Destination Earth: Hydrology

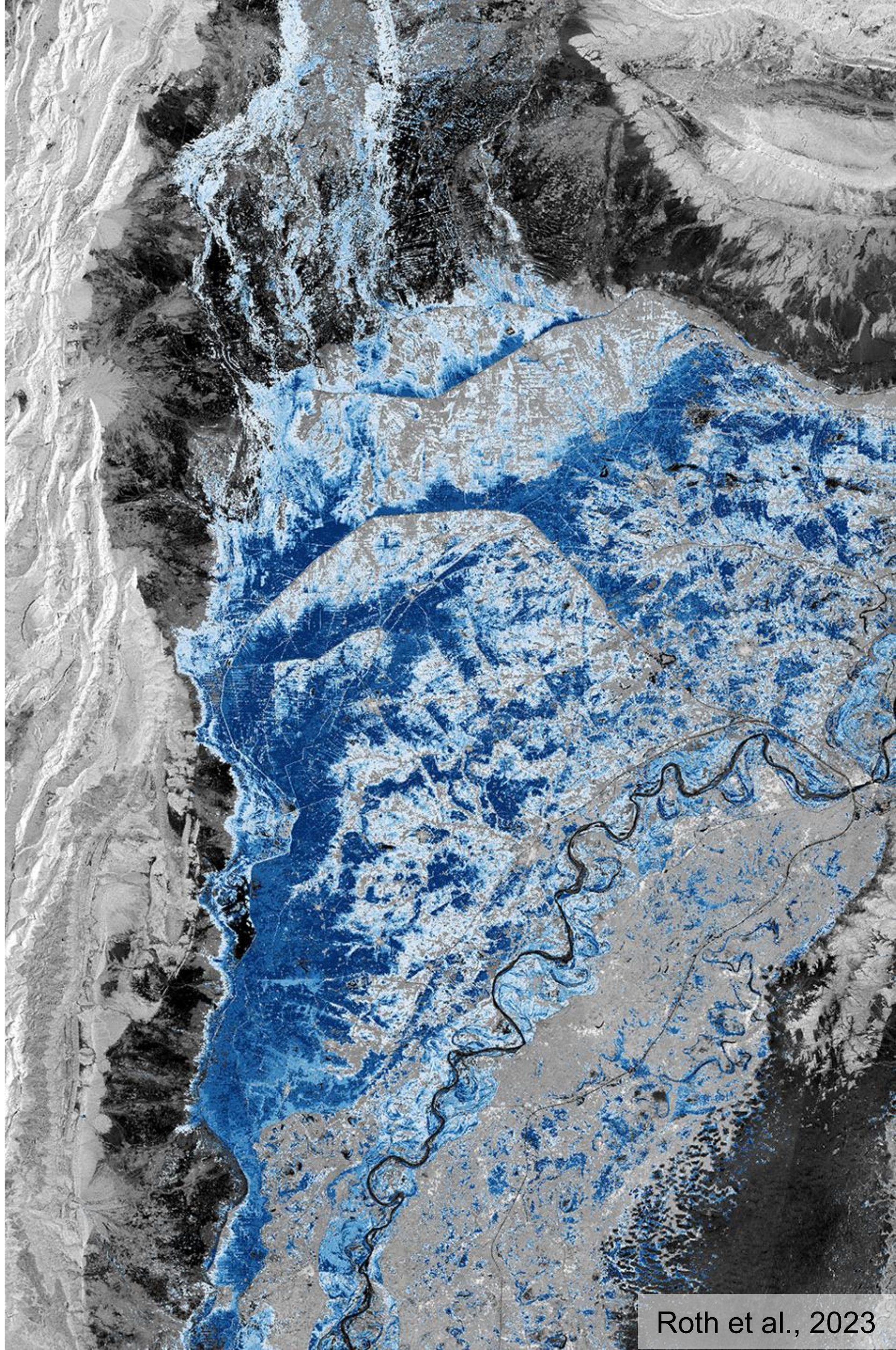
Training course: Training course: Machine learning and Destination Earth

Presented by Maria Luisa Taccari
with contributions and materials from Kenza Tazi, Oisín M. Morrison, Andreas Grafberger, Juan Colonese and Corentin Carton de Wiart

marialuisa.taccari@ecmwf.int

Talk outline

1. Introduction
    a. The rising threat of floods
    b. Challenges of hydrological data
    c. Operational global flood forecasting models
2. AIFL
    a. Architecture
    b. Training strategy
    c. Model performance
    d. Case study: 2024 Luxembourg province
3. Operational pipeline
    a. Efficient operations
    b. Effective communication
4. Future work
5. Key takeaways



The rising threat of floods

Floods, deadliest and most expensive natural disaster

- 55 million people directly affected in 2022
- More frequent and severe with climate change
- USD 5.6 trillion in projected economic losses by 2050

Predictions, vital to mitigation and response strategies

- 6x lower mortality
- 12-hours ahead ⇔ up to 60% reduction in economic damages
- Most-needed in low-income countries with greater flood risks

Zhengzhou, Henan province, China
July 20, 2021 © Reuters

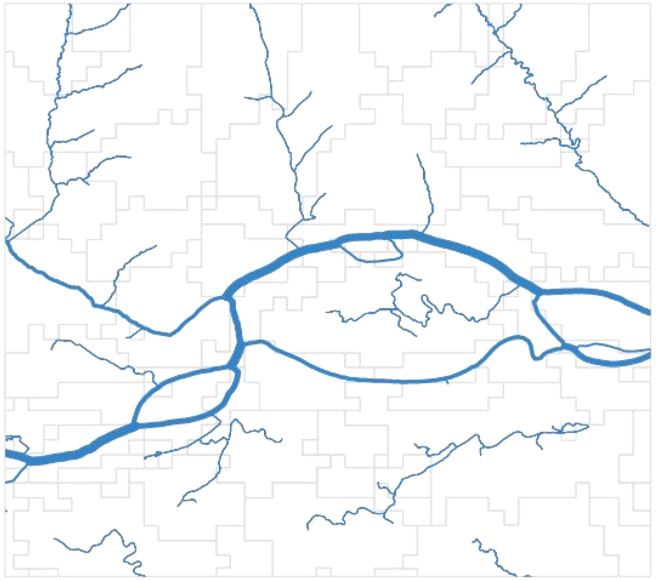
Challenges of hydrological data

Complexity

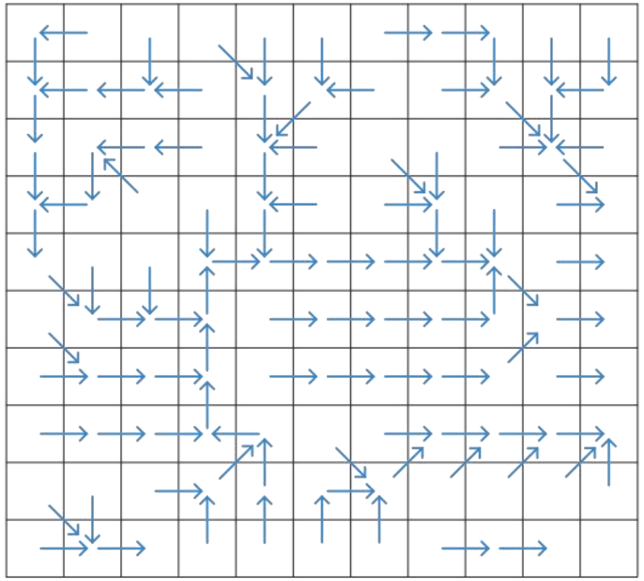
- Causal unidirectional network dependency
- No interpolation
- Approximation through idealised and often non-comparable river networks

Challenges of hydrological data

Rio Negro and Solimões River



GRIT network



GloFAS network

Challenges of hydrological data

Complexity

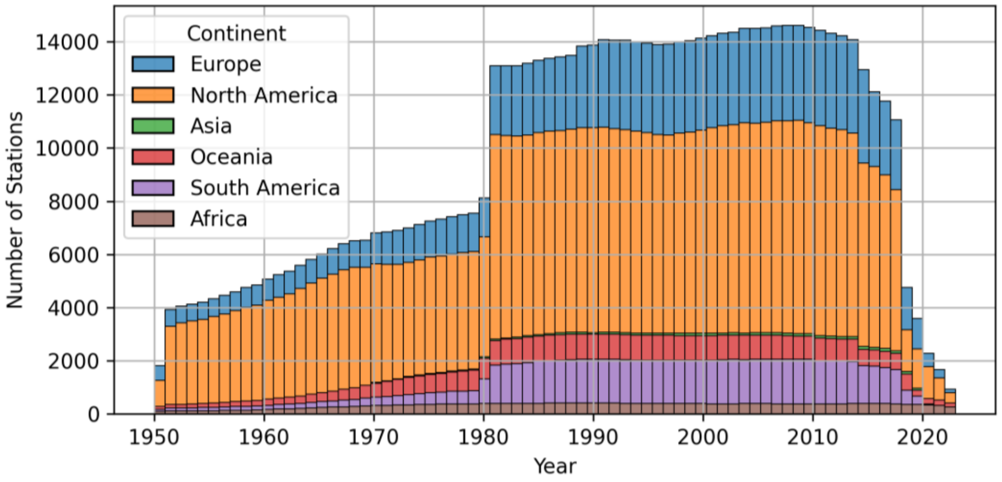
- Causal unidirectional network dependency
- No interpolation
- Approximation through idealised and often non-comparable river networks

