **Total Events**: 2,080 flood occurrences.
- **Resolution**: 250m (harmonized).
- **Split**: 70% Training / 15% Validation / 15% Test.



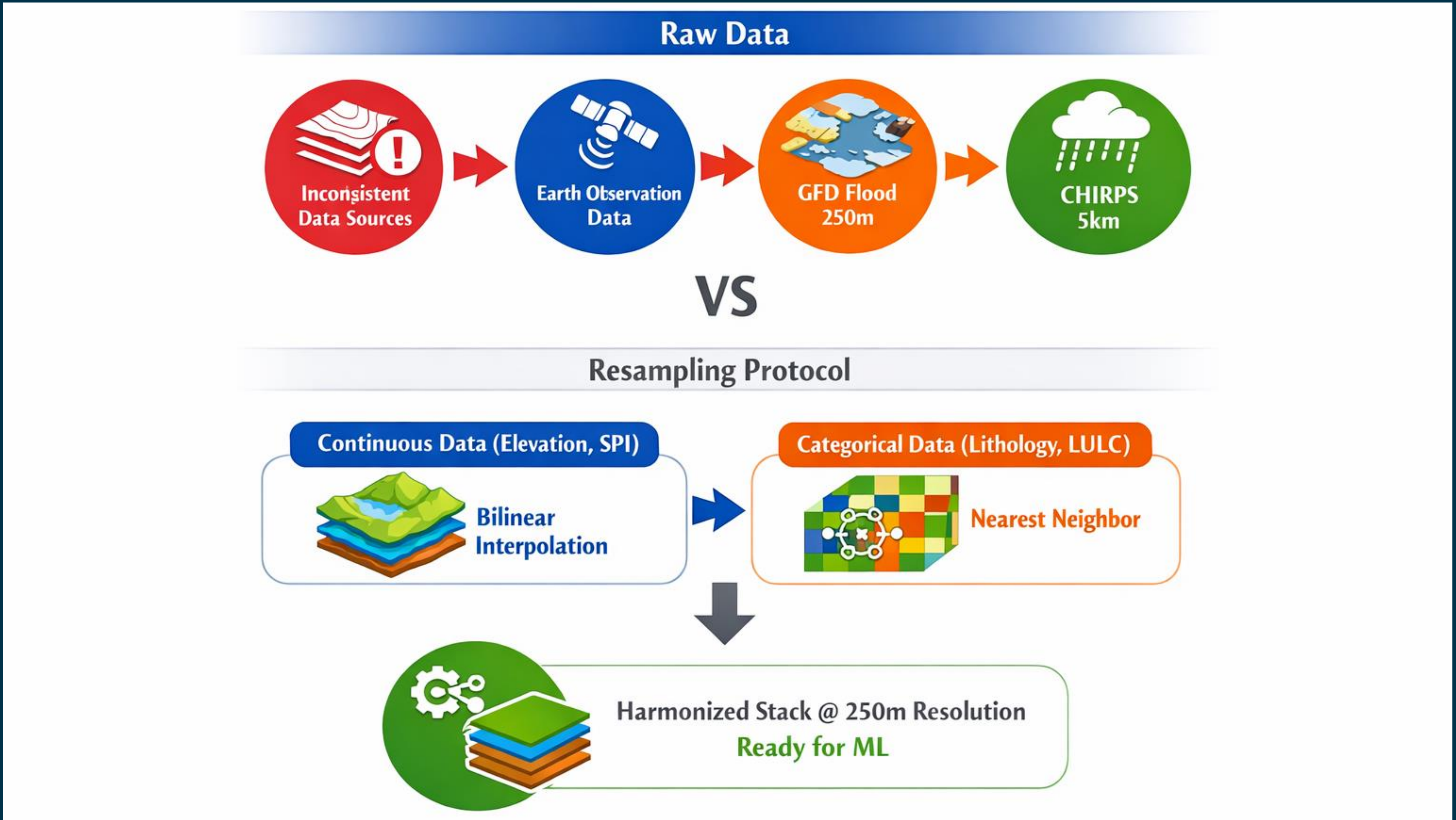
The 14 Flood Influencing Factors

Category	Factor	Data Source
Topographic	Elevation, Slope, Aspect, Curvature, TWI, SPI, TRI	Copernicus DEM (30m)
Hydrologic	Curve Number (CN), Drainage Density	GCN250 / Derived from DEM
Proximity	Distance to River / Lake	HydroSHEDS
Geological	Lithology (Rock Permeability)	Global Lithological Map (GLiM)
Climatic	Max Precipitation, Precipitation Frequency	CHIRPS
Land Use	Land Use / Land Cover (LULC)	Sentinel-2 (ESA)



