AI Weather Quest

ECMWF

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

DJF Awards Webinar

ECMWF

Destination Earth

WORLD METEOROLOGICAL ORGANIZATION

Agenda

- **Presentations**
  - ✓ Introduction by Frédéric Vitart
  - ✓ DJF Period participation overview
  - ✓ DJF Period evaluation overview and cold spell case studies
  - ✓ Presentation from team MicroEnsemble
  - ✓ Presentation from team LP
  - ✓ Presentation from team AIFS
  - ✓ Key milestones and actions

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 take the floor during the Q&As or ask questions in the chat.

AI Weather Quest ECMWF

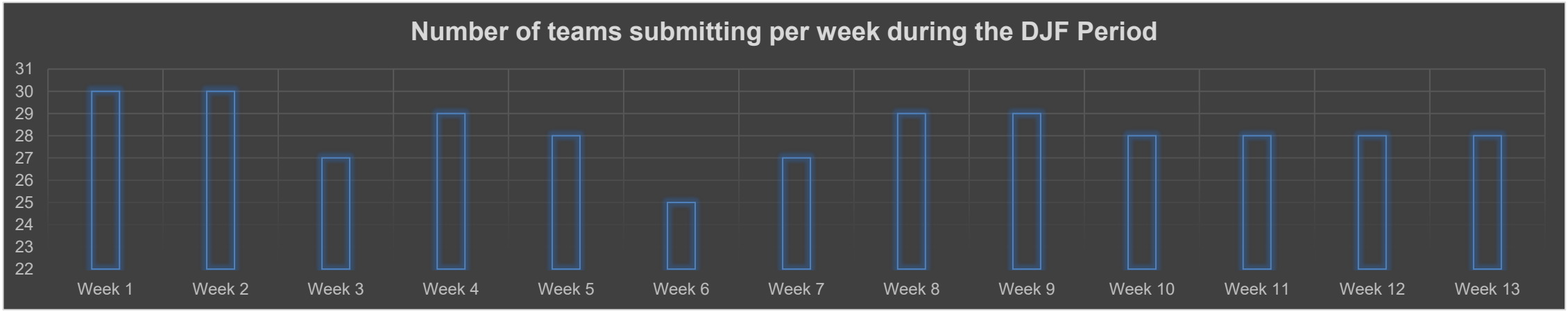
Introduction by Frédéric Vitart

Head of sub-seasonal forecasts research, ECMWF

ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

DJF Period participation overview

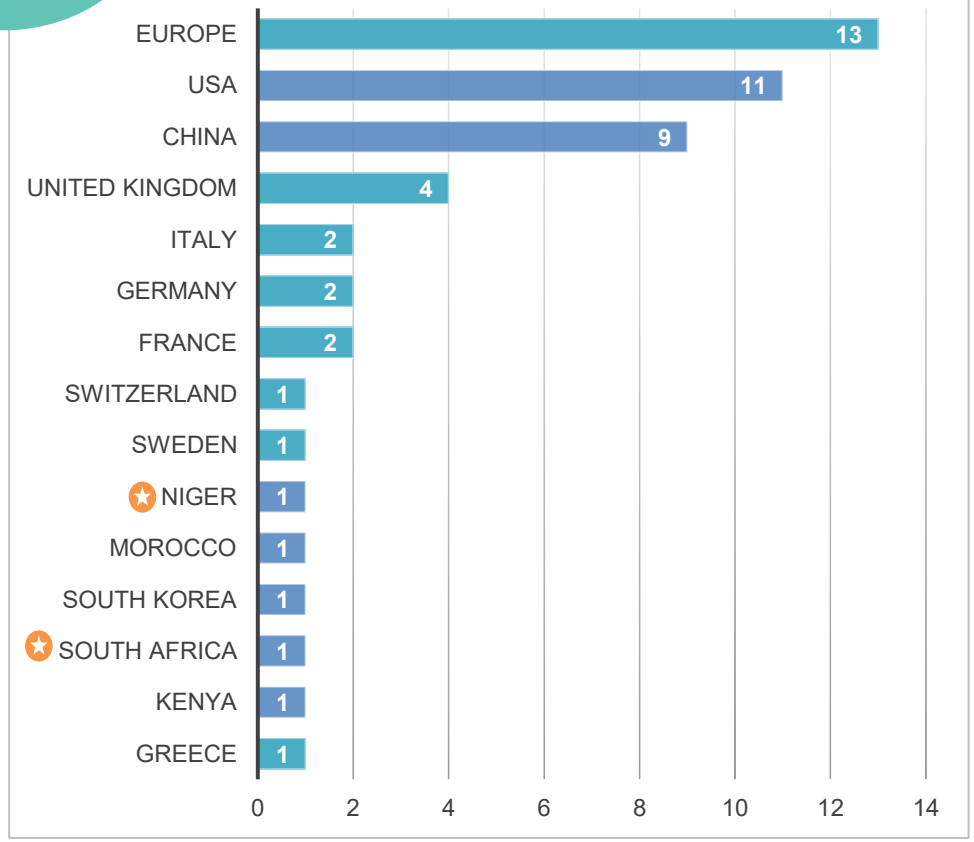
	Number of competitors to the Quest	Number of teams	Number of models
Accumulated participation since the start of the competition	194	42	74
Accumulated over current DJF period	177	37 (including 7 teams not competing in SON)	66
Accumulated over SON period	166	35 (including 5 teams not competing in DJF)	60



