Hydrological modelling

Kenza Tazi

ML weather forecasting training

19th November 2026

The rising threat of floods

Floods, most common natural disaster

- 55 million people directly affected in 2022 [1]
- More frequent and severe with climate change
- Projected economic loss of USD 5.6 trillion by 2050 [2]

Predictions, vital to mitigation and response strategies

- 6x lower mortality [3]
- 12-hours ahead ⇔ up to 60% reduction in economic damages [4]
- Most-needed in low-income countries with greater flood risks



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

Outline

1. Challenges and state-of-the art in flood forecasting
2. Introduction to state space models
3. ECMWF's AI for Floods (AIFL) model
4. Hydrology and the AIFS
5. Towards coupled models



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

1 | Challenges and state-of-the art

Challenges of hydrological data

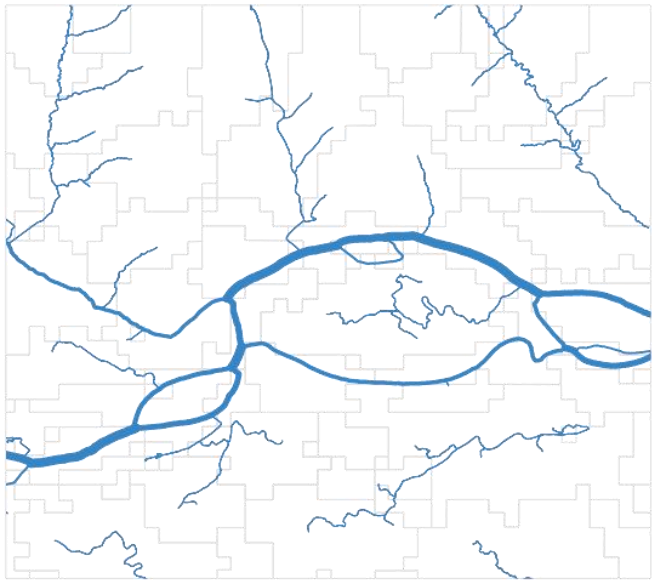
Complexity

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

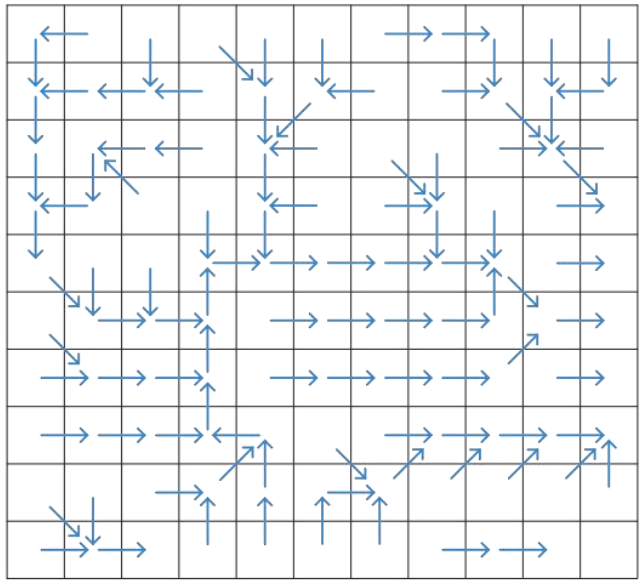
Challenges of hydrological data



Rio Negro and Solimões River



Vector-based network



Grid-based network

Challenges of hydrological data

Complexity

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

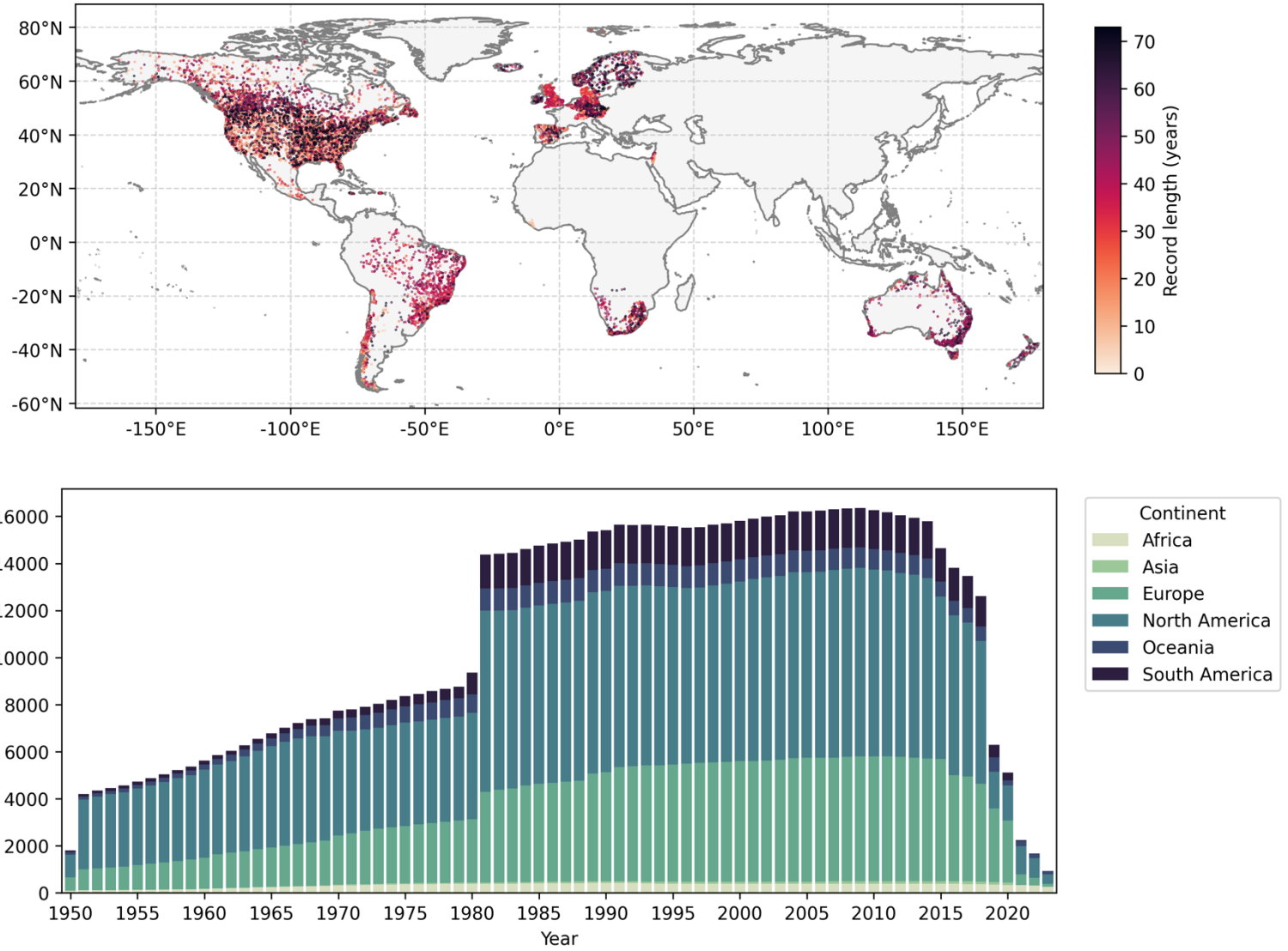
Availability

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

Challenges of hydrological data

Spatiotemporal distribution of the Caravan dataset [5]

[5] Kratzert et al. "Caravan – A global community dataset for large-sample hydrology." Scientific Data 10.1 (2023): 61.



Challenges of hydrological data

Complexity

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

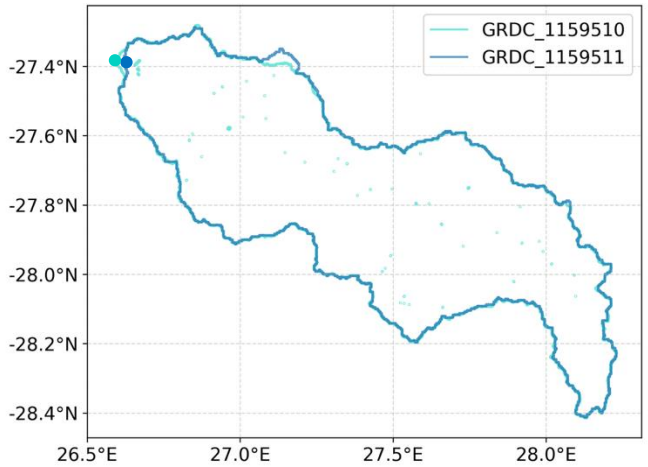
Availability

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

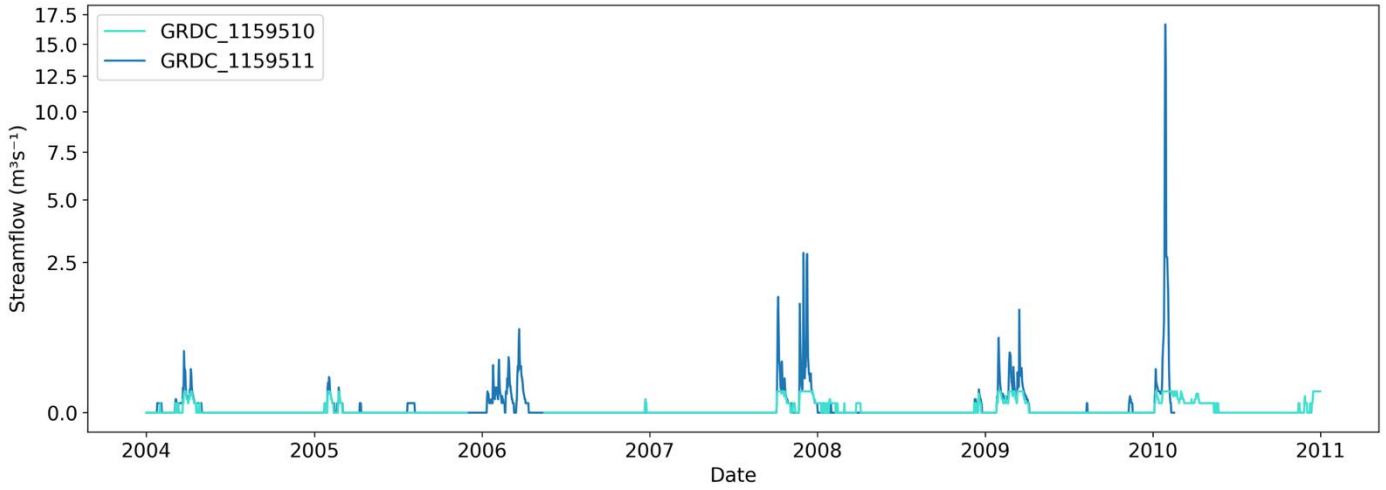
Non-stationarity

- Gauge network expansion/ shrinkage, with irregular updates
- Hydrosphere changes naturally and under human influence

Challenges of hydrological data





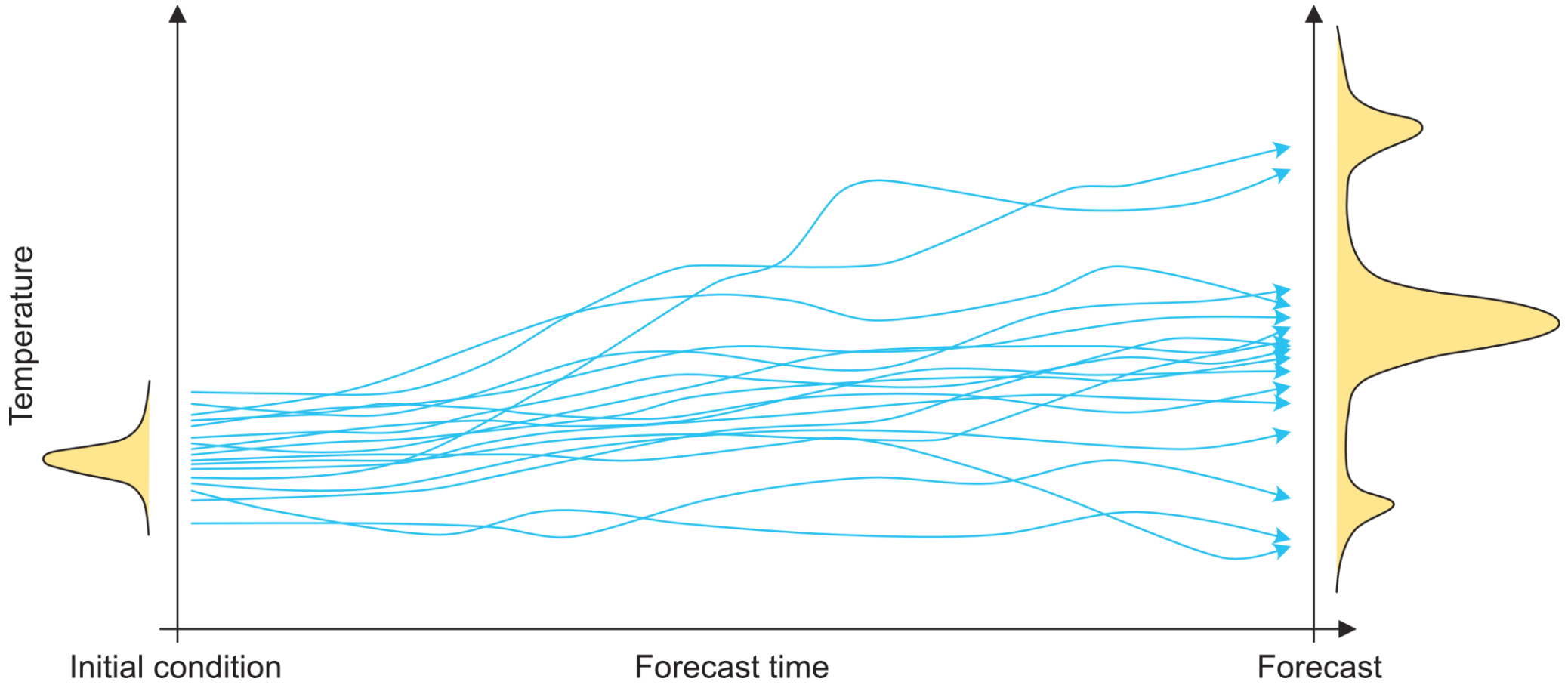


Global operational flood forecasting models

Three global operational models

	GloFAS	GEOGLOWS	Flood Hub
Developed by	Joint Research Commission	Group on Earth Observation (GEO)	Google
Operated by	ECWMF	Many incl. 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, GraphCast
Output	Ensemble of streamflow predictions	Ensemble of streamflow predictions	Mean of parameterised asymmetric Laplacian streamflow distribution + water level