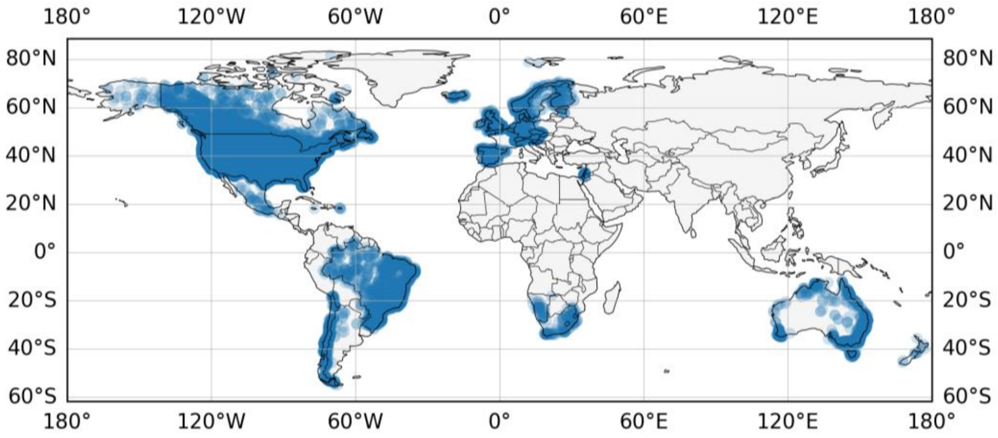
Availability

- Sparse and heterogeneous observations
- Uncertain streamflow observation quality
- Missing hydrological forcing data

Challenges of hydrological data

Spatiotemporal distribution of the CARAVAN dataset





Challenges of hydrological data

Complexity

- Causal unidirectional network dependency
- No interpolation
- Approximation through idealised and often non-comparable river networks

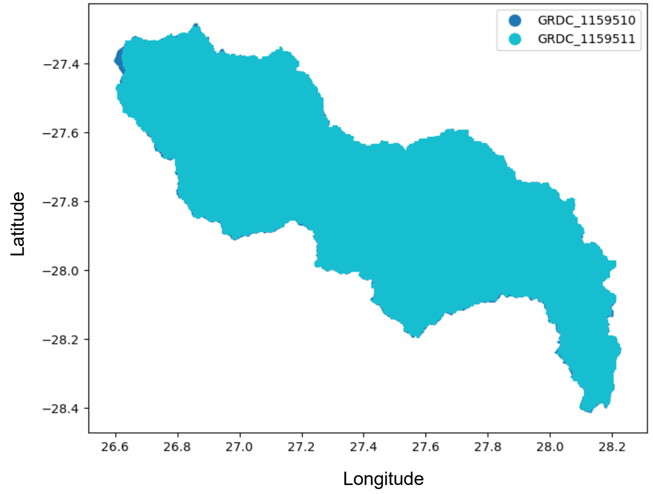
Availability

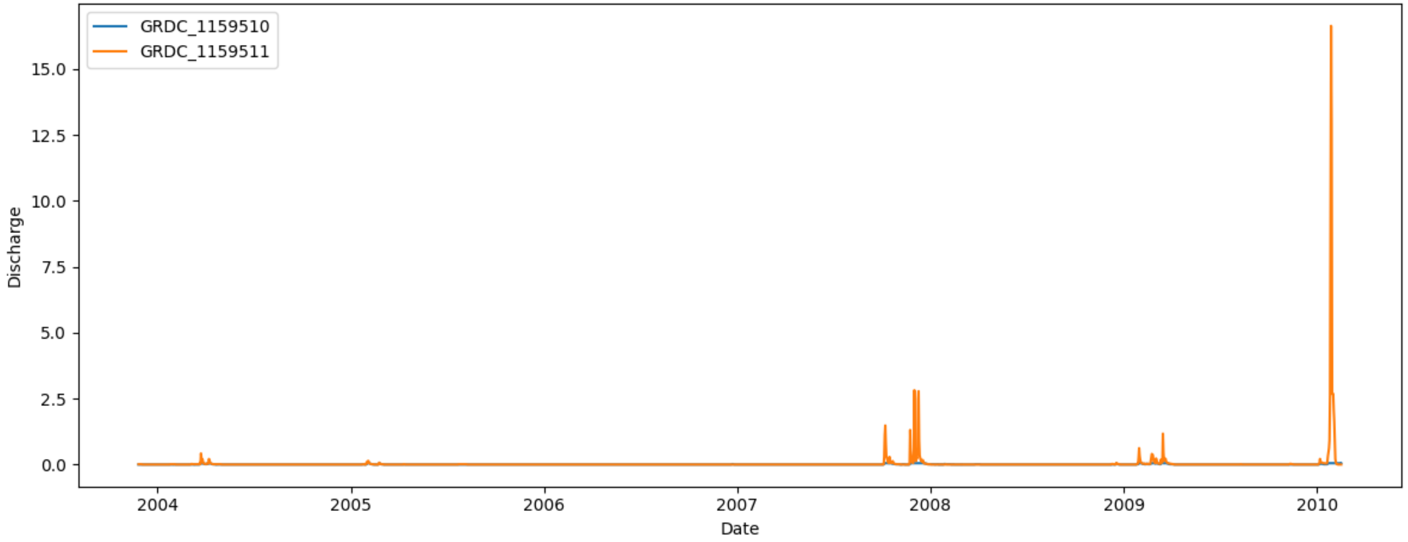
- Sparse and heterogeneous observations
- Uncertain streamflow observation quality
- Missing hydrological forcing data

Non-stationarity

- Gauge network expansion/ shrinkage, with irregular updates
- Hydrosphere is changing with climate and human influence

Challenges of hydrological data



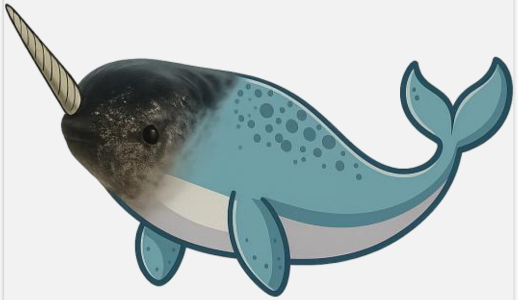


Common challenges among the other Earth system components

- Data quality & consistency
- Limited validation data
- Capturing extremes
- Scale mismatch
- Incomplete observability

Atmosphere

ERA5 reanalysis product
¼°, 6 hourly, resolution
1979-2025
~ 7 TB



Ocean and sea-ice

ORAS6 reanalysis product
¼°, 6 hourly, resolution.
2005-2025
~3 TB

Waves

Hindcast of wave model
with altimeter assimilation
¼° resolution.
1979-2025
~700 GB

Land

ERA5 reanalysis product
¼°, 6 hourly, resolution.
1979-2025
~1 TB

Hydrology

CARAVAN (global
large-sample hydrology)
(observed streamflow +
ERA5/IFS)
Catchment scale (area-
averaged)
1980-2024
~50 GB

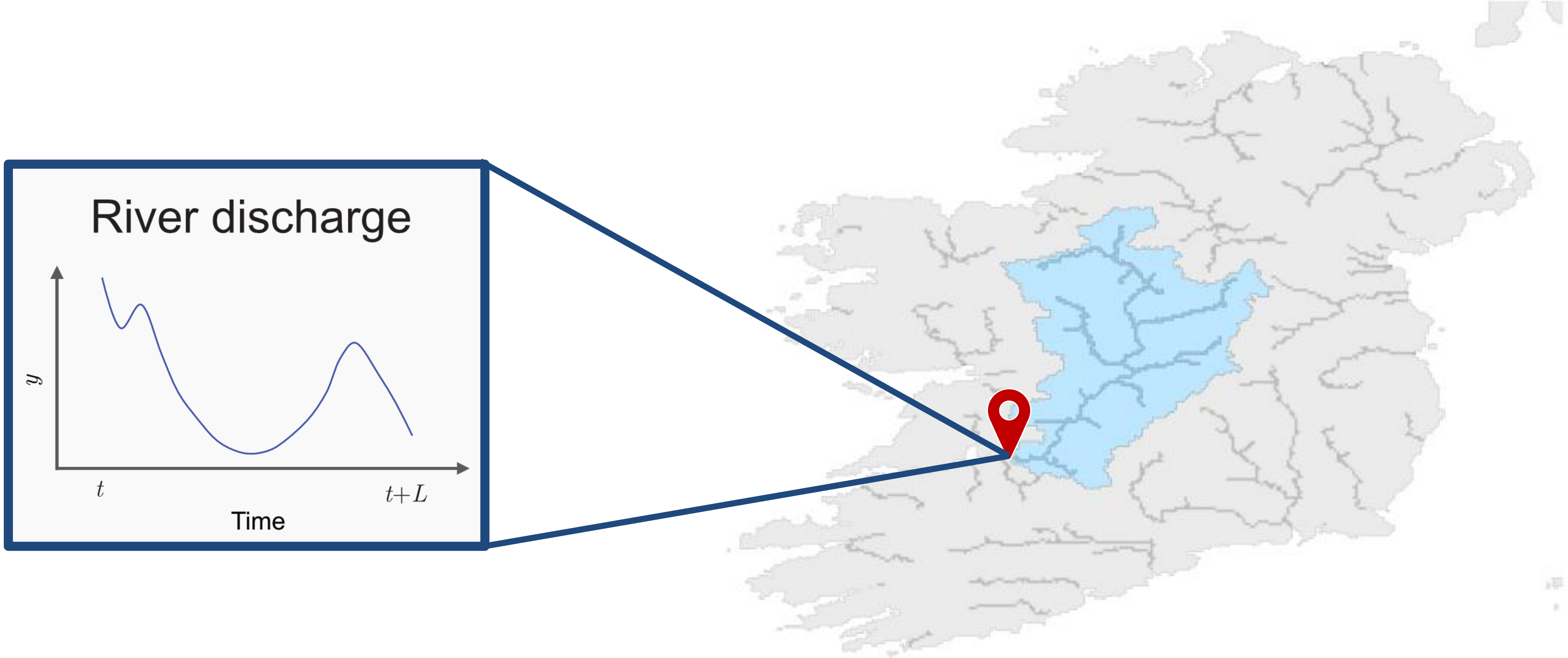
Global operational flood forecasting models

Three global operational models

	GloFAS	GEOGLOWS	Flood Hub
Developed by	Joint Research Commission	Many incl. ECMWF	Google
Operated by	ECWMF	ECMWF	Google
Operates on	Grid (distributed)	Per reach (distributed)	Per catchment (lumped)
Base model	LISFLOOD (physical)	CaMaFLOOD (physical)	LSTM (machine learning)
Forcings	ECMWF IFS meteorology	ECMWF IFS land processes	• ECMWF IFS meteorology • NASA IMERG • NOAA CPC
Output	Ensemble of streamflow predictions	Ensemble of streamflow predictions	• Mean of parameterised asymmetric Laplacian streamflow distribution • Water level

💫 Create our own operational machine learning model!

