AI-based rainfall forecasts for Early Warnings over the Greater Horn of Africa

SHRUTI NATH, F. Cooper, D. Macleod, M. Gudoshava, A. Amdihun, N. Kalladath, J. Kinyua, H. Kimani, D. Koros, Z. Mwai, A. Teshome, I. Obai, J. Mason, M. Chantry, F. Pappenberger, A. Weisheimer, T. Palmer

Resilience to Natural Hazards through AI solutions,
Session 2: AI models, methods and added value for tracking and forecasting | 08.05.2025

Objectives

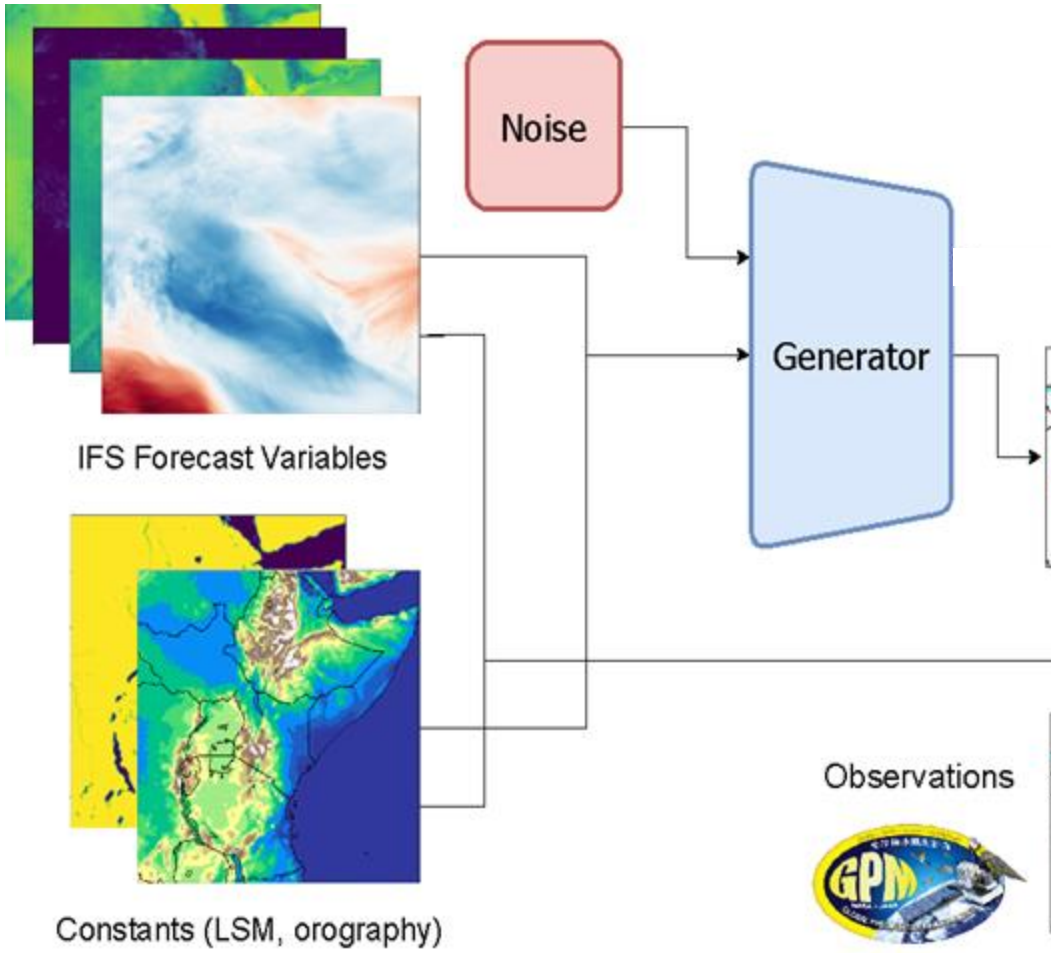
Exploring the **potential of cloud-optimized AI** based approaches to improve accuracy of **operational rainfall forecasts** and provide better **early warnings** over the **Greater Horn of Africa**

