Imperial College London

SnooCODE

Flood Probability Prediction Based on Machine Learning

Speaker Chaofan Wu

September 2023

OUTLINE

BACKGROUND

1. Flood Risks in Ghana

1. Annual Floods: Every year from May to July, Ghana faces floods that are risky to people and property [1, 2, 3, 4, 5].

2. Causes of Floods:
   - **Weather:** Heavy rain during the wet season [1, 2, 3].
   - **Geography:** Low land and clayey soil add to flood risks [1, 2, 3].
   - **Human Factors:** Poor drainage and waste disposal make floods worse [1, 2, 3].

2. Modelling Approaches

1. SWE Models:
   - **1D-SWEs**: Good for basic flood simulation, but not for complex scenarios [6].
   - **2D-SWEs**: Better for complex floods but costly and complicated [6, 7].

2. ML Models:
   - **Quick:** Fast predictions and adaptable [6, 8, 9].
   - **Multi-Dimensional:** Considers weather, land, and human factors [6, 8, 9].

Imperial College London

SnooCODE

OBJECTIVES

This project aims to develop a **Machine Learning model** capable of predicting the probability of flood occurrence.

The specific objectives of the project include:

- **Objective 1 (Data Collection):** Historical flood data, Rainfall data, and Geospatial information.
- **Objective 2 (Data Processing):** Exploratory data analysis (EDA), Feature extraction.
- **Objective 3 (Model Construction):** Experiments with various machine learning models, Evaluations and performance comparisons, Get the best model.
- **Objective 4 (Model Application):** Flood predictions and evaluate the performance, Generate flood maps for the target regions.