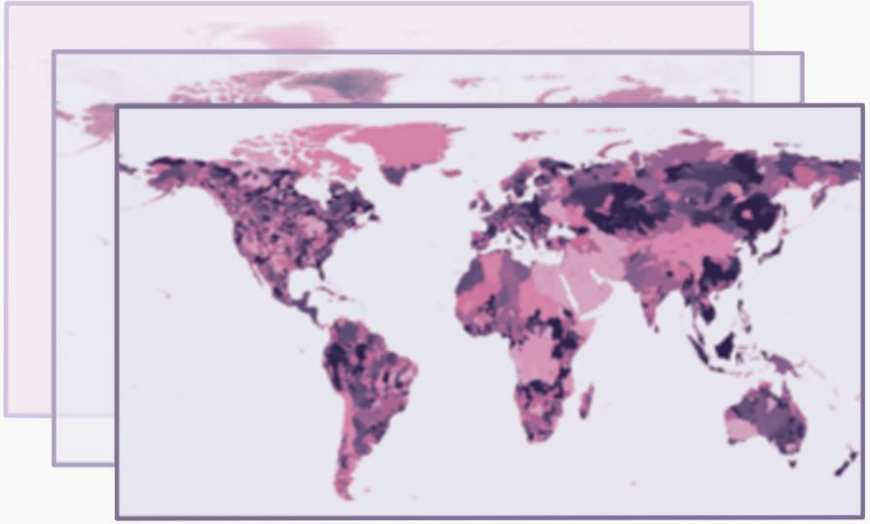
Quick overview of the workflow



Gather Inputs

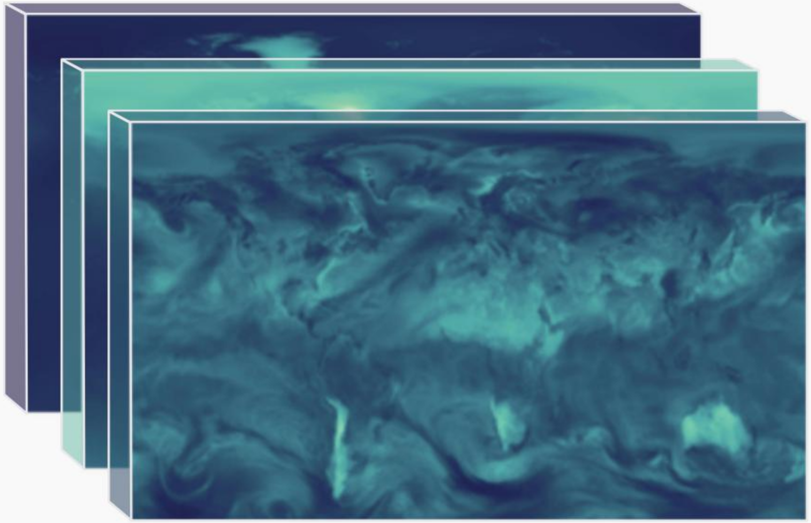
Static catchment attributes

208 HydroATLAS/ ERA5 Land fields

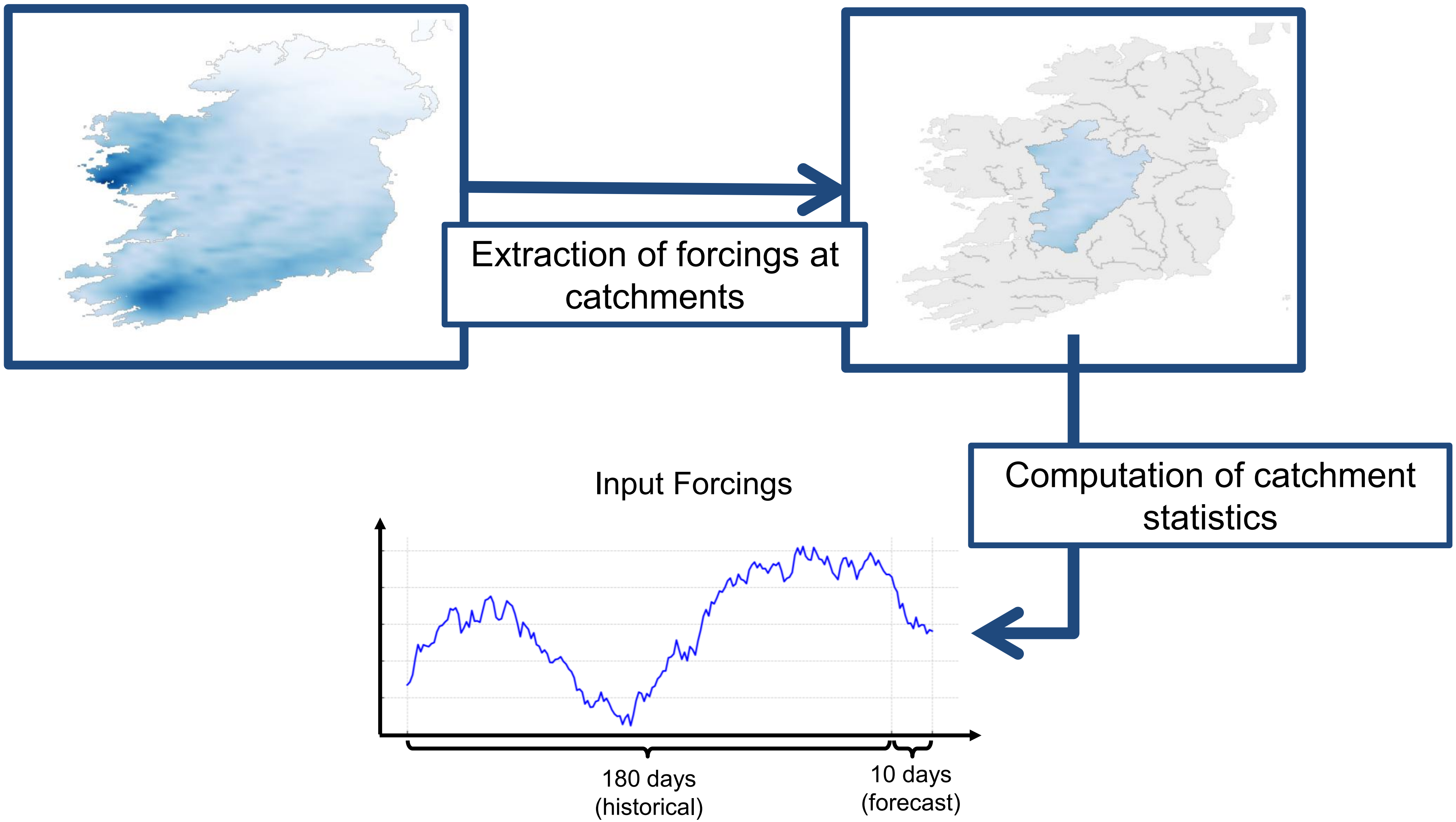


Dynamic forcings

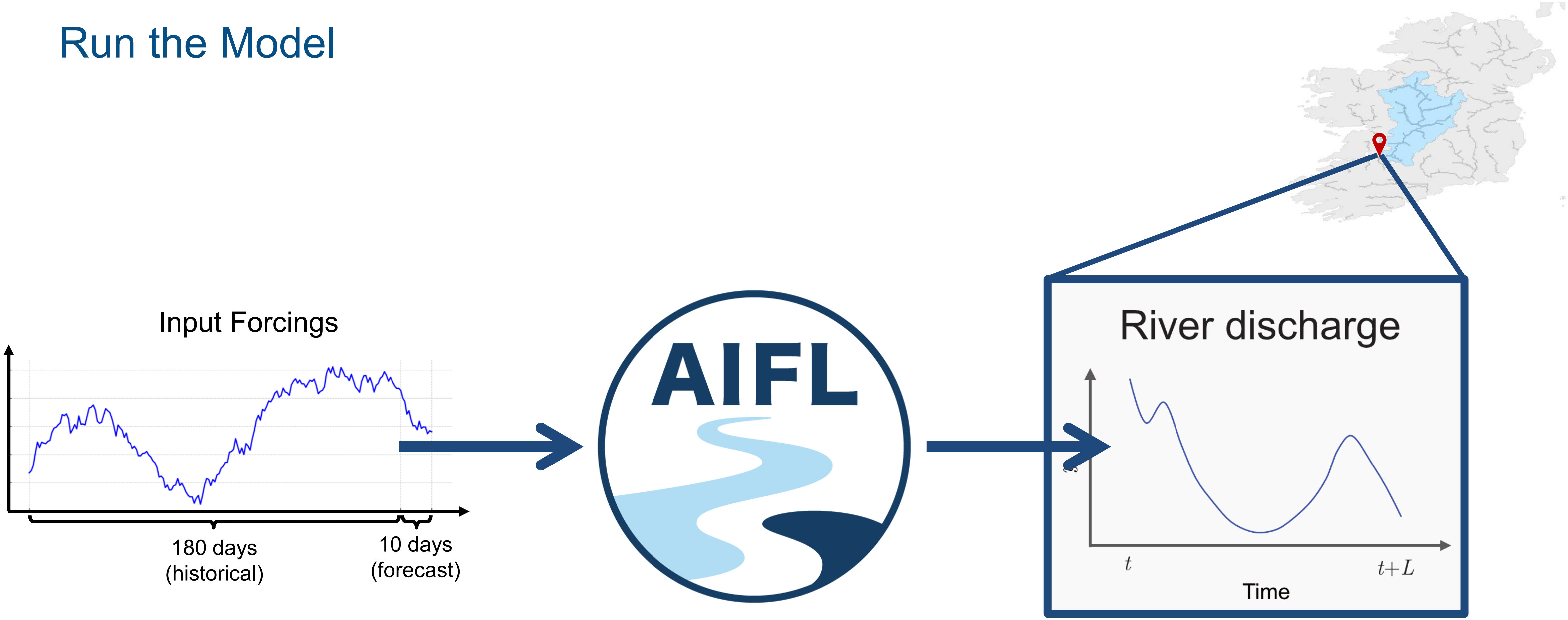
5 ERA5 Land/IFS HRES fields
+ 4 periodic time inputs