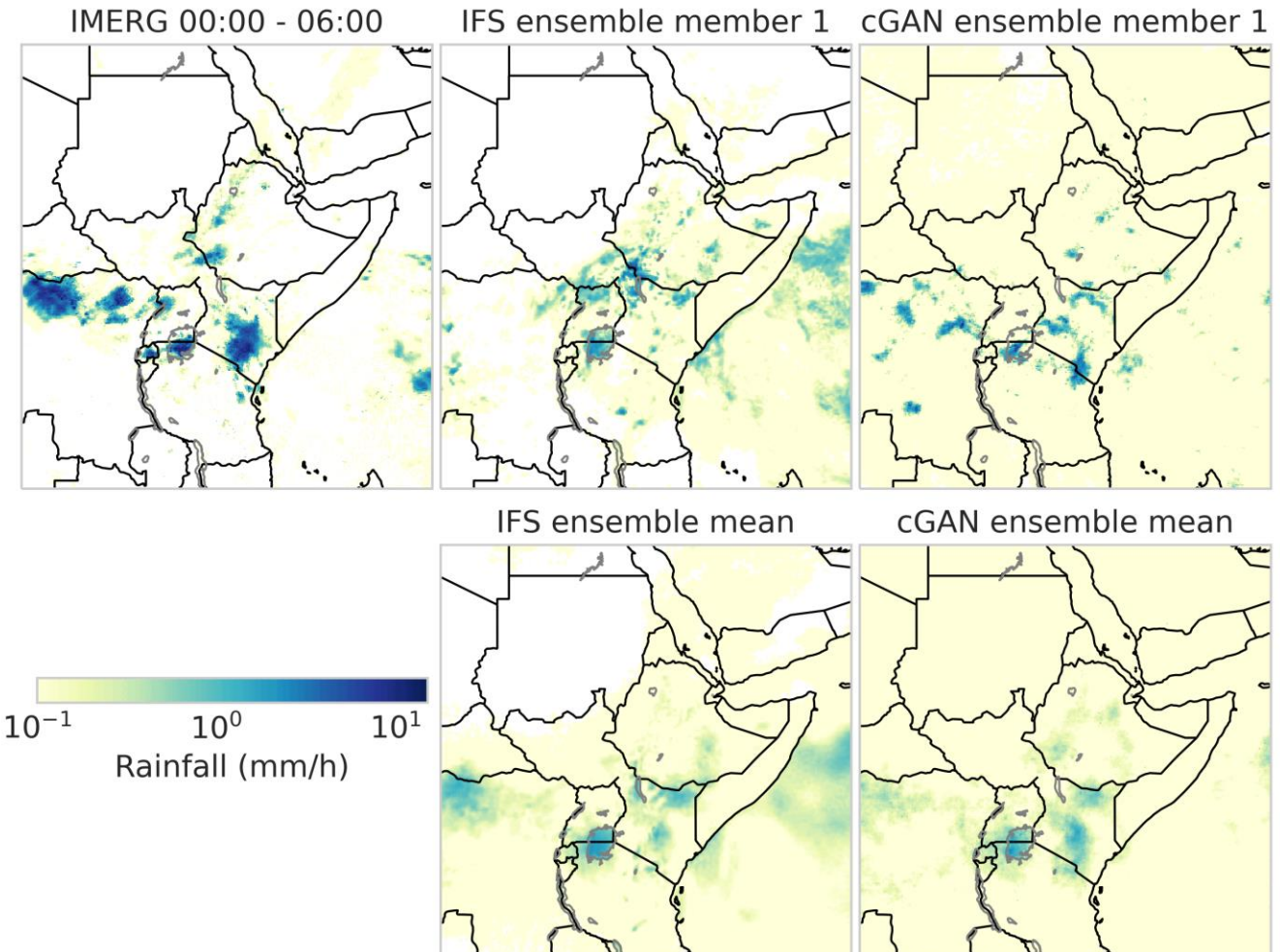
Low-cost forecast generation

Minimal infrastructure and data download

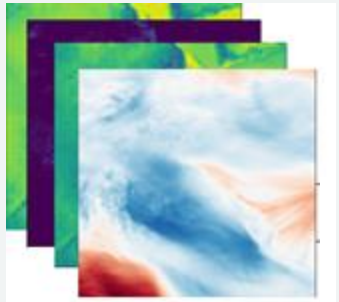
Mainly talk about journey towards evaluating **value added**

Current model set-up

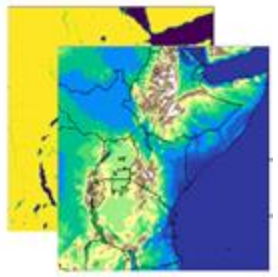




End-to-end early warning generation



IFS Forecast Variables



Constants (LSM, orography)

