AI Weather Quest

ECMWF

Enhancing sub-seasonal predictions with AI/ML: A competition by ECMWF, endorsed by WMO

March-May Awards Webinar

ECMWF

Destination Earth

WORLD METEOROLOGICAL ORGANIZATION

Agenda

AI Weather Quest ECMWF

Presentations

- ✓ Introduction by Gabriela Aznar, MeteoSwiss
- ✓ MAM participation overview
- ✓ MAM evaluation overview
- ✓ East Africa rainfall case study
- ✓ Presentations from teams CMAandFDU, SBUHybrid and IgnisNeuralis42
- ✓ Key milestones and an announcement

This session is being recorded.
The recording will be made available online after the webinar. If you do not wish to appear, please turn off your camera.

Please mute your microphone.
Please keep yourselves muted during presentations. You are welcome to ask questions during the Q&As.

14/08/25 13/11/25 12/02/26 14/05/26 13/08/26

Submit global weekly forecasts of near-surface temperature, precipitation, and mean sea level pressure for week 3 and week 4 ahead

SON period	DJF period	MAM period	JJA period

ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

Introduction by Gabriela Aznar

Co-Lead Machine Learning Integration Team,
Federal Office of Meteorology and Climatology MeteoSwiss

MAM Period participation overview

AI Weather Quest ECMWF

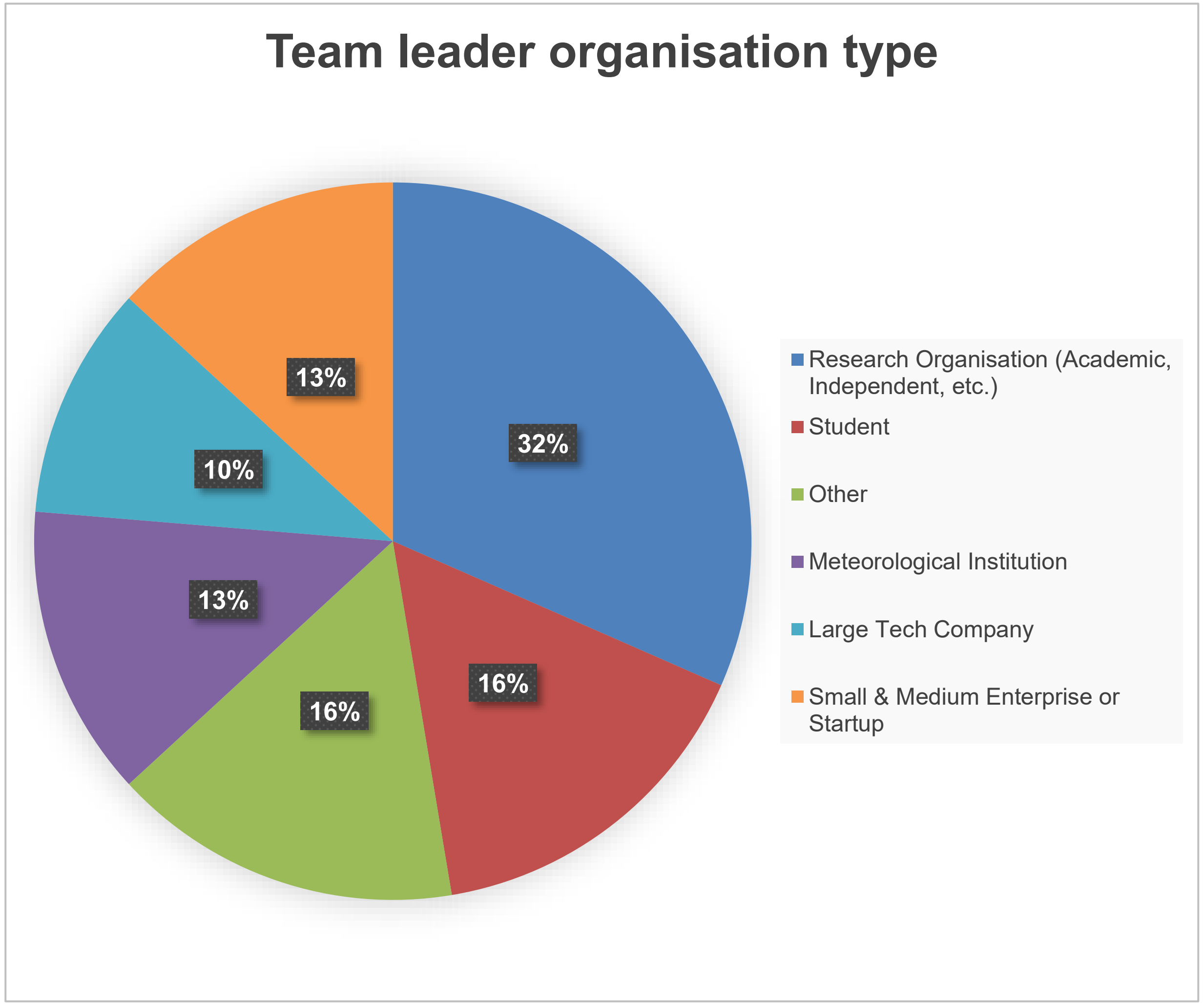
	Number of competitors to the Quest	Number of teams	Number of models
Accumulated participation since the start of the competition	194	42	74
Accumulated over MAM	170	37	64
Accumulated over DJF	177	37	66
Accumulated over SON	166	35	60

Affiliations of teams that accepted identities public display

ECMWF · University of Cambridge · NSF NCAR · Nanjing University · IGAD ICPAC · neuralio artificial intelligence · Fudan University · Imperial College London · University of Washington · GIST · the climate data factory · Harvard · 南京大学 Nanjing University · Microsoft · impa Instituto de Matemática Pura e Aplicada · ETH zürich · cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici · ACMAD · MIT · NVIDIA · University of Toronto · 中国科学院青藏高原研究所 Institute of Tibetan Plateau Research Chinese Academy of Sciences · Caltech · EPFL · rhiza research · SDSC · 米塔碳 Meta Carbon · KIT Karlsruher Institut für Technologie · WKU · Universidad de Alcalá · celest.science Anticipating climate risk & Empowering adaptation · Stony Brook University · China Meteorological Administration · WindBorne Systems · Ghent University · Tsinghua University · INPE

ECMWF · Destination Earth · World Meteorological Organization

MAM Period participation overview



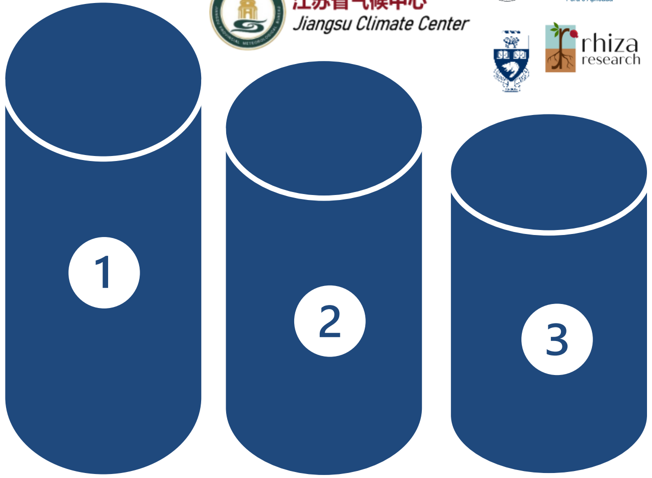
WORLD METEOROLOGICAL ORGANIZATION

AI Weather Quest ECMWF

MAM Period evaluation: Top-performers overview for **week 3** ahead

18 Teams eligible for variable-averaged, period-aggregated scores

11 Teams above climatology



江苏省气候中心 Jiangsu Climate Center

4) the climate data factory, cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici, Universidad de Alcalá

5) ECMWF

6) NVIDIA

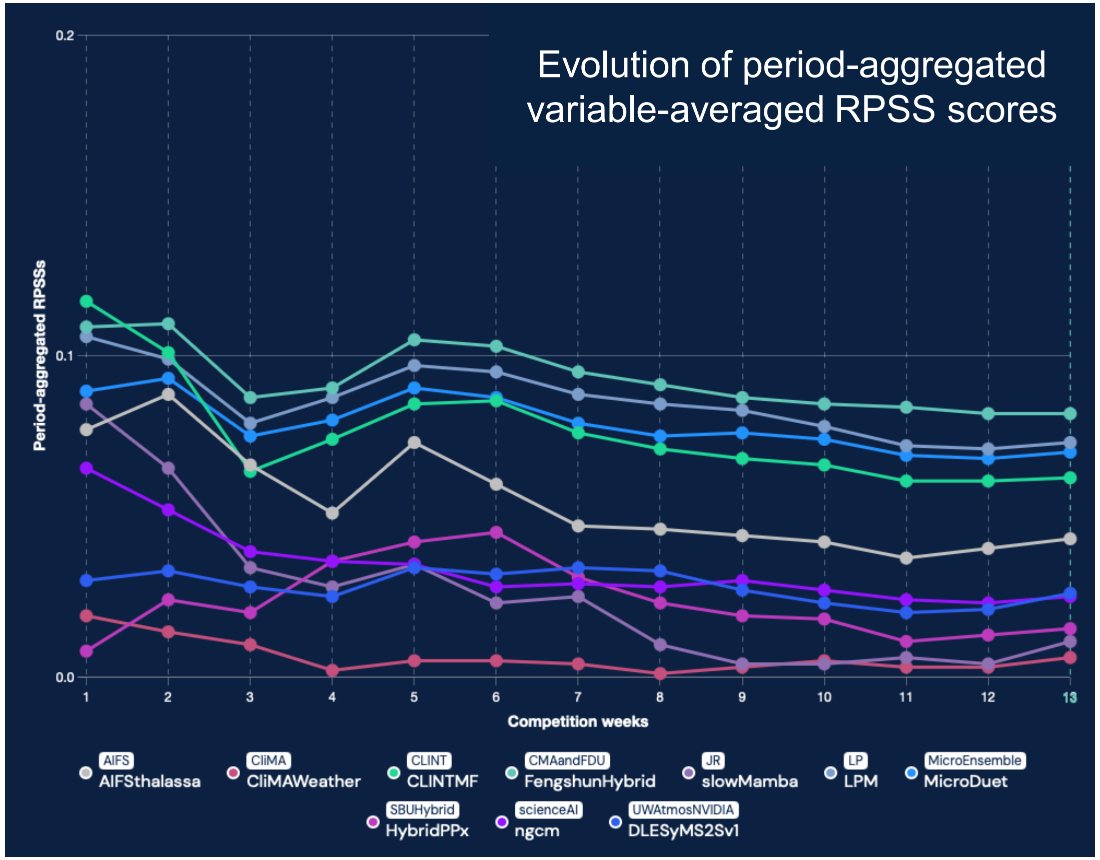
7) Anonymous

8) Stony Brook University

9) Individual

10) CliMA

11) cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici



Best performing team

Five-minute explainer on model update + Q&A

MAM Period evaluation: Top-performers overview for **week 4** ahead

AI Weather Quest ECMWF

18 Teams eligible for variable-averaged, period-aggregated scores

10 Teams above climatology

1

2

3

江苏省气候中心
Jiangsu Climate Center

4) the climate data factory · cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici · Universidad de Alcalá