Imperial College London

SnooCODE

PROPOSED METHODOLOGY

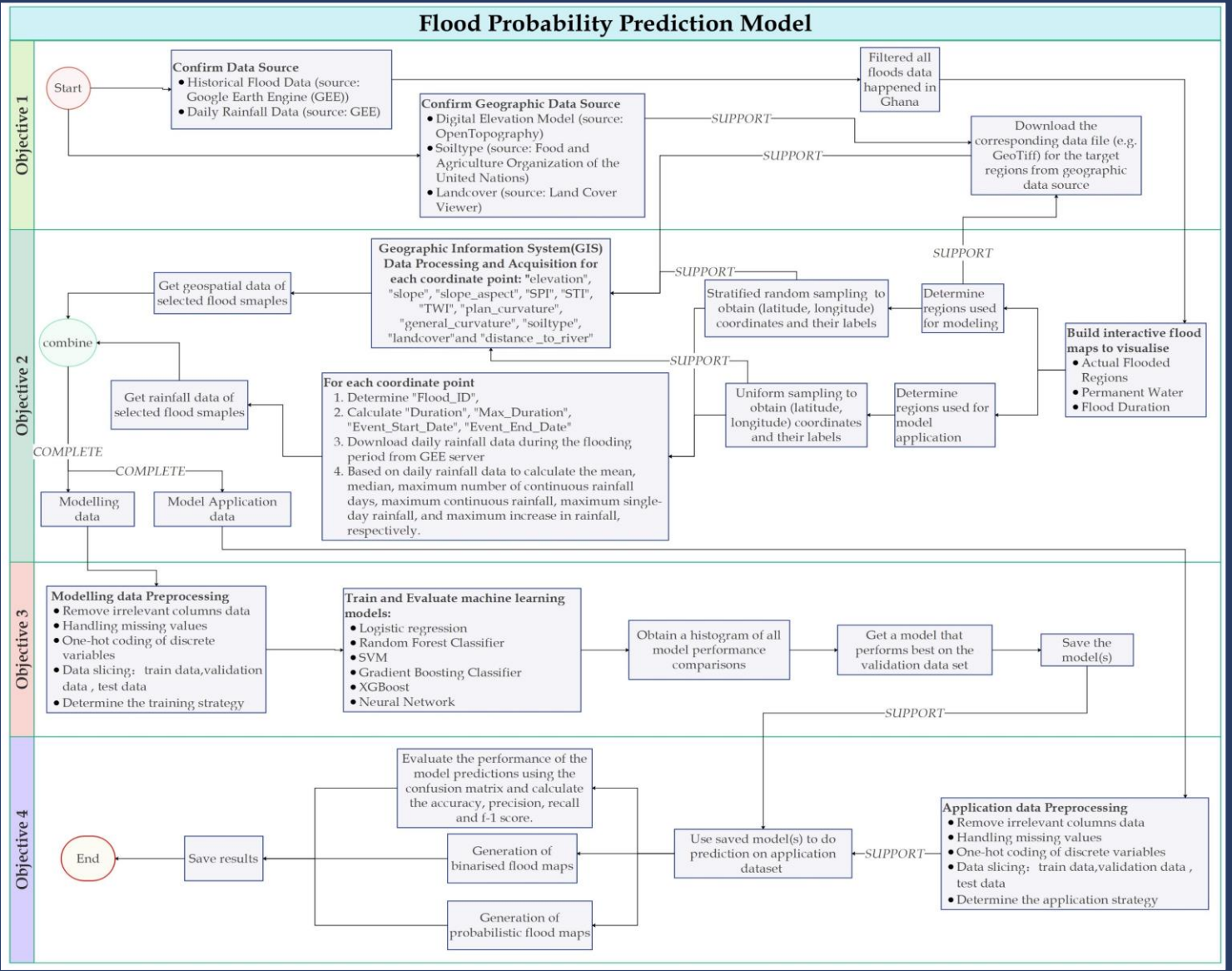


Figure 1: Flow Diagram of the Proposed Methodology.

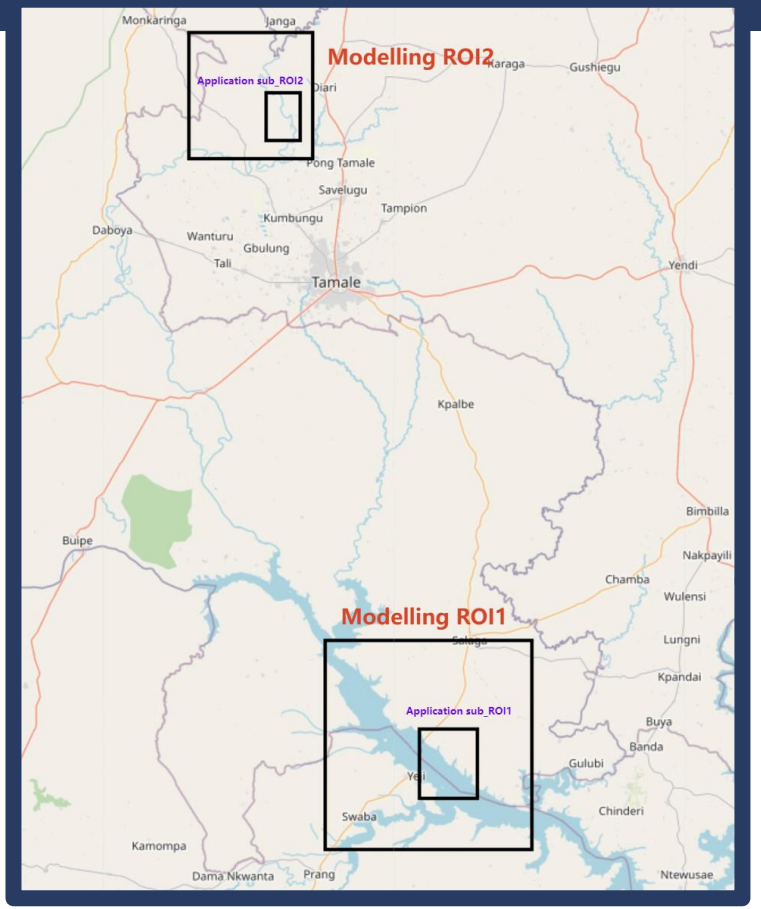


Figure 2: Geographical Representation of Modelling and Application ROI (region of interest)
Areas. Legend: Large Rectangles = Modelling Areas; Small Inner Rectangles = Application Areas.