Transform Inputs

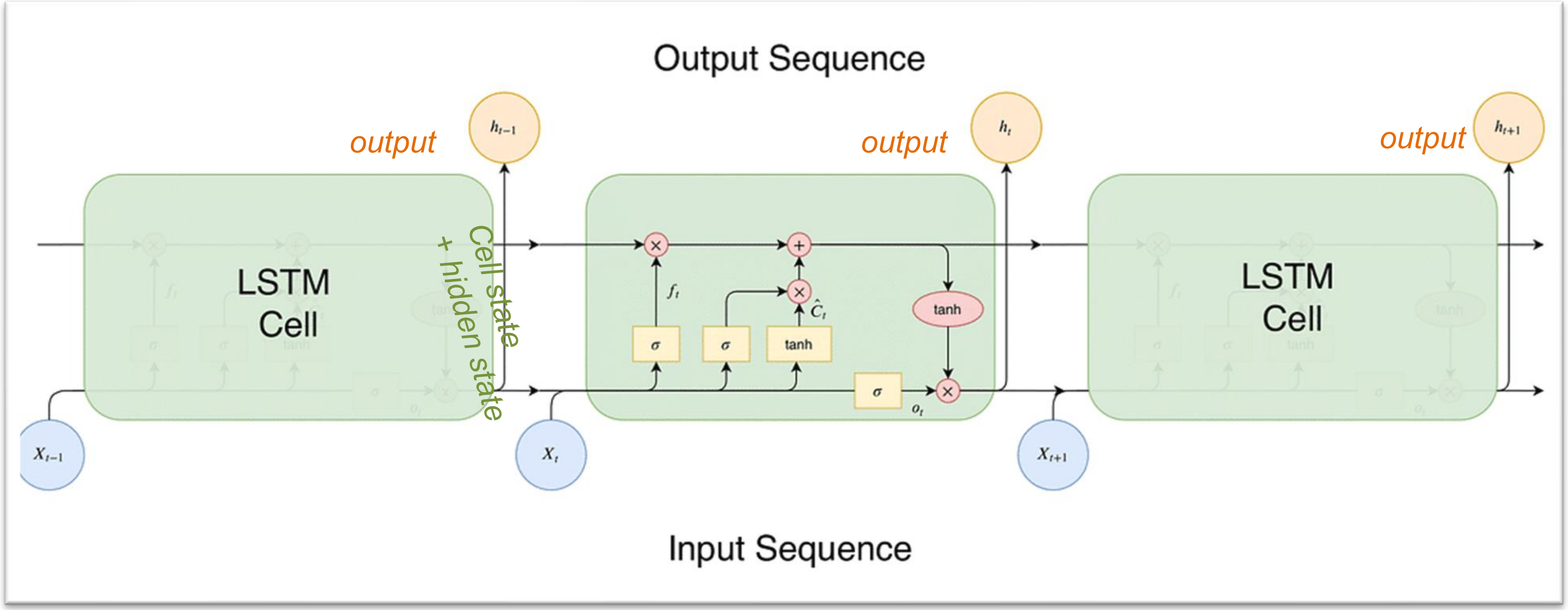


Run the Model

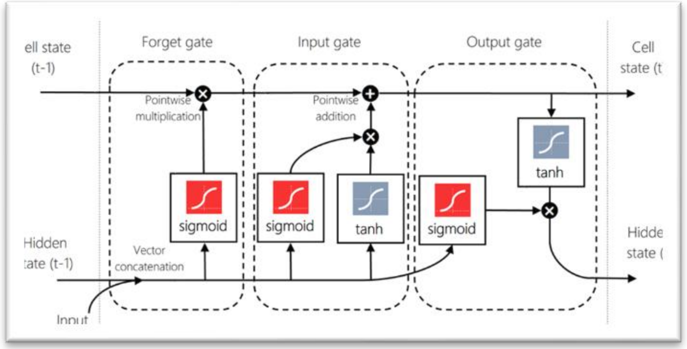


… and repeat for thousands of stations!

Background: LSTM model

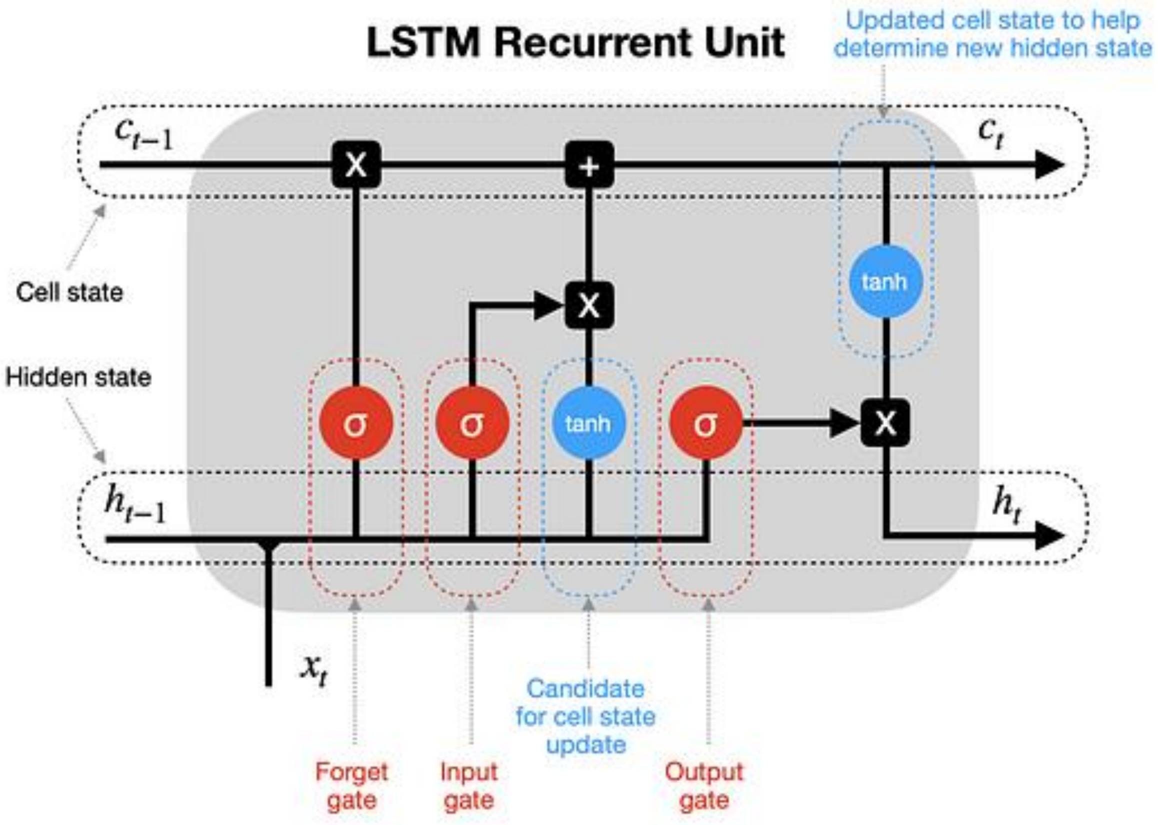


An LSTM model consists of a sequence of repeated LSTM cells, each processing one time step and sharing the same parameters across time.