5) ECMWF

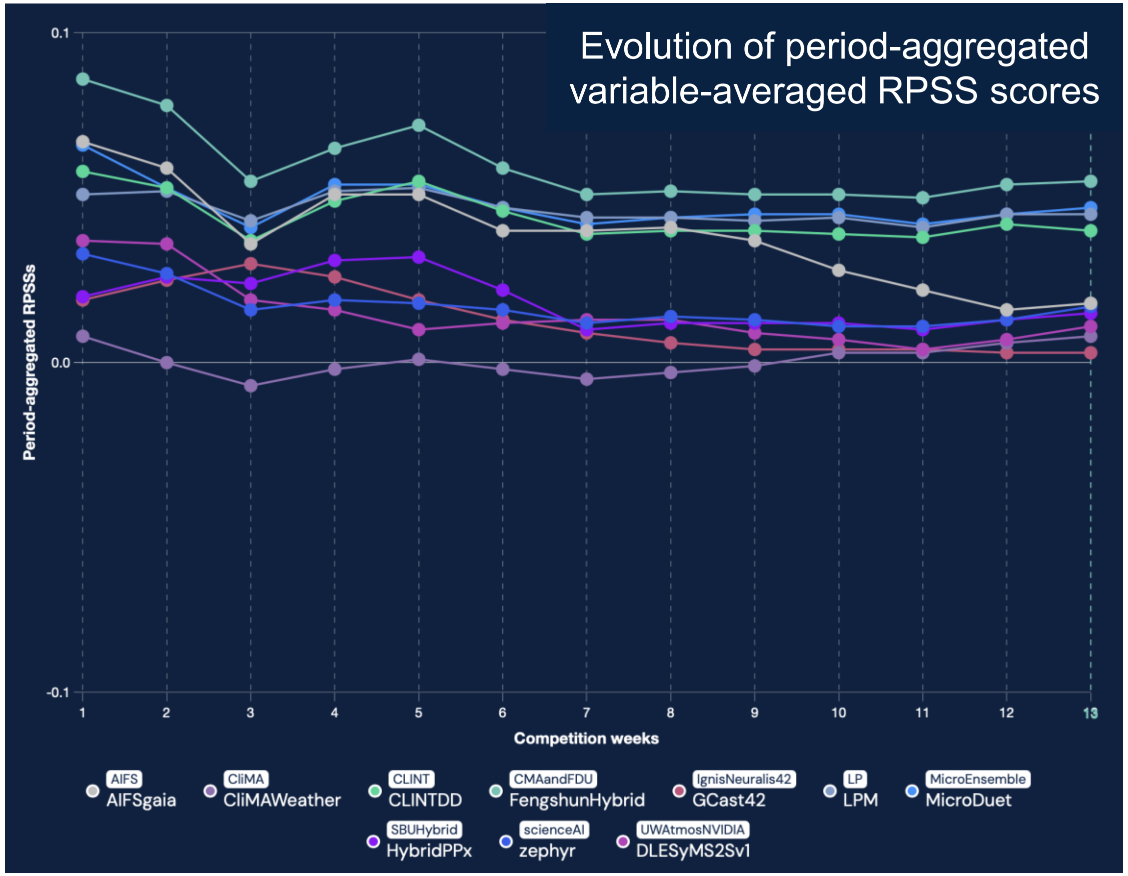
6) Anonymous

7) Stony Brook University

8) NVIDIA

9) CliMA

10) cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici · INPE



Highlight academic / institutional contributions

Stony Brook University / cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici · INPE

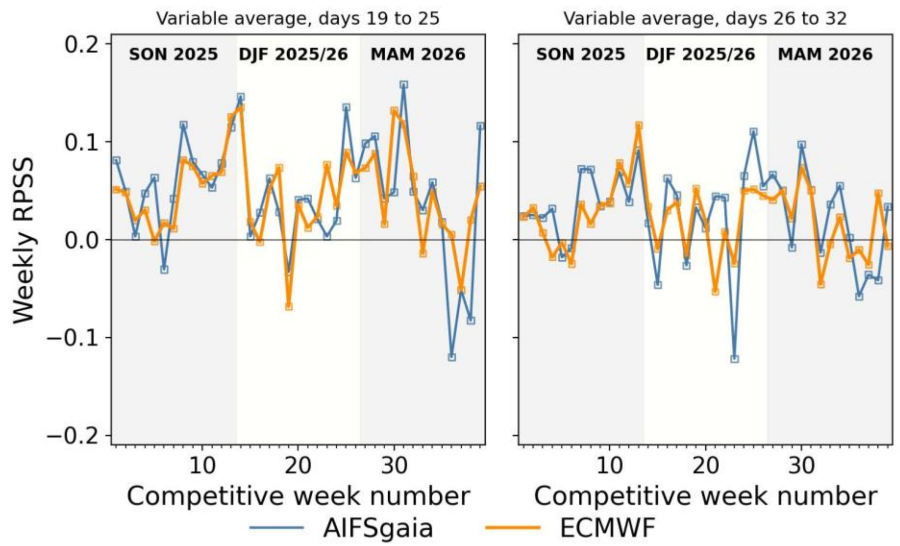
Five-minute presentations + Q&A

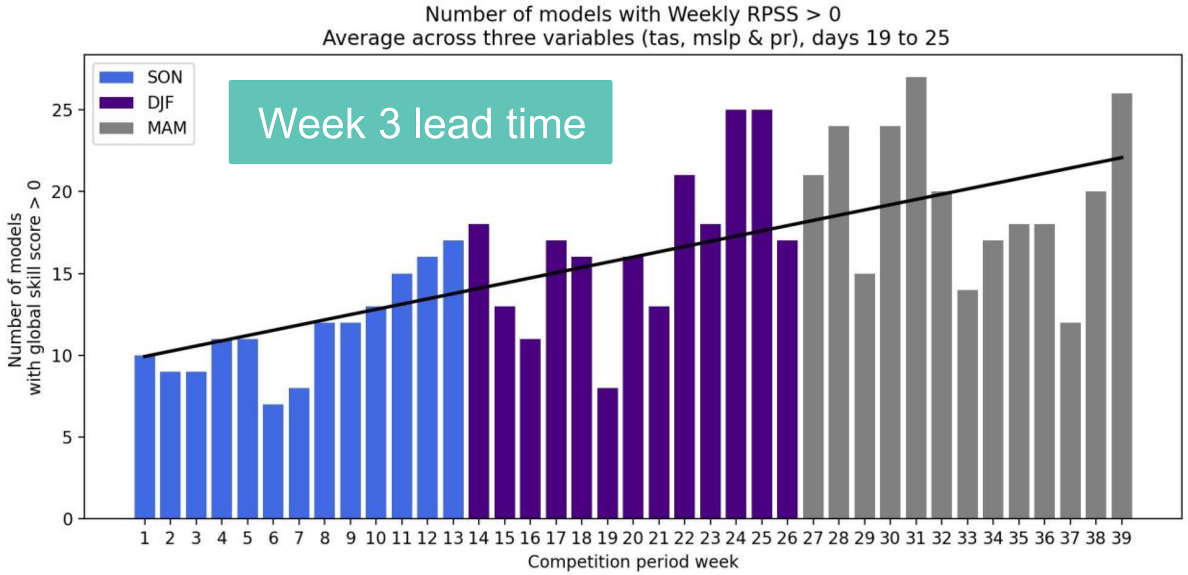
MAM Period evaluation: Scientific insights

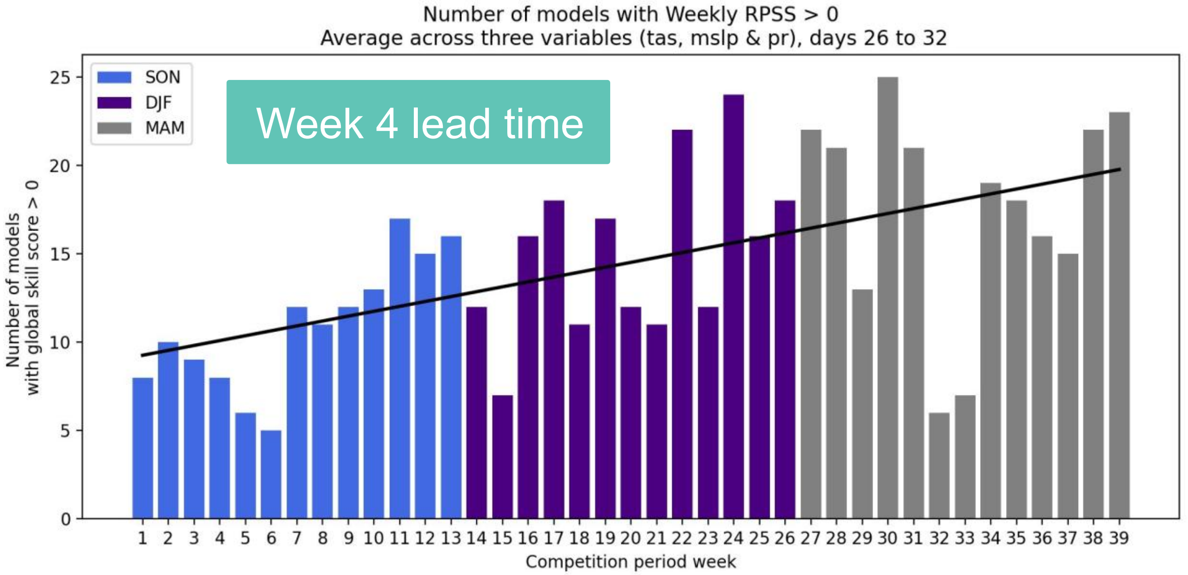
- Increasing number of teams and models outperforming climatology.
- No corresponding trend in intrinsic predictability (inferred from IFS and AIFSgaia benchmarks), suggesting improvements are driven by model development.

Week 3: 5/7/11 Week 4: 5/7/10

Number of teams scoring above climatology for SON/DJF/MAM (period-aggregated, variable-average).