Uncertainty through ensembles



Global operational flood forecasting models

Three global operational models

	GloFAS	GEOGLOWS	Flood Hub
Developed by	Joint Research Commission	Group on Earth Observation (GEO)	Google
Operated by	ECWMF	Many incl. 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, GraphCast
Output	Ensemble of streamflow predictions	Ensemble of streamflow predictions	Mean of parameterised asymmetric Laplacian streamflow distribution + water level

Global operational flood forecasting models

Three global operational models

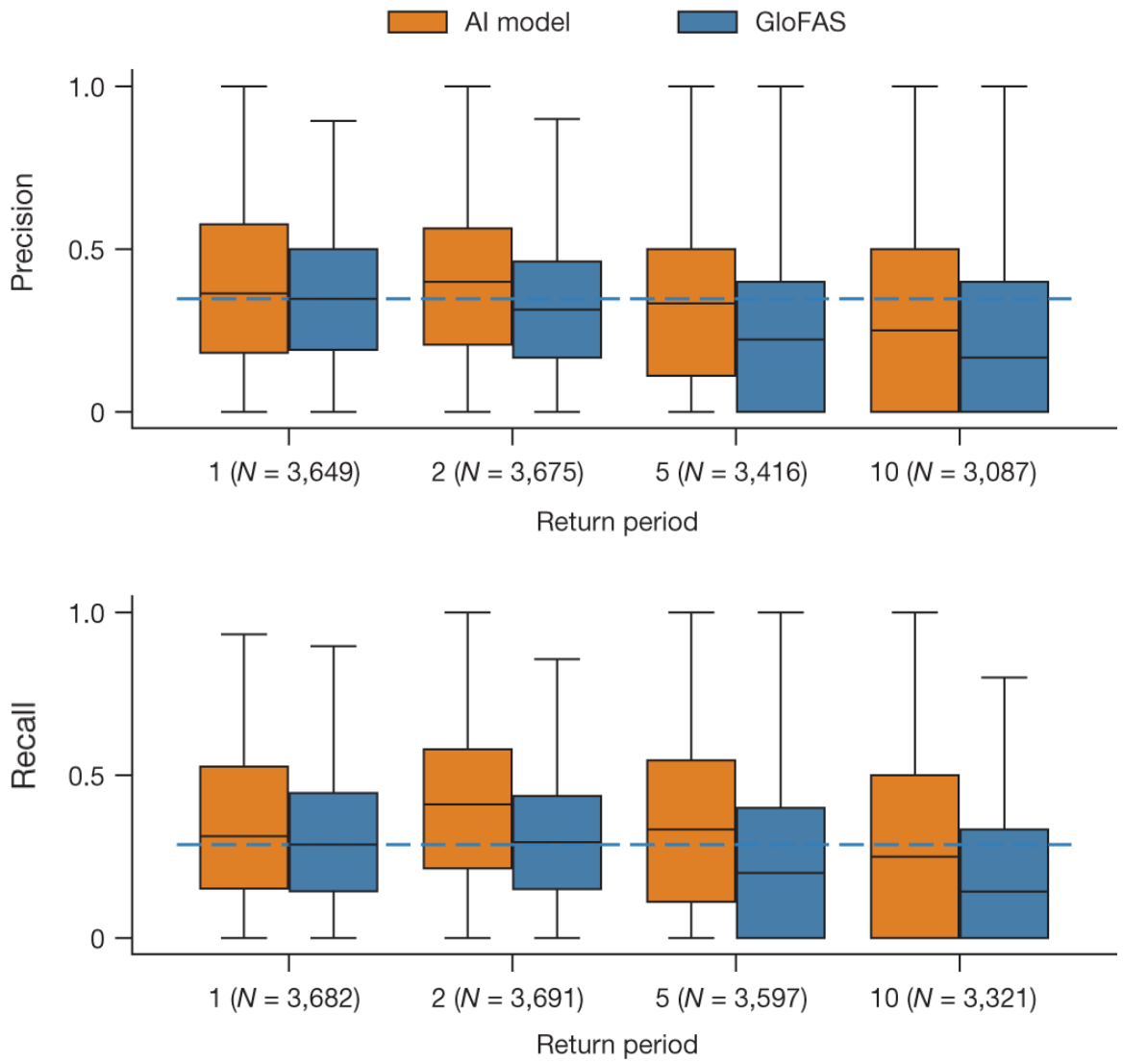
	GloFAS	GEOGLOWS	Flood Hub
Developed by	Joint Research Commission	Group on Earth Observation (GEO)	Google
Operated by	ECWMF	Many incl. ECMWF	Google
Operates on	Grid (distributed)	Per reach (distributed)	~~Per catchment (lumped)~~ > Semi-distributed
Base model	LISFLOOD (physical)	CaMaFLOOD (physical)	LSTM (machine learning)
Forcings	ECMWF IFS meteorology	ECMWF IFS land processes	ECMWF IFS meteorology, NASA IMERG, NOAA CPC, GraphCast
Output	Ensemble of streamflow predictions	Ensemble of streamflow predictions	Mean of parameterised asymmetric Laplacian streamflow distribution + water level

Machine learning for hydrology

Many examples of machine learning for streamflow forecasting

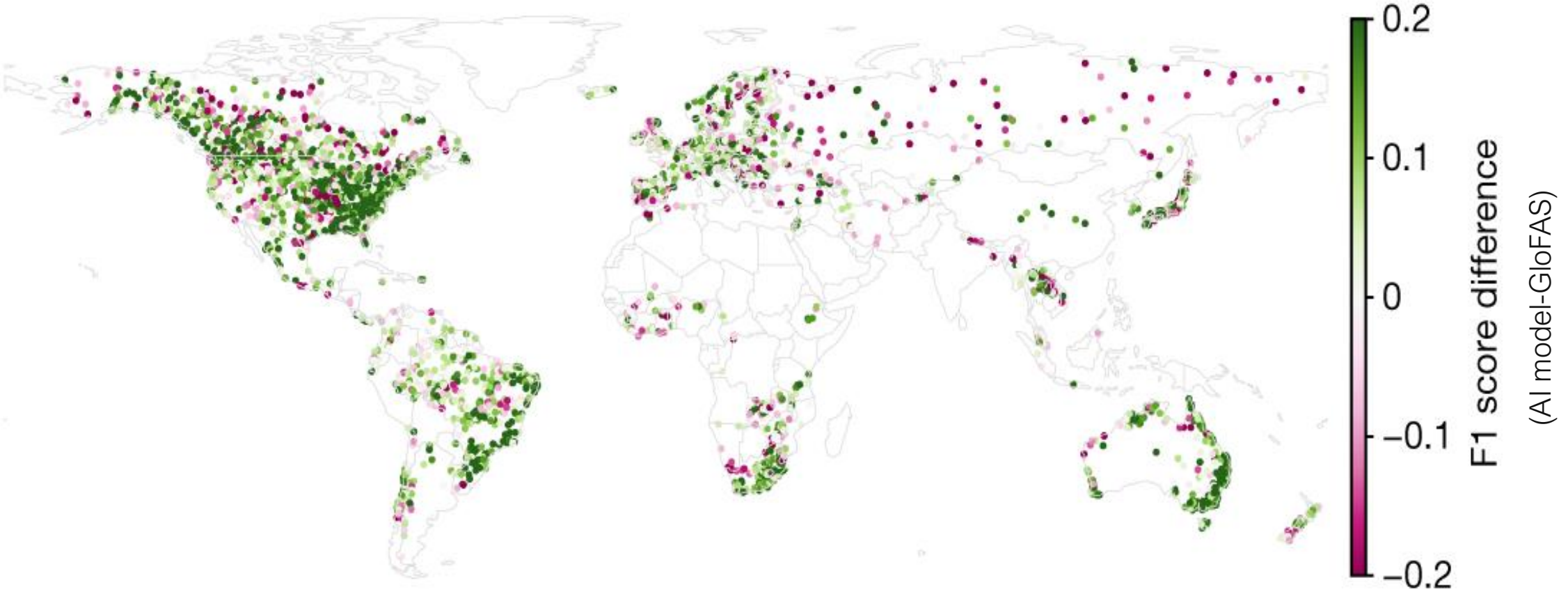
- Inductive bias helps leverage limited data
- Larger variety of data helps generalisation
- Enforcing physics doesn't always help

→ In 2024, Google model outperformed GloFAS in many aspects [6]



Machine learning for hydrology

Non-uniform performance



2 | Introduction to state space models

Memory in the machine

An introduction to state space models

THE STATE SPACE IDEA

u[t] inputs

e.g. precipitation, temperature, radiation

x[t] state

e.g. snow, soil water, groundwater

y[t] output

e.g. streamflow at station

$$\mathbf{x}[t] = f(\mathbf{x}[t-1], \mathbf{u}[t]) \quad \# \textit{ state update}$$

$$\mathbf{y}[t] = g(\mathbf{x}[t]) \quad \# \textit{ readout}$$

Memory in the machine

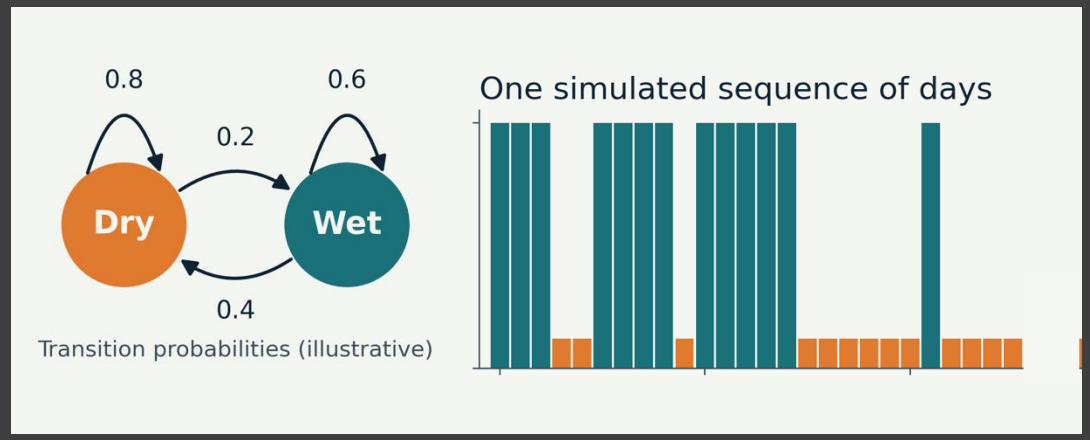
An introduction to state space models

SIMPLE CASES

Markov chain

$$\mathbf{x}[t] = \mathbf{A}\cdot\mathbf{x}[t-1]$$

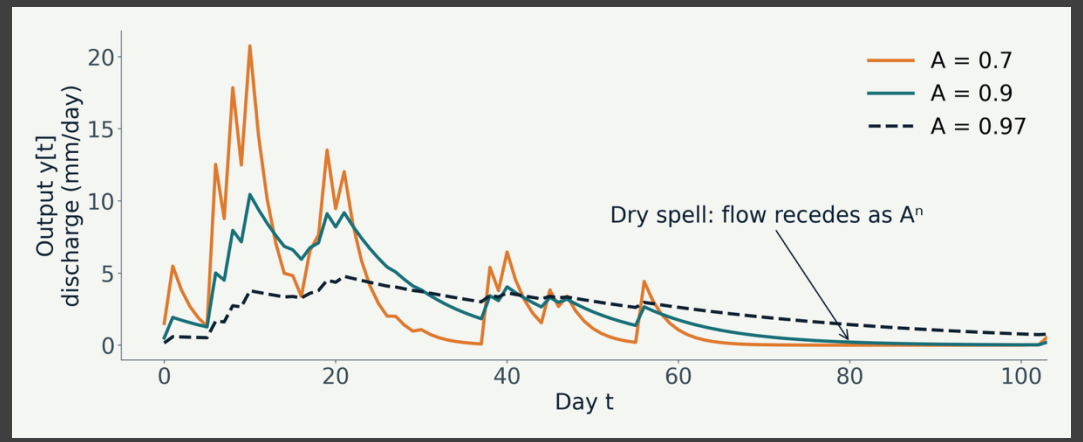
$$\mathbf{y}[t] = \mathbf{C}\cdot\mathbf{x}[t]$$



Linear SSM

$$\mathbf{x}[t] = \mathbf{A}\cdot\mathbf{x}[t-1] + \mathbf{B}\cdot\mathbf{u}[t]$$

$$\mathbf{y}[t] = \mathbf{C}\cdot\mathbf{x}[t]$$



Memory in the machine

An introduction to state space models

Linear reservoir drains 10% of its storage each day. What is its state factor?

QUIZ 1 OF 2

A 0.1

B 0.9

C 1.1

D It depends on the rainfall

Memory in the machine

An introduction to state space models

Linear reservoir drains 10% of its storage each day. What is its state factor?

QUIZ 1 OF 2

A 0.1

B 0.9

C 1.1

D It depends on the rainfall

> Answer: B. 1 − 0.1 = 0.9, so 90% of storage is kept each day. Memory decays as 0.9^n: after 30 days only about 4% of today's water is left.

Memory in the machine

An introduction to state space models

RECURRENT NEURAL NETWORKS – Learning the state update from data

$$\mathbf{x}[t] = \tanh(\mathbf{A}\cdot\mathbf{x}[t-1] + \mathbf{B}\cdot\mathbf{u}[t] + \mathbf{b})$$

$$\mathbf{y}[t] = \mathbf{C}\cdot\mathbf{x}[t]$$

What's new

SSM is wrapped in a non-linearity and A, B, and C become learned weight matrices.

The catch

Gradients pass through **A** and tanh at every step, so they can shrink or explode exponentially. Memory beyond a few dozen steps is hard to learn.

Memory in the machine

An introduction to state space models

LONG SHORT-TERM MEMORY (LSTM) – Protected memory

$\mathbf{g}[t] = \tanh(\mathbf{A}_g \cdot h[t-1] + \mathbf{B}_g \cdot \mathbf{u}[t] + \mathbf{b}_g)$ # candidate information

$\mathbf{c}[t] = \mathbf{f}[t] \odot \mathbf{c}[t-1] + \mathbf{i}[t] \odot \mathbf{g}[t]$ # how much to keep or forget in long term memory

$\mathbf{h}[t] = \mathbf{o}[t] \odot \tanh(\mathbf{c}[t])$ # update short term memory

Forget gate f

$\sigma(\mathbf{A}_f \cdot \mathbf{h} + \mathbf{B}_f \cdot \mathbf{u} + \mathbf{b}_f)$

How much of the cell state **c** to keep

Input gate i

$\sigma(\mathbf{A}_i \cdot \mathbf{h} + \mathbf{B}_i \cdot \mathbf{u} + \mathbf{b}_i)$

How much of the candidate **g** to write in

Output gate o

$\sigma(\mathbf{A}_o \cdot \mathbf{h} + \mathbf{B}_o \cdot \mathbf{u} + \mathbf{b}_o)$

How much of the cell state **c** to expose as the hidden state **h**

Memory in the machine

An introduction to state space models

Snow falls in November and melts in April. Which gate must stay close to 1 all winter for the LSTM to remember it?