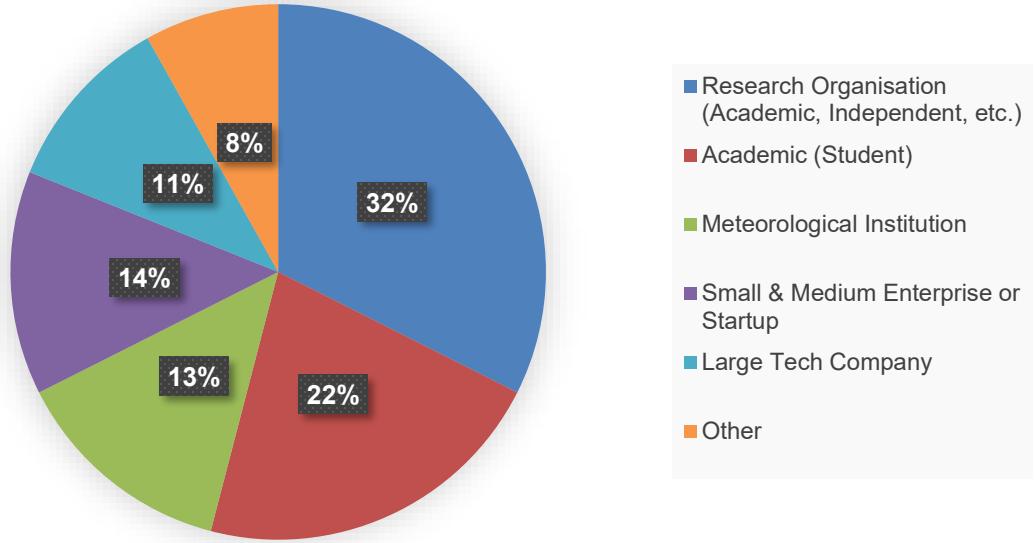
DJF Period participation overview

From 14 countries

Team leader organisation location



Team leader organisation type

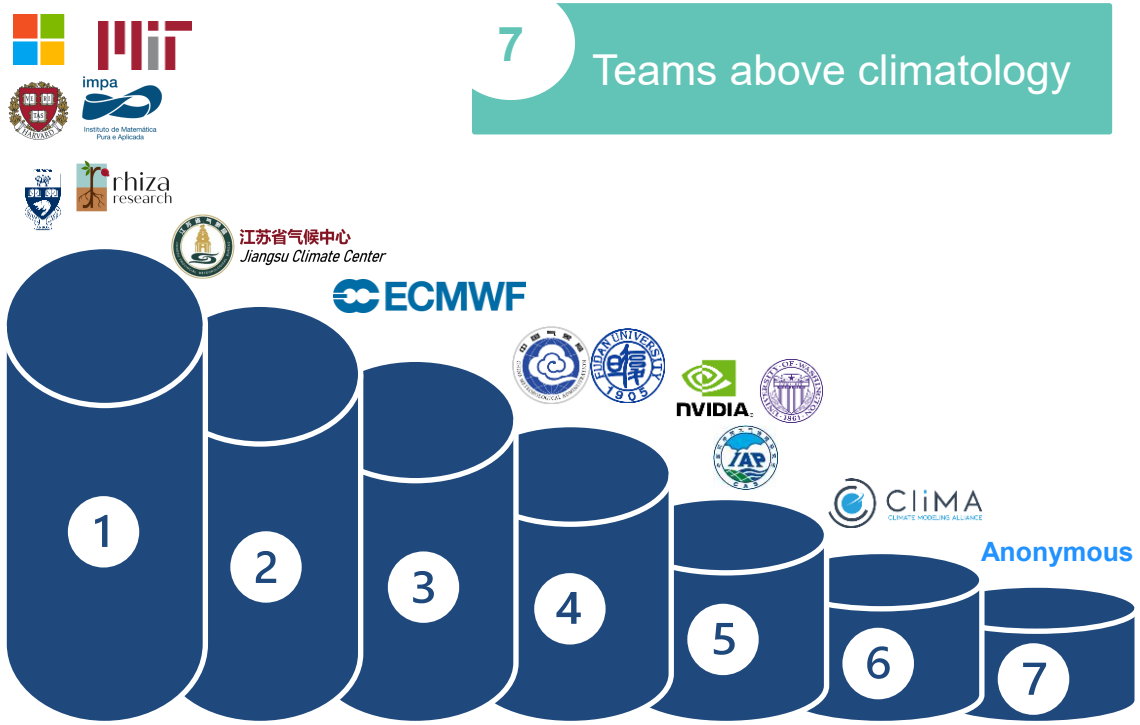


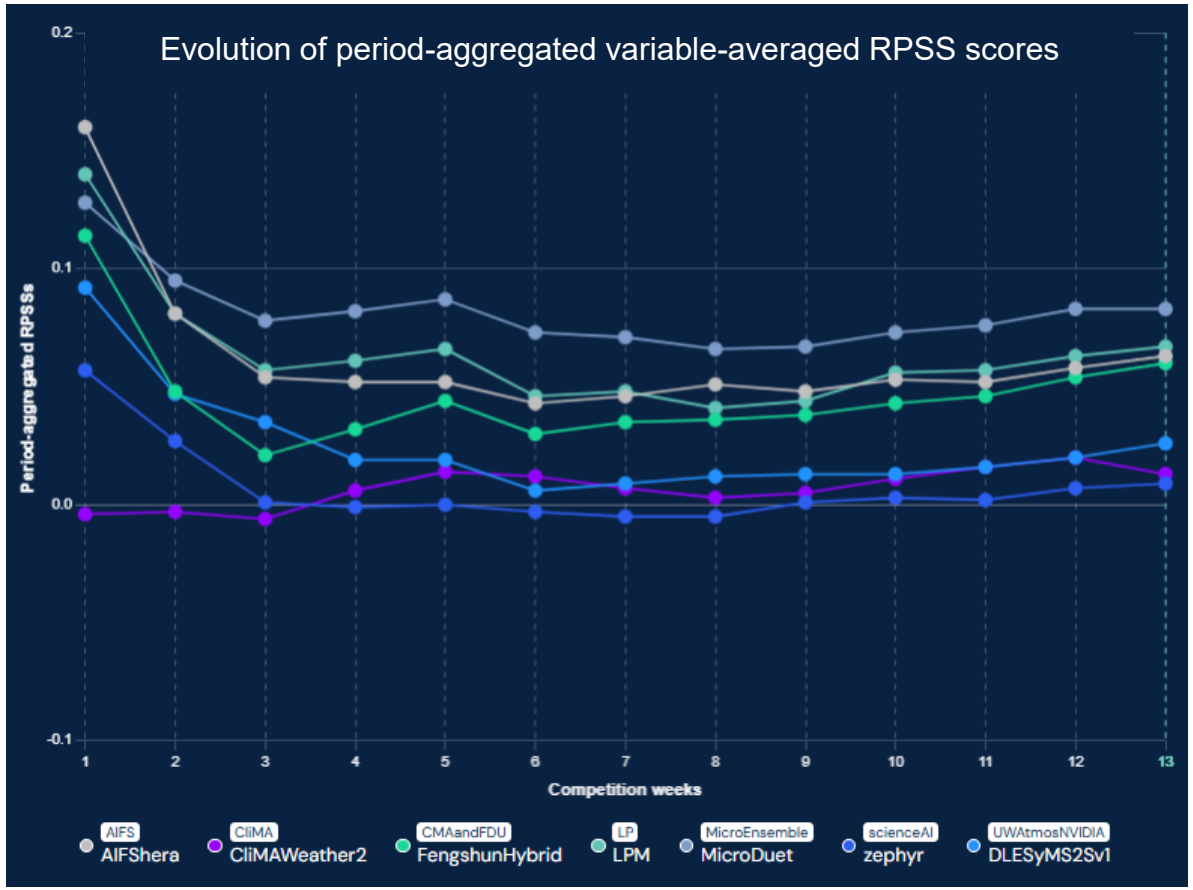
- New countries represented (★)
- Increasing participation from the United States
- Shift in participant profile, with a decreasing proportion of student-led teams and a growing share of professional participants across all organisation types