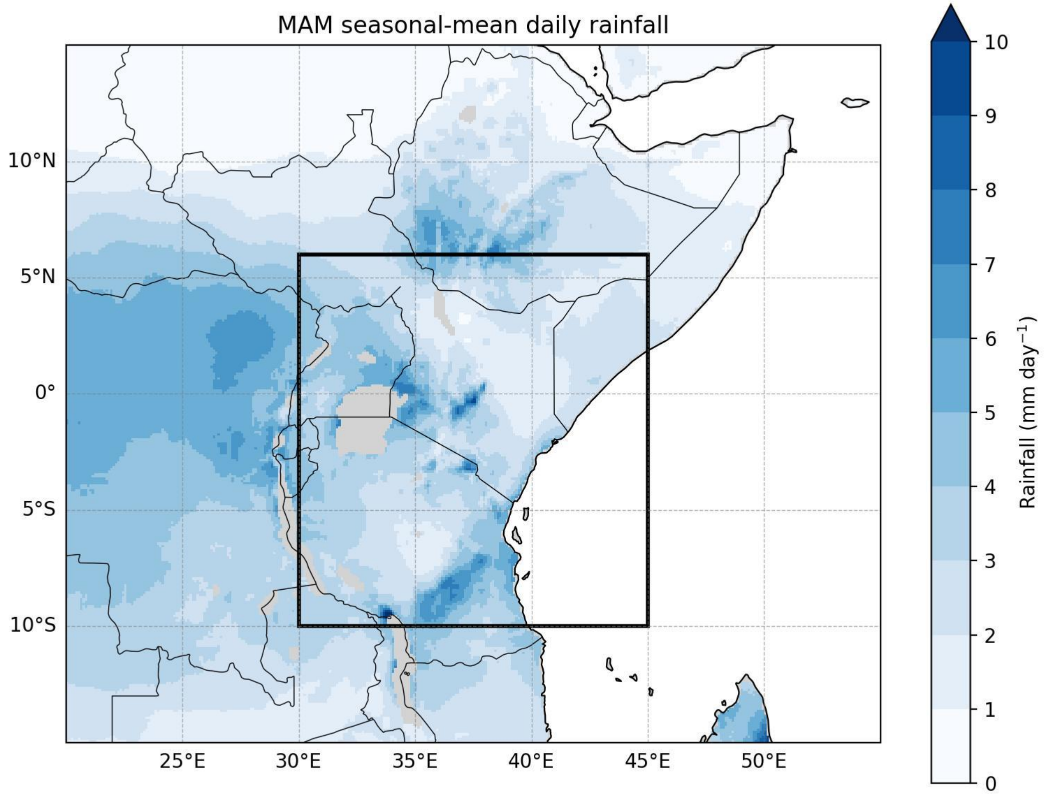




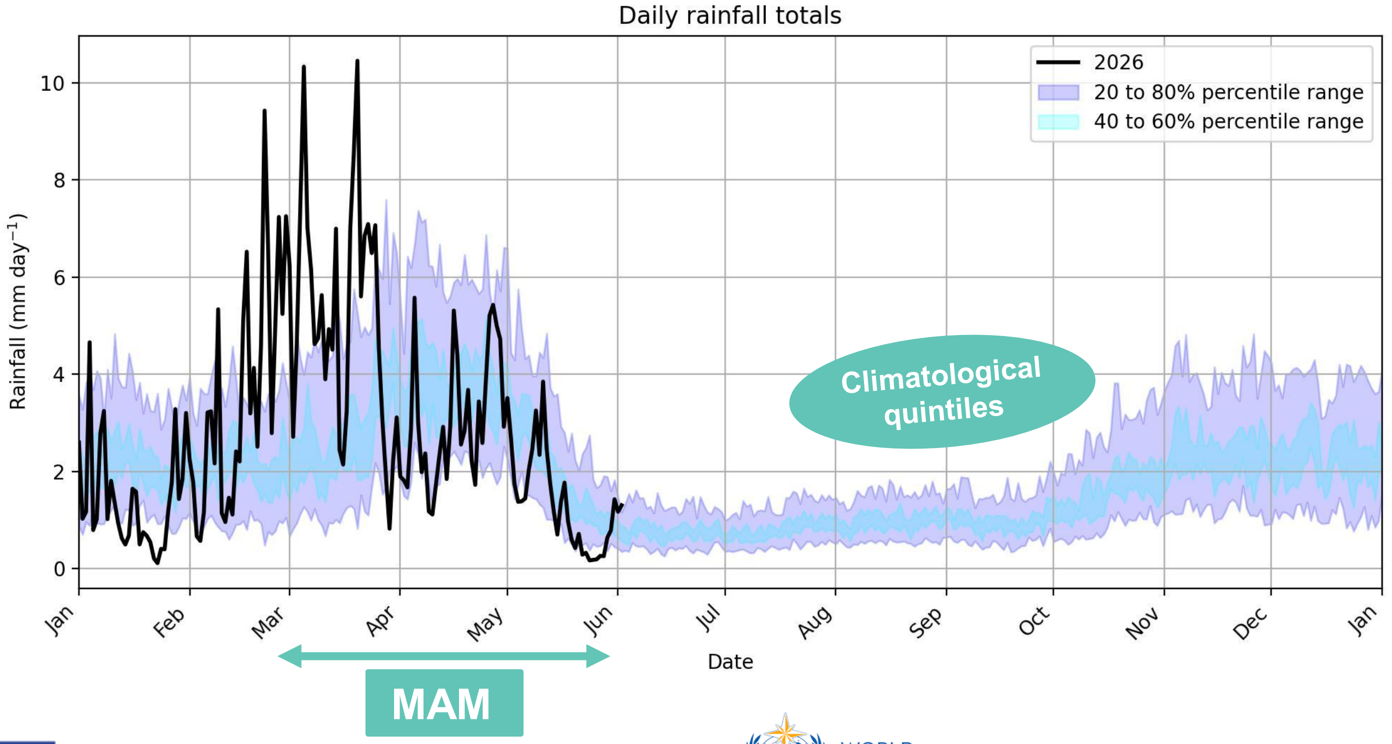
ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

East African MAM rainfall case study

- East African rainfall is characterised by a primary wet season during March to May (MAM), and a secondary, shorter and weaker wet season during October to December (OND).
- MAM 2026 exhibited pronounced intraseasonal variability, with above-normal rainfall at onset followed by below-normal conditions during peak and cessation.



Multi-Source Weighted-Ensemble Precipitation (MSWEP). Combines gauge, satellite and model precipitation estimates (Beck et al., 2019).

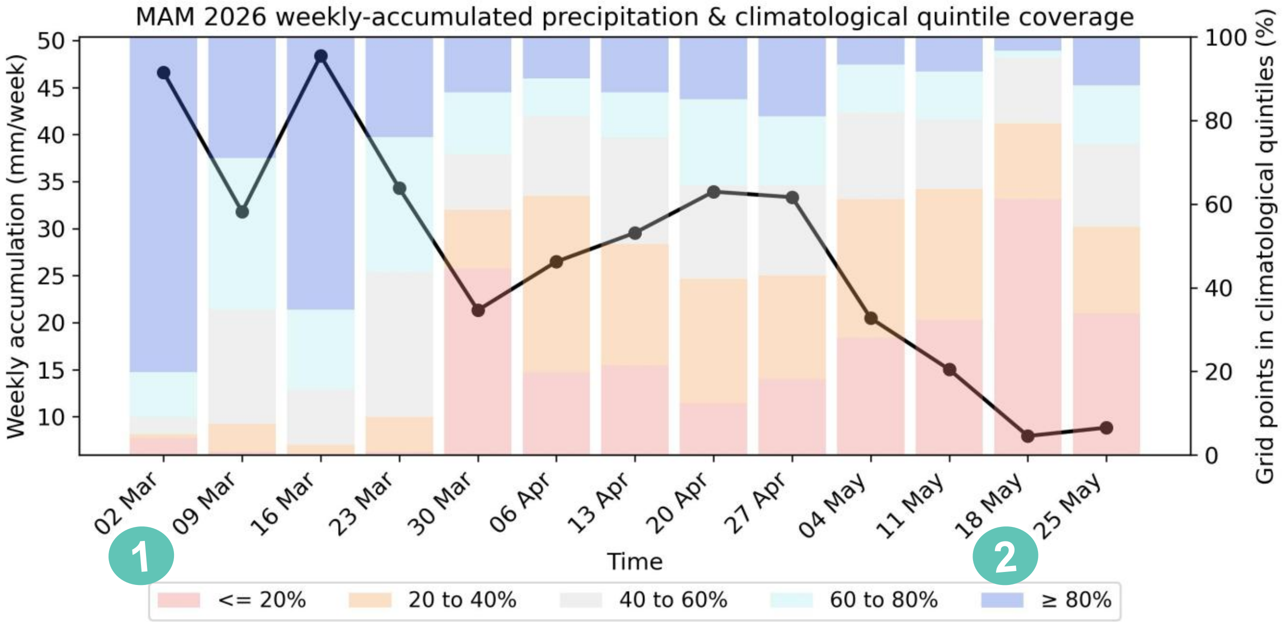


ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

East African MAM rainfall case study

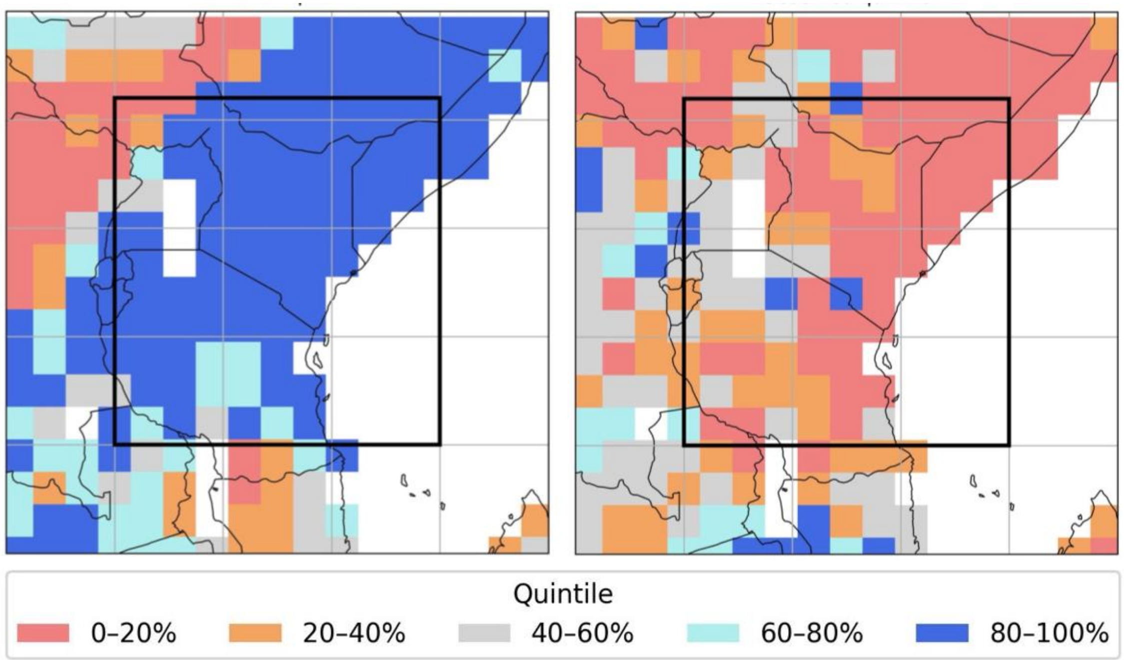
- The first four weeks of MAM were dominated by widespread above-normal rainfall (≥ 60% of climatology).
- Conditions transitioned in April and May, with roughly half the grid points experiencing below-normal rainfall.