AI-model

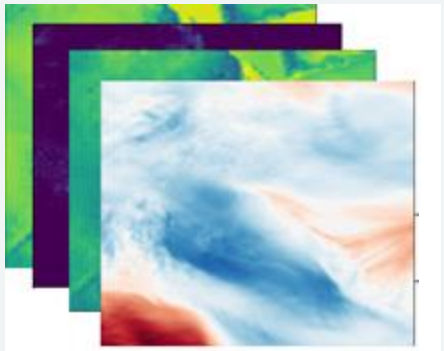
Warning Production

We still need to communicate probability as a **yes/no** for user-defined **rainfall levels**

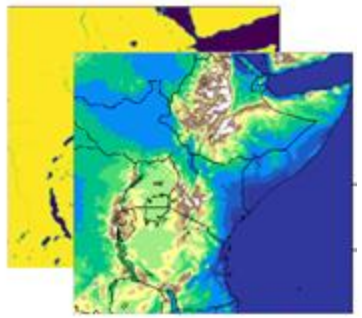
Triggering

Weigh **cost of action vs inaction,** for no-regret AA

Taking a step back



IFS Forecast Variables



Constants (LSM, orography)

cGAN

Criteria for method:

- Retains **latent representation of underlying distribution**
- Calibrated on scores **that cannot be hedged** (proper scores)
- Allows easy scanning of the economic value added under **different cost-loss ratios**

Additional task-specific **distillation step** for user-defined thresholds

Jack of all thresholds: Brier score optimized over **continuum of thresholds**

Master of a few

Cost-loss ratios?

	Event doesn't happen	Event happens
No preparation made	0	L
Preparation made	C	C

Economic value of:

- Our forecast:

$$E = (TP + FP) \cdot C + FN \cdot \mathrm{L}$$

- The default i.e. best out of always or never acting

$$\mathrm{E}_{clim} = \min\{(TP + FN) \cdot L, C\}$$

- A perfect forecast:

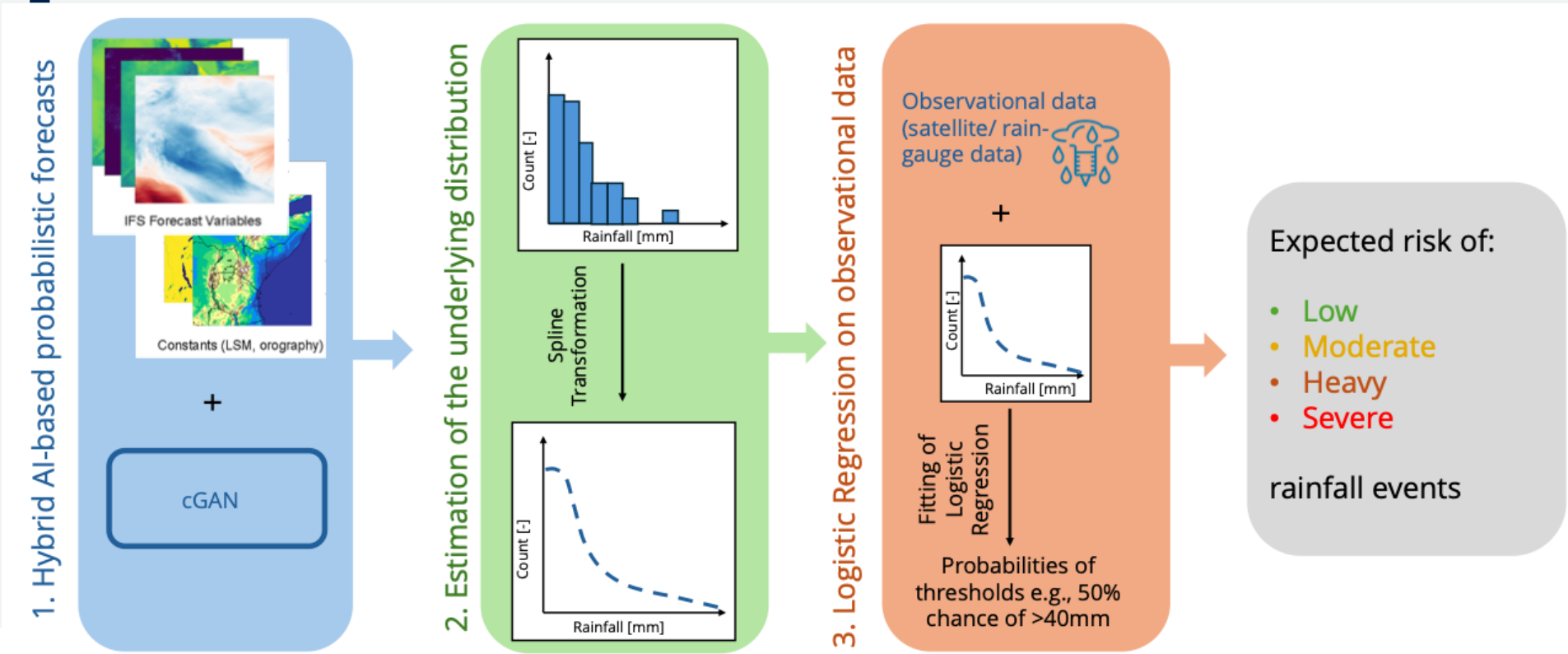
$$E_{perfect} = (TP + FN) \cdot C$$

At each C/L ratio, the relative economic value can be quantified:

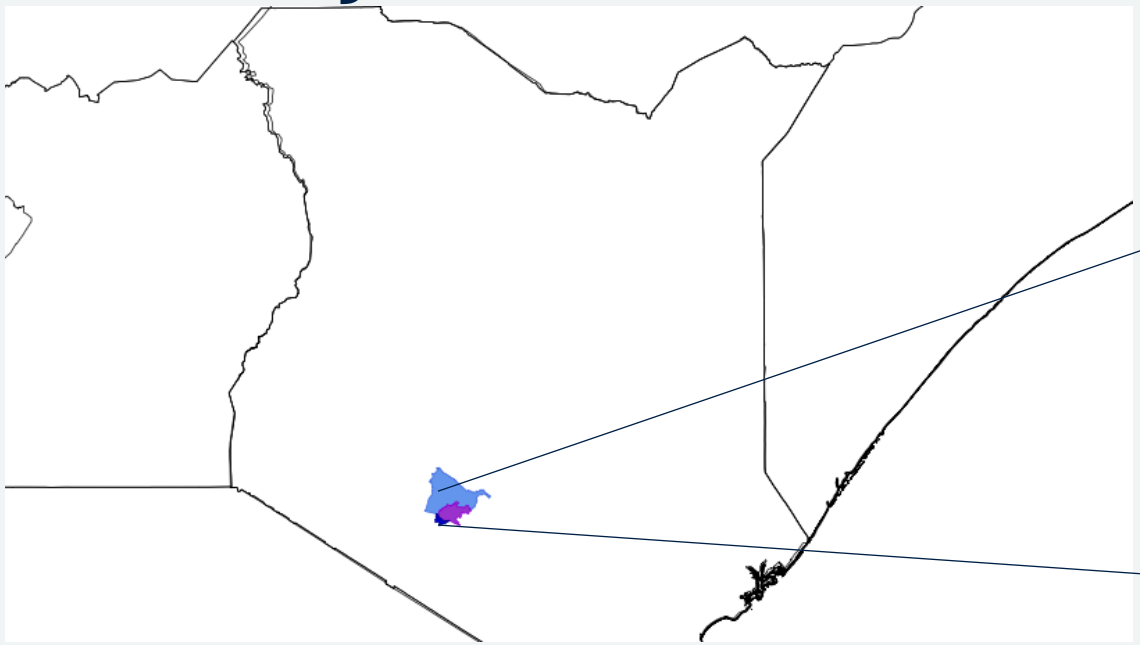
$$REV\left(\frac{C}{L}\right) = \frac{E_{clim} - E}{E_{clim} - E_{perfect}}$$

→ Well-calibrated p means we can set p=c/L!