QUIZ 2 OF 2

A Input gate i

B Forget gate f

C Output gate o

D None: h[t] stores it

Memory in the machine

An introduction to state space models

Snow falls in November and melts in April. Which gate must stay close to 1 all winter for the LSTM to remember it?

QUIZ 2 OF 2

A Input gate i

B Forget gate f

C Output gate o

D None: h[t] stores it

> Answer: B. With f ≈ 1, c[t] ≈ c[t−1] and the storage is kept. A low output gate helps too: the snow is stored but not yet released as runoff.

Memory in the machine

An introduction to state space models

MODERN STATE SPACE MODELS – Learned and selective linear updates

$$\mathbf{x}[t] = \mathbf{A}\cdot\mathbf{x}[t-1] + \mathbf{B}\cdot\mathbf{u}[t] \quad \# (\mathbf{A}, \mathbf{B} \text{ discretised from continuous time with step } \boldsymbol{\Delta})$$

$$\mathbf{y}[t] = \mathbf{C}\cdot\mathbf{x}[t]$$

S4

- **A** is structured and initialised for long memory (Gu et al., 2022)
- Linear in **x**, so whole sequence is computed at once as a convolution or parallel scan

Mamba

B, **C** and **Δ** depend on the input, so the model chooses what to store and forget (Gu & Dao, 2023)

Memory in the machine

An introduction to state space models

SIDE BY SIDE

	Linear SSM	RNN	LSTM	S4 / Mamba
State update	Linear	Non-linear	Gated	Linear (+ input-selective)
Long memory	Set by **A**	Poor	Good	Very good
Parallel training over time	Yes	No	No	Yes
Interpretability	High	Low	Medium	Medium
In hydrology	Linear reservoirs, Kalman filters	Rarely used now	Current workhorse	Emerging

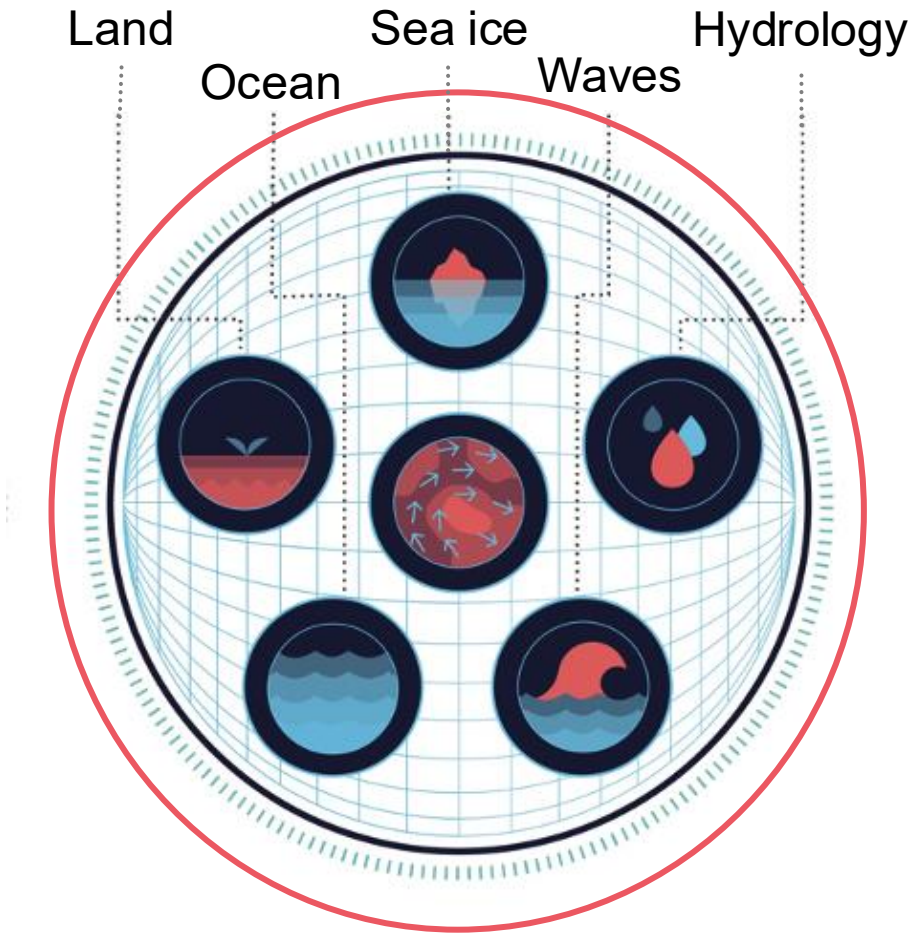
3 | ECMWF’s AI for floods (AIFL) model

Hydrology and the AIFS

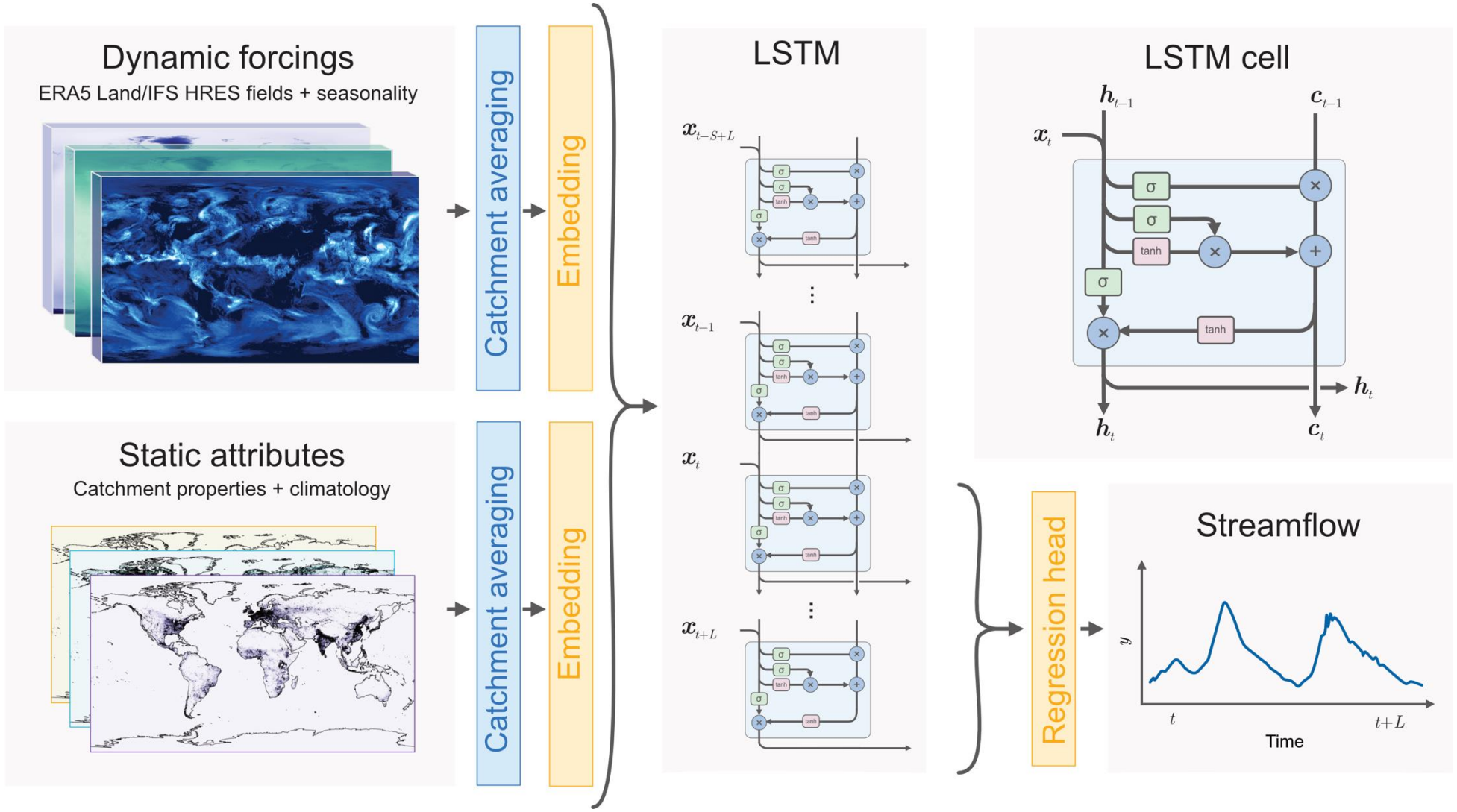
Hydrology is one of five ML Earth system components (land, ocean, waves, sea ice) being developed independently from the AIFS.

We are exploring trade-offs between

1. Standalone hydrological model
   → AIFL
2. Training a single model with both components together
   → Hydrology and the AIFS
3. Using one model to drive the other
   → Towards coupled models



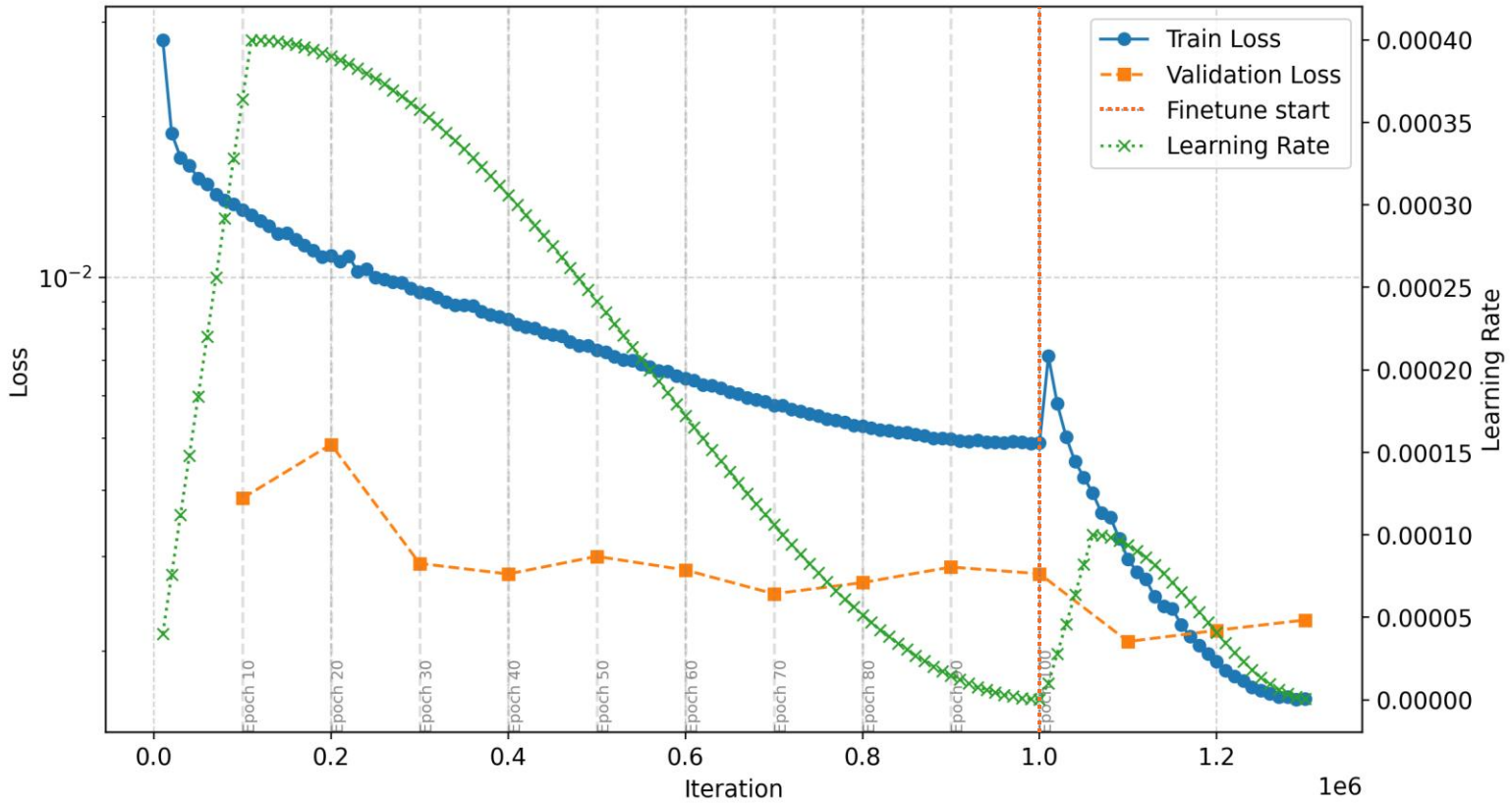
AIFL – AI for Floods



Training strategy

Fine-tune to overcome distributional shifts

- Pretrain on ERA5 Land between 1980 and 2022
- Finetune on IFS HRES between 2016 and 2022