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

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

7 Teams above climatology



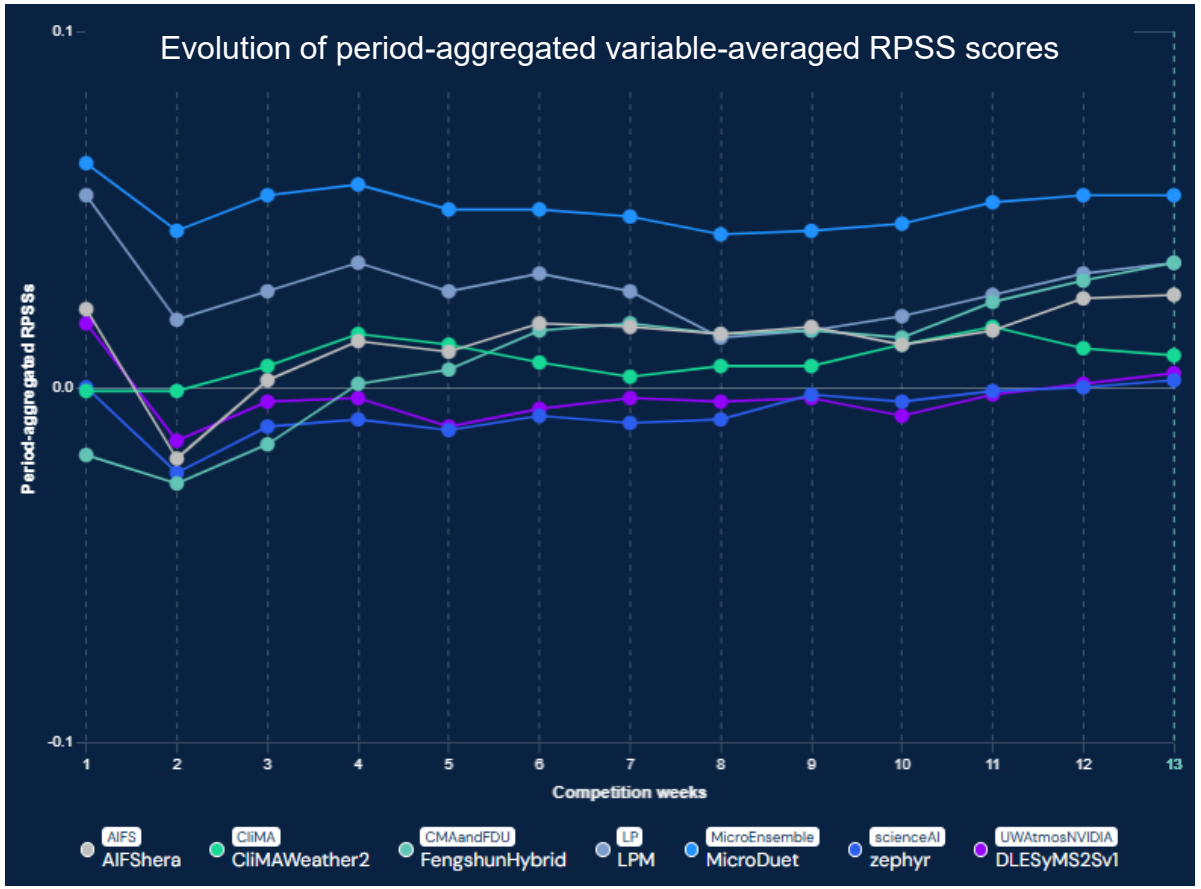


ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

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

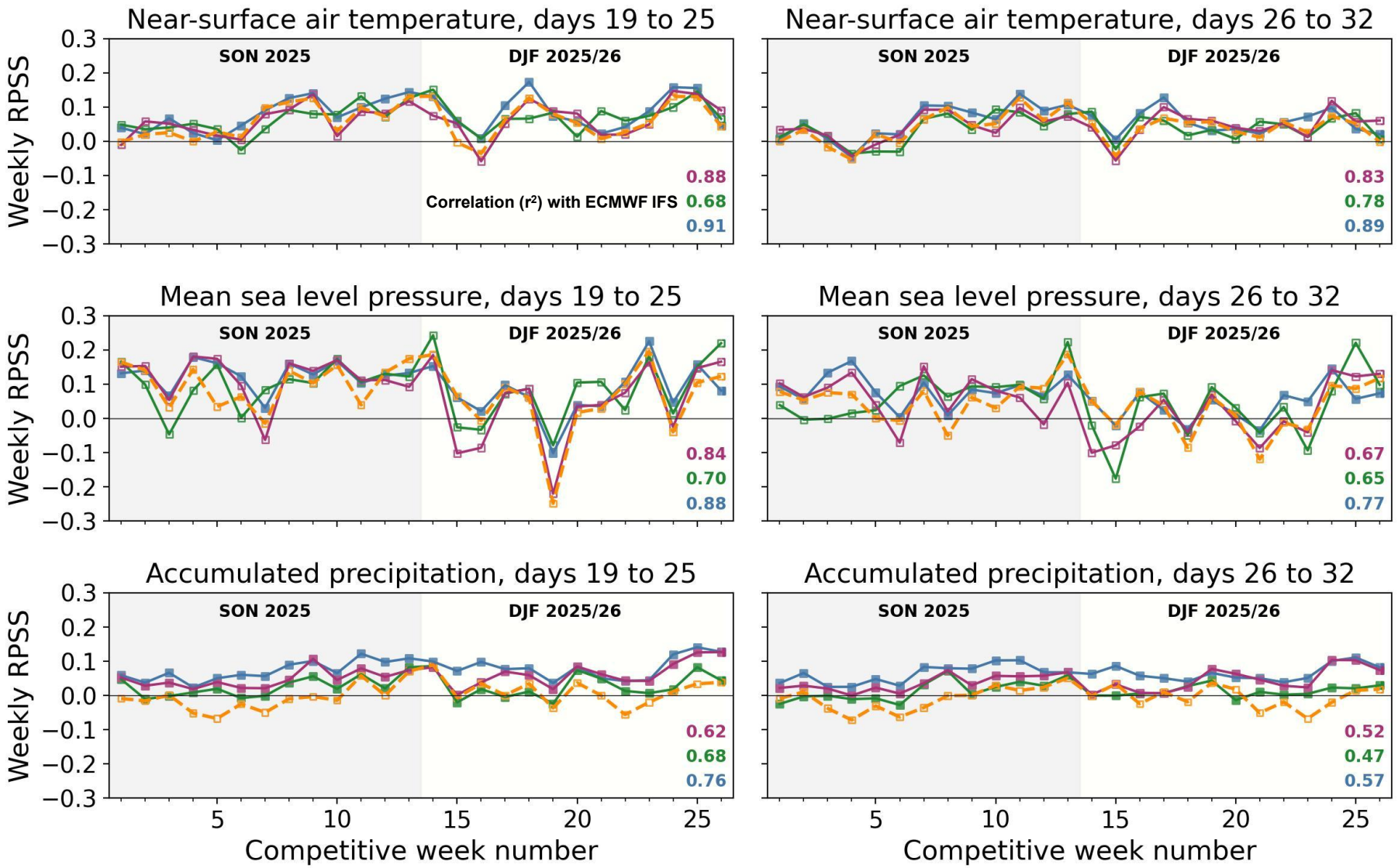
18 Teams eligible for variable-averaged, period-aggregated scores

7 Teams above climatology



ECMWF Destination Earth WORLD METEOROLOGICAL ORGANIZATION

DJF Period evaluation: First scientific insights



Best model in each category

MicroDuet	AIFShera	FengshunHybrid	ECMWF
Post-processing	**Data-driven**	**Hybrid**	**Dynamical**

Weekly RPSSs for best performing data-driven, post-processing, and hybrid model. Filled markers when significant across all weeks.

- For all variables and lead times, best post-processing model, significantly outperforms ECMWF IFS.
- Best data-driven and hybrid models only significantly outperforms ECMWF IFS for precipitation.
- Development still required for data-driven techniques to significantly outperform ECMWF IFS.