LSTM CELL

Background: LSTM cell

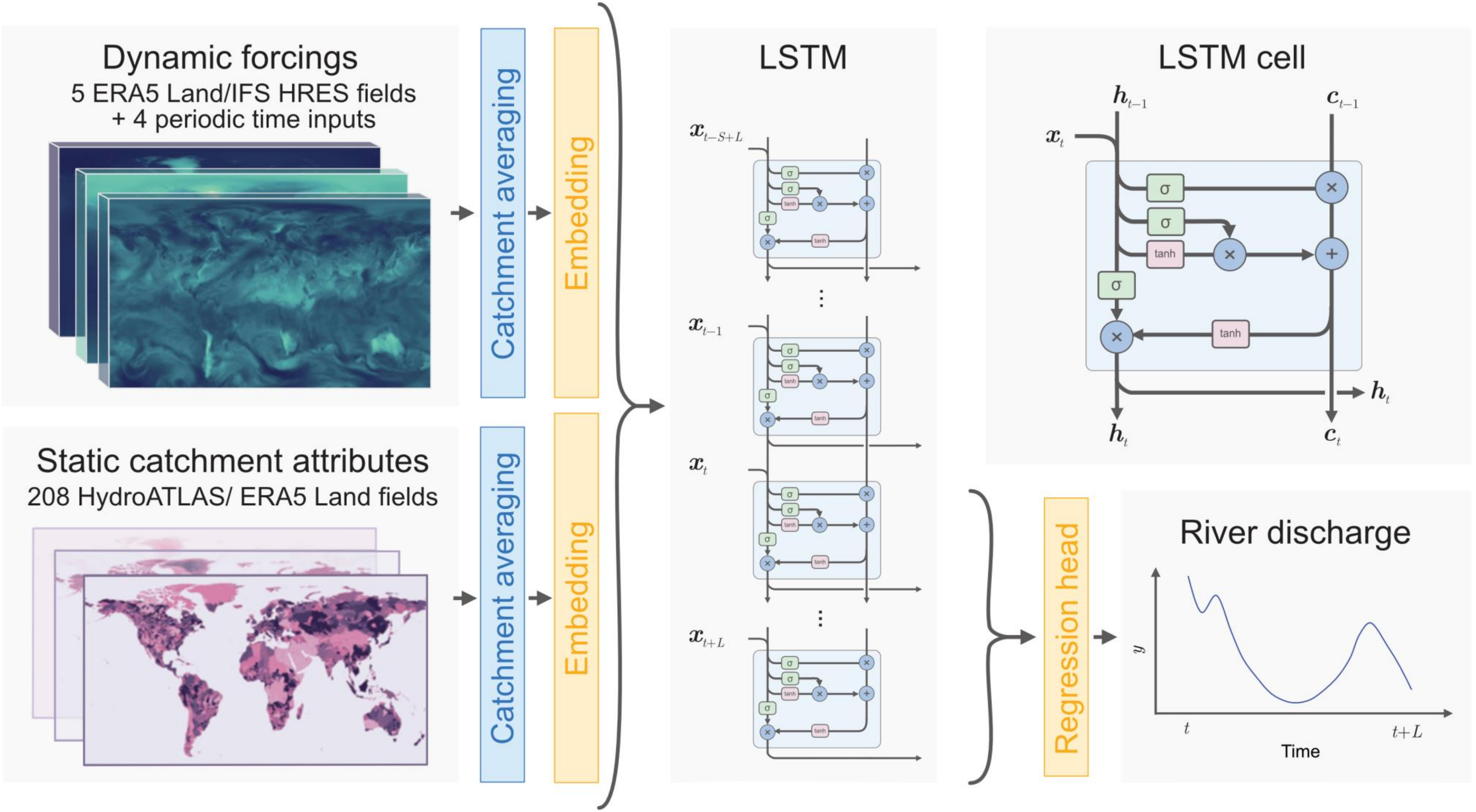


LSTM Architecture | Photo based on Towards Data Science

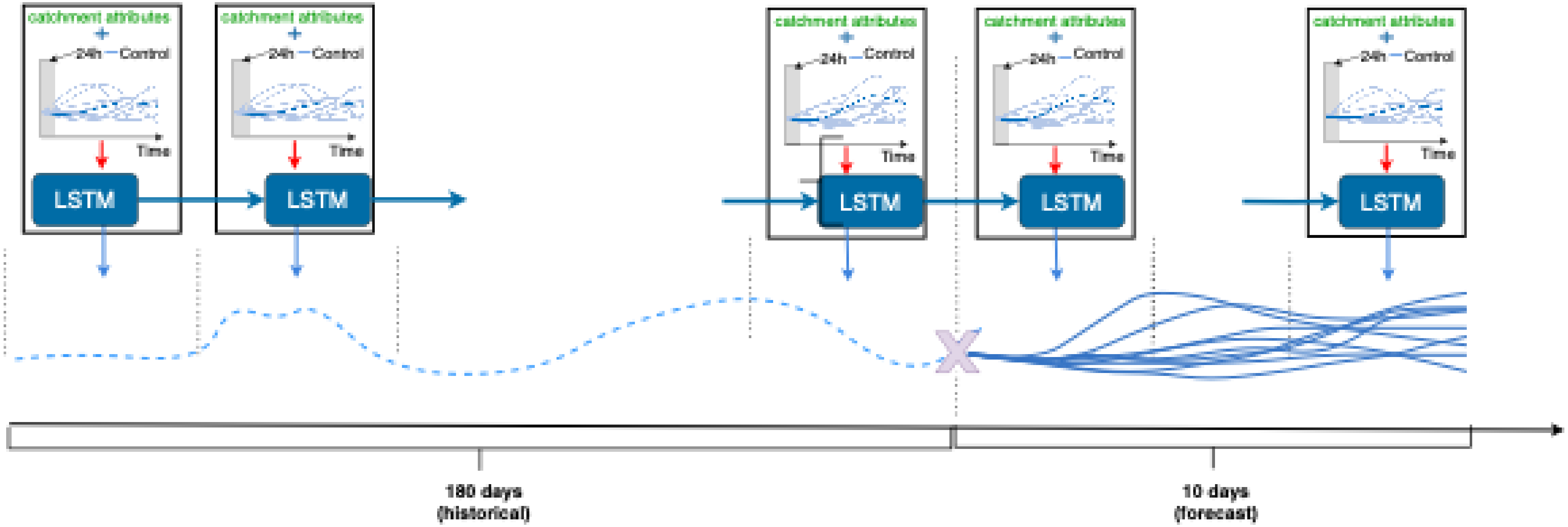
- **Forget Gate**: $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$
- **Input Gate**: $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$
- **Candidate Cell State**: $\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$
- **Current Cell State**: $C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$
- **Output Gate**: $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$
- **Hidden State**: $h_t = o_t \cdot \tanh(C_t)$

- x_t: Current input
- h_{t-1}: Previous hidden state
- C_{t-1}: Previous cell state
- C_t: Current cell state
- $\tilde{C}_t$: Candidate cell state
- W_f, W_i, W_C, W_o: Weight matrices for the input
- U_f, U_i, U_C, U_o: Weight matrices for the hidden state
- b_f, b_i, b_C, b_o: Biases
- σ: Sigmoid function
- $\tanh$: Hyperbolic tangent function

AIFL architecture



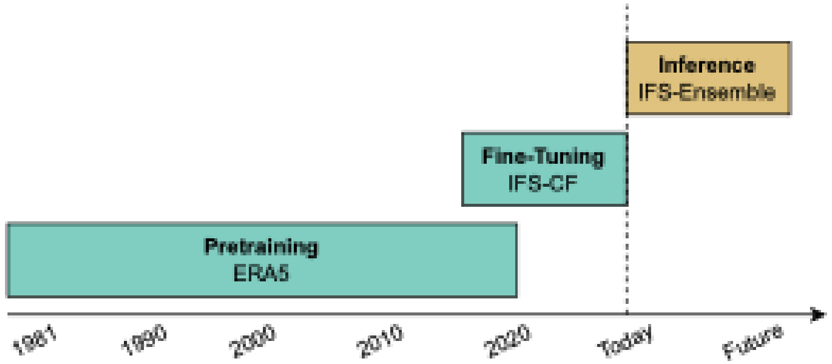
Kenza Tazi

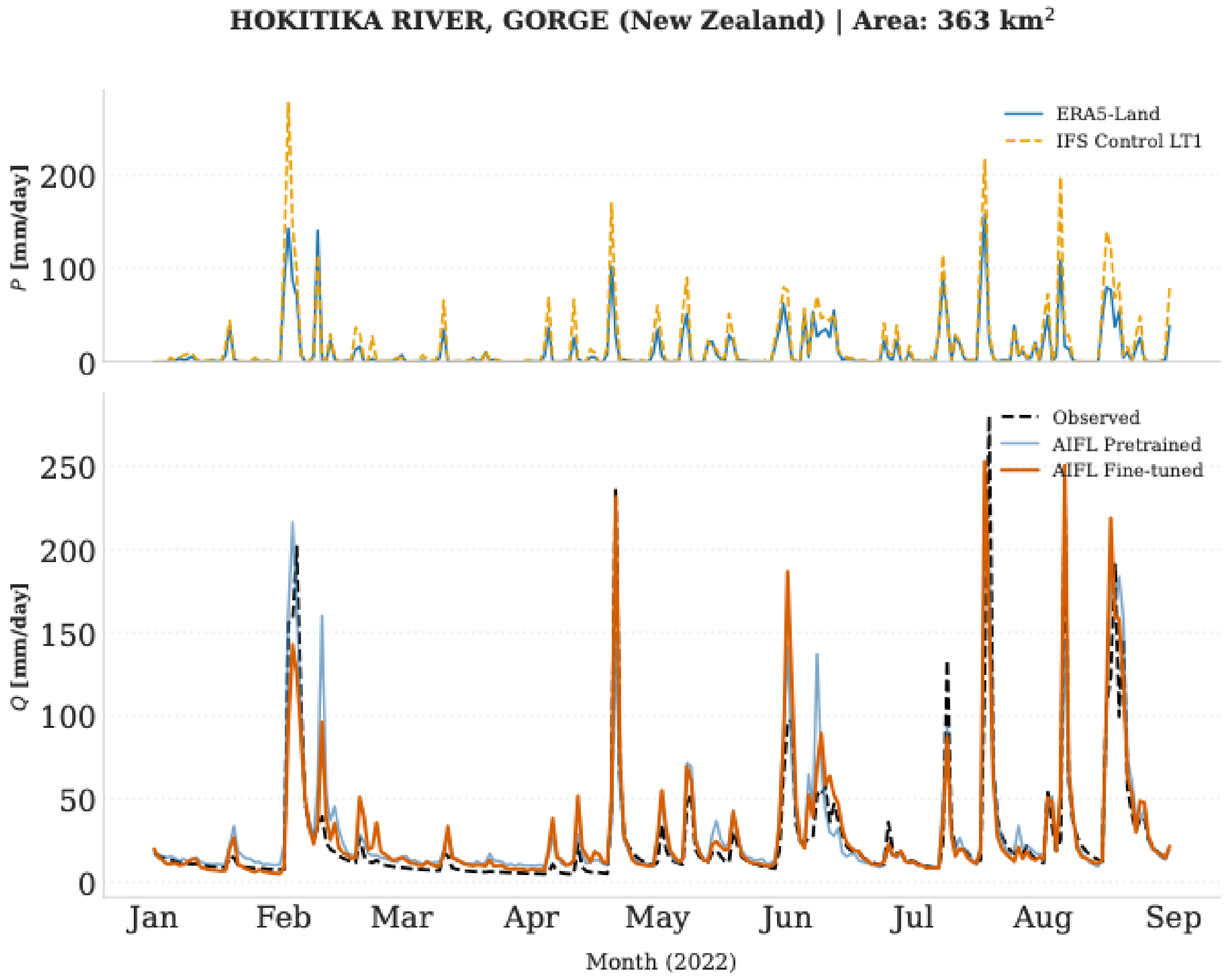

catchment attributes
+
24h
Control
Time
LSTM
180 days
(historical)
10 days
(forecast)