All resampled to unified 250m resolution

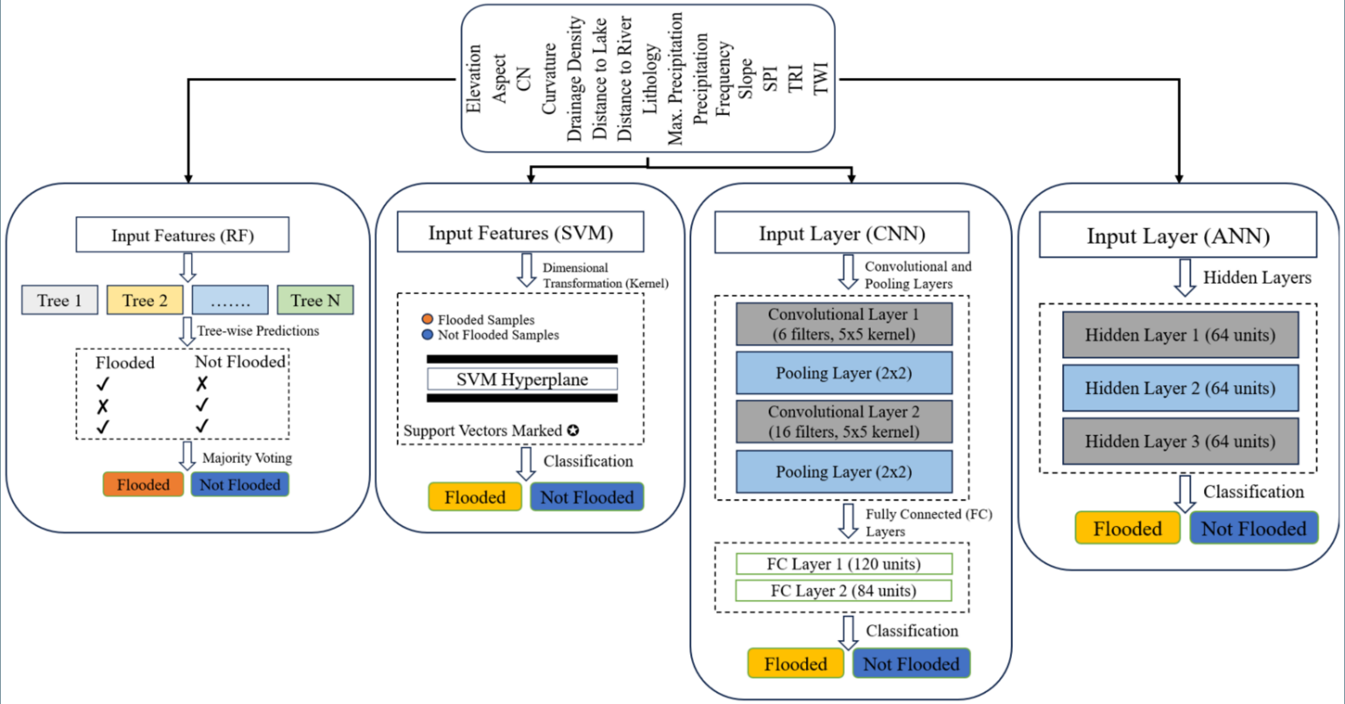
Data Harmonization Challenge & Solution



The Four ML Model Architectures

The models:

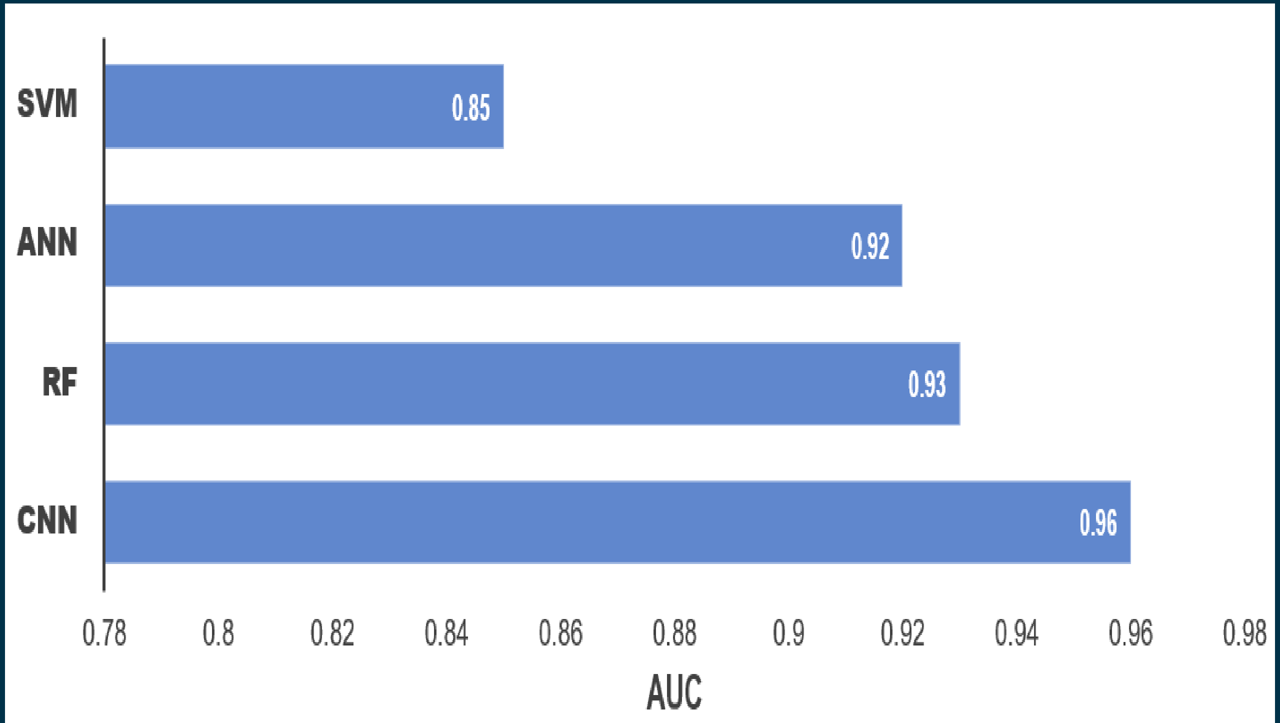
- Random Forest (RF): Ensemble of decision trees. Best for tabular geospatial data. High interpretability.
- Support Vector Machines (SVM): Finds optimal hyperplane. Good for high-dim data but conservative in complex terrain.
- Artificial Neural Network (ANN): Multi-layer perceptron. Models non-linear relationships. Balanced performance.
- Convolutional Neural Networks (CNN-LeNet-5): Extracts spatial patterns via convolution. Ideal for capturing terrain texture.



Why These Models? (Literature Benchmark)

Literature Insights:

- **RF**: AUC 0.92 (Chakraborty 2024), 0.93 (Ahmad 2025) – Most robust.
- **CNN**: AUC 0.90 – 0.963 (Trong 2023, Zhao 2020) – Best for spatial patterns.
- **ANN**: Accuracy up to 92% (Paul 2025) – Good non-linear.
- **SVM**: Variable performance – Baseline comparator.