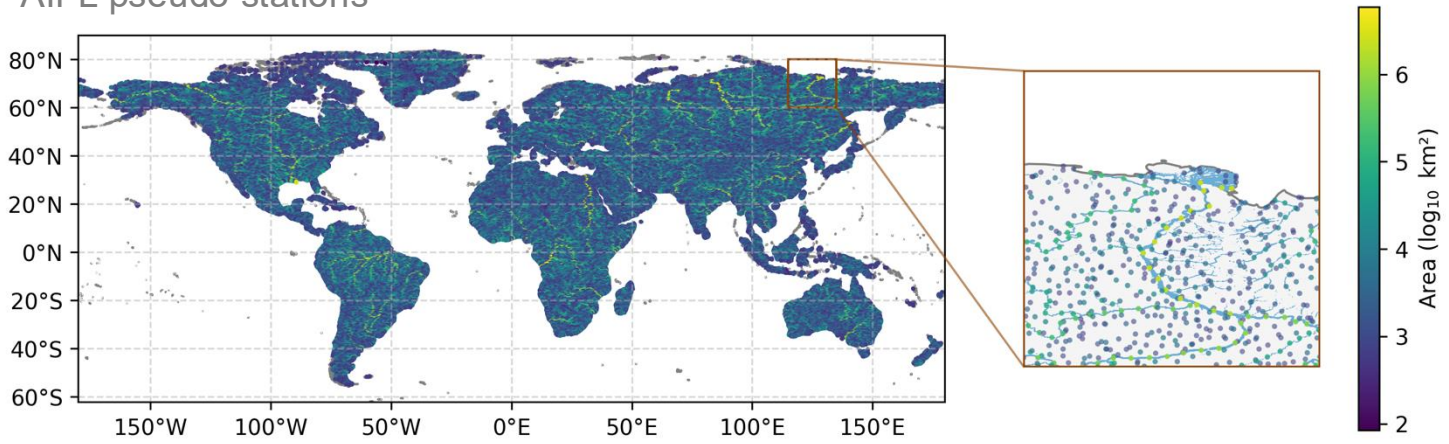
Maria Luisa Taccari

AIFL global 2.0

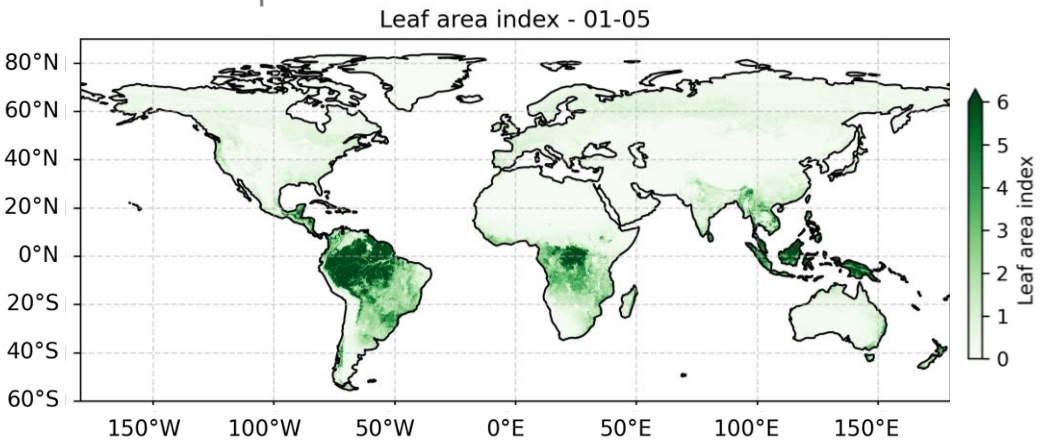
Key updates

- Extend forecast from 10 to 15 days
- Extend forecast spatial coverage
  - Pseudo stations for upstream areas > 500km^2 every 100km
  - Model inference at observation and pseudo stations (73 000 locations)
- Expand and standardise streamflow data
  - Mapped station observations on Cama-Flood networks
  - UTC-aligned observations
- Replace Hydro Atlas attributes with 1 arc-minute IFS climate fields
- Strict independence requirements during model development

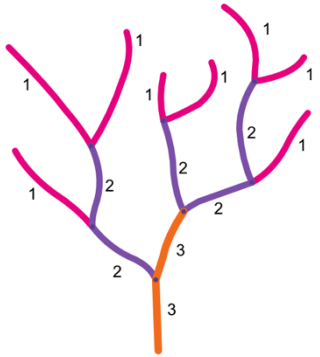
AIFL pseudo stations



IFS climate maps

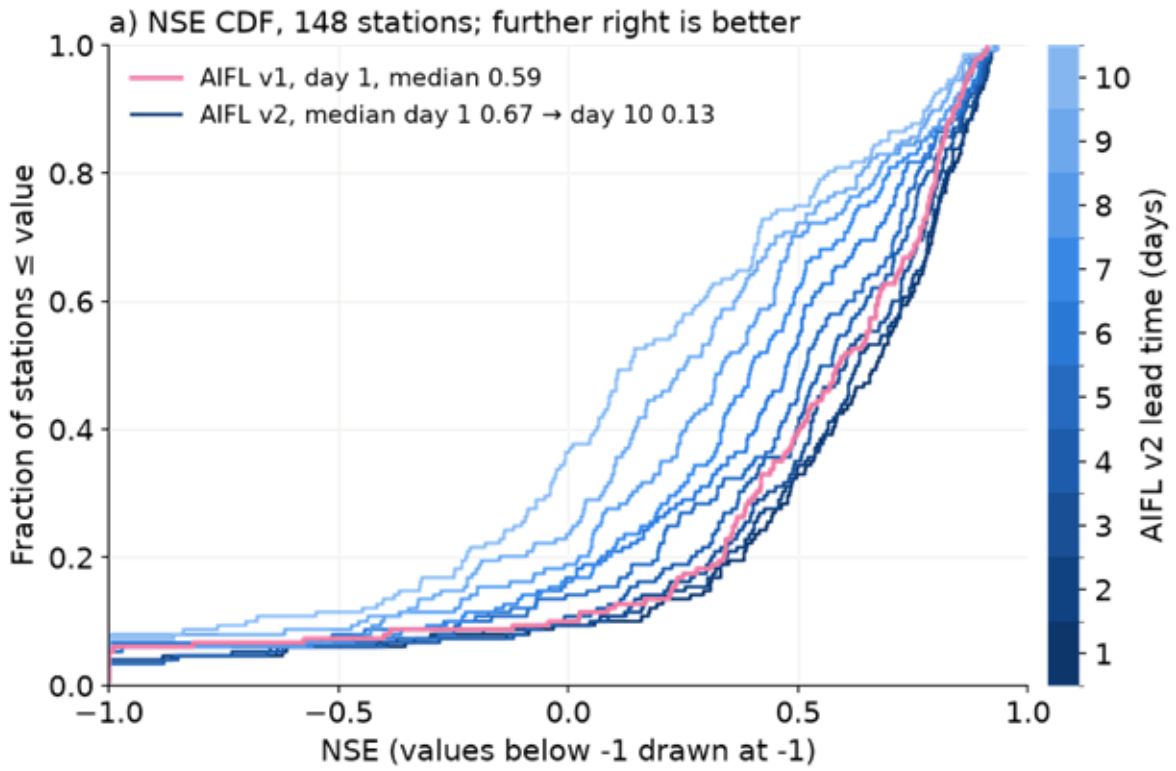


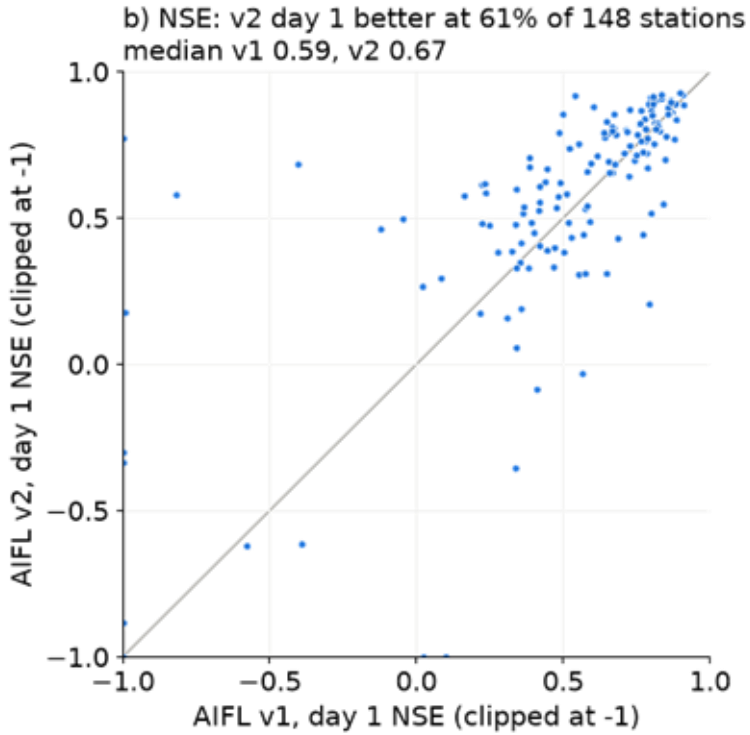
Topological buffer



AIFL v2 global performance

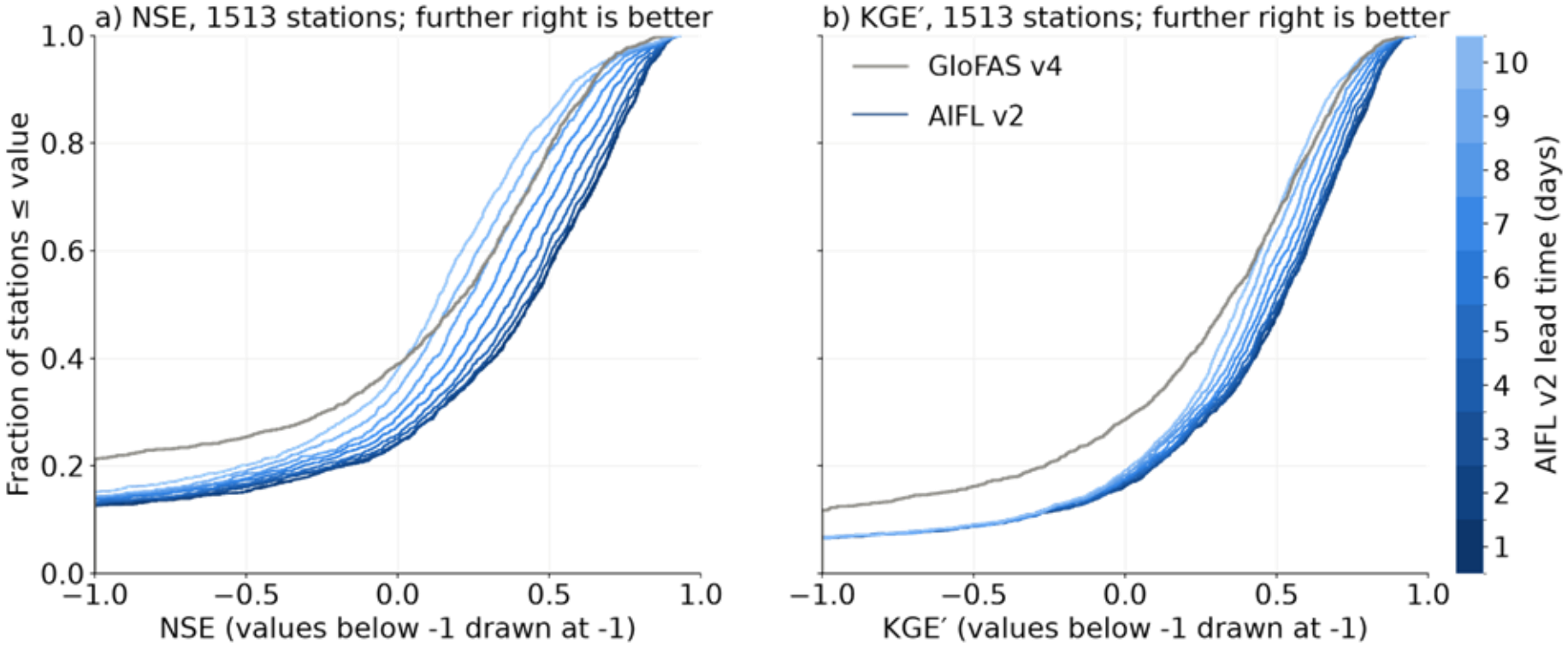
Improvement over AIFL v1 at new times





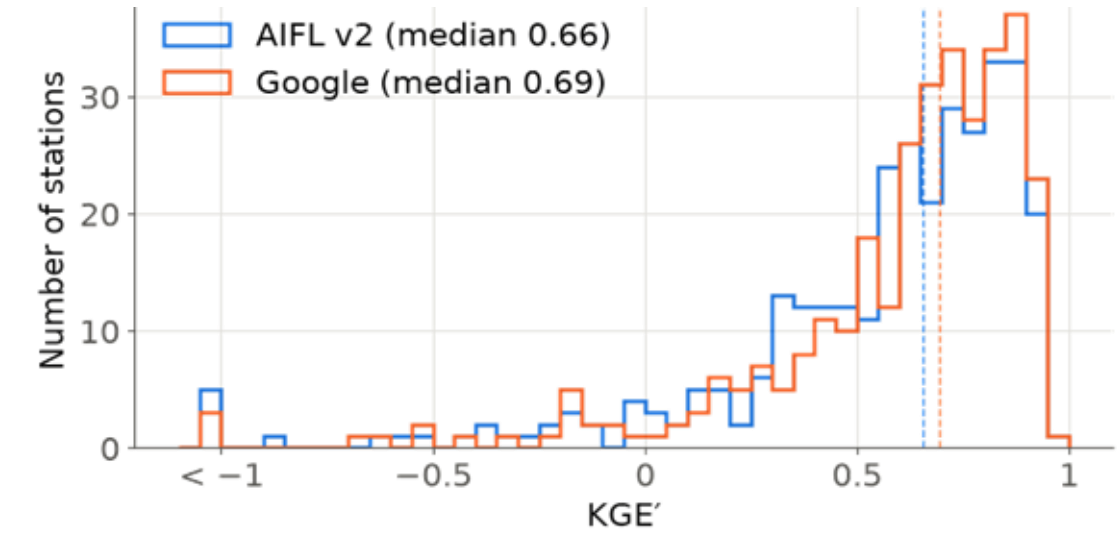
AIFL v2 global performance

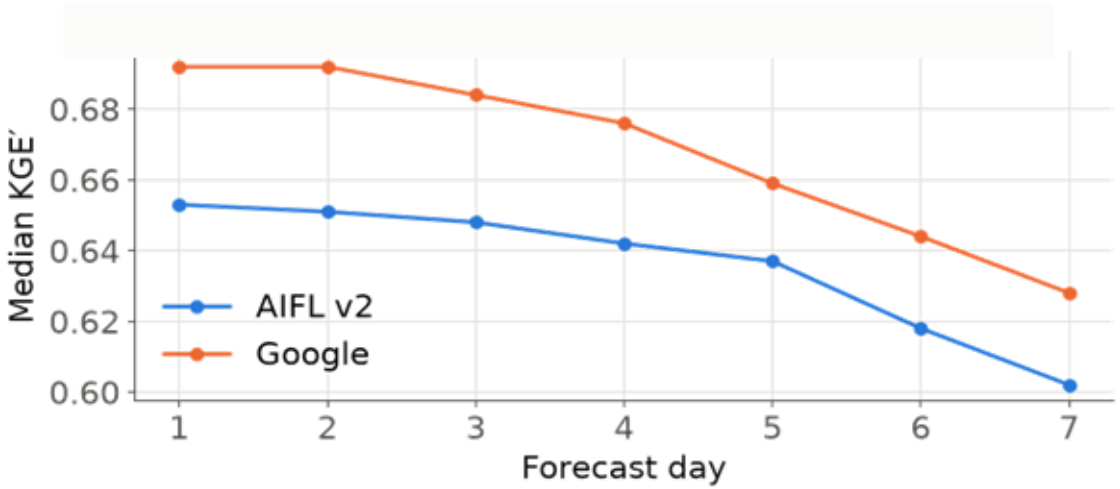
Outperforms GloFAS at unseen times and locations