DJF Period evaluation: Northern Hemisphere Extratropical cold spell case studies

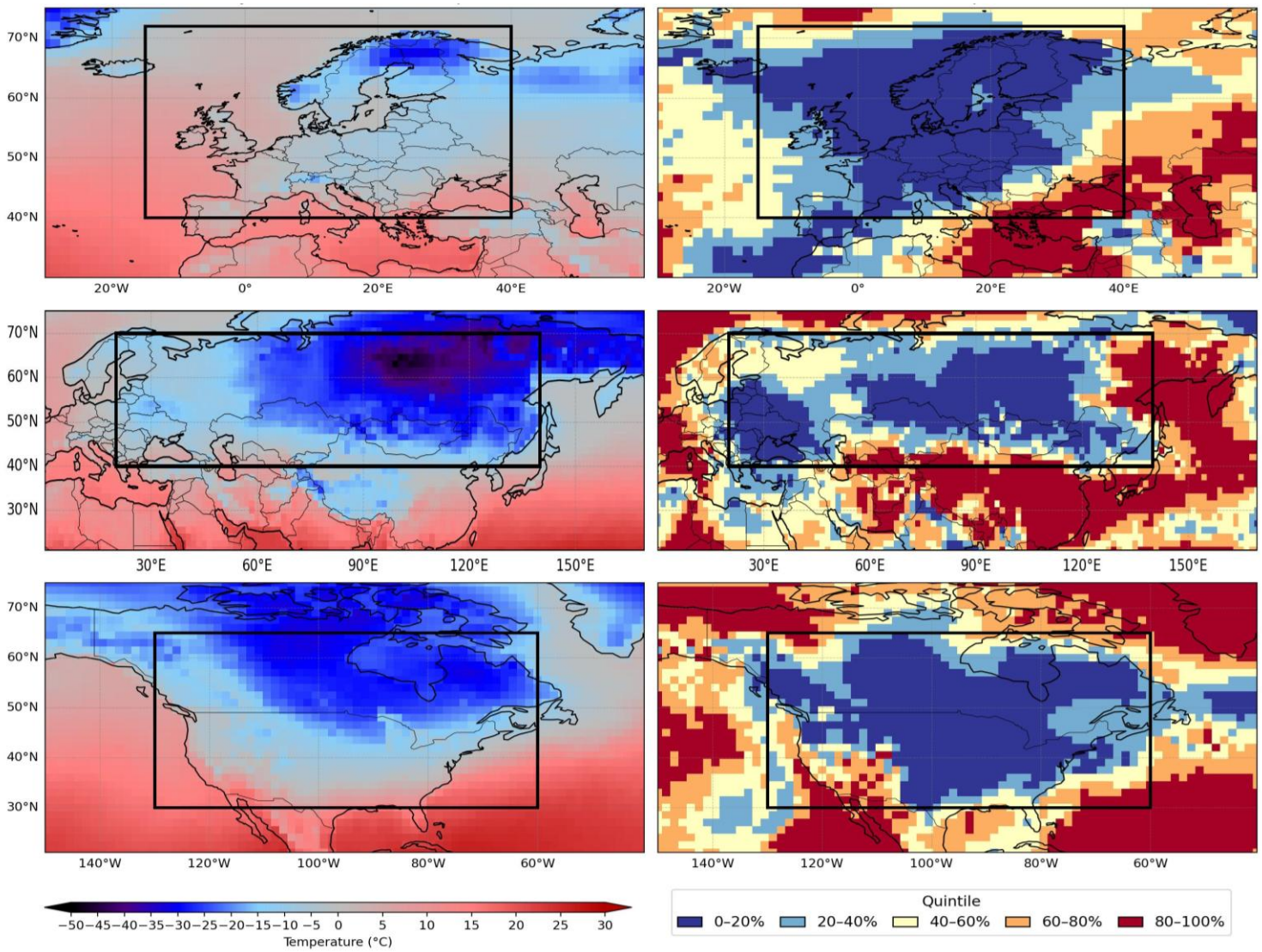
AI Weather Quest ECMWF

Absolute weekly-mean near-surface air temperature (ERA5T)	Quintile bin for observed temperature

5th to 11th January 2026
European Cold Spell

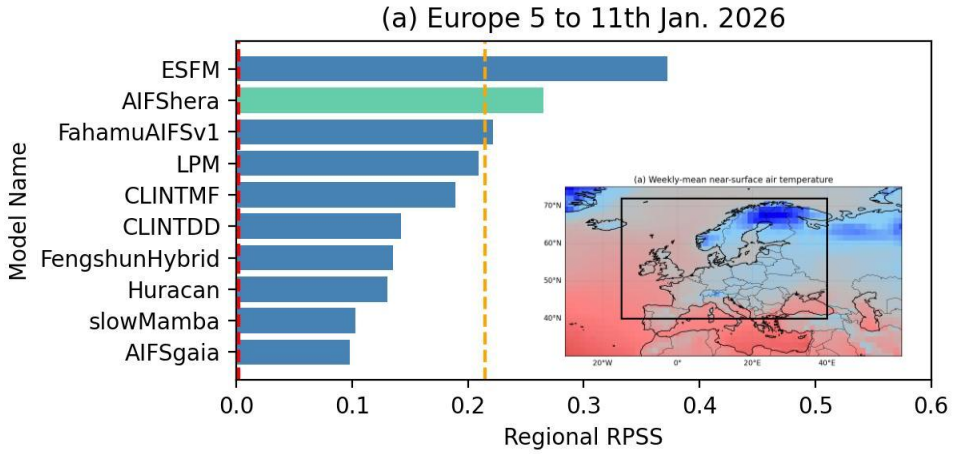
12th to 18th January 2026
Eurasia Cold Spell

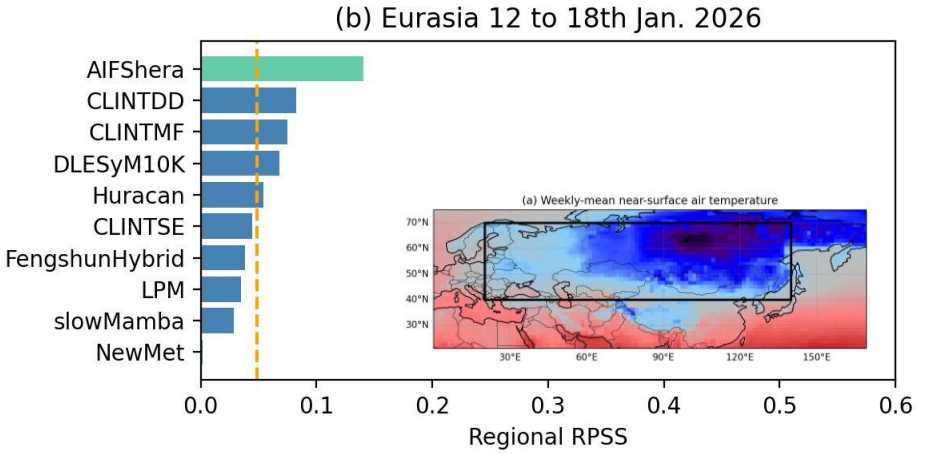
19th to 25th January 2026
North American Cold Spell