Training strategy

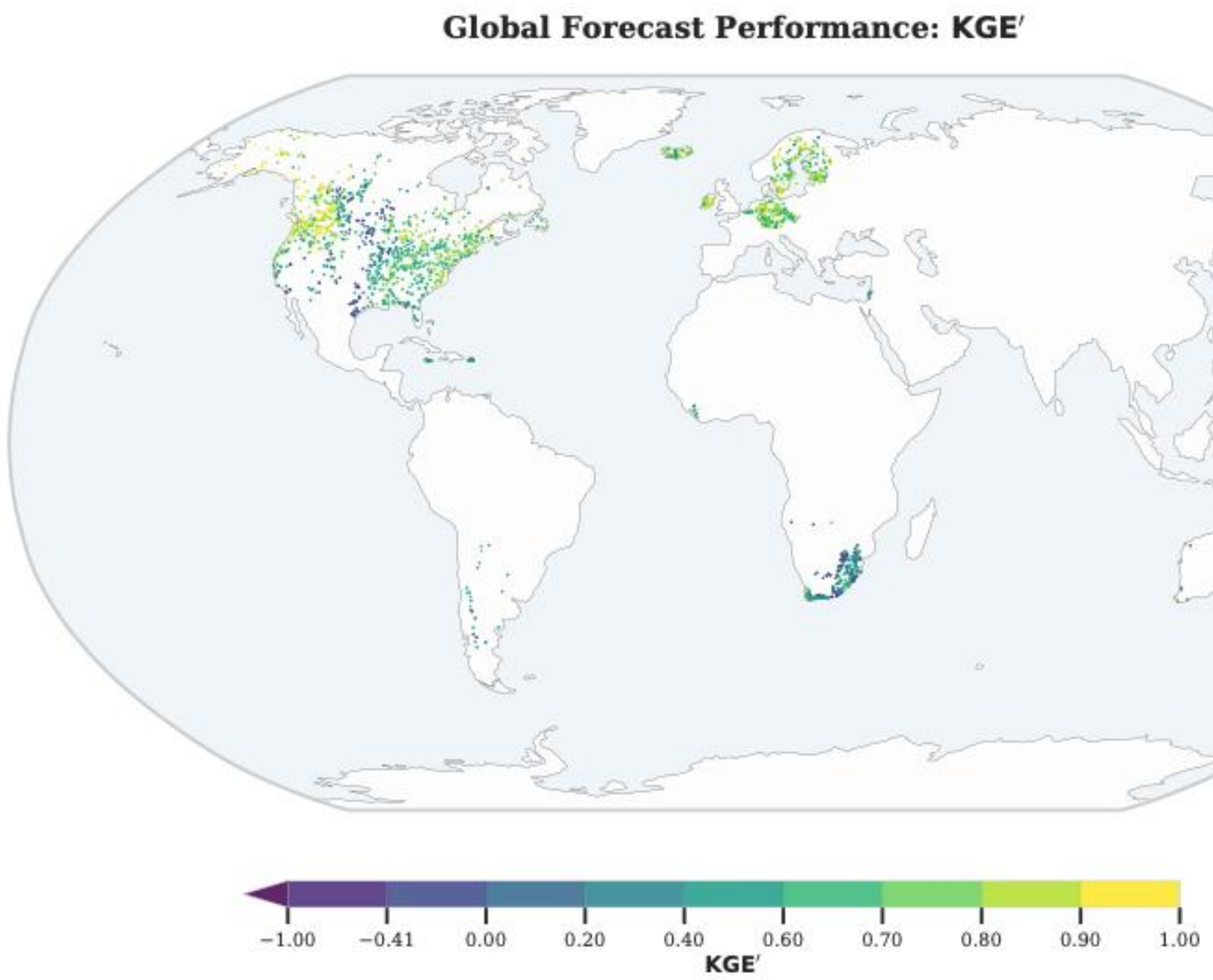
Fine-tune to overcome distributional shifts

- Pretrain on ERA5 Land between 1980 and 2019
- Finetune on IFS HRES between 2016 and 2019

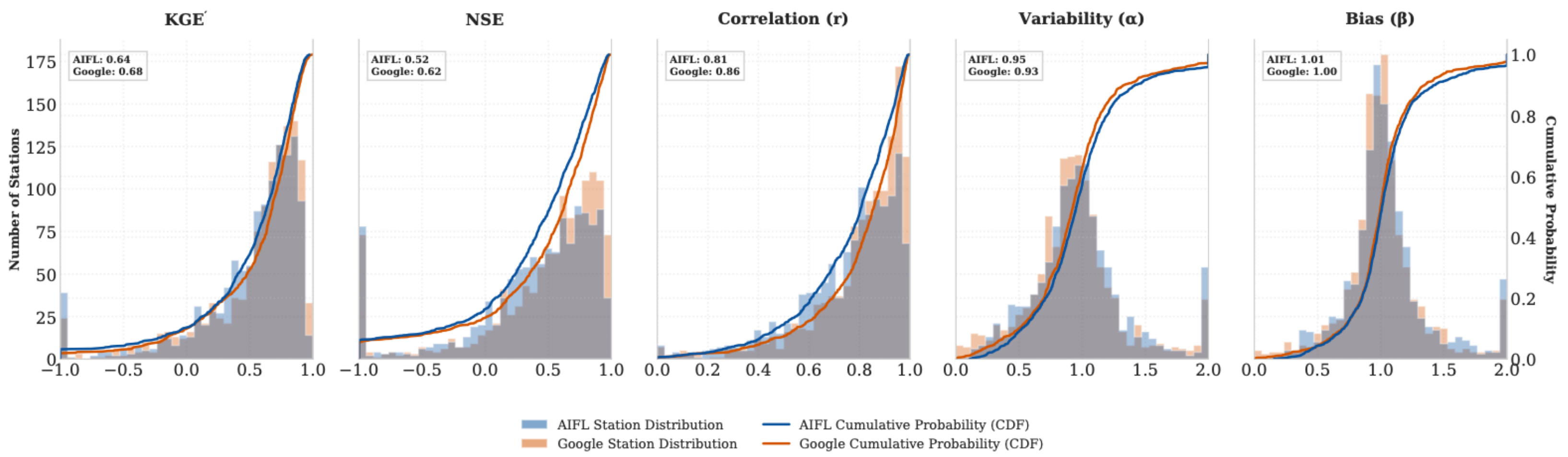




AIFL performance



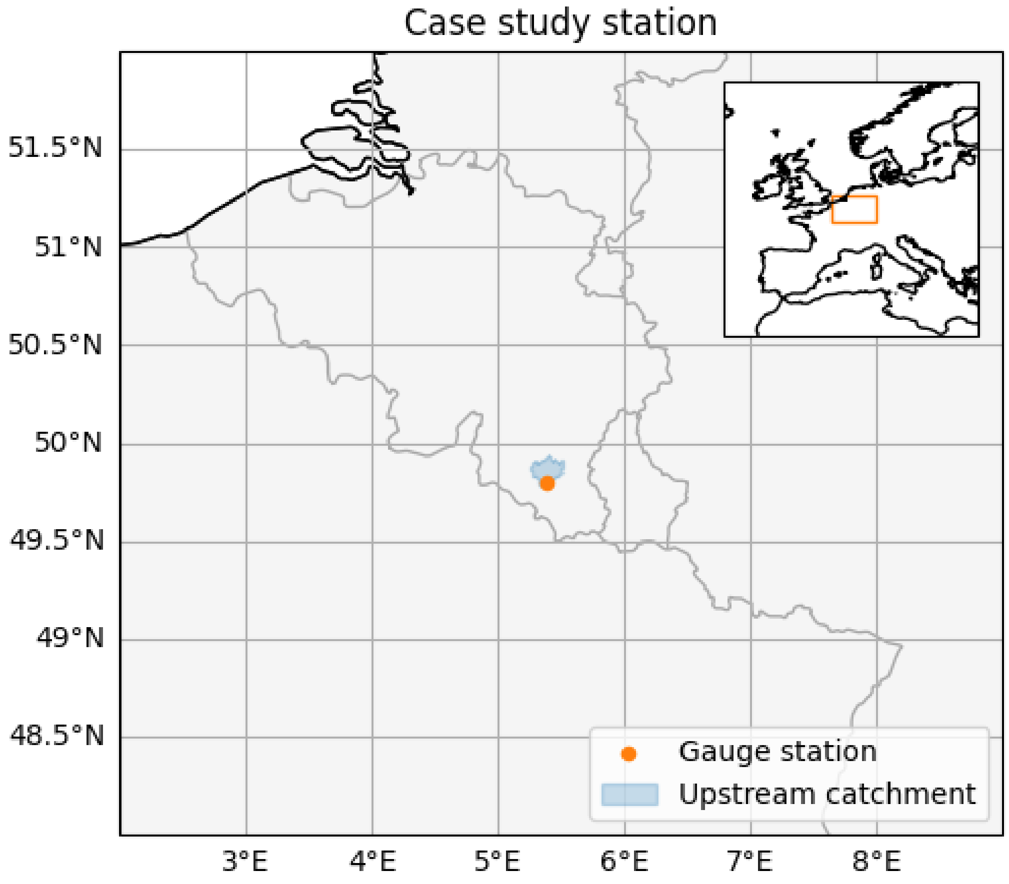
AIFL performance



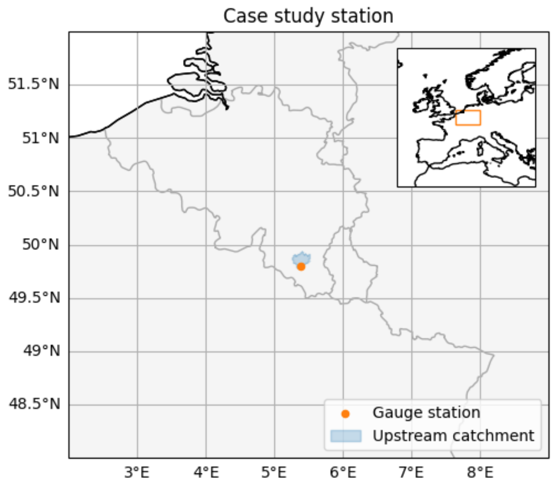
	Google*	AIFL
NSE	**0.624**	0.518
KGE’	**0.678**	0.636
r	**0.857**	0.808
α	0.927	**0.955**
β	**1.000**	1.007