AIFL v2 global performance

Competitive with Google Flood Hub (Nearing et al. 2024) at unseen times and locations





Looking ahead for floods

TOOLING

earthkit-hydro

v1.0

Fast river network statistics with GPU acceleration



Open AIFL code and weights

soon

Train your own LSTM model or adapt AIFL to your basins





RESEARCH

AIFL sub-daily

Six hourly forecasts outperforming EFAS

AIFL local

Data assimilation from real-time observations

AIFL probabilistic

Flow trajectories conditioned on uncertain atmosphere and catchment conditions

Coupling with AIFS

Drive AIFL directly with AIFS ENS forecasts

AIFL local: a first prototype for Europe

AIFL global

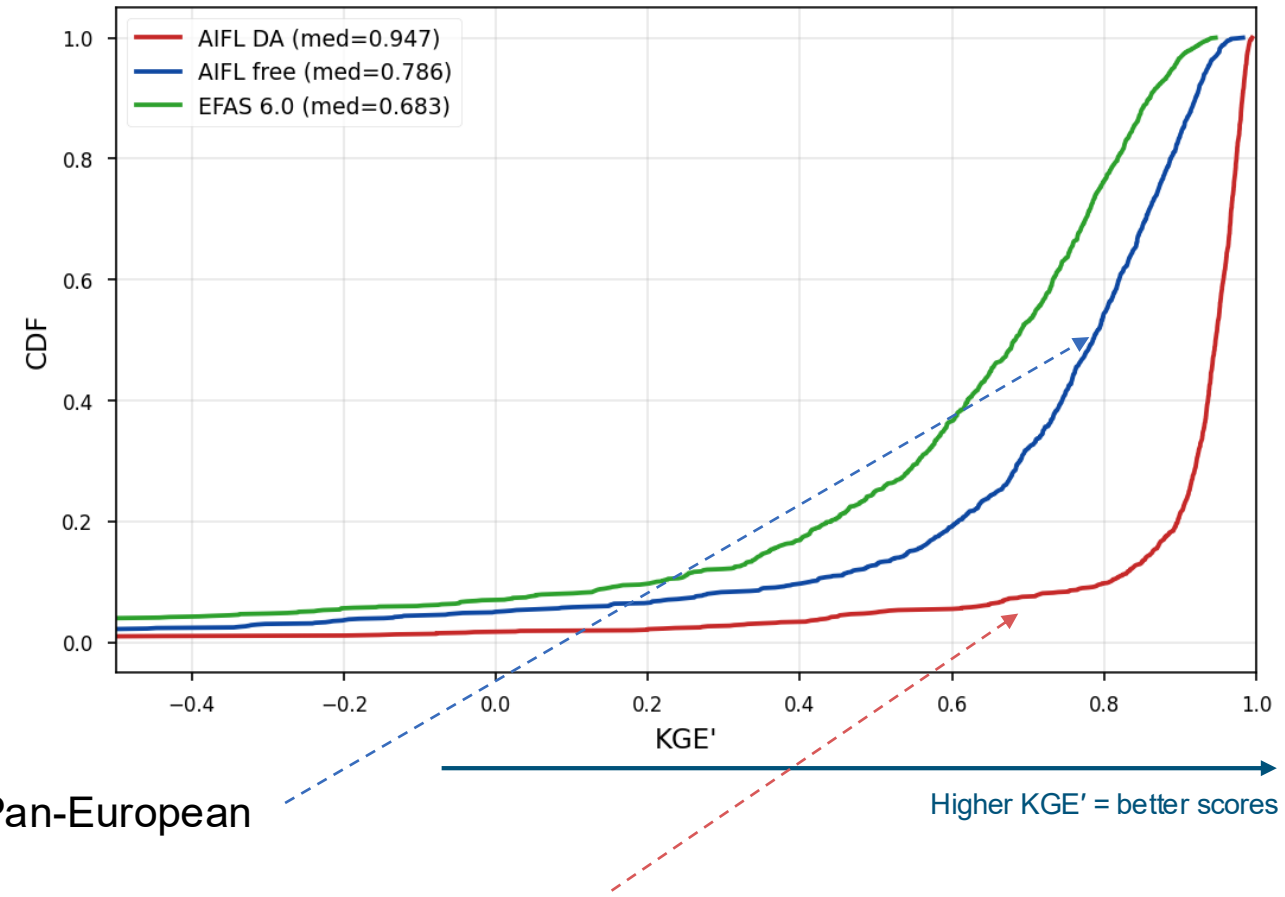
pre-trained daily model · 185-day context

regional adaptation
gauges + local data

AIFL local

6-hourly regional forecasts to T+5 days

- Sub-daily time step
- Assimilates real-time discharge observations
- Integrates local and regional datasets

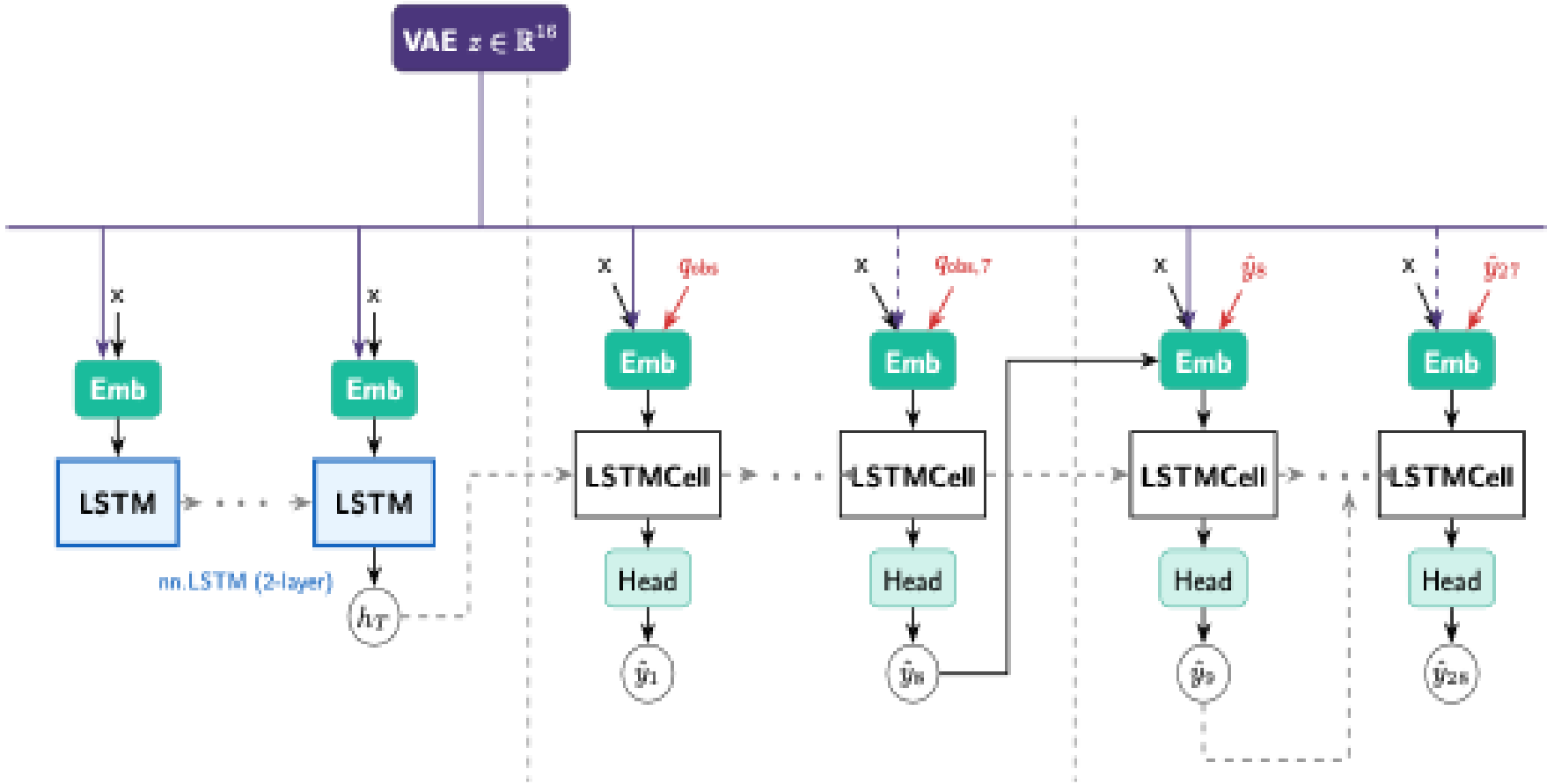


Sub-Daily Model: AIFL adapted to 6-hourly Pan-European catchments outperforms EFAS v6.0

Real-time data integration: Integrating even a small fraction of JRC's real-time discharge data (HYDRO) yields an accuracy boost over EFAS v6.

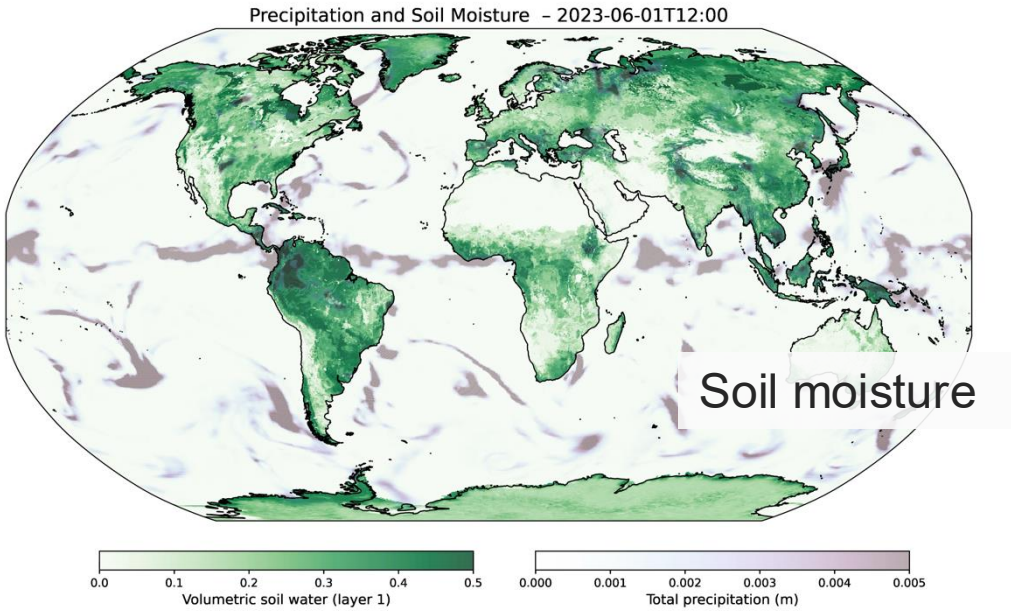
Maria Luisa Taccari

AIFL local: a first prototype for Europe

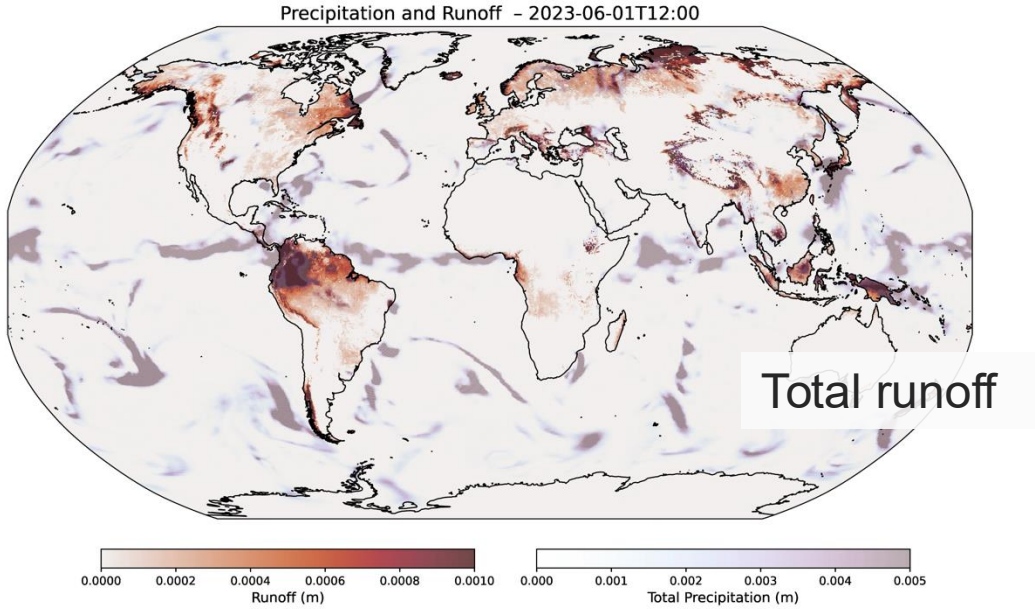


4 | Hydrology and the AIFS

Land variables in the AIFS



Soil moisture



Total runoff

AIFSv1
- soil moisture (lvls 1-2)
- soil temperature (lvls 1-2)
- runoff

AIFSv2
- snow cover
- snow depth in water equiv.