ERA5T



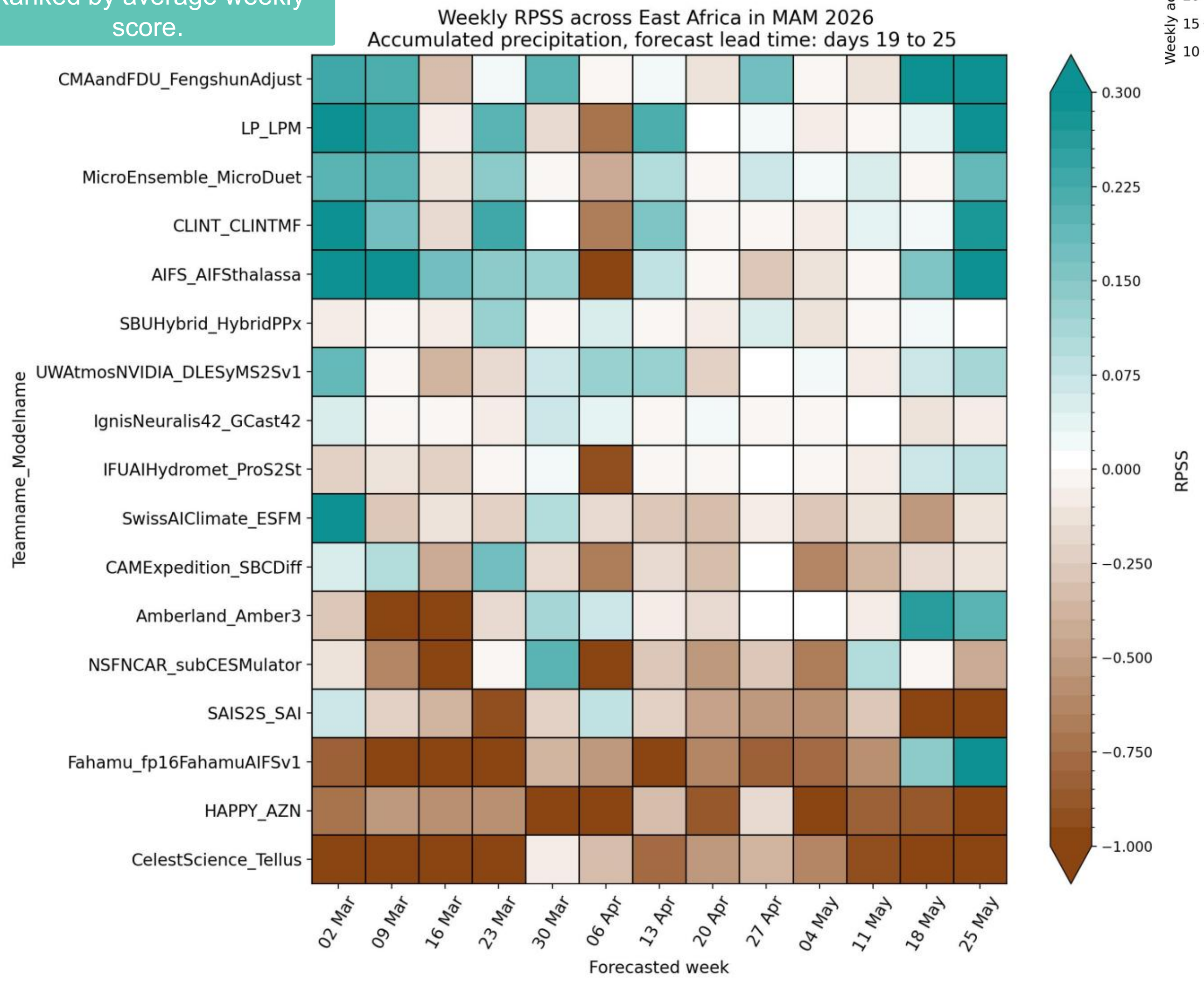
(1) 2nd March 2026

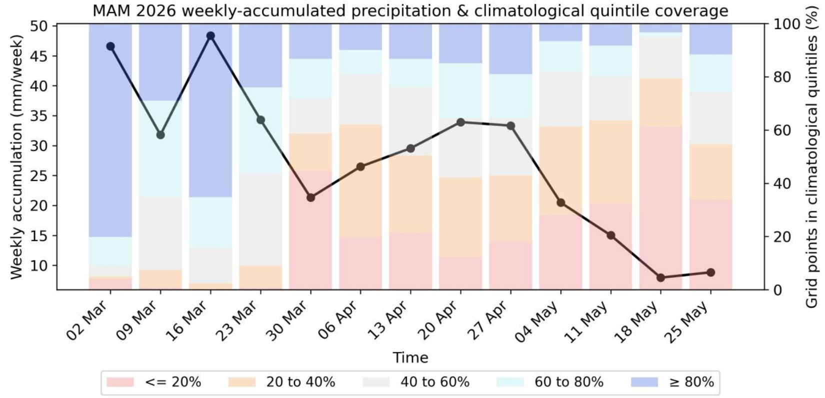
(2) 18th May 2026



East African MAM rainfall case study

Best model from each team. Ranked by average weekly score.





- Skill peaks at the beginning and end of the season.
- Post-processing/hybrid techniques excel during anomalously wet onset, but are outperformed by data-driven methods mid-season.
- A notable skill reduction corresponds to a period (20th April to 11th May) of limited anomalous signal.

Future planned developments and analysis

- Reliability graphs for wet and dry spell representation.
- Conduct region-specific evaluations to pinpoint model weaknesses.
- Perform global extreme event analyses using focused case studies.

AI Weather Quest ECMWF

ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

Presentation by team CMAandFDU

Best ranked-team of the MAM Period for variable-averaged, period-aggregated scores, for both 1st and 2nd forecast windows

Fengshun Family by CMAandFDU

Team Leaders： Bo Lu & Hao Li

Team Members: Z. Dou, L. Chen, Y. Zhao, X. Zhong, J. Hu, Q. Qian, J. Wu, C. Zhao, C. Zhou, C. Wang, L. Du, Z. Shu, Y. Xin

Our Team Members

Meteorologist

AI Expert

CMA

University

Three Submitting Models:

Data-Driven

- Fengshun (AI-light weight)
- FengshunAdjust

AI+NWP

- FengshunHybrid (AI+NWP)

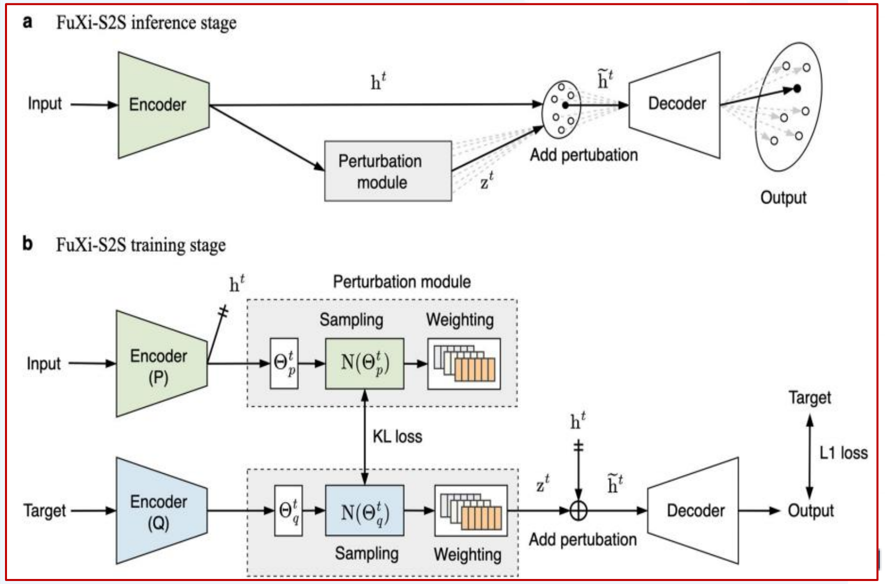
Fengshun

Chen et al., 2024——FuXi-S2S

nature communications

Article https://doi.org/10.1038/s41467-024-50714-1

A machine learning model that outperforms conventional global subseasonal forecast models



1. Purely Data-Driven (ERA5);
2. Forecast weekly quintiles directly (tas, pr, slp);
3. Easy to update (no need to calculate climatology)

Use Monday initial conditions only, due to ERA5 delay

Week number	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	Day of week						
0				1	2	3	4
1	5	6	7	8	9	10	11
2	12	13	14	15	16	17	18
3	19	20	21	22	23	24	25
4	26	27	28	29	30	31	32
5	33	34	35	36	37	38	39
6	40	41	42	43	44	45	46

Forecast submission window | First forecast period | Second forecast period | Publication of evaluation results

FengshunAdjust Updates

An ensemble model for multiple AI forecasts

Base AI Models

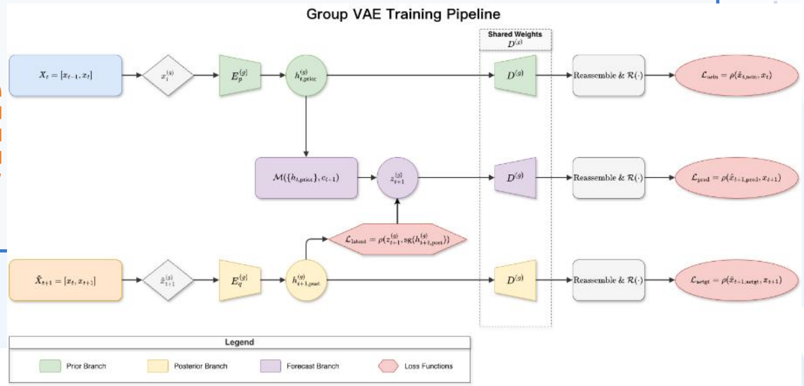
- **FS** — **Fengshun**
- **FD** — **FengshunDiffusion**
- **GV** — **Group VAE + Mamba**

New in MAM

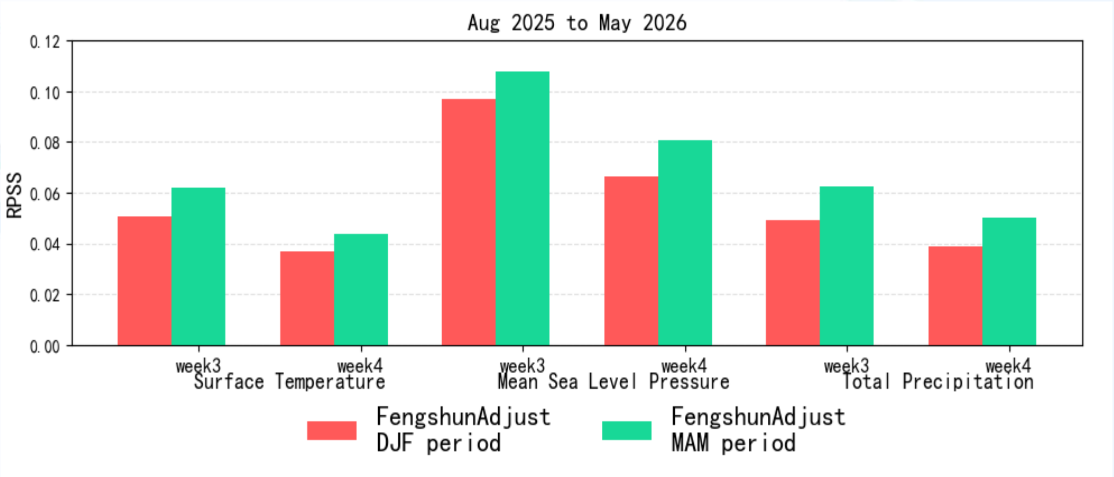
State-space transition modeled with Mamba