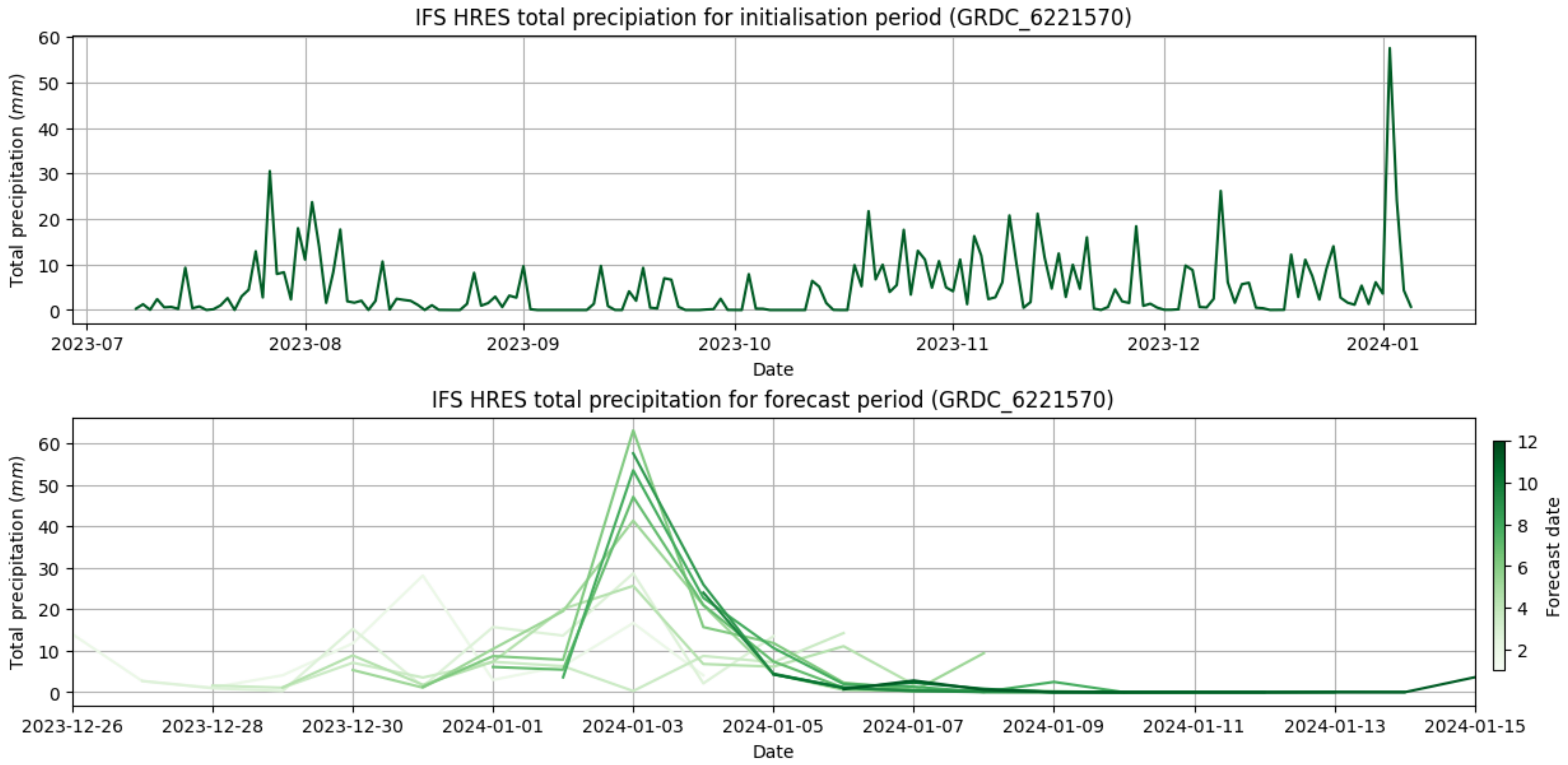
* Nearing et al. "Global prediction of extreme floods in ungauged watersheds." *Nature* 627.8004 (2024): 559-563

Case study: 2024 Luxembourg province flood



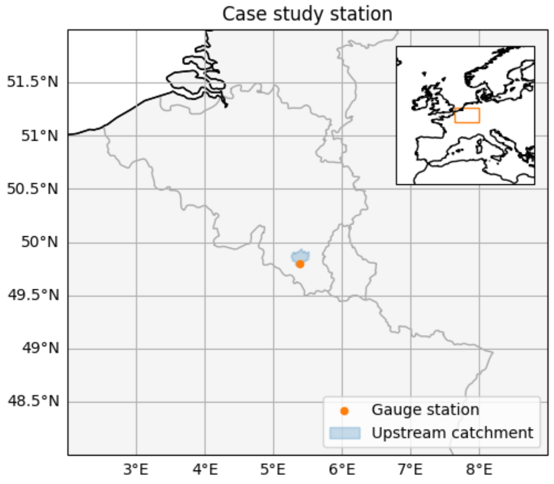
Case study: 2024 Luxembourg province flood

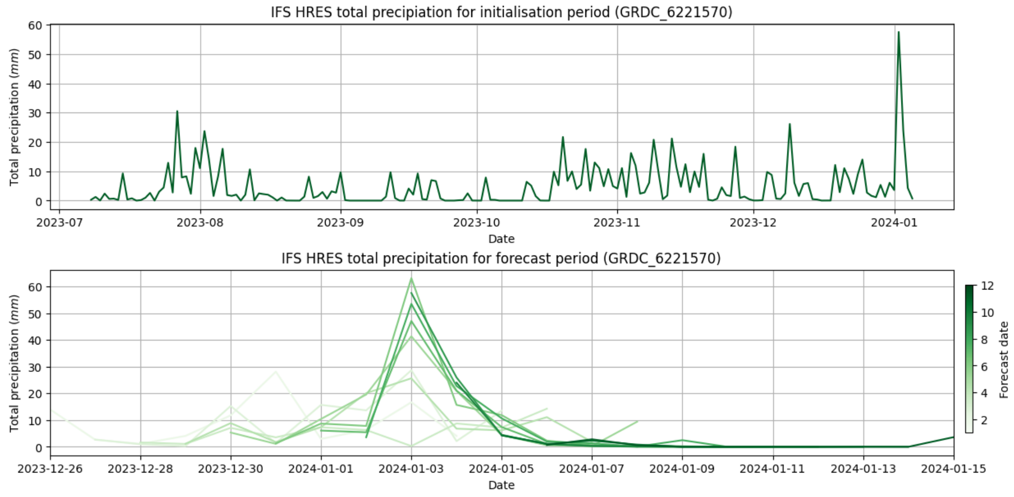


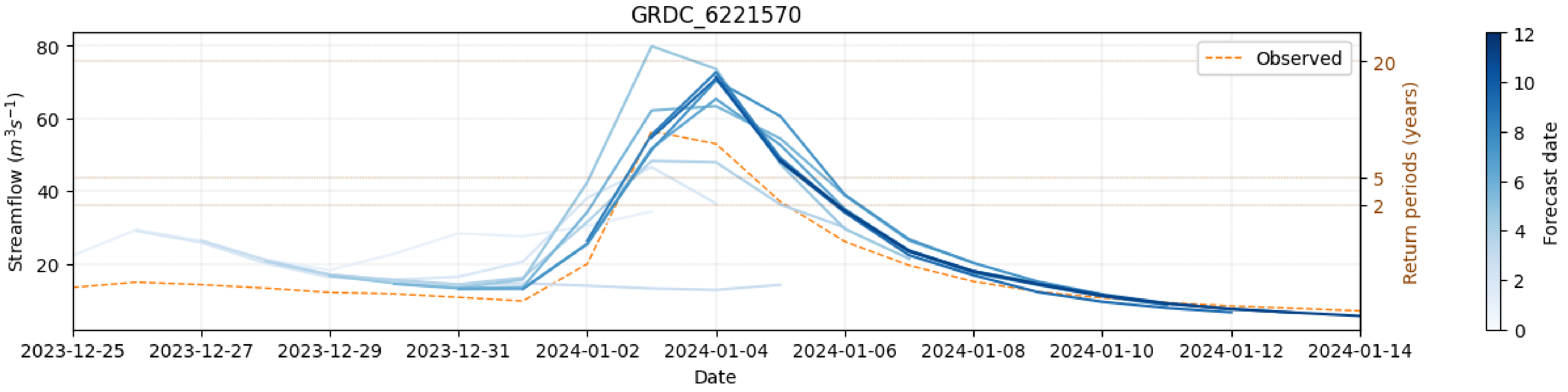


Kenza Tazi

Case study: 2024 Luxembourg province flood



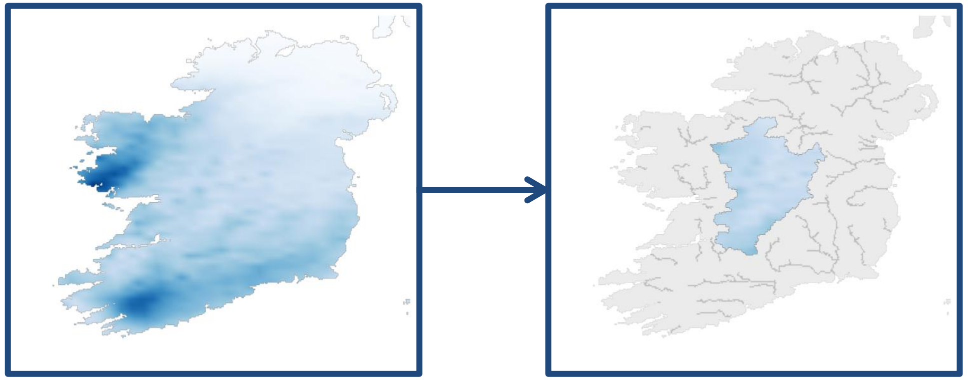




Kenza Tazi

Efficient operations

- Need to run the model every day at the same for each of the IFS ensemble member
- Most onerous task is the dynamic forcing extraction
- Use `earthkit-hydro`
  - Quick access to popular river networks: EFAS, GloFAS, CaMa, and GRIT
  - Efficient extraction of catchments and fast statistic and metric calculations
  - High (`xarray`) and low-level interfaces (`numpy/torch/cupy`)
  - Bottleneck acceleration using `rust`
  - Designed for machine learning (GPU acceleration, automatic differentiation, etc.)
- Speed up data processing from hours to minutes