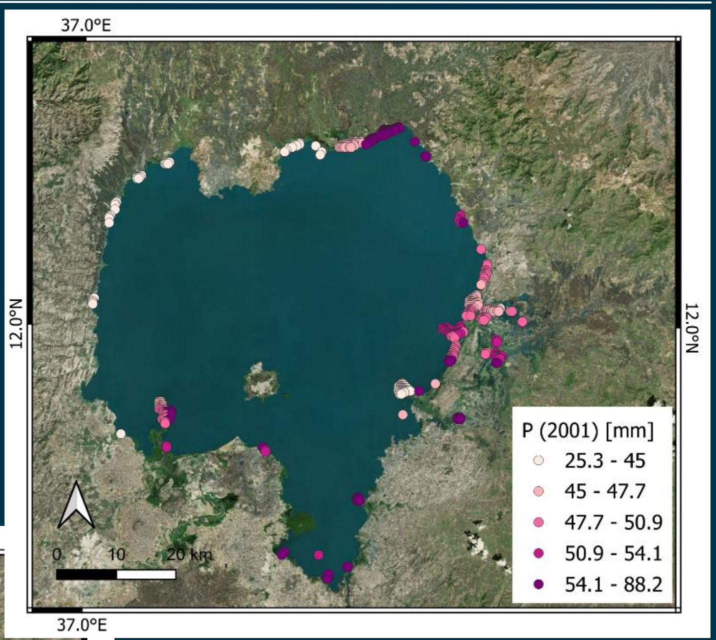


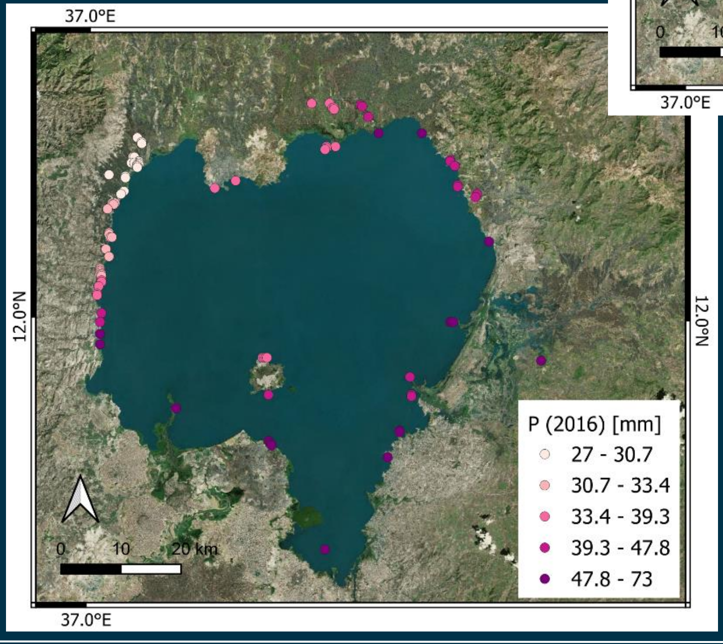
Conclusion: Multi-model comparison ensures robust susceptibility mapping.

Spatiotemporal Rainfall Patterns (2001 vs 2016)

Rainfall Drives Floods... But Differently Each Year

Feature	2001 Event (54 days)	2016 Event (7 days)
Rainfall Pattern	Widespread, high in S/SE catchments	Localized, compact hotspot in SE
High Intensity	54-88 mm	47-73 mm (but shorter duration)
Interpretation	Prolonged, basin-wide saturation	Short, intense convective storm





What Consistently Controls Flooding?