Fengshun Adjust
multi-model correction / calibration

Weekly Quintile forecasts

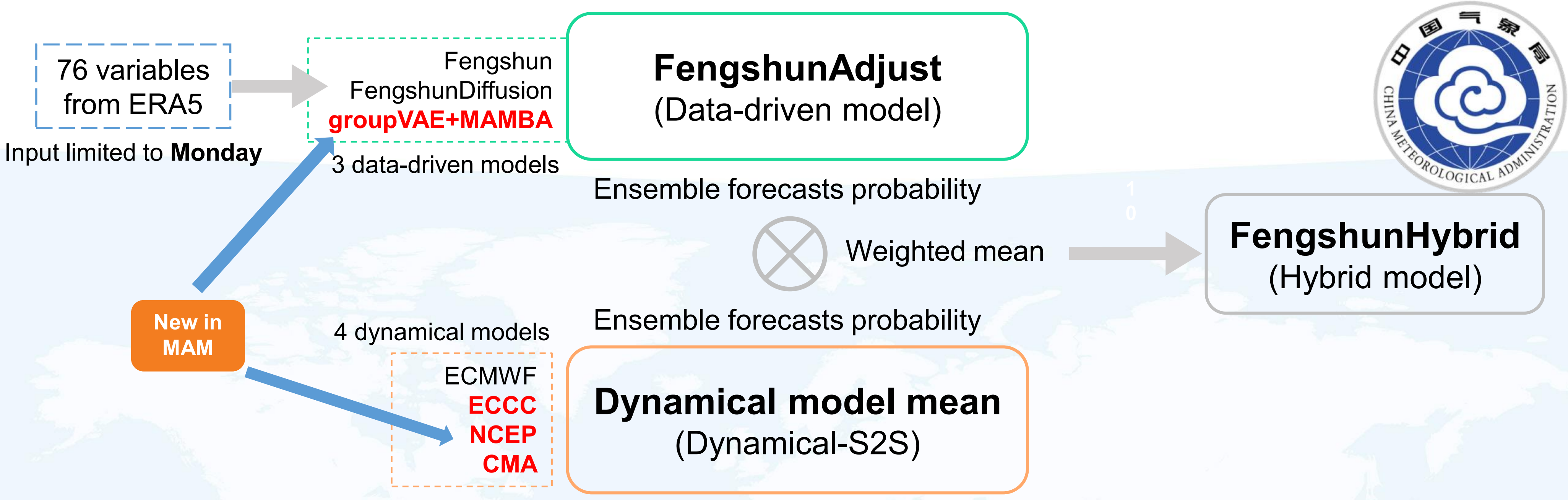


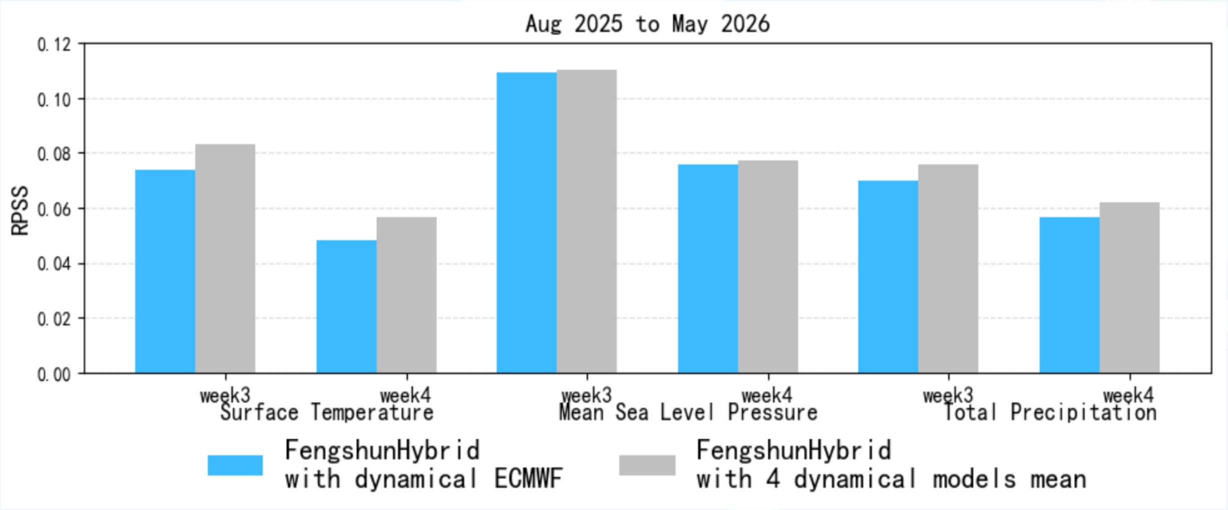
- For the MAM season, we extended the original Fengshun and FengshunDiffusion framework by adding a **new model with Group VAE** and **Mamba for state-space transition modeling**

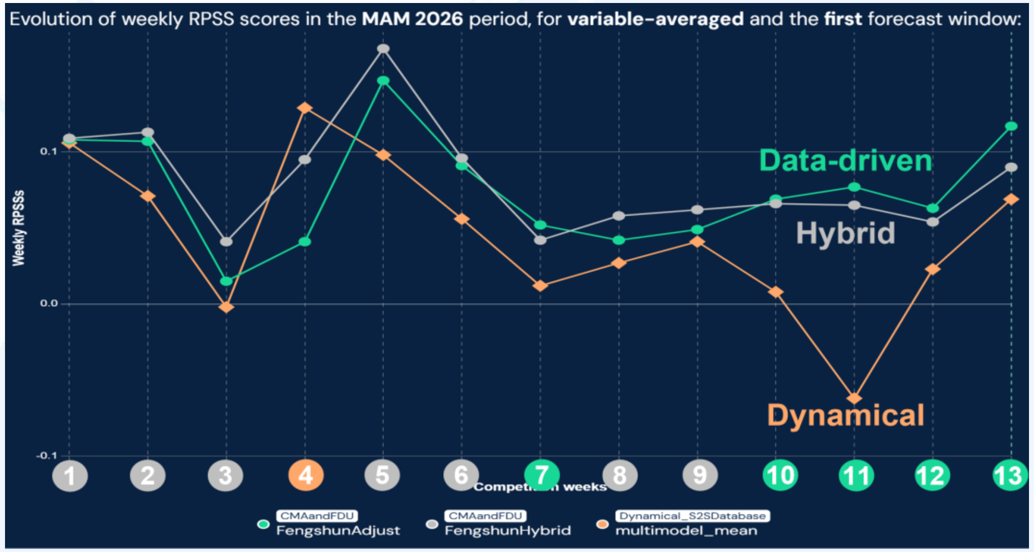


Key update in MAM:

- \+ **Group VAE base model**
- \+ **Mamba state-space transition**
- \+ **Better forecast skill**



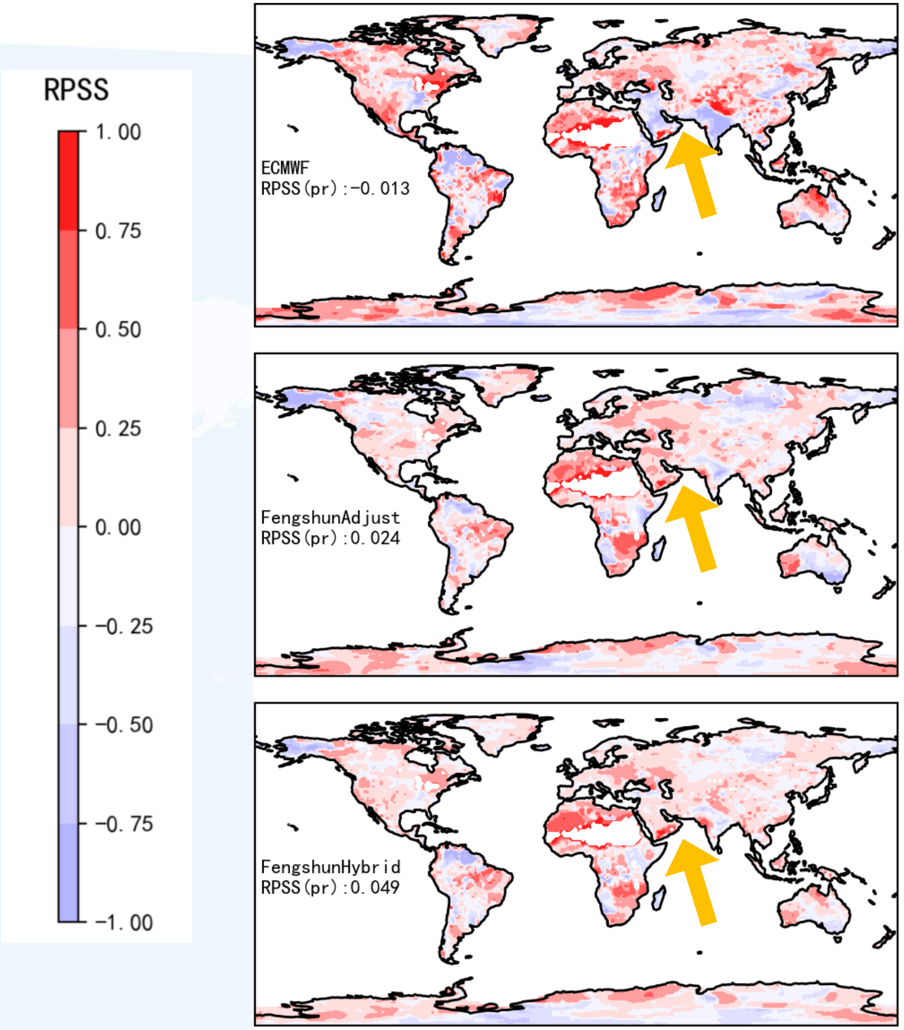




FengshunHybrid Updates

A combination of a data-driven model and dynamical models

Variables: total precipitation
Base time: 26 February, 2026
Forecast period: week-3



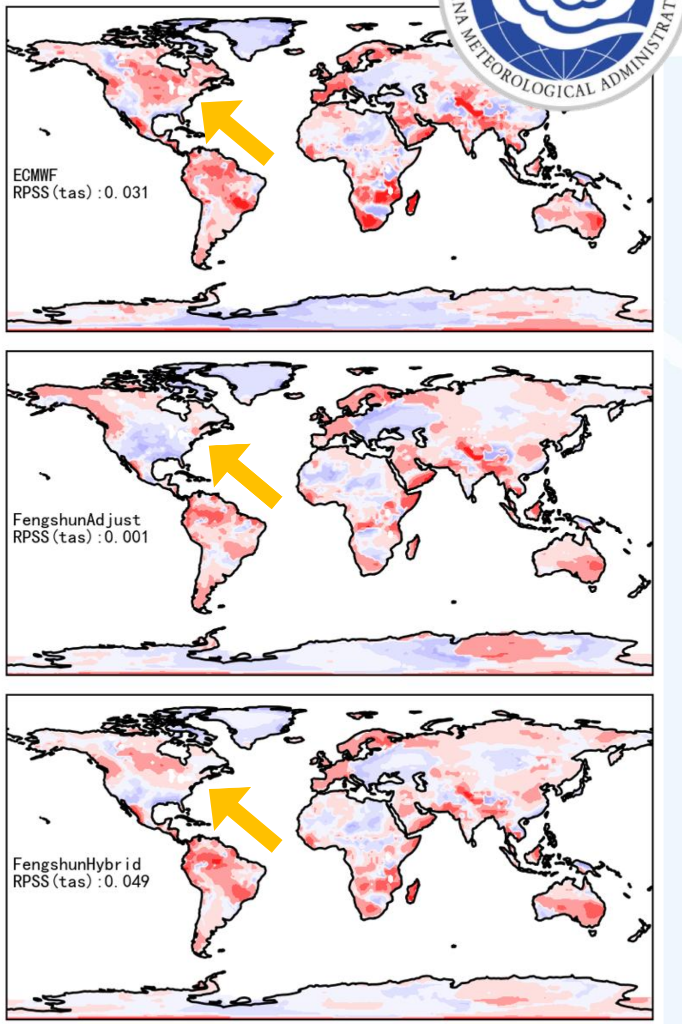
In case 1, the data-driven model (FengshunAdjust) outperforms the dynamical_S2S model (ECMWF)

surface temperature
09 April, 2026
week-3

ECMWF
(Dynamical-S2S)