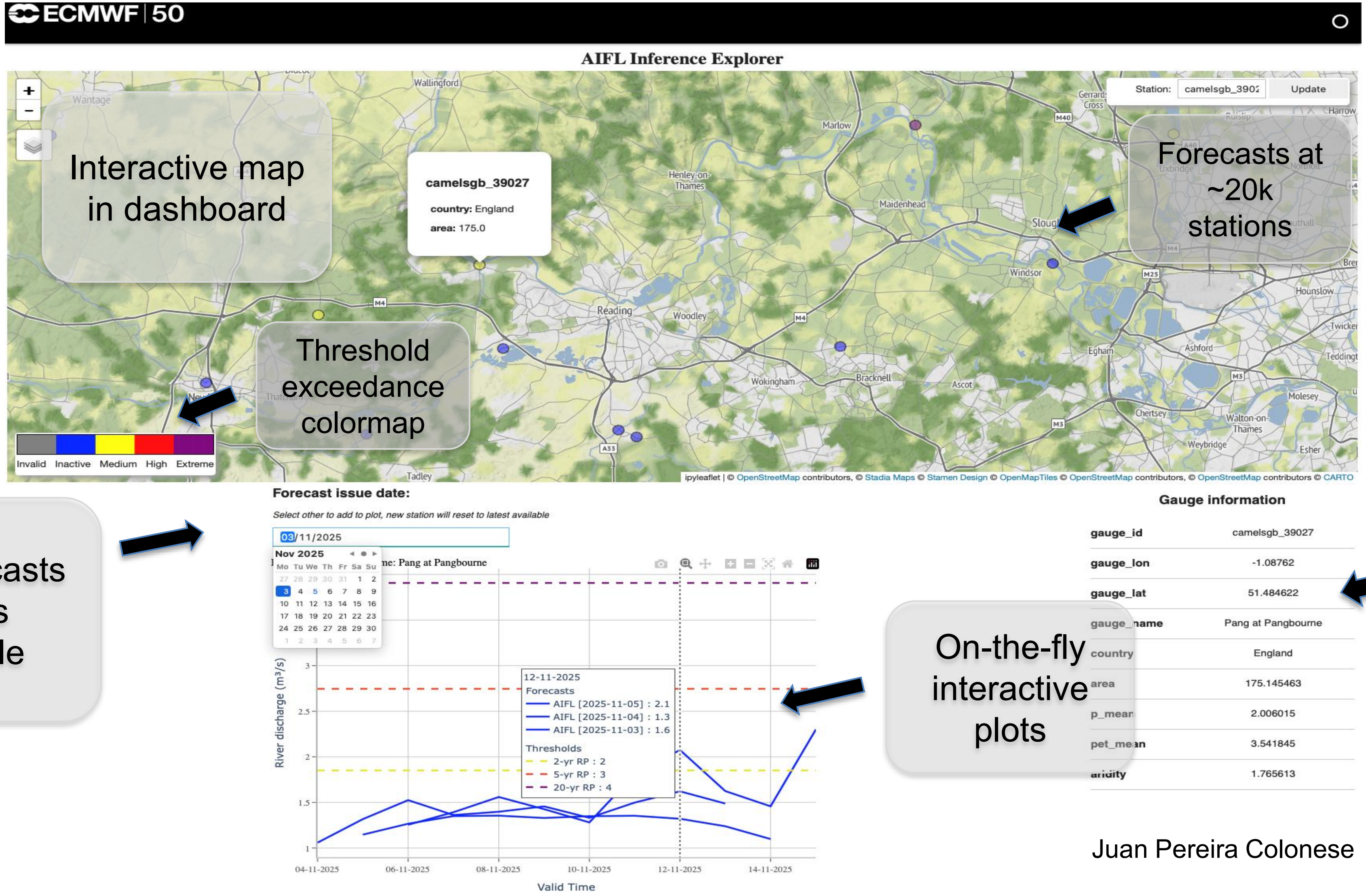
Osìn Morrison

earthkit hydro

Effective communication

Station text query box

Layer or basemap selection



Interactive map in dashboard

Forecasts at ~20k stations

Threshold exceedance colormap

Past forecasts always available

On-the-fly interactive plots

Station attributes table

Juan Pereira Colonese

Future work

Flood forecasting models

- Improved model with 14-day horizon (AIFL v2) for which we will
  - More dynamic fields
  - Custom static fields
  - Different architectures
  - Tailored loss functions
  - Pretraining on GloFAS
  - and more!
- Probabilistic model global forecasting model
- European 6-hourly model

Operations

- Ensemble operational pipeline to visualise distribution
- Real-time verification
  - Day-to-day comparison with other operational models
  - Event-based analysis against proxy datasets

Events

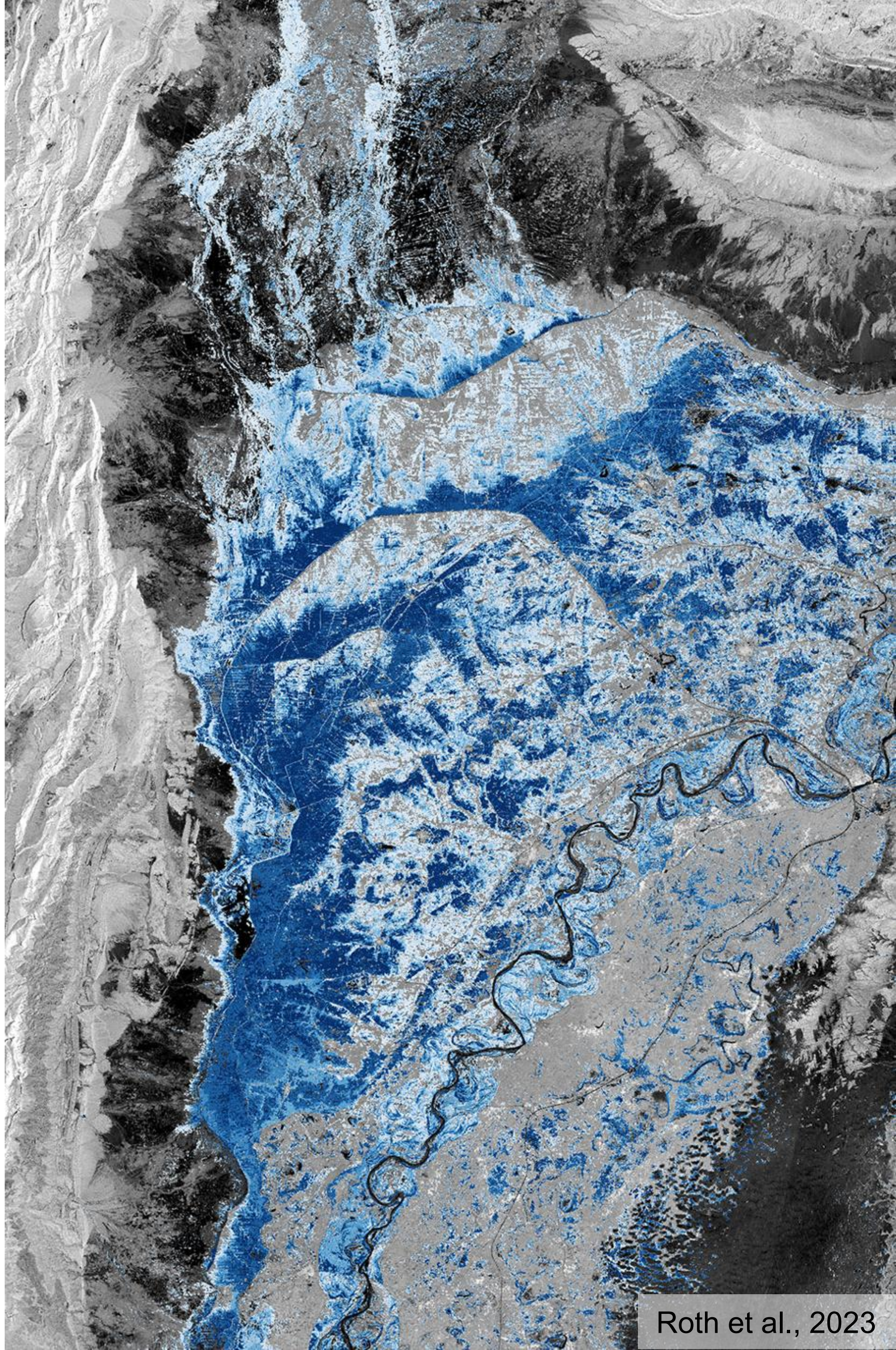
- 2026 Workshop on Machine Learning for Land and Hydrology
- EGU session: HS3.5 Deep Learning in Hydrology
- Code for Earth

Key takeaways

1. Flood forecasting is extremely valuable but remains challenging

2. AIFL, a LSTM-based streamflow model comparable to Google's Flood Hub

3. Efficient operations and effective communication also needs to be considered



Roth et al., 2023

Thank you for your attention!