AIFSv3
- urban & lake cover
- **runoff partitioning**

AIFS experiment setup

Runoff partition

runoff (ro) = surface runoff (sro)+ subsurface runoff (ssro)

with moderate loss function weighting (much higher than surface winds, much lower than pressure levels)

Key training figures

N320 grid

ERA5 1979-2022

1.3 days

Hypotheses

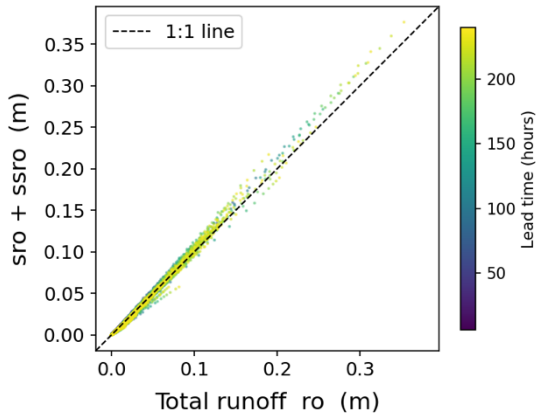
- better runoff
- better soil moisture evolution
- better surface energy partitioning
- better 2m and skin temperatures

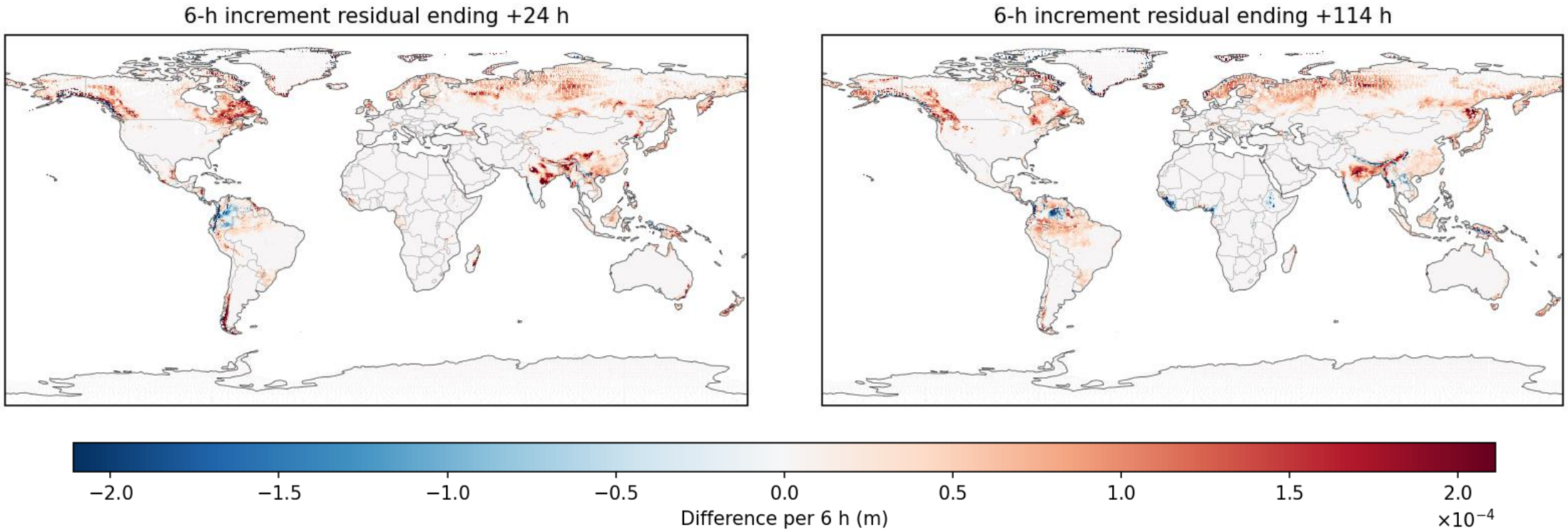
Same as with total precipitation (tp) and total cloud cover (tcc), fraction bounding ensures the sum of sro and ssro equally ro

```yaml
- _target_: anemoi.models.layers.bounding.FractionBounding # fraction of tp
  variables:
  - cp
  - sf
  min_val: 0
  max_val: 1
  total_var: tp
- _target_: anemoi.models.layers.bounding.FractionBounding # fraction of ro
  variables:
  - sro
  - ssro
  min_val: 0
  max_val: 1
  total_var: ro
- _target_: anemoi.models.layers.bounding.FractionBounding # fraction of tcc
  variables:
  - lcc
  - mcc
  - hcc
```

Is the runoff partition consistent?

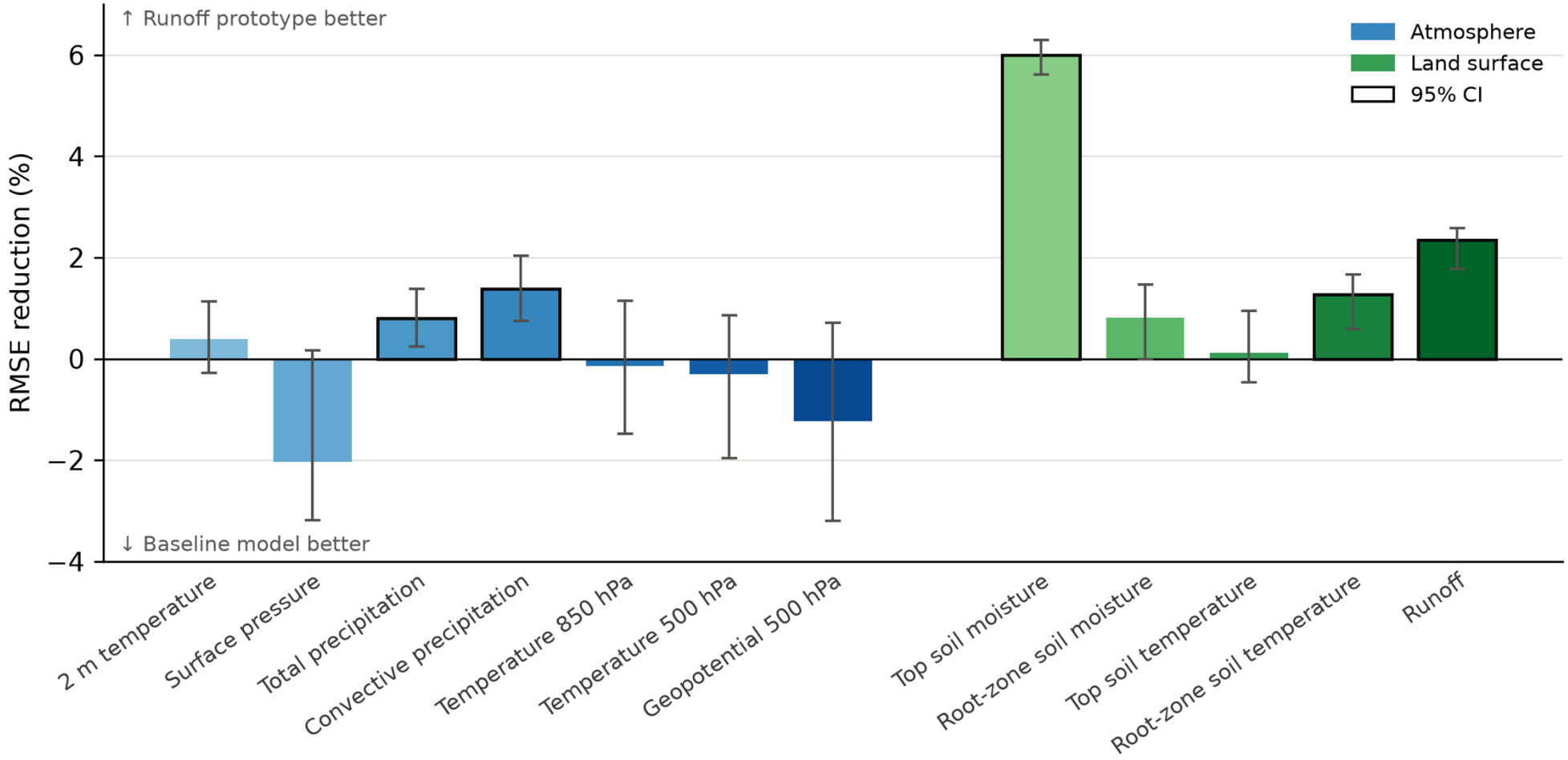
For the 1st July 2025 forecast, the runoff partition is consistent, with residuals concentrated in areas driven by strong convective precipitation or snowmelt





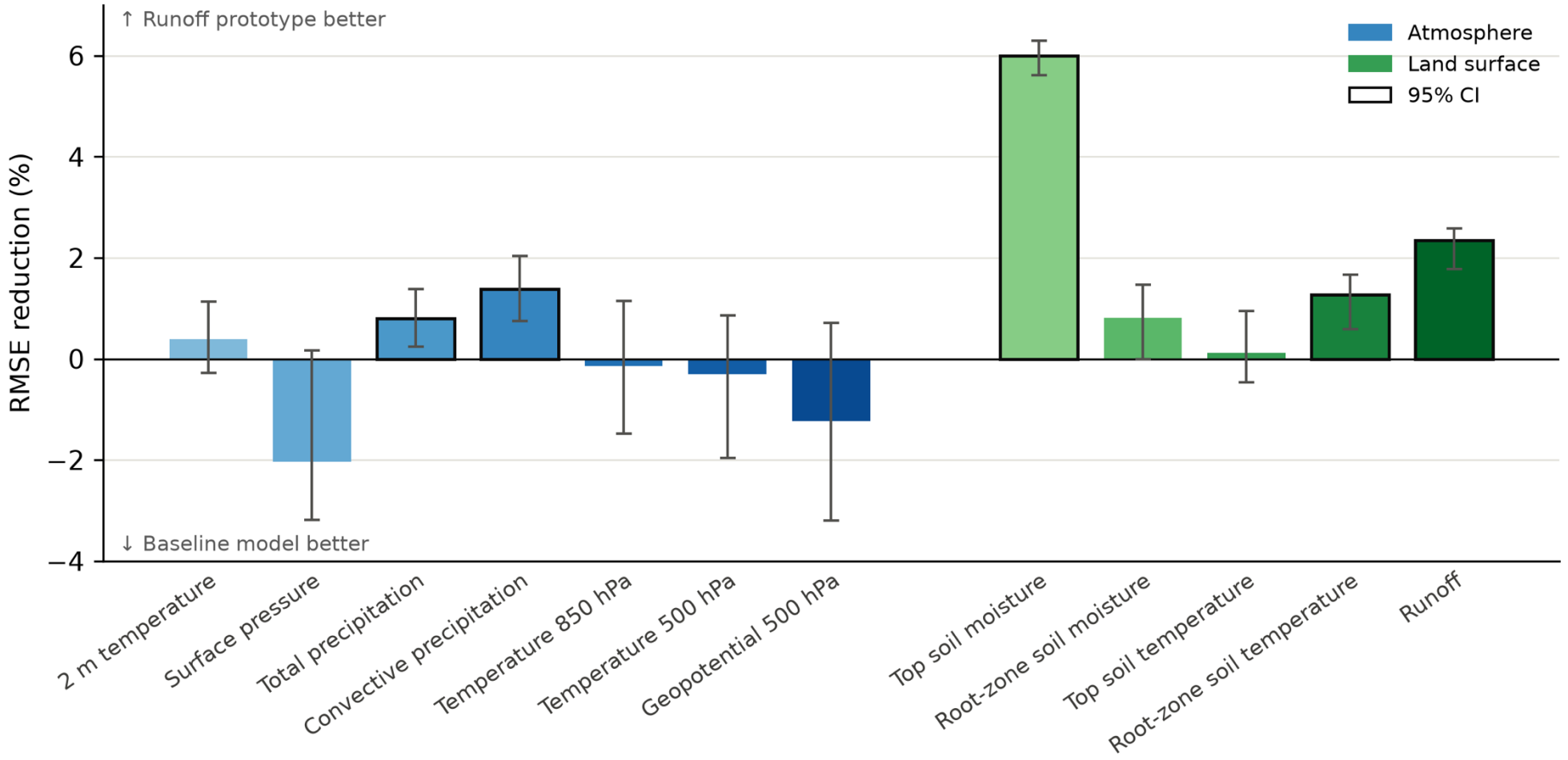
Does the runoff partition affect the atmosphere?

With respect to IFS analysis, neutral to positive impact on atmosphere vs AIFS 2 for JJA 2025



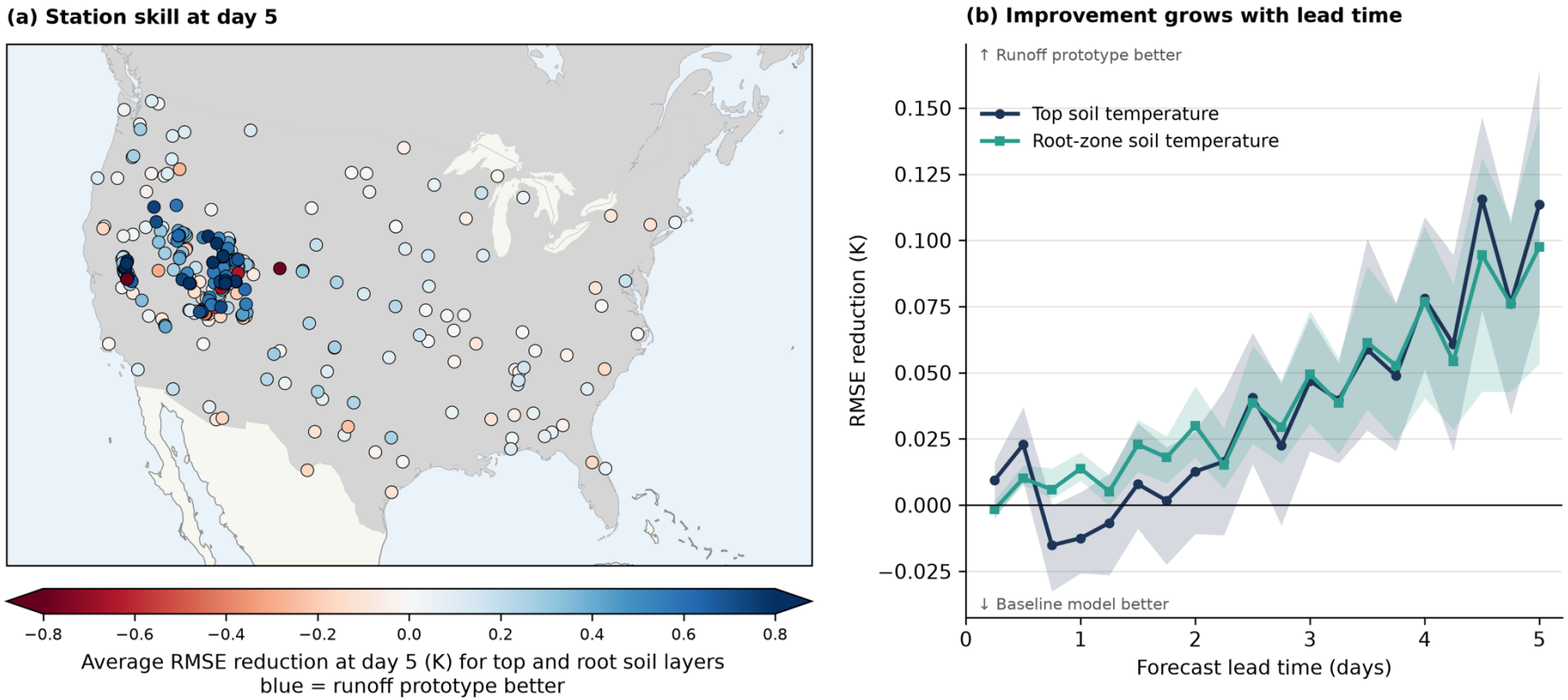
What surface processes does the AIFS learn?