FengshunAdjust
(Data-Driven)

FengshunHybrid
integrates the advantages of both and balances their respective weaknesses

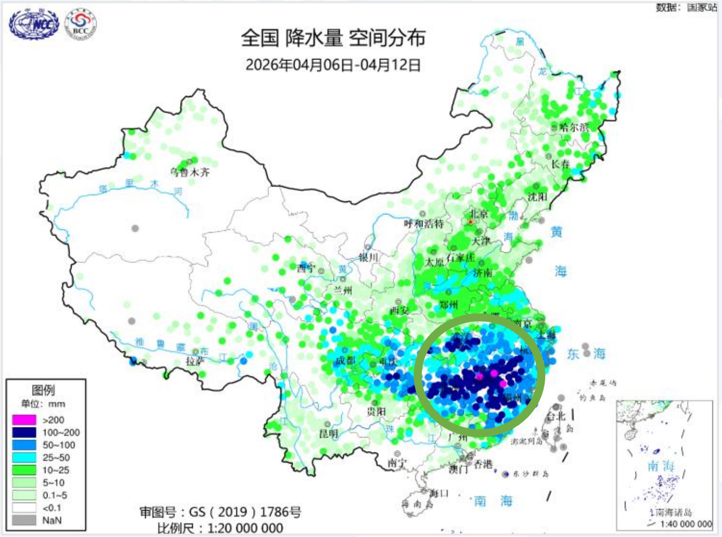


In case 2, the dynamical_S2S model (ECMWF) outperforms the data-driven model (FengshunAdjust)

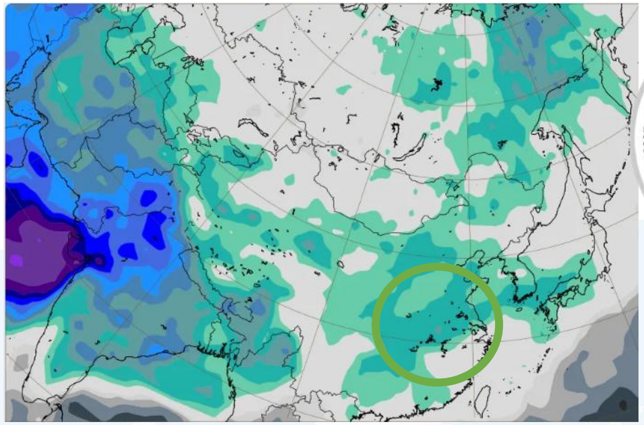
FengshunHybrid
A combination of a data-driven model and dynamical models

An abnormally northward and strong subtropical high caused heavy rainfall across the Jiangnan region in early April.

Observed

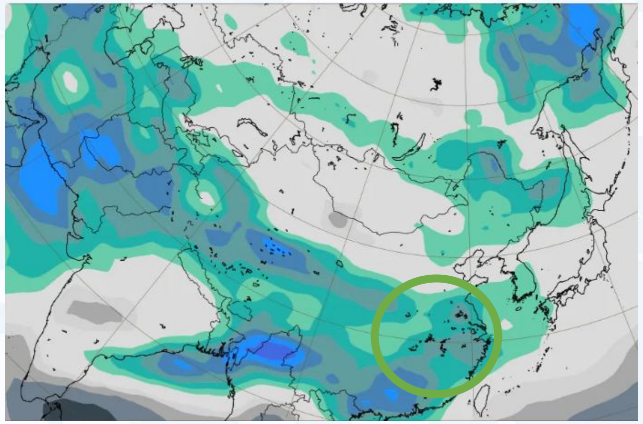


Base time	Valid time	Area	Quintile interval
Thu 19 Mar 2026	Mon 06 Apr 2026 - Sun 12 Apr 2026	Eastern Asia	>= 80%



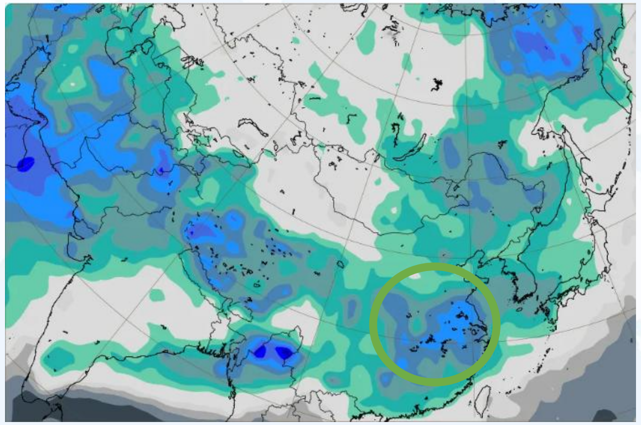
Dynamical model mean

Dynamical-S2S



FengshunAdjust

Data-driven

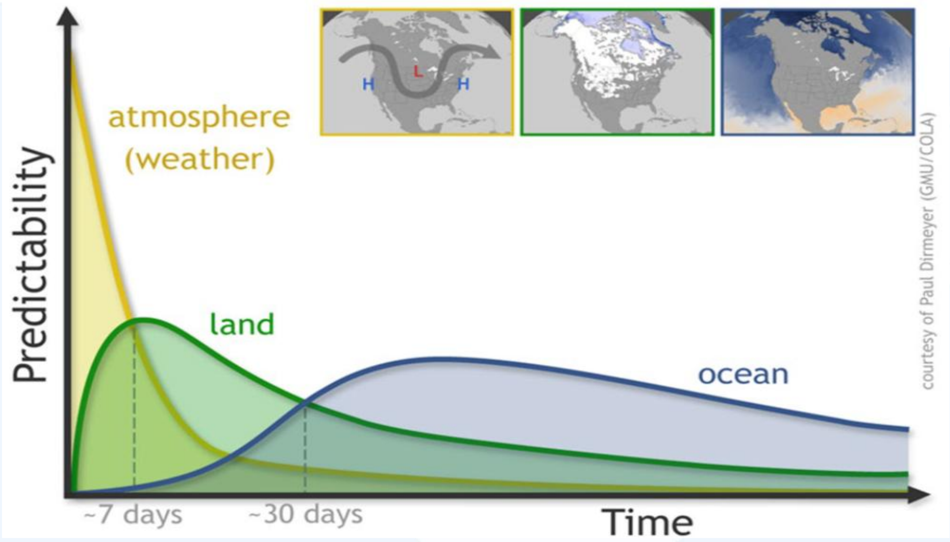


FengshunHybrid

Hybrid-model

Case of 06 to 12 April 2026, heavy rainfall on Jiangnan, China

What's next

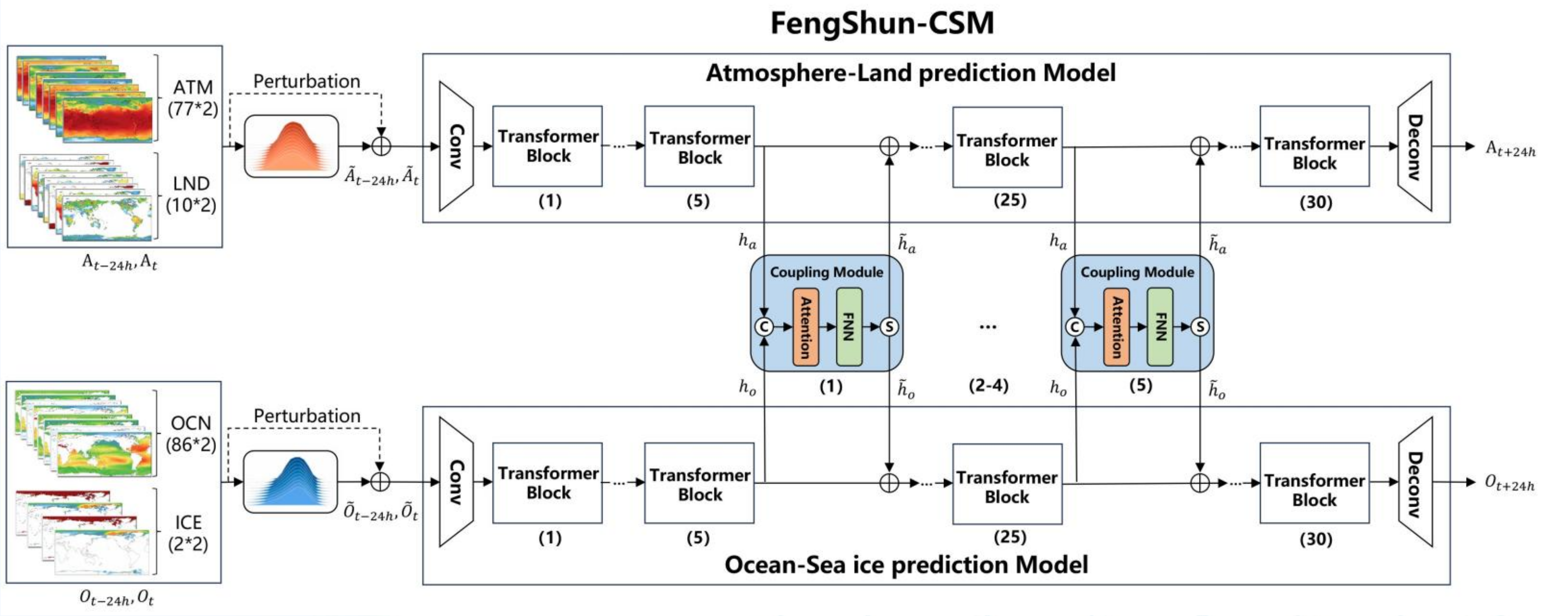


1. Incorporating multi-sphere variables (ocean, land) for S2S prediction

Given the higher predictability of ocean and land surfaces, additional variables from these spheres can be jointly leveraged during model training.

2. Enforcing physical constraints in the loss function for long-term influences

Design physics-informed loss functions to embed the prolonged memory effects of ocean and land on the atmosphere into autoregressive training.



3. Exploring coupling schemes for multi-sphere information

Investigate effective ways to couple information across different Earth system components.