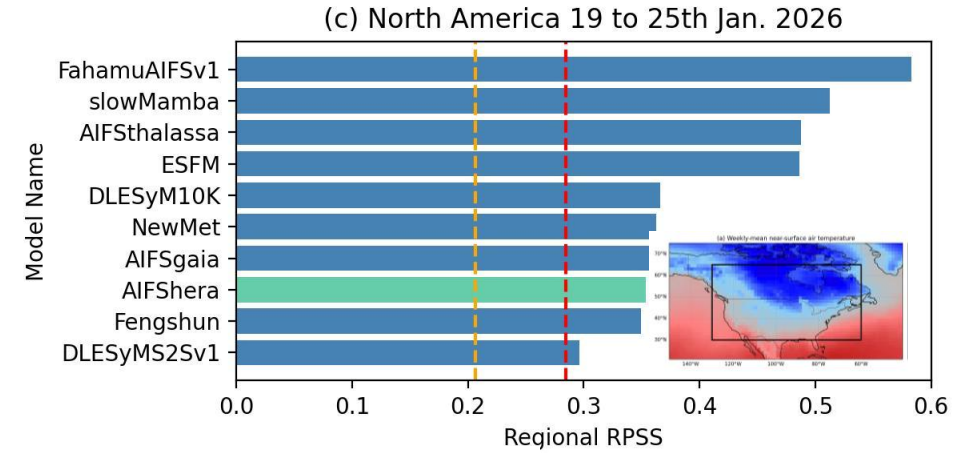


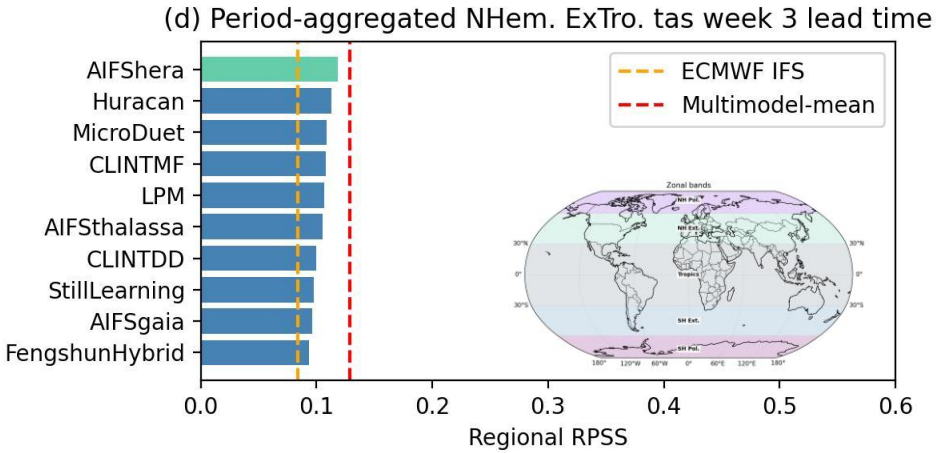
= Evaluated region

DJF Period evaluation: Northern hemisphere extratropical cold spell case studies









Regional case study near-surface air temperature RPSSs (top 10 models shown)

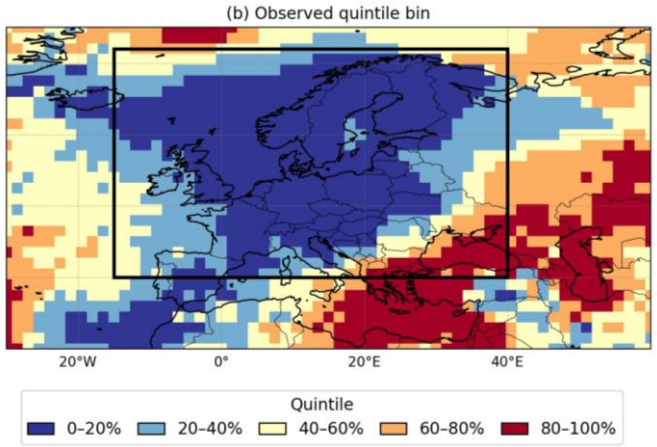
- AIFShera scoring relatively well for all three case studies.
- AIFShera also illustrates strong scores for N. Hemi. ExTro. RPSSs
- AIFS contributions are only best data-driven forecasting system.

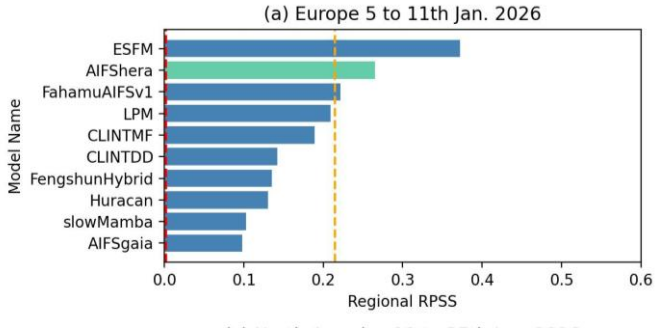
Data-driven spotlight

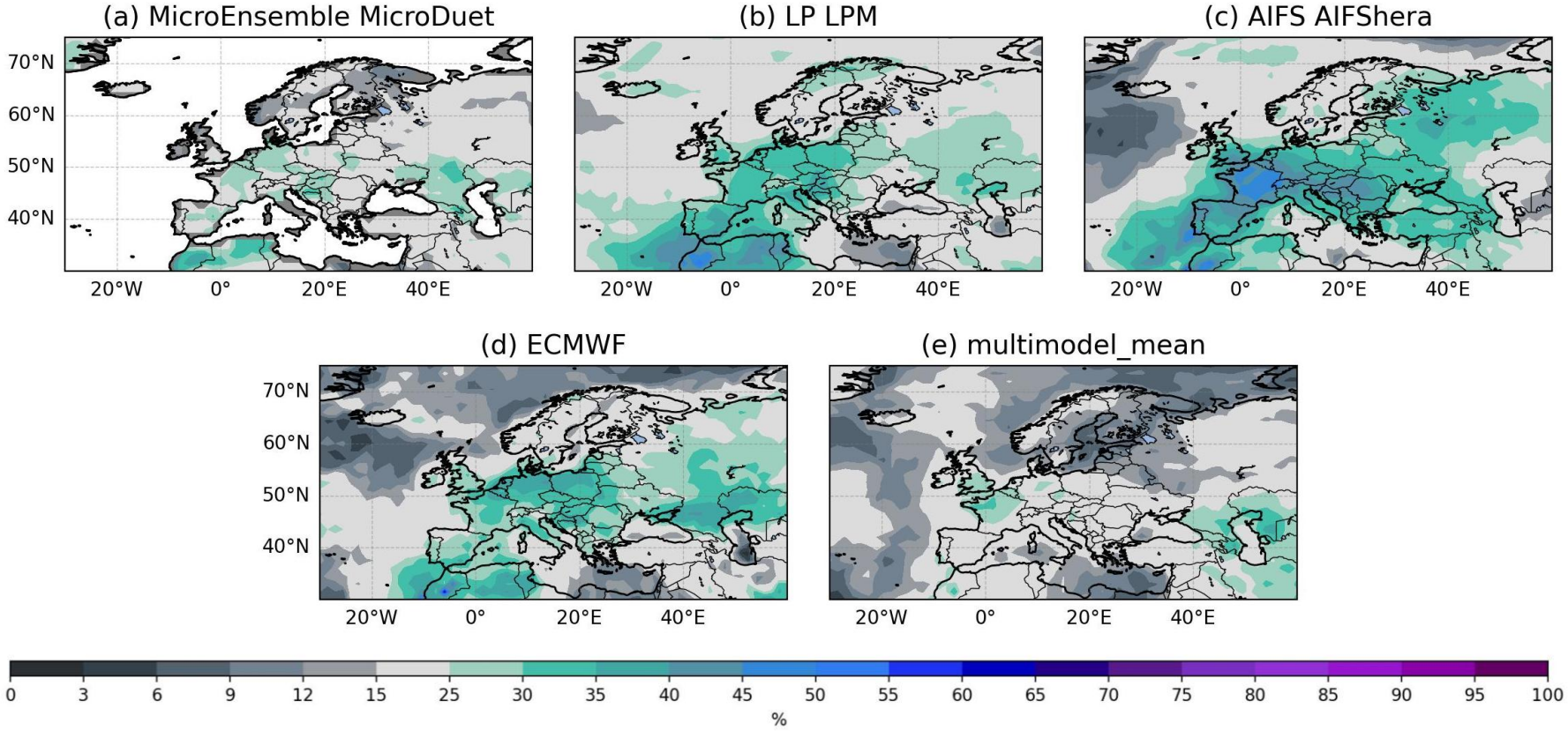
AIFS will be highlighted as the most skillful data-driven model, particularly for its strong performance in forecasting cold spells.