PROPOSED METHODOLOGY

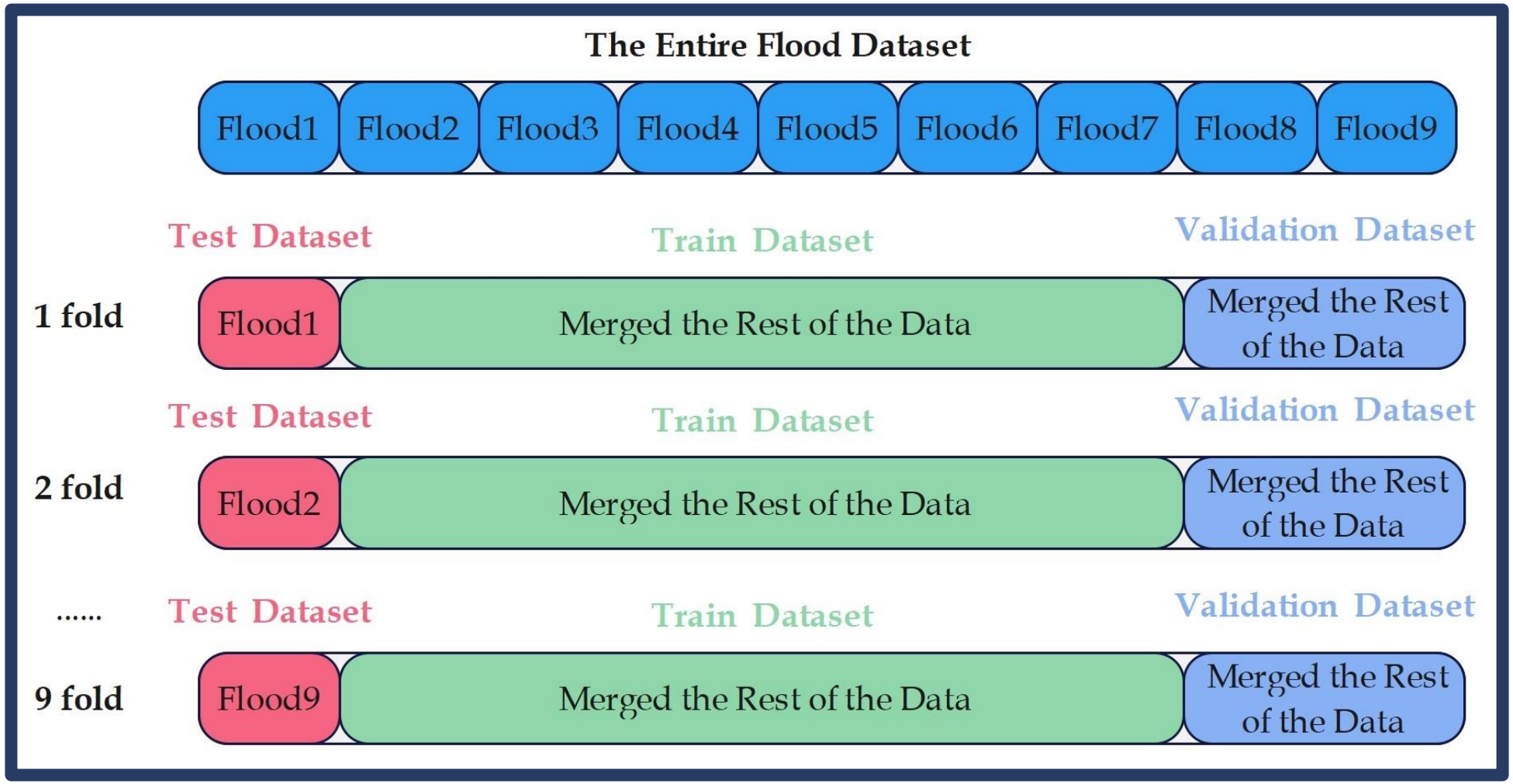