The END

Stony Brook University

AI Weather Quest ECMWF

Presentation by team SBUHybrid

Spotlighted as an outstanding academic team

ECMWF

Destination Earth

WORLD METEOROLOGICAL ORGANIZATION

Team SBUHybrid

AI Weather Quest MAM Webinar

Cheng Zheng and Edmund Chang, Jun 18, 2026

Team Members

Cheng Zheng
Research Scientist

Edmund Chang
Professor

Academic research group
School of Marine and Atmospheric Sciences, Stony Brook University (SBU)

Stony Brook University: The largest campus of the State University of New York (USA)

SBUHybrid Model

Training Stage

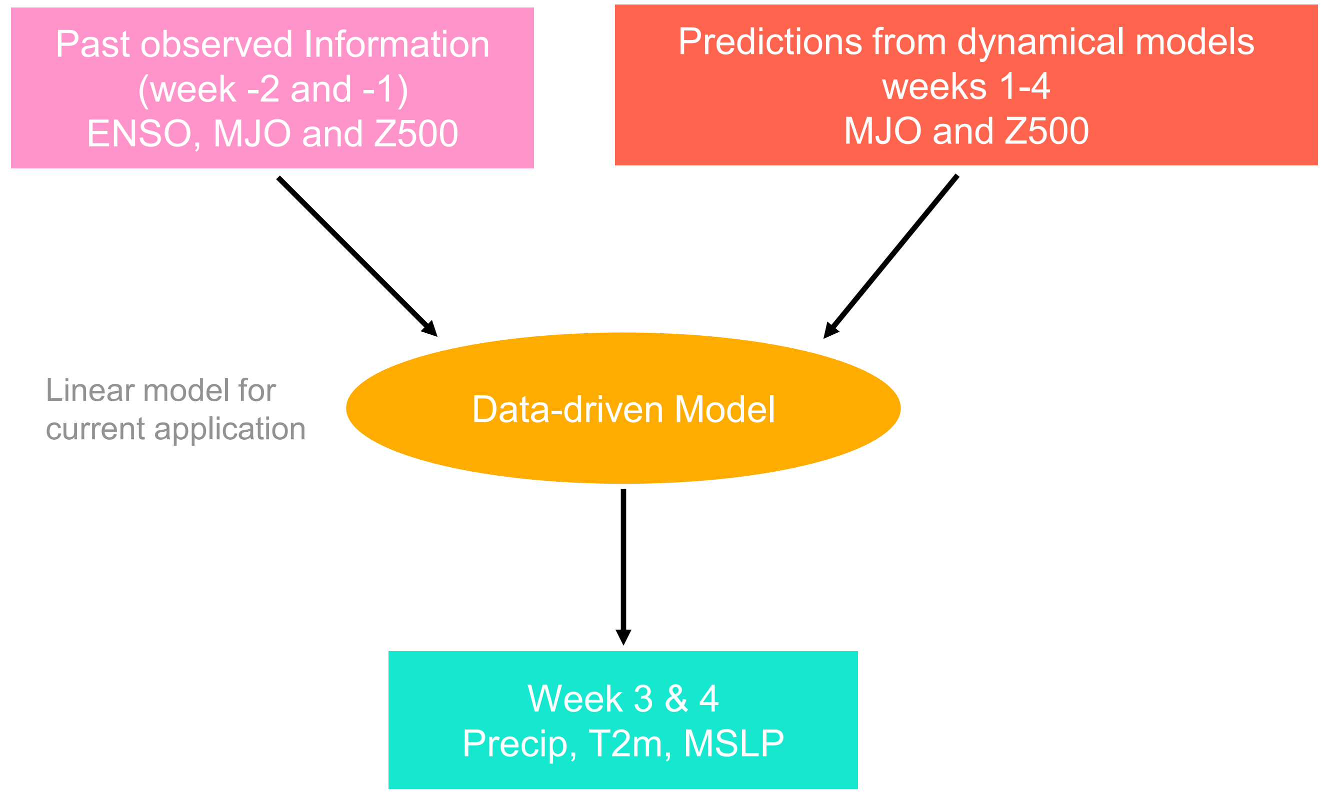
Application/Tuning (reforecast)

Application (real-time)

A (hybrid) method incorporating data-driven tools, climate model simulations, and dynamical model predictions

Input:

Output:



ENSO, MJO (Madden Julian Oscillation), and Z500 (500-hPa geopotential height) represented by indices

The data-driven model is a reconstruction/prediction model for precip, T2m and MSLP

Computer resources: HPC clusters at Stony Brook University

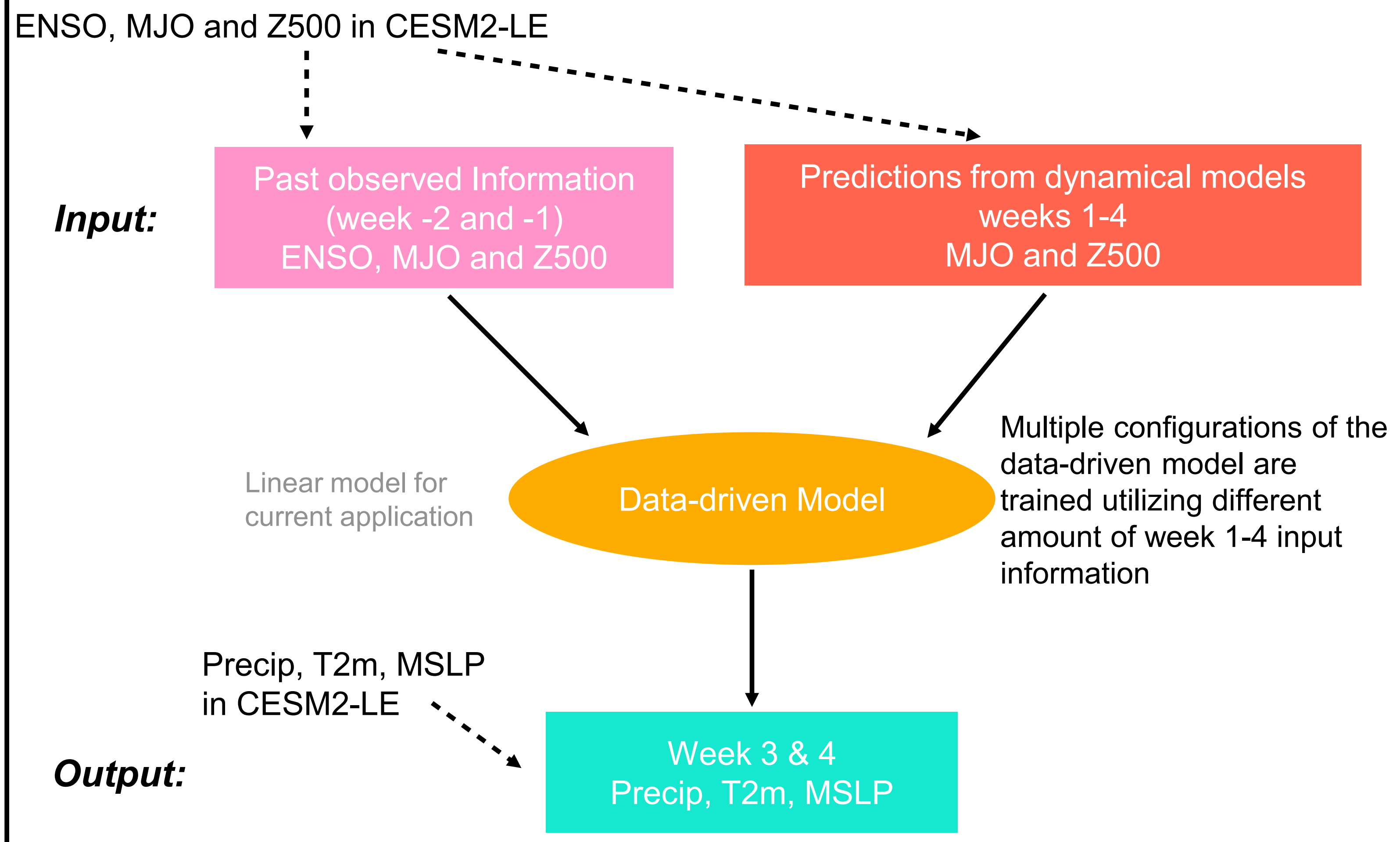
SBUHybrid Model

Training Stage

Training with climate model data

Application/Tuning (reforecast)

Application (real-time)



Training data: NCAR Community Earth System model version 2 Large Ensemble (CESM2-LE)
CMIP6-style simulations (100 ensemble members, 1950-2022)
No Observational/dynamical forecast data used

Computer resources: HPC clusters at Stony Brook University

SBUHybrid Model

Training Stage

Application/Tuning (reforecast)

Application (real-time)

ENSO, MJO and Z500 in ERA5

MJO and Z500 in IFS-CY49r1
subseasonal reforecast

Input:

Past observed Information
(week -2 and -1)
ENSO, MJO and Z500

Predictions from dynamical models
weeks 1-4
MJO and Z500

Linear model for
current application

Data-driven Model

Evaluate the performance of
each configuration in
reforecast

Output:

Week 3 & 4
Precip, T2m, MSLP

Prediction verified against ERA5 data
Identify the optimal configuration (amount of dynamical model information to be included)
Estimate 5-category probabilistic threshold in reforecast period

Precip, T2m, MSLP from dynamical predictions not required as inputs

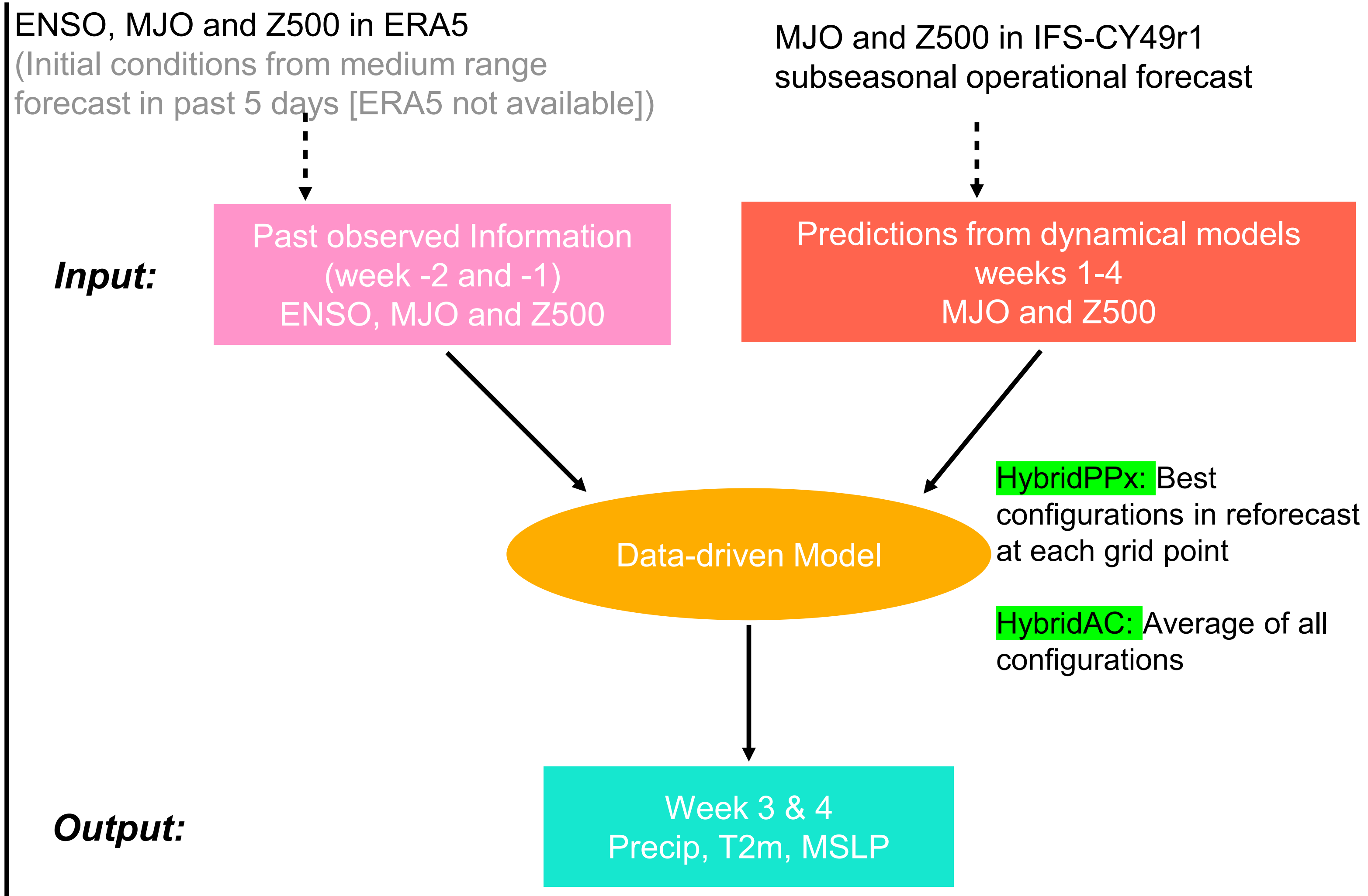
Computer resources: HPC clusters at Stony Brook University