DJF Period evaluation: Northern hemisphere extratropical cold spell case studies

Examples of colder-than-normal quintile (< 20%) forecasts



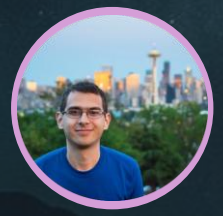




Presentation by team MicroEnsemble

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

MicroEnsemble



Jonathan Weyn

Microsoft

Hannah Guan

Harvard University

Soukayna Mouatadid

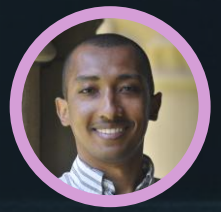
University of Toronto

Paulo Orenstein

Instituto de Matemática Pura e Aplicada

Judah Cohen

MIT



Lester Mackey

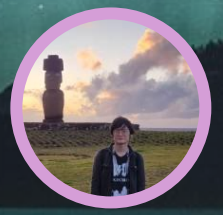
Microsoft Research

Alex Lu

Microsoft Research

Genevieve Flaspohler

Rhiza Research



Zekun Ni

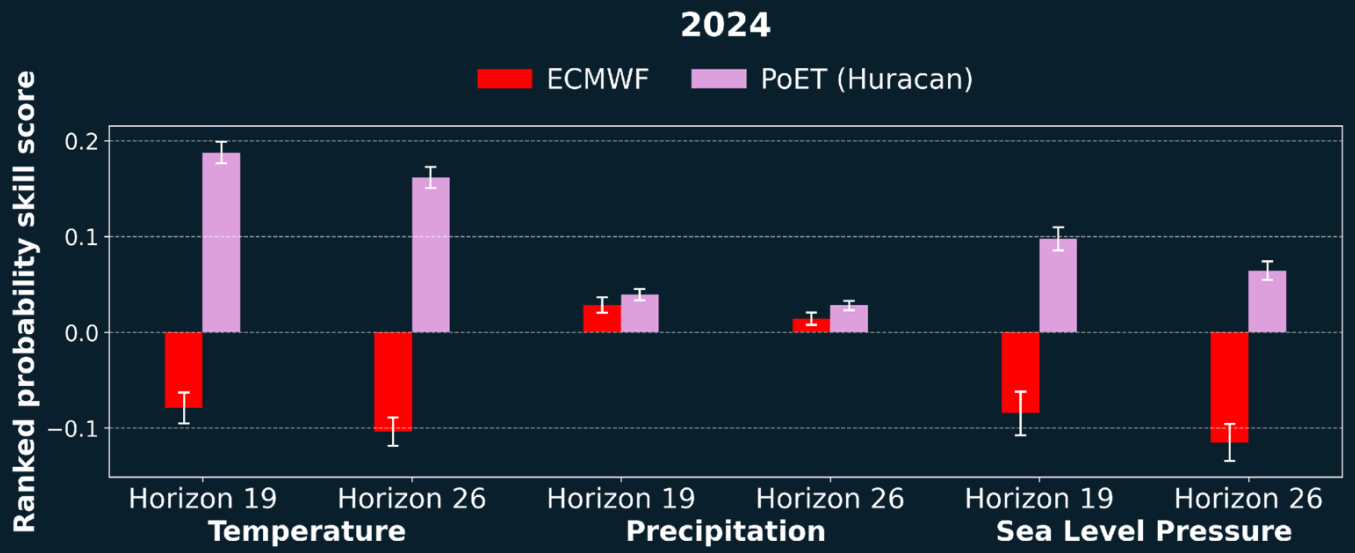
Microsoft

Haiyu Dong

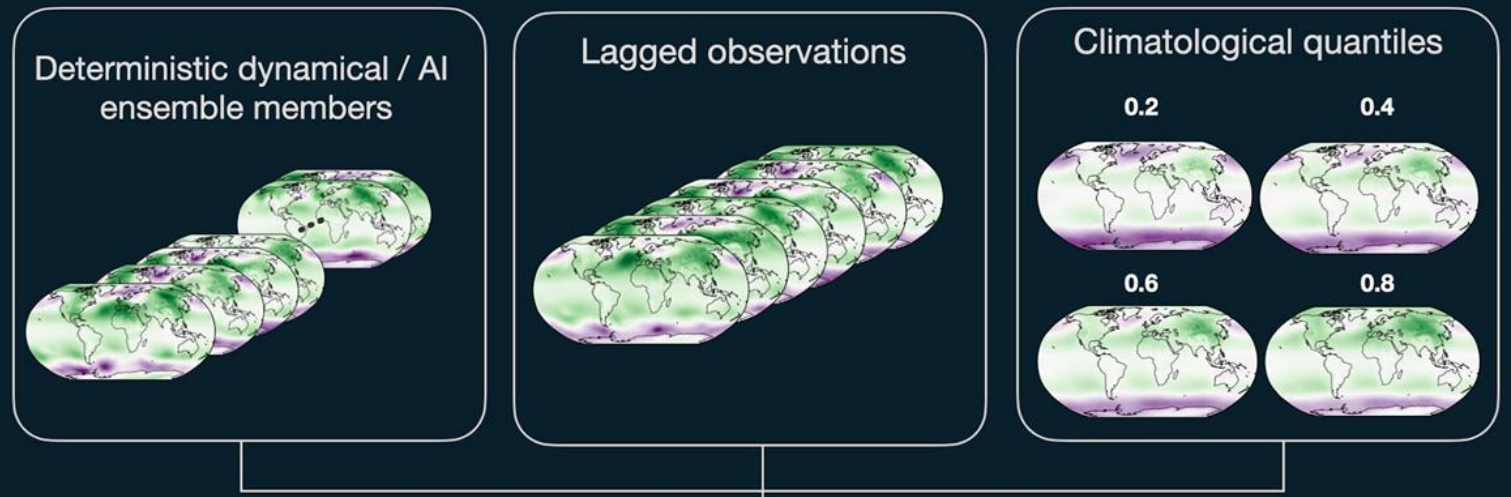
Microsoft

Huracan
Post-processing with ensemble transformer (PoET)

- Improving Medium-Range Ensemble Weather Forecasts with Hierarchical Ensemble Transformers (Bouallegue et al., 2024)
- Inputs consist of all 85 parameters available in S2S plus 37 prescribed static variables
- 10-member perturbed hindcasts for training
- The model is agnostic to ensemble size so inference is done on the 100-member operational ensemble