Globally, improved runoff, soil moisture, and soil temperature



What surface processes does the AIFS learn?

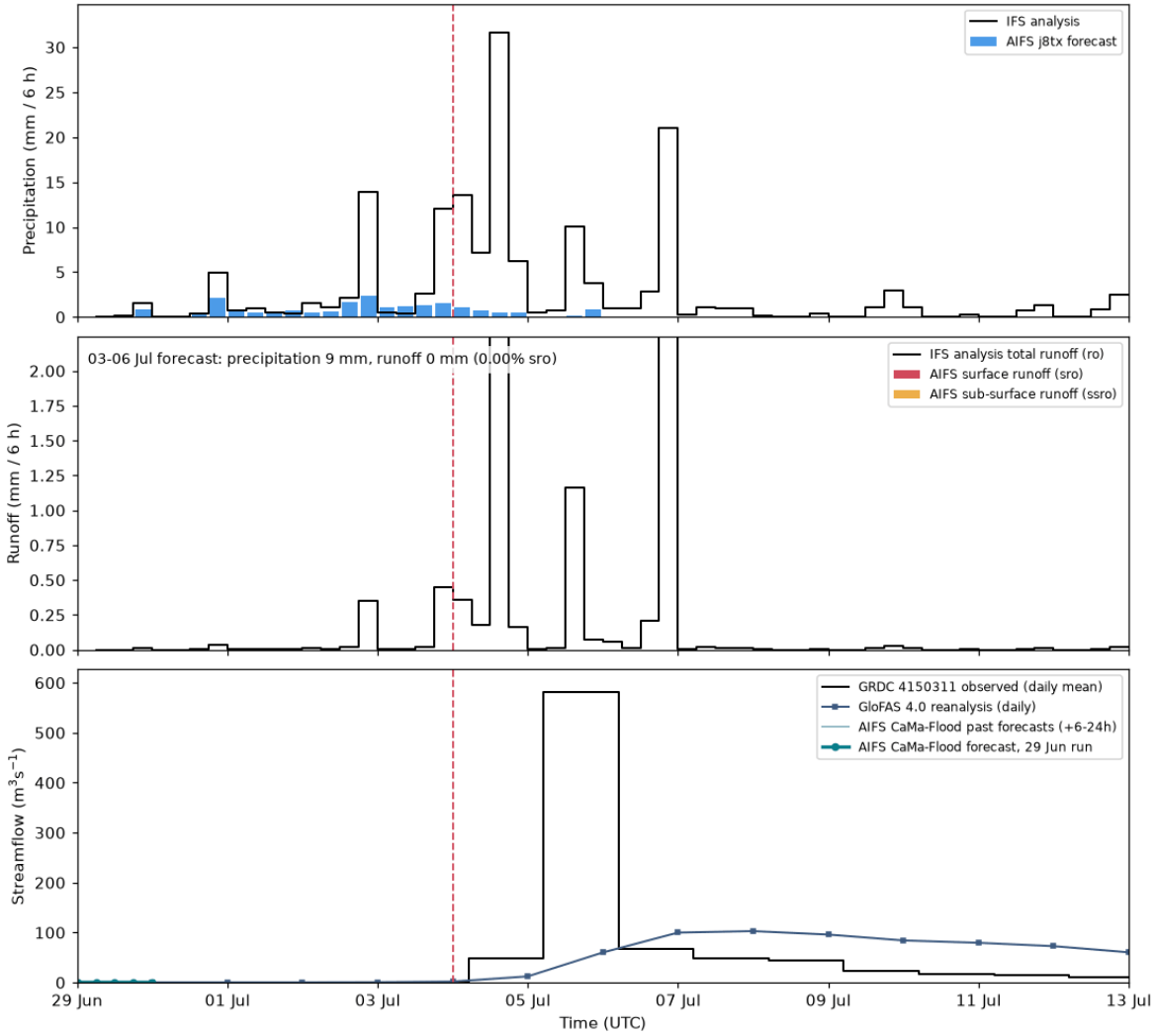
Surface temperature is also improved against observations



New pathways for evaluation and verification

Texas flash floods, July 2025

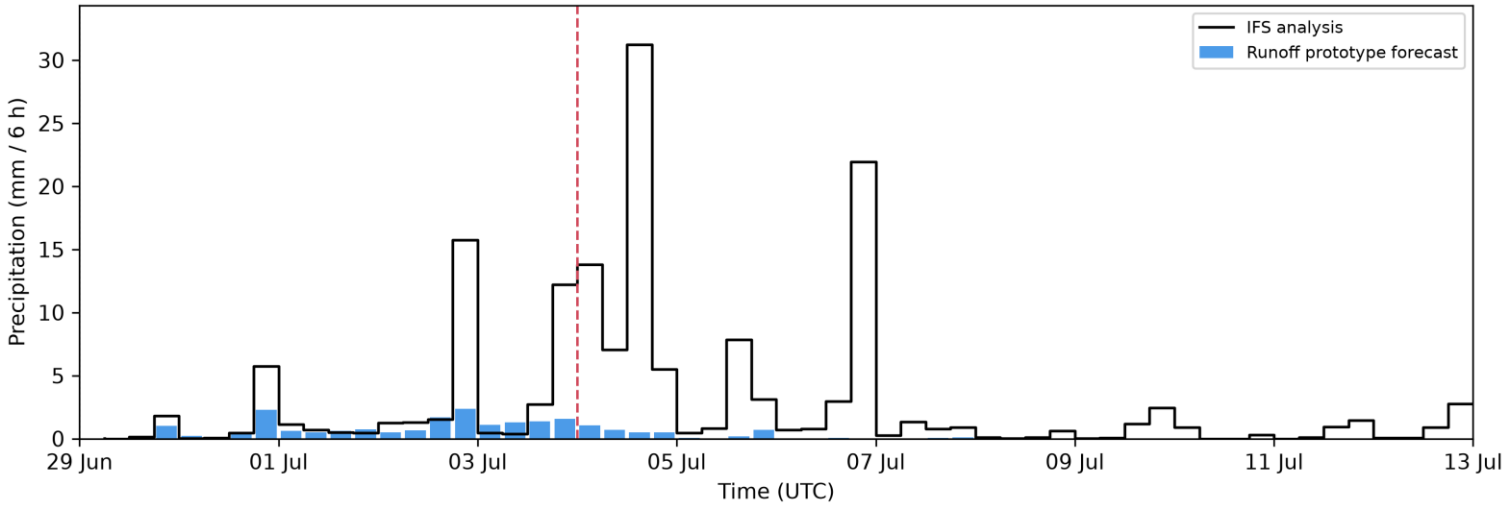
- Record-breaking rainfall and flooding on the Guadalupe River, 137 dead
- Run CaMa-Flood with new sro and ssro outputs at 6-arcmin resolution
- Correctly routes the runoff to surface
- Peak streamflow timing improved over GloFAS reanalysis v4

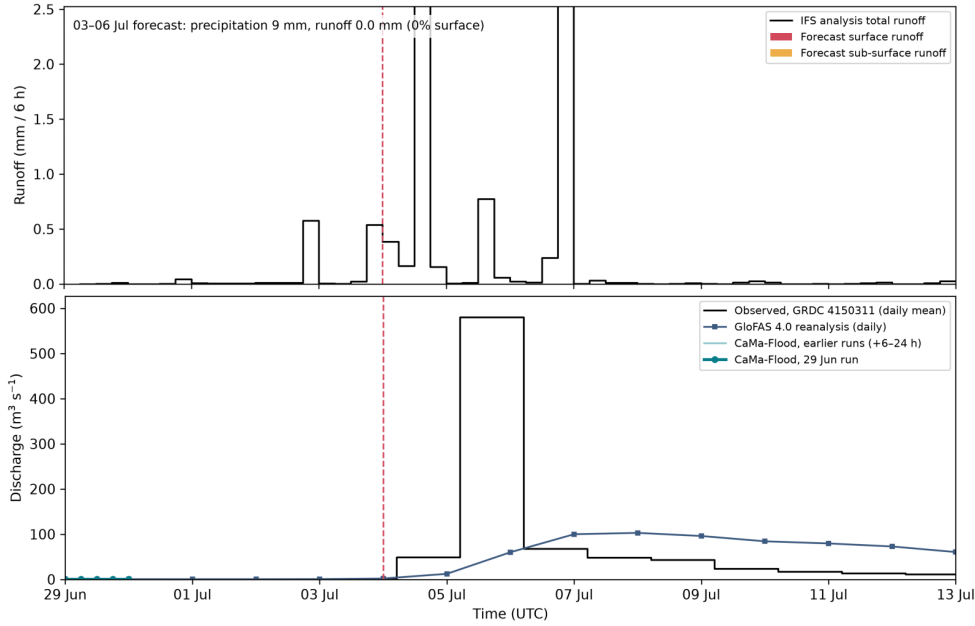


AIFS with runoff partition

Texas flash floods, July 2025

- Record-breaking rainfall and flooding on the Guadalupe River, 137 dead
- Run CaMa-Flood with new sro and ssro outputs at 6-arcmin resolution
- Correctly routes the runoff to surface
- Peak streamflow timing improved over GloFAS reanalysis v4