Proposed framework



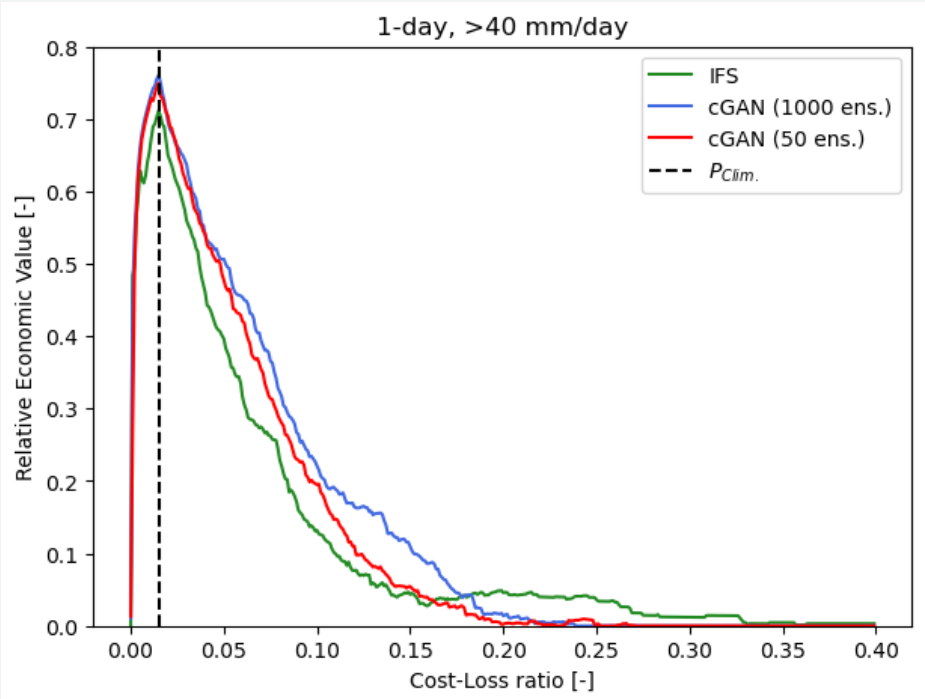
Case study

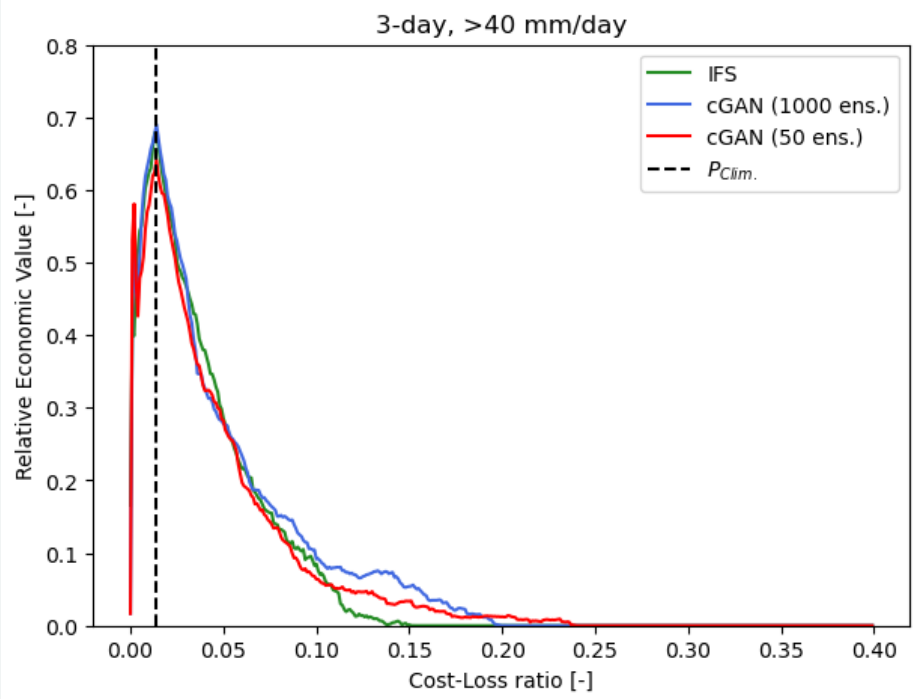


- Event: >40 mm/day
- Test period of focus: 2021,2023,2024

Urban flooding, personal camera, 22.04.2024

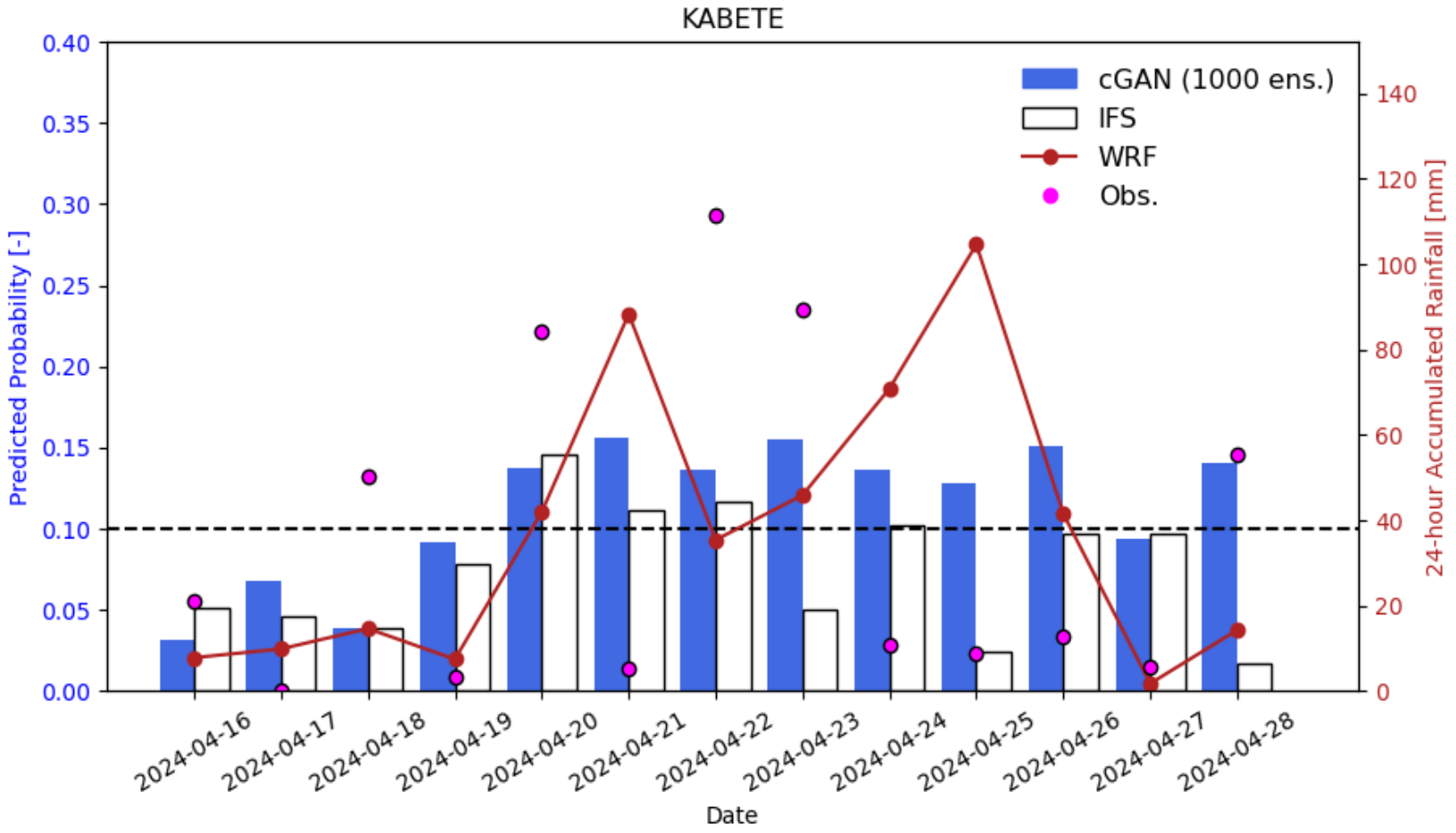
Economic value of forecast probabilities obtained





Calibration on individual rain gauge data

- cGAN is **betting on wetter periods**
- IFS tries to give more **precise probabilities**
- WRF also shows **elevated rainfall** over period
- Precisely pinpointing exact time of heavy rain may be tricky for convective active periods → **probabilities are important**!



Summary

- Operational value of probabilistic (AI) forecasts demonstrated by translation to “yes/no triggers” based on **event-specific probability thresholds** chosen via **cost-loss frameworks**

- Going from probabilistic forecasts to triggers requires **mapping** from the **true underlying distribution** to **well-calibrated probabilities** of threshold exceedance

- AI-generated, low-cost large ensembles can provide **better calibrated probabilities**

Contact me: shruti.nath@physics.ox.ac.uk