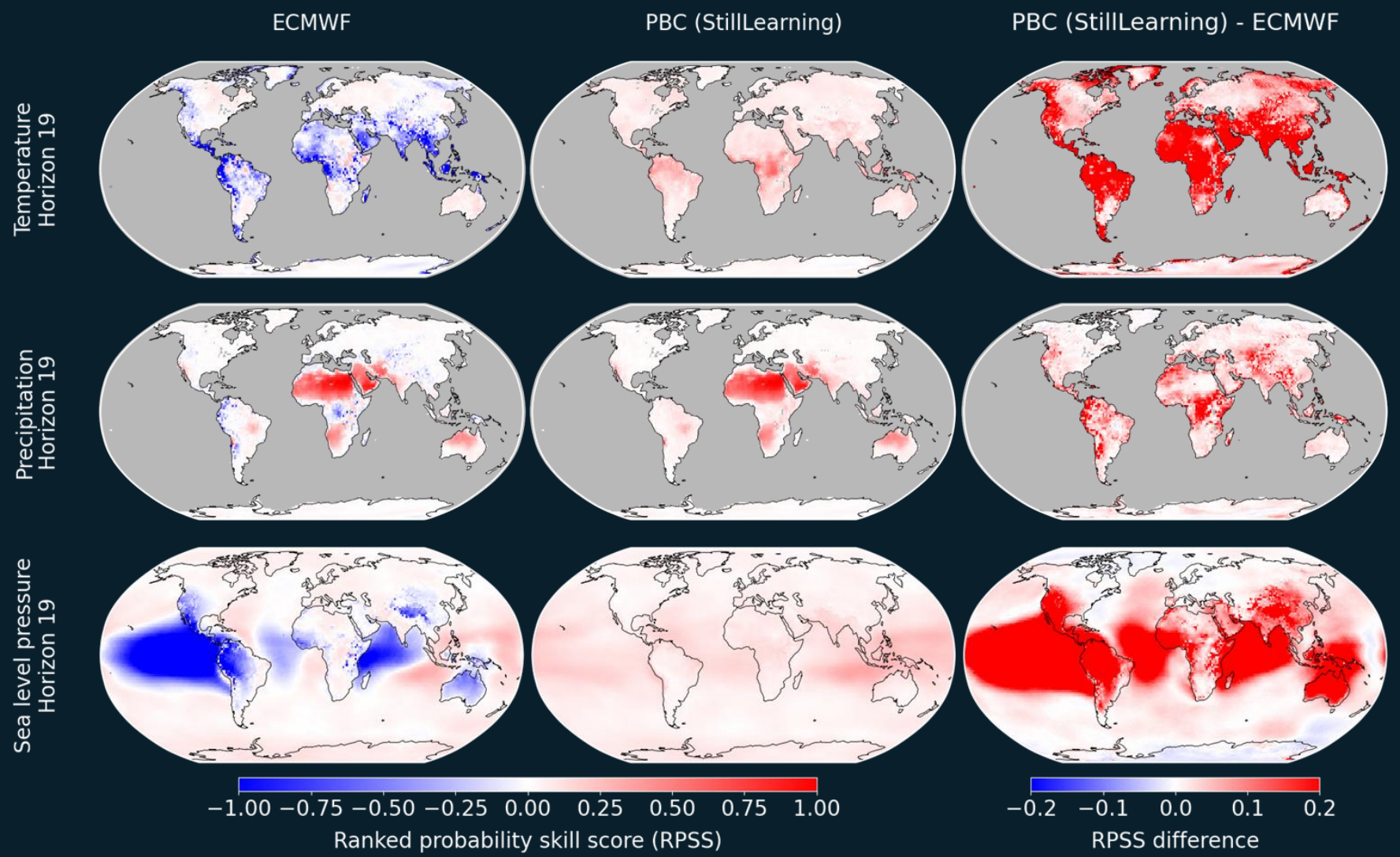
StillLearning
Probabilistic bias correction (PBC)



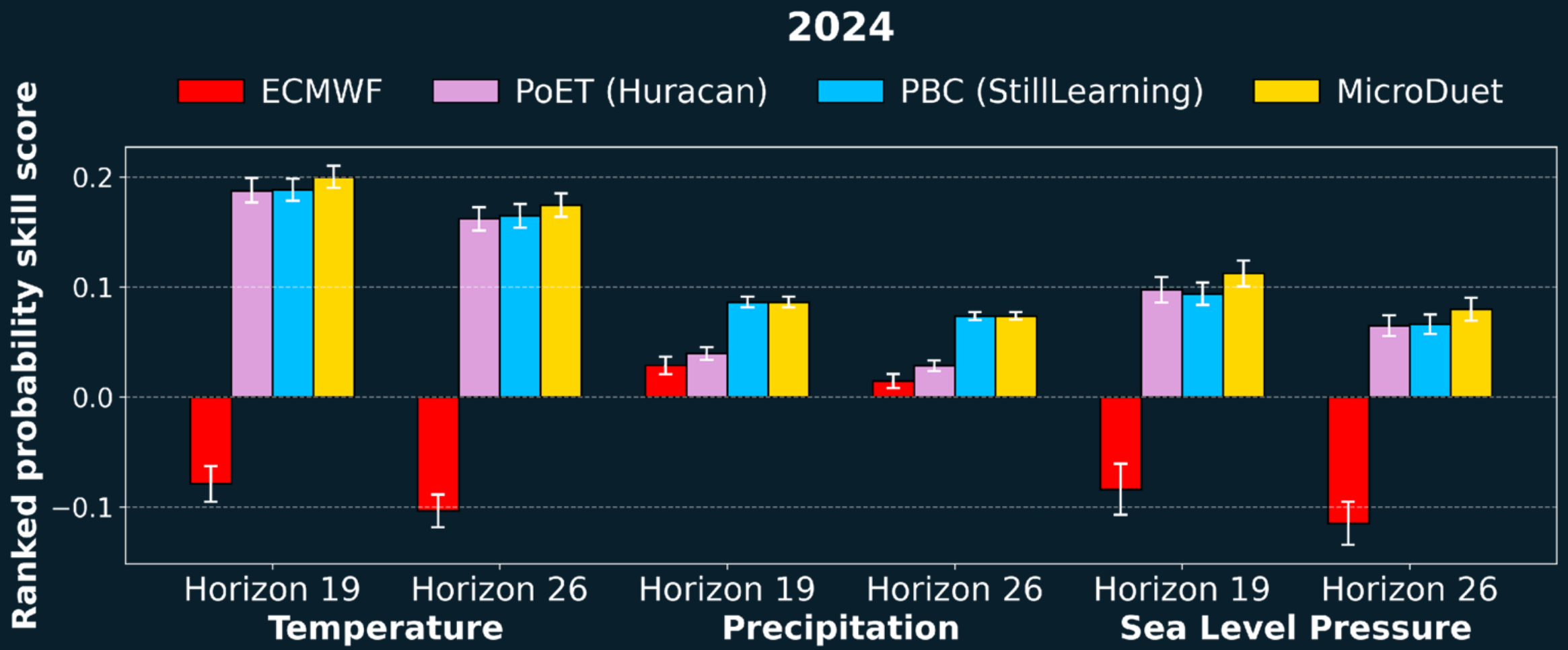
Inspired by:
Adaptive bias correction for improved subseasonal forecasting
(Mouatadid et al.,
Nature Communications, 2023)

StillLearning
Probabilistic bias correction (PBC)