SBUHybrid Model

Training Stage

Application/Tuning (reforecast)

Application (real-time)



Probabilistic prediction using 5-category threshold estimated in reforecast

Computer resources: HPC clusters at Stony Brook University

Discussion

Not a post-processing model:

Climate model simulations used as training data;
Dynamical predictions of target variable (e.g., precipitation) not used as inputs of the model.

Works well for precipitation:

For a certain weekly-mean large-scale circulation field (e.g., represented by Z500), precipitation is determined by evolution of synoptic scale circulation under such weekly mean circulation field.
If we consider weekly-mean large-scale circulation is somewhat predictable (e.g., predictability from teleconnections) on subseasonal time scales, then weekly-mean precipitation prediction depends on sampling possible synoptic scale circulation evolution under the weekly-mean circulation field.
Dynamical model use about 10-100 ensemble members for such sampling.
The data-driven model is trained with thousands of years of climate model simulations. Applying the data-driven model is similar to using a very large ensemble to explore possible synoptic scale circulation evolution under a certain weekly-mean large scale circulation field.

The method works particularly well if the target variable is noisy (precipitation), as a larger ensemble size benefits for prediction of noisy fields

For T2m and MSLP:

1) Less noisy so our hybrid model is less advantageous;
2) More input information (in addition to ENSO, MJO, and Z500) may be required for better predictions

Zheng, C., and E. K. M. Chang, 2026 (submitted): Skillful Subseasonal Precipitation Forecasts Integrating Climate Simulations with Numerical Weather Predictions over the CONUS.

Challenges & Future Developments

Challenges:

- Producing probabilistic forecast using data-driven models trained with climate simulations
    - No forecast spread/uncertainty information in climate simulations
- Relatively small reforecast dataset for fine-tuning.

Future Developments:

- Adding more input (predictors) into the model
    - Incorporating moisture as an additional predictor leads to higher skill
        - Already implemented (SBUHybridStillDeveloping)
- Applying machine-learning method (instead of linear method) in the data-driven model
- Using data from other climate models for training
- Improving tuning procedures

Presentation by team IgnisNeuralis42

Spotlighted as an outstanding institutional team

GCast42: Latent-Space Subseasonal Forecasting on the Sphere and Next Steps

Team IgnisNeuralis42 • CMCC Foundation (Italy) & INPE (Brazil)

ECMWF AI Weather Quest — MAM Awards Webinar, 18 June 2026

The Team

IgnisNeuralis42 — a CMCC (Italy) × INPE (Brazil) collaboration


Ilenia Manco (Team Leader)
CMCC Foundation, Italy


Otavio M. Feitosa
INPE, Brazil / CMCC Foundation, Italy


Paola Mercogliano
CMCC Foundation, Italy


Antonio Navarra
CMCC Foundation, Italy


Saulo Ribeiro de Freitas
INPE Instituto Nacional de Pesquisas Espaciais, Brazil


Haroldo Fraga de Campos Velho
INPE Instituto Nacional de Pesquisas Espaciais, Brazil

Competition entry: GCast42 (data-driven) — submitted weekly to the ECMWF AI Weather Quest. Eligible for variable-averaged, period-aggregated scores and scoring above climatology.

Training Data

ERA5 reanalysis on an equal-area HEALPix sphere

Source: ERA5 reanalysis, 1950–2016 (67 years), daily fields. The longest consistent global record, giving the model decades of recurring seasonal cycles to learn from.

Grid: HEALPix equal-area, nside=64 (~1 deg). No polar singularity and no latitude weighting in the loss; hierarchical nesting maps onto encoder-decoder levels.

Variables: 138 ERA5 fields: upper-air (5 vars × 13 levels), surface, soil moisture at 4 depths, sea ice, the full radiation budget, CAPE/CIN and stratospheric ozone — the slow forcings that carry S2S signal.

Conditioning: Day-of-year climatology (mean + std), z-score normalized. Anchoring each forecast to the seasonal baseline is what keeps skill above climatology at these lead times.

DATASET AT A GLANCE

67 years of ERA5, 1950–2016

138 atmospheric & forcing variables

768 px HEALPix nside=8 (~815 km)

daily temporal resolution

For the competition, forecasts are regridded to 1.5° and submitted as quintile probabilities at days 19–25 and 26–32, scored by RPSS against ERA5.

Model Architecture

GCast42 — forecasting in a learned latent space

The competition model forecasts in a learned latent space rather than on the raw grid. Three stages:

Encoder : compresses the 138-variable HEALPix state into a compact latent representation, keeping the large-scale structure that drives sub-seasonal signal.

Latent forecast : an autoregressive model advances the state entirely within latent space out to a 45-day horizon. Working in low dimension keeps each step cheap and stable over long rollouts, and lets us draw an ensemble.

Decoder : the latent trajectory is reconstructed into physical fields, from which calibrated quintile probabilities are produced for submission.

Trained on H100 GPUs; a full 45-day forecast runs in seconds.

INPUT ERA5 analysis (HEALPix, 138 vars)

ENCODER
138-var state → compact latent z

FORECAST IN LATENT SPACE
autoregressive rollout, z(t) → z(t+1)
45-day horizon • ensemble members

DECODER
latent z → physical fields

OUTPUT quintile probabilities

Challenges & Next Steps

Where we are, and where we are going

Key challenges

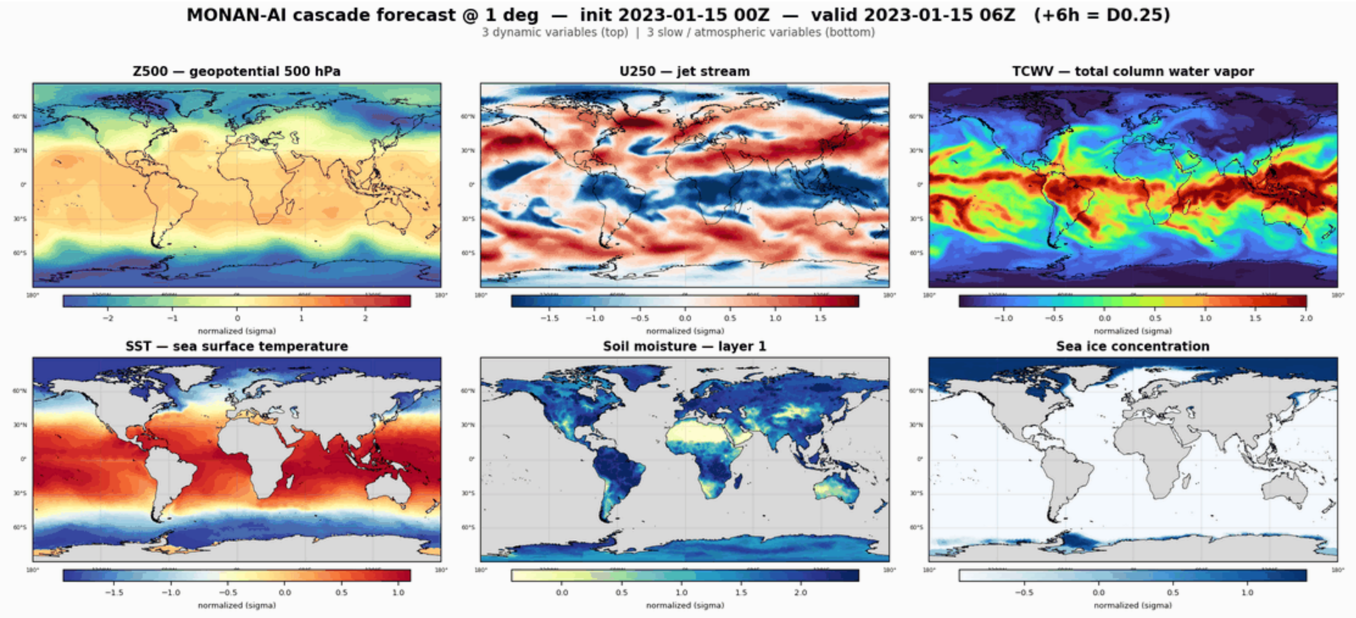
- **The predictability desert.** Skill collapses fastest in weeks 3/4, and precipitation is the hardest field of all.
- **Coarse resolution.** captures planetary scales but not mesoscale or orographic precipitation, which caps T2M and Z500 skill.
- **Probabilistic calibration.** Keeping ensemble spread matched to forecast error across long rollouts.
- **Initial-condition dependence.** We still initialize from ERA5, so the system is not yet self-contained.

Planned developments

Next-generation generative model. A generative model based on diffusion / Conditional Flow Matching (CFM), currently in training and being prepared for publication.

What our experience tells us

ML forecast skill follows the classical predictability literature: the signal comes from slow forcings and coupled Earth-system components: ocean, land, sea ice, stratosphere, not the atmosphere alone.



Thank you for your attention.

Any questions?

Key milestones and an announcement: the Quest continues!

14/08/25				13/08/26	12/11/26			
Initial framework				**Bridge period**	**Expanded framework**			
September - November	**December - February**	**March - May**	**June - August**	**September - November**	**December - February**	**March - May**	**June - August**	**September - November**

Access MAM data
Access the complete MAM forecast datasets and participating teams' models' summaries

First JJA leaderboards

07-11/09/26
Join us at
S2S2D, Reading

16/07/26
Join us for to the *AI Weather Quest: What's next?* webinar
Building on the success of the first forecasting year, the Quest will continue as a long-term benchmarking and learning infrastructure, with an expanded scientific framework and new improvements. Join us to learn more!

AI Weather Quest ECMWF

ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

Thanks!

To everyone involved in the organisation of the AI Weather Quest and everyone involved in its journey!