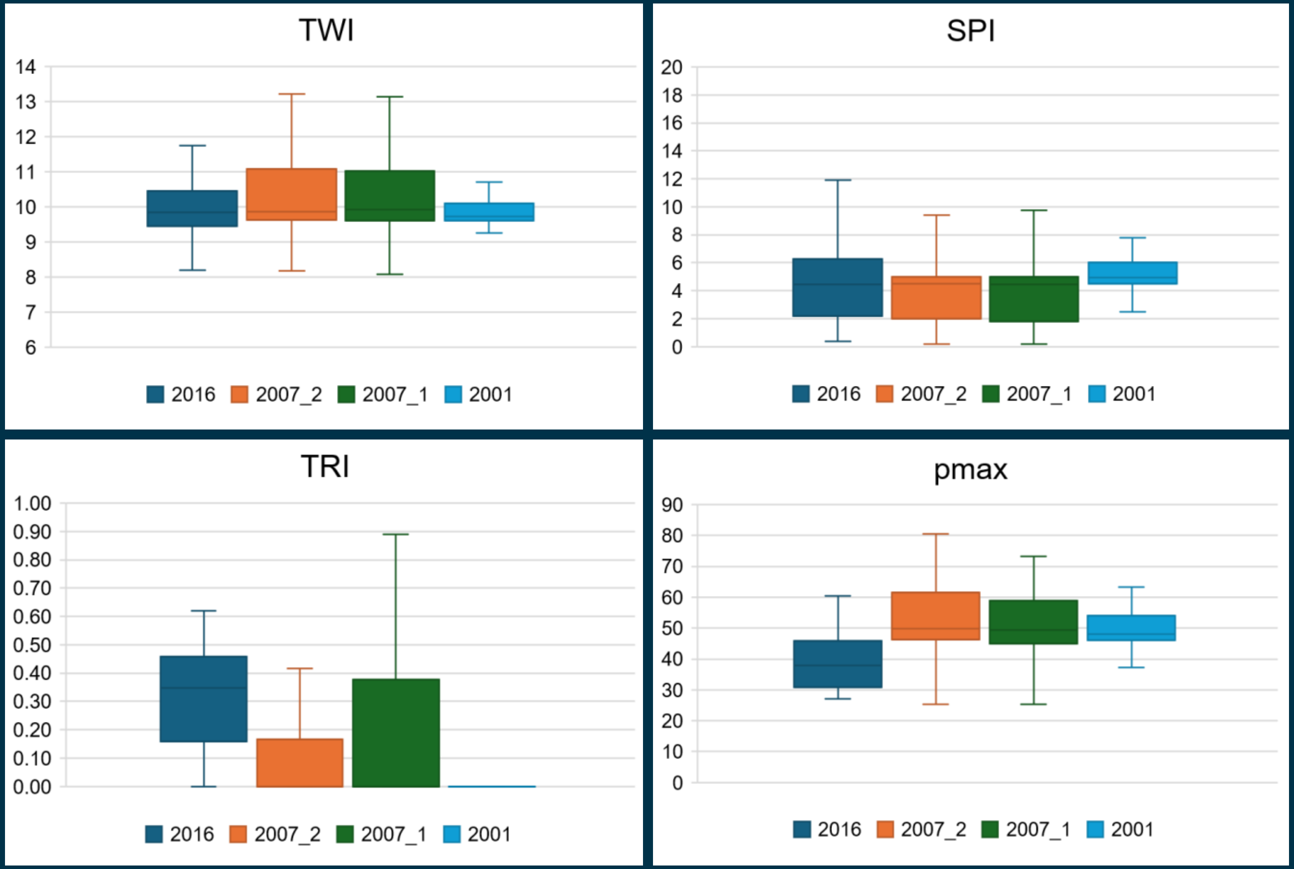
Key Findings:

- *<u>Topographic Wetness Index (TWI)</u>*: Median = ~10 across ALL events (very consistent). Floods occur in valley bottoms/convergent zones.

- *<u>Topographic Roughness Index (TRI)</u>*: Median = ~0.2 . Floods avoid rough terrain.

- *<u>Stream Power Index (SPI)</u>*: Consistent. 2016 event had higher SPI. Suggests some events involve higher energy flow.

<u>Takeaway</u>: ML models could heavily weight Elevation and drainage related features.