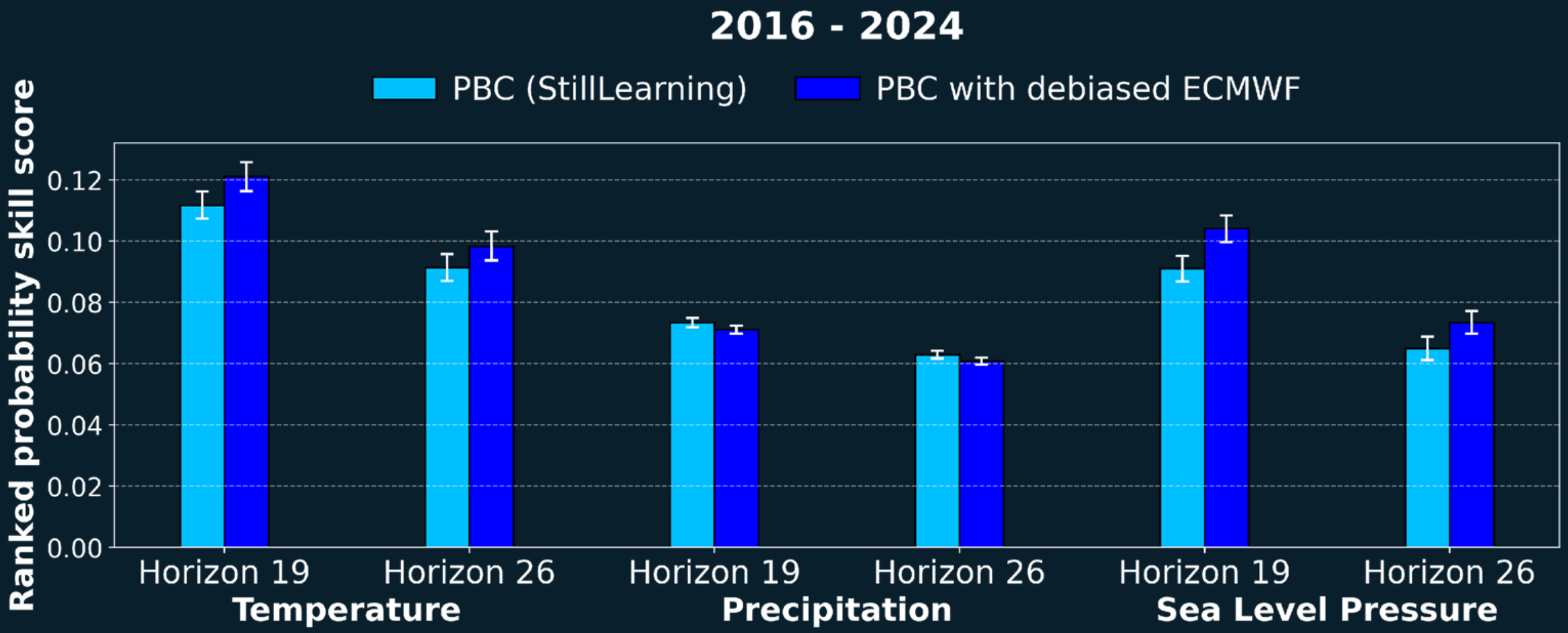
2016 – 2024



MicroDuet RPSS



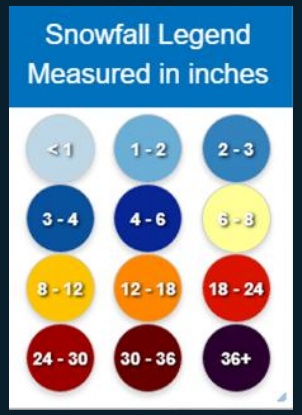
Ongoing Model Development

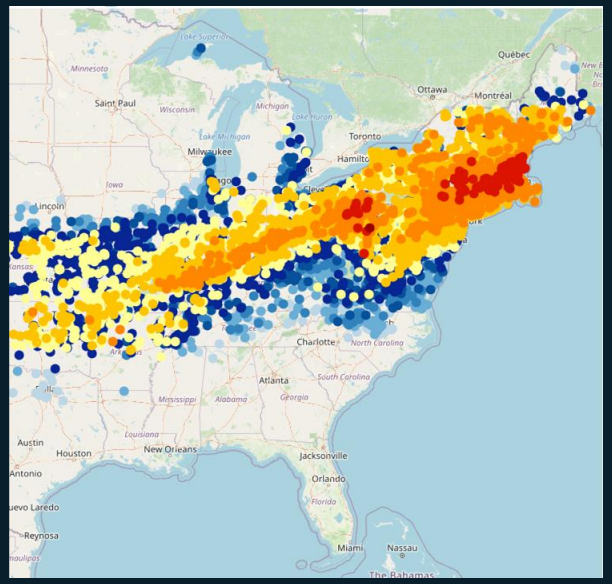


January 2026 Snowstorm

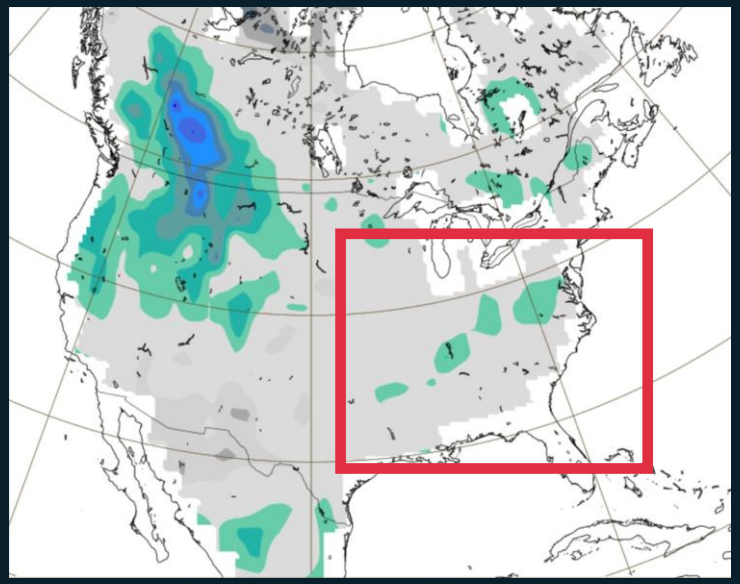
Mon 19 Jan 2026 - Sun 25 Jan 2026 — Quintile interval : >= 80%

Observed

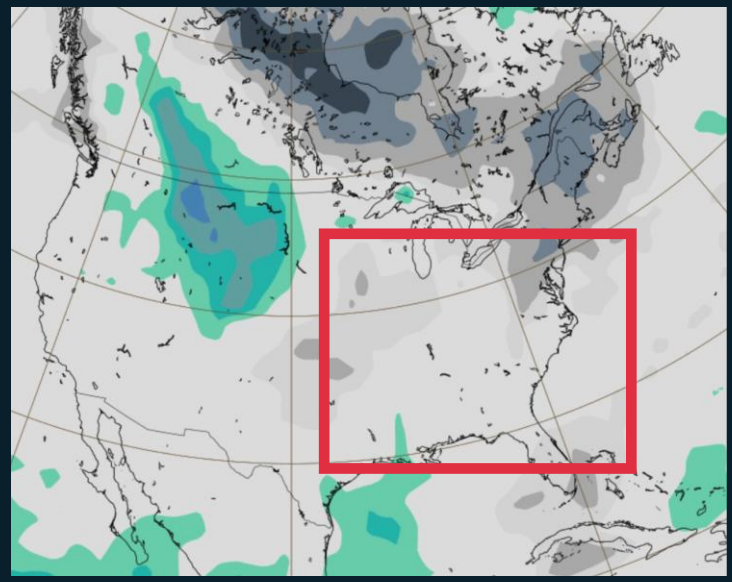




MicroDuet



AIFSgaia



Accumulated precipitation quintile probabilities

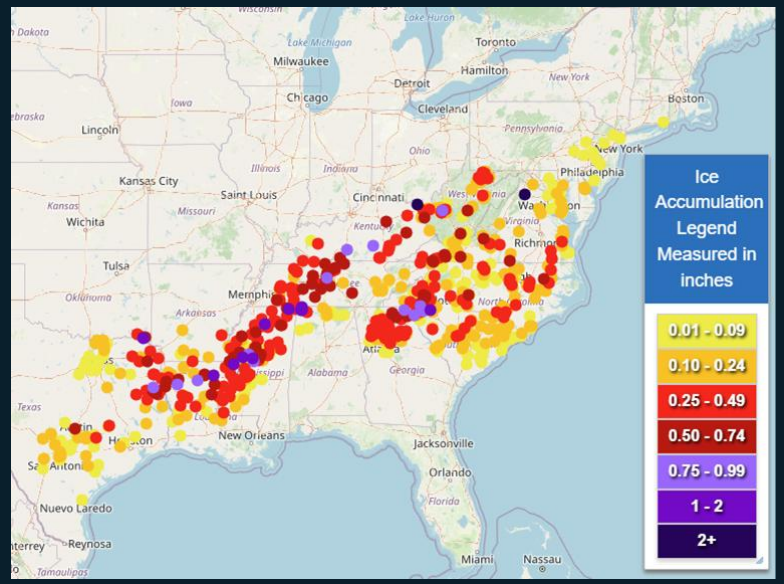
0 3 6 9 12 15 25 30 35 40 45 50 55 60 65 70 75 80 85 90 0.95 100%

Source: www.ecmwf.int