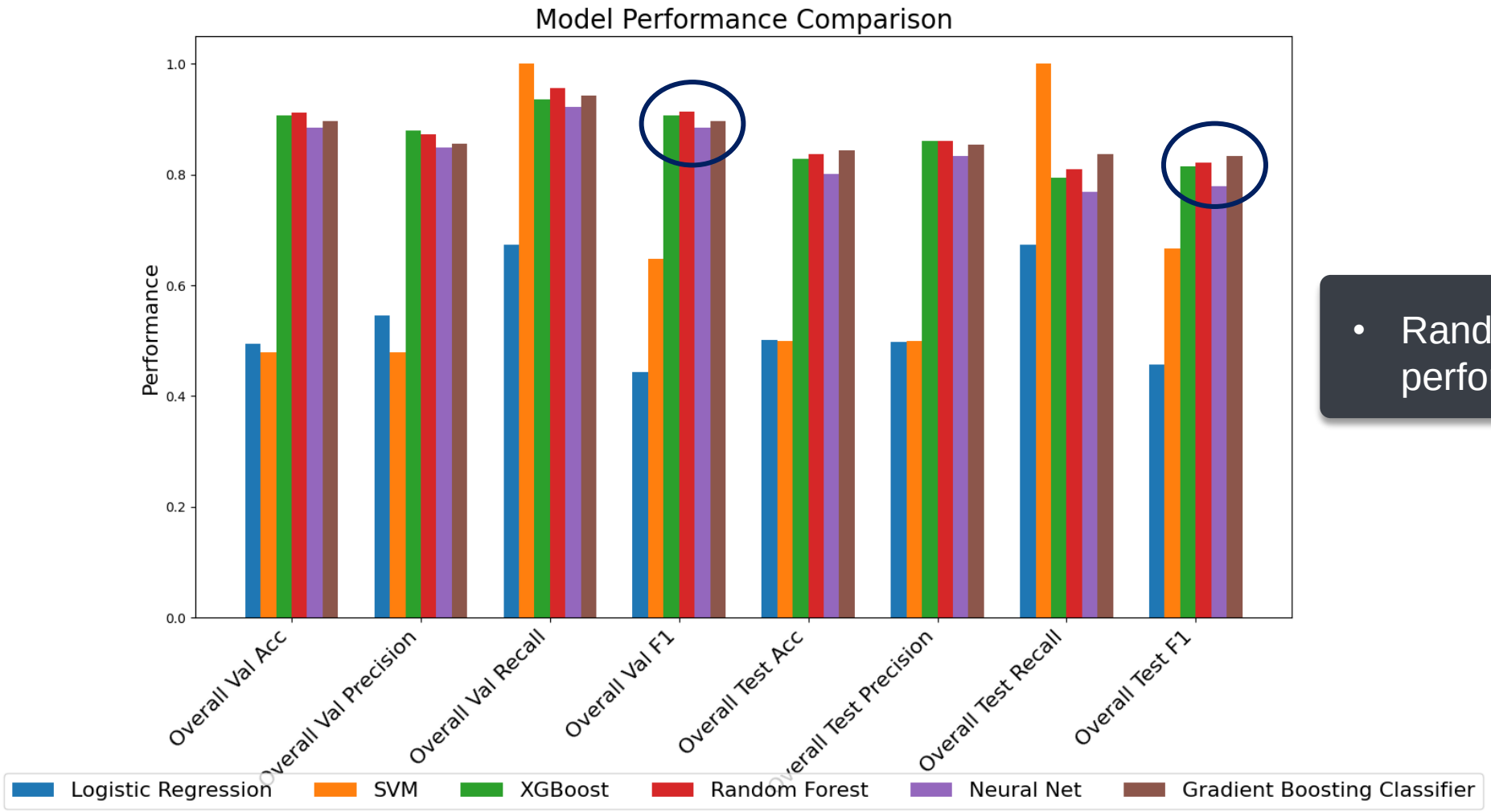