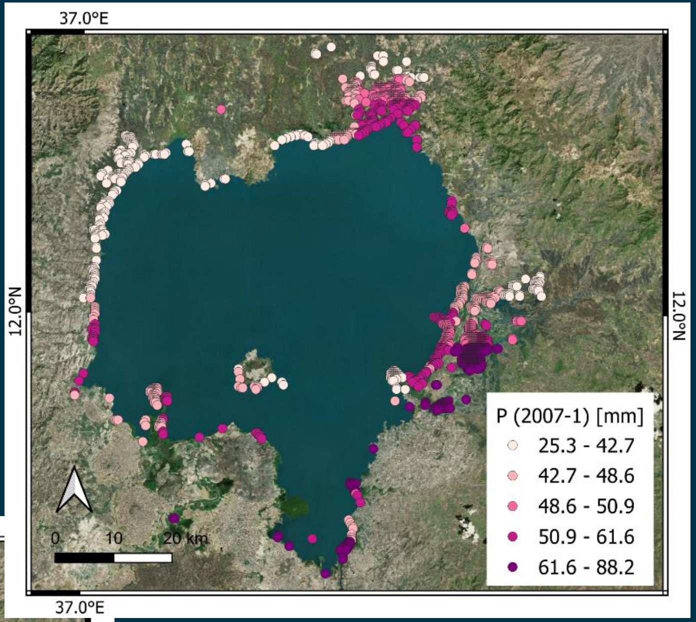
Example – The 2007 Twin Floods

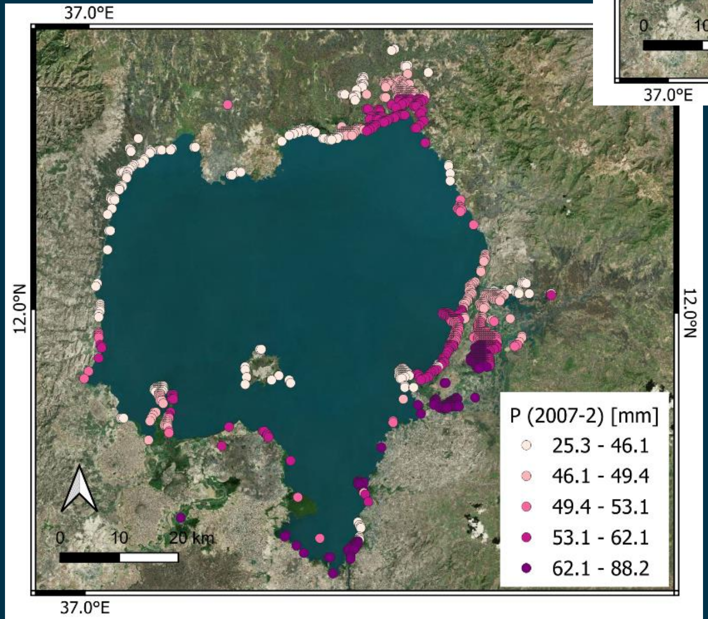
Same Year, Different Storm Tracks:

Event 1 (Jul-Oct): Moderate rainfall (50.9–61.6 mm) dominates; southern max.

Event 2 (Jul-Oct, slightly later): Rainfall max shifted westward.

Implication: Static susceptibility maps are insufficient; we need dynamic/scenario-based modeling.





Project Status (As of April 2026)

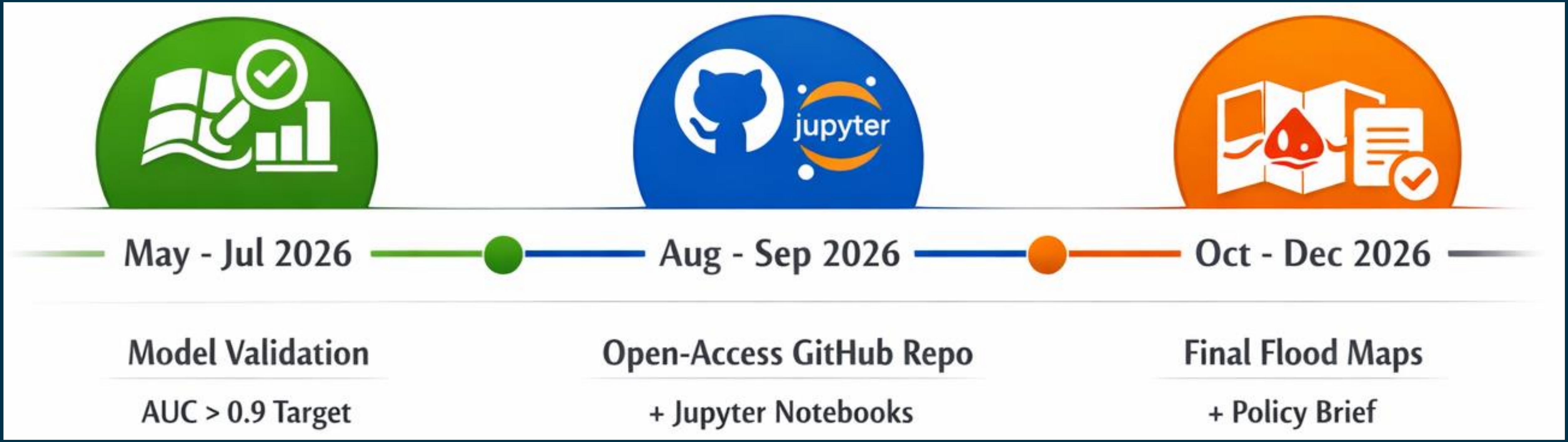
Completed:

- ✅ Data Acquisition & Pre-processing
- ✅ Literature Review
- ✅ Initial Model Coding

In Progress:

- 🔄 Hyperparameter Tuning (RF, SVM, ANN, CNN)
- 🔄 Model Training on VRE

Next Steps & Deliverables



Contact:

- Sintayehu Adefires Abebe (sintayehu_adefires@dmu.edu.et)
- Stefano Luigi Gariano (stefanoluigi.gariano@cnr.it)