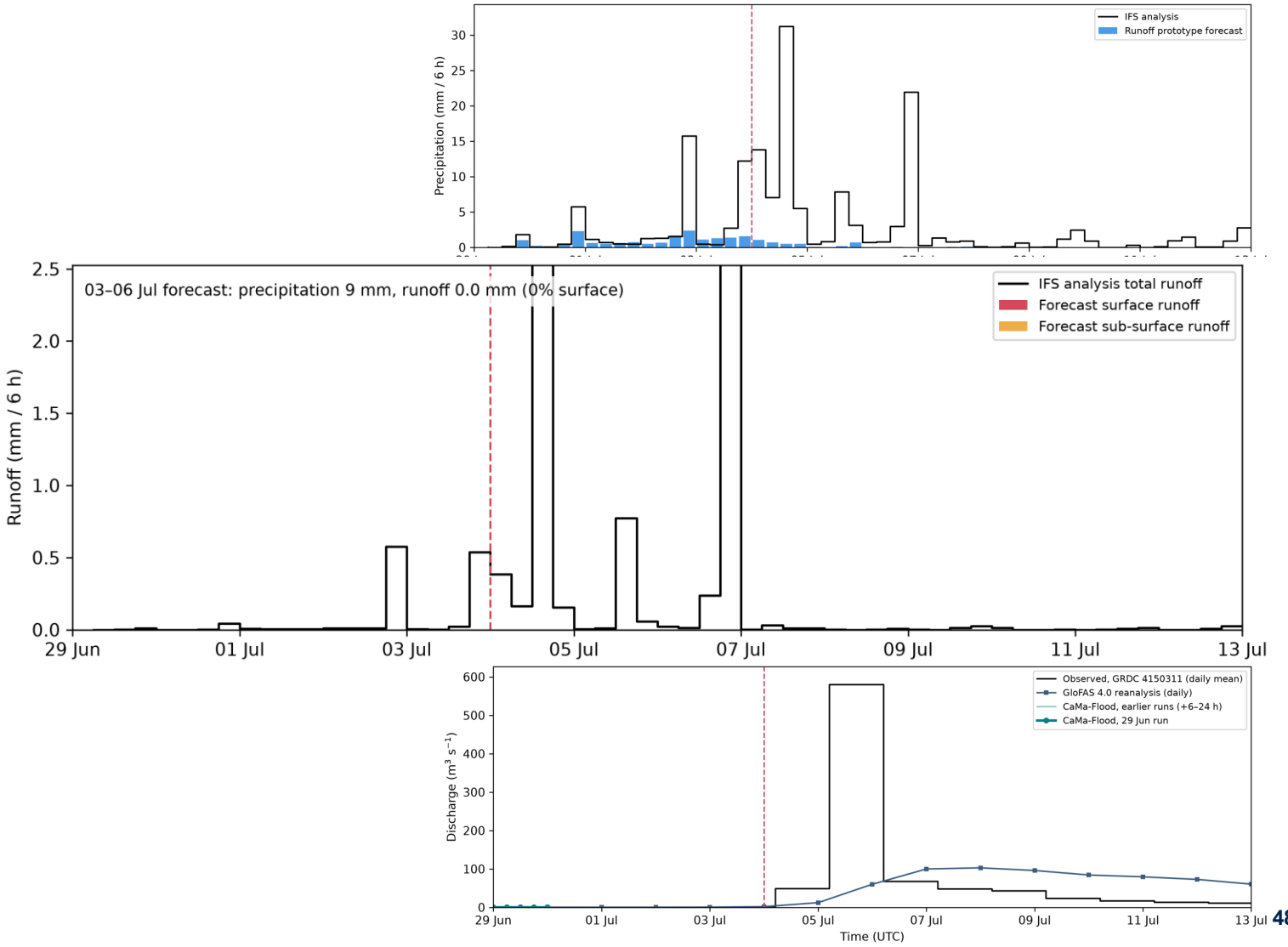




AIFS with runoff partition

Texas flash floods, July 2025

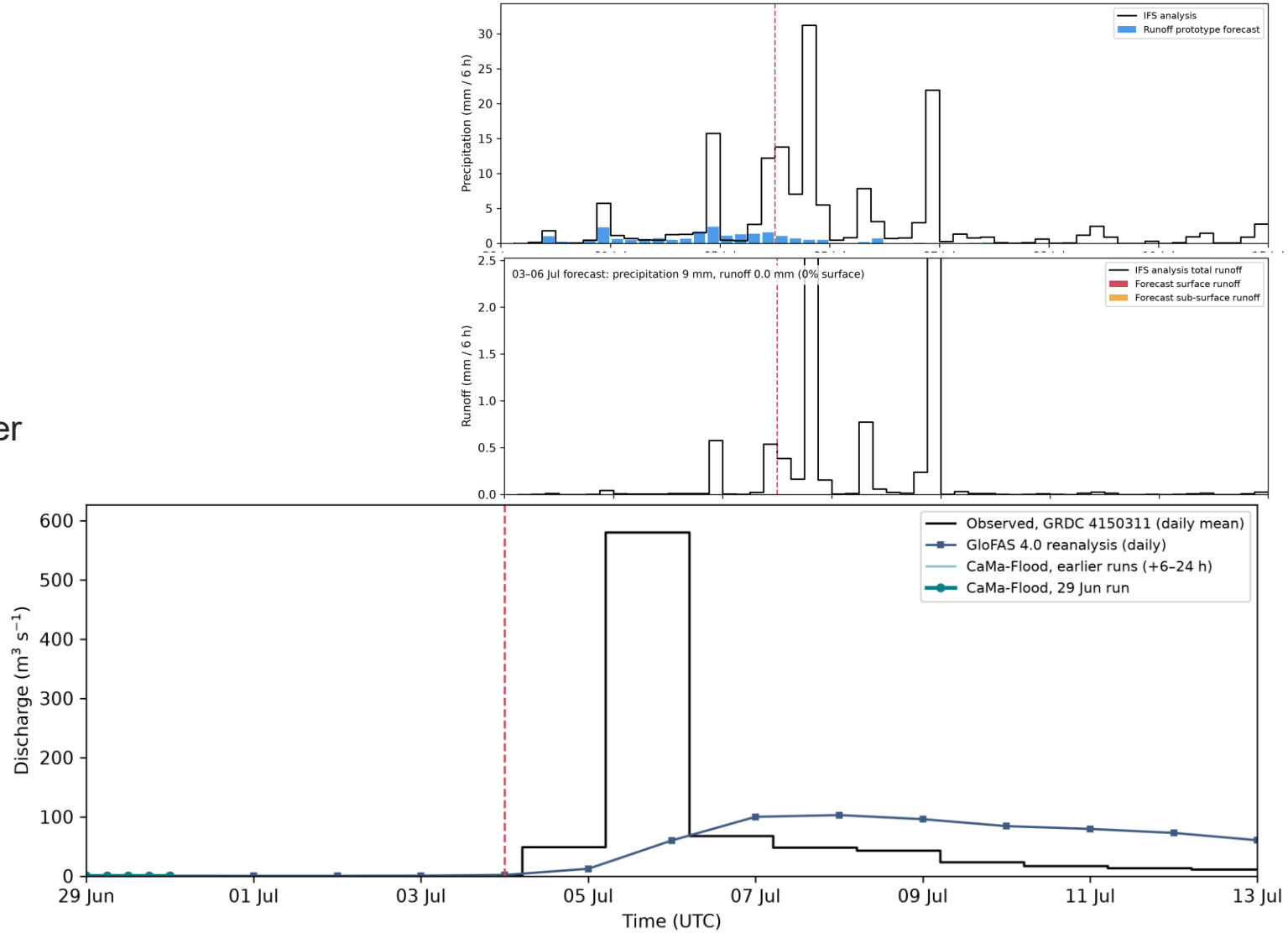
- Record-breaking rainfall and flooding on the Guadalupe River, 137 dead
- Run CaMa-Flood with new sro and ssro outputs at 6-arcmin resolution
- Correctly routes the runoff to surface
- Peak streamflow timing improved over GloFAS reanalysis v4



AIFS with runoff partition

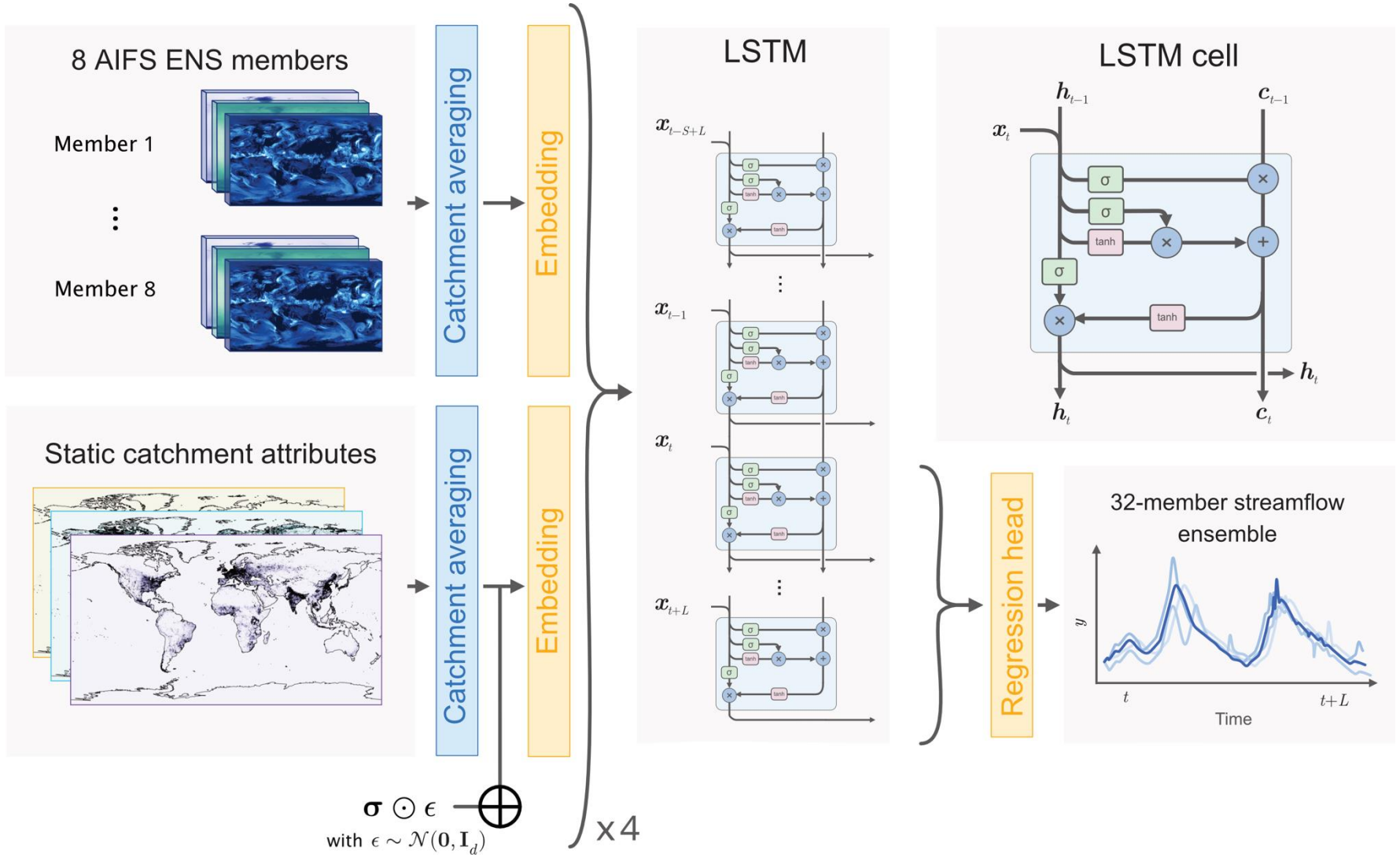
Texas flash floods, July 2025

- Record-breaking rainfall and flooding on the Guadalupe River, 137 dead
- Run CaMa-Flood with new sro and ssro outputs at 6-arcmin resolution
- Correctly routes the runoff to surface
- Peak streamflow timing improved over GloFAS reanalysis v4



5 | Towards coupled models

Probabilistic coupling to AIFS

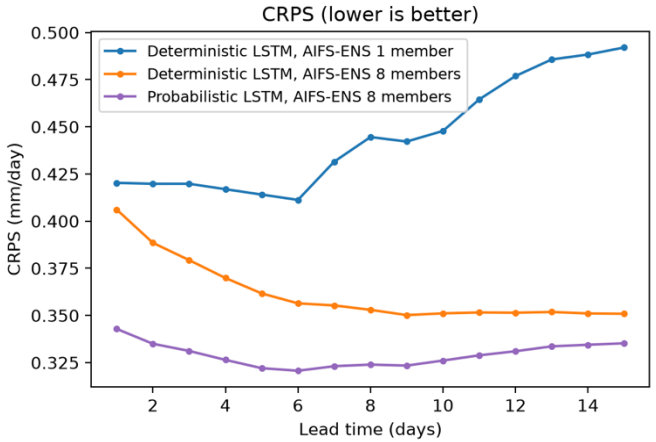


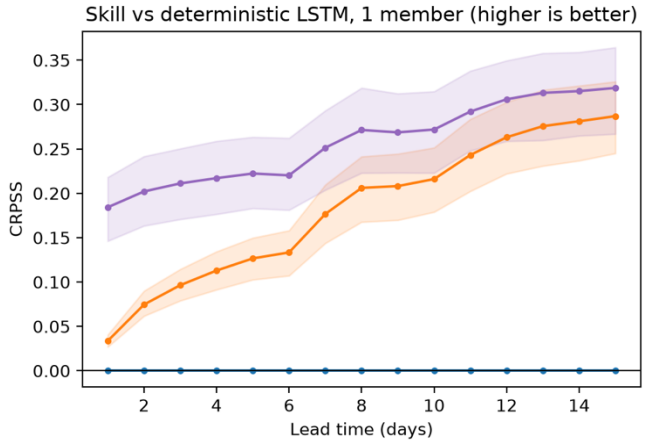
Probabilistic coupling to AIFS

Work in progress: AIFS ENS to force probabilistic AIFL

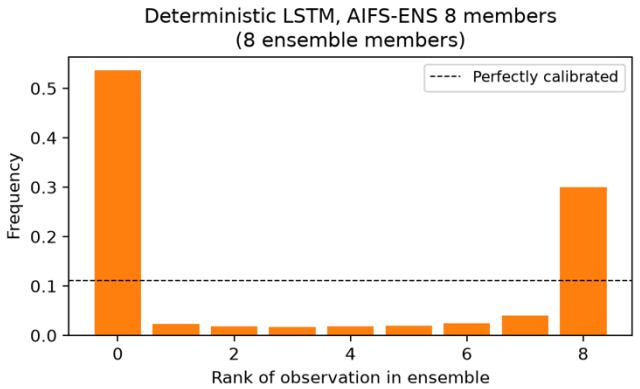
- Training
  1. ERA5 Land forcing with deterministic loss (NSE)
  2. AIFS ENS forcing with deterministic loss (NSE)
  3. AIFS ENS forcing with probabilistic loss (CRPS loss + noise injected to static attributes)
- Inference
  - 8 AIFS ensemble members
  - 4 AIFL members
  - 32 total members
- Initial results
  - AIFS ENS distribution complements AIFL uncertainty over the static catchment attributes
  - Still very under-dispersed, more work needed to create calibrated uncertainty bounds

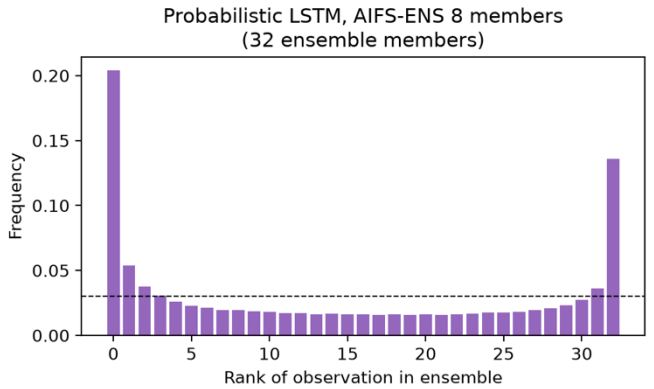
CRPS with lead time





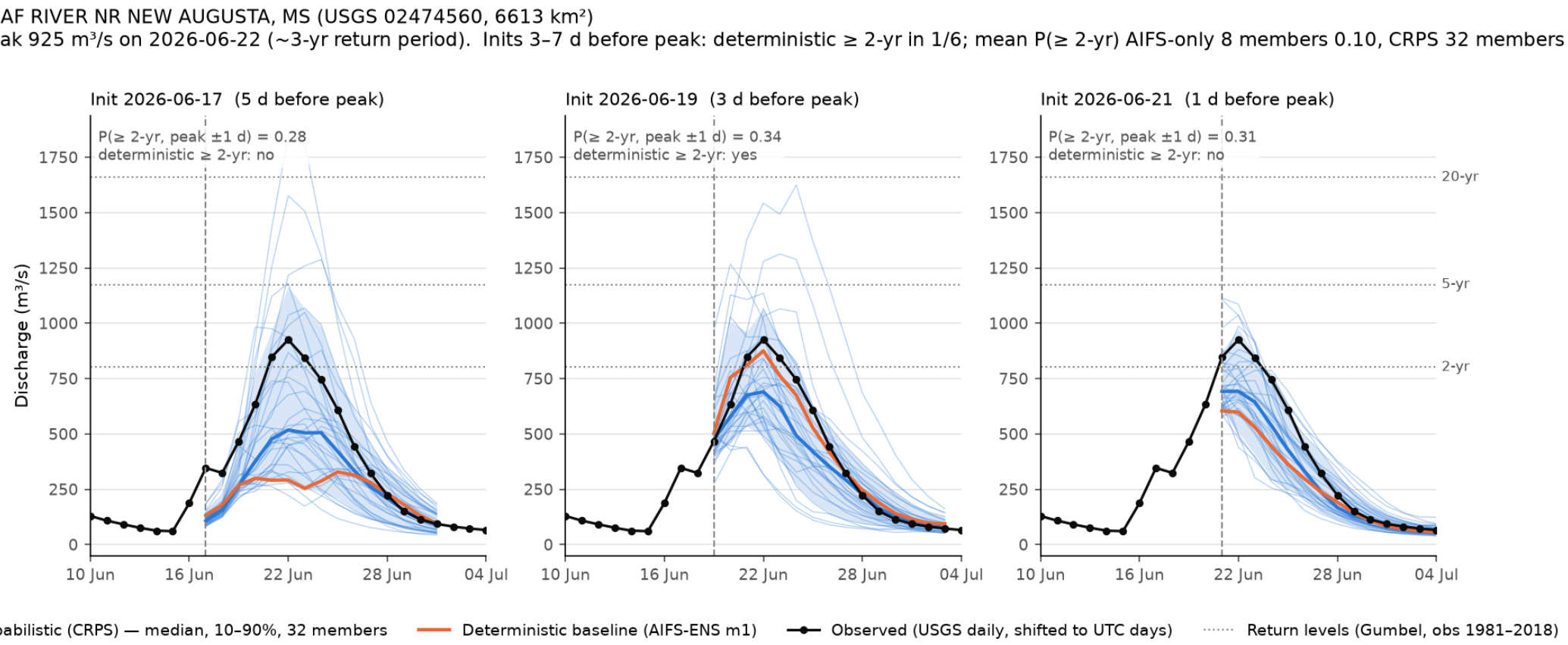
Rank histogram at day 5





Probabilistic coupling to AIFS

Leaf River case study



Key takeaways

Key takeaways



Flood forecasting saves lives, but river data are sparse, noisy and changing



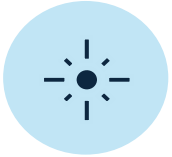
State space models as methods for encoding memory



AIFL, ECMWF’s LSTM model, rivals leading global flood forecasts



Hydrology as a tool for better accuracy and diagnosis



Probabilistic models are the next step for flood forecasting