Figure 3: Illustration of Leave-One-Out Cross Validation Strategy. Legend: Test Dataset = Final Model Testing Dataset; Train/Validation Dataset = Model Training and Validation Dataset.

Imperial College London

SnooCODE

RESULTS



- Random Forest had the best overall performance

Figure 4: Comparative Analysis of Model Prediction Performances.

RESULTS

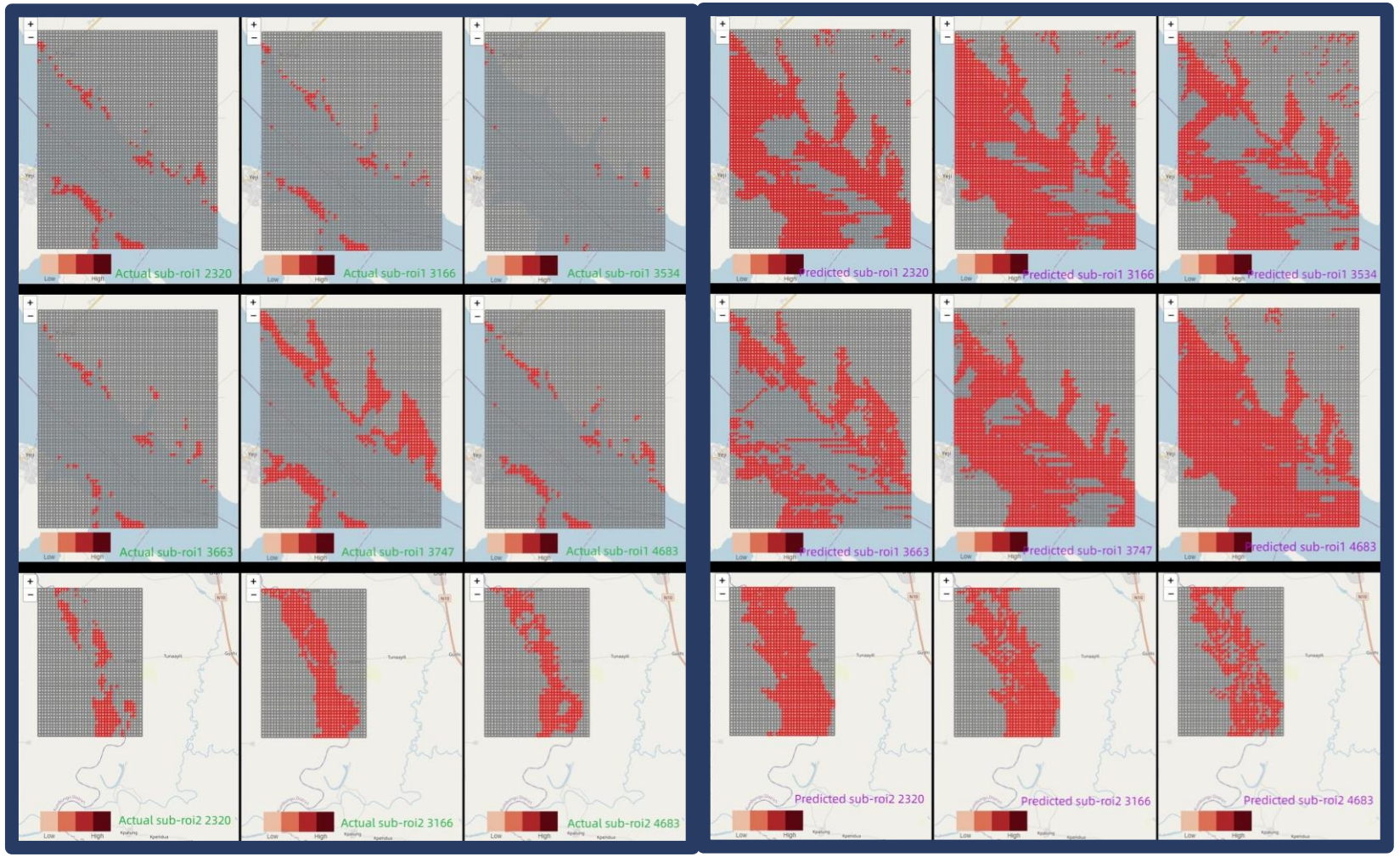


Figure 5: Comparison of Visualised Binary Flood Maps, (left) Actual, (right) Predicted.

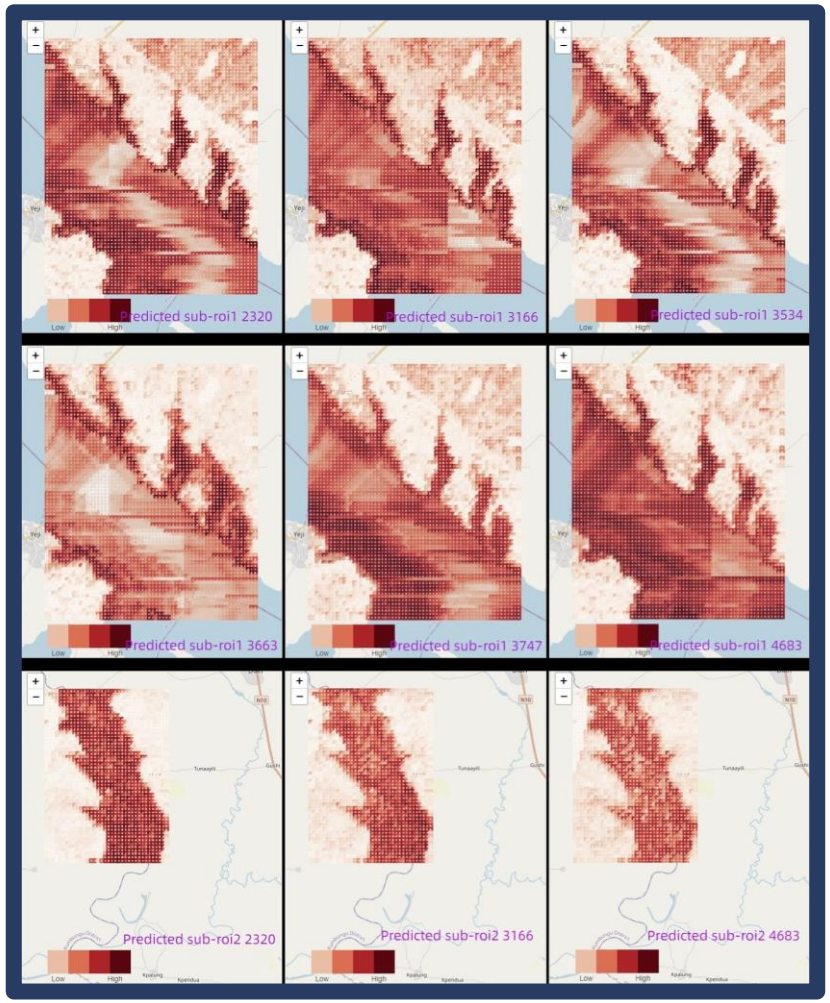


Figure 6: Probabilistic Predicted Flood Maps

Imperial College London

SnooCODE

RESULTS

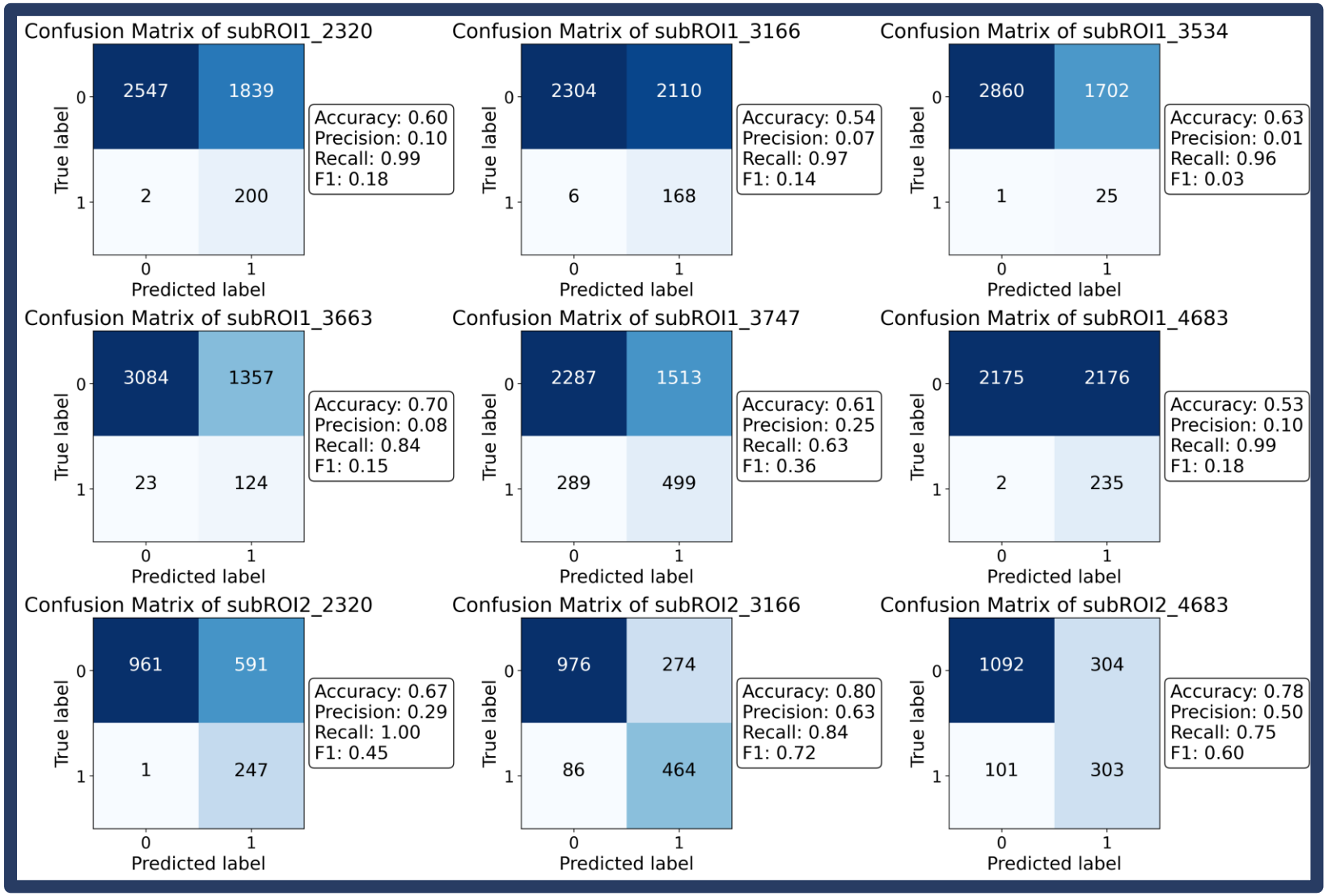


Figure 7: Confusion Matrices for Binary Flood Prediction

It is very clear to see that the model performs exceptionally well on **RECALL** but poorly on **PRECISION**

DISCUSSION

1. Model Strengths & Comparison

- **High Recall:** Better at identifying dangerous flood areas.
- **Data Combo:** Uses both geographical and rain data for more accuracy.
- **Old Models:** Less practical due to focus on single data type.

2. Precision & Resources

- **Low Precision:** Can misallocate resources.
- **Balance:** Need better precision without losing recall.
- **Regional Issues:** Varies in performance across areas.

3. Real-time Limits & Fixes:

- **Old Data:** Cannot predict floods in real-time.
- **Timing:** Needs better timing for upcoming floods.
- **Solution:** Adding rainfall predictions improves but complicates.

4. Scalability & Probability

- **Data acquisition Issues:** very slow (Batch and network)
- **No Truth Map:** Hard to validate probability estimates.

TO SUM UP…

- Objectives Met: All 4 goals achieved using Random Forest as the final model.
- Model Performance: High in recall but needs better precision.
- Study Impact: Helps identify likely flood zones.

REFERENCE

1. S. K. M. Agblorti D. Babanawo, P. A. D. Mattah, E. K. Brempong, M. M. Mattah, and D. W. Aheto. Local indicator-based flood vulnerability indices and predictors of relocation in the ketu south municipal area of ghana. Sustainability, 2022.
2. J. Abazaami C. Anab and S. A. Achanso. Moving beyond the ad hoc responses in flood management to a localization approach in ghana. Ghana Journal of Development Studies, 2022.
3. Y. A. Twumasi and R. Asomani-Boateng. Mapping seasonal hazards for flood management in accra, ghana using gis. In IEEE International Geoscience and Remote Sensing Symposium, volume 5, pages 2874–2876, 2002.
4. S. A. Adinku. Disaster preparedness: A sociological study of the flood problem in the odaw catchment in accra. Google Statistics, 1994.
5. Y. A. Twumasi and R. Asomani-Boateng. Planning and managing urban organic solid waste in an african city: linking organic solid waste composting to urban cultivation in accra, ghana. Google Statistics, 1999.
6. L. Cea and P. Costabile. Flood risk in urban areas: Modelling, management and adaptation to climate change. a review. Hydrology, 9, 2022.
7. B. Jamali, E. Haghighat, A. Ignjatovic, J. P. Leit ˜ao, and A. Deletic. Machine learning for accelerating 2d flood models: Potential and challenges. Hydrological Processes, 35, 2021.
8. Z. Wang, X. Chen, Z. Qi, and et al. Flood sensitivity assessment of super cities. Scientific Reports, 13:5582, 2023.
9. W. Fenglin, I. Ahmad, M. Zelenakova, and et al. Exploratory regression modeling for flood susceptibility mapping in the gis environment. Scientific Reports, 13:247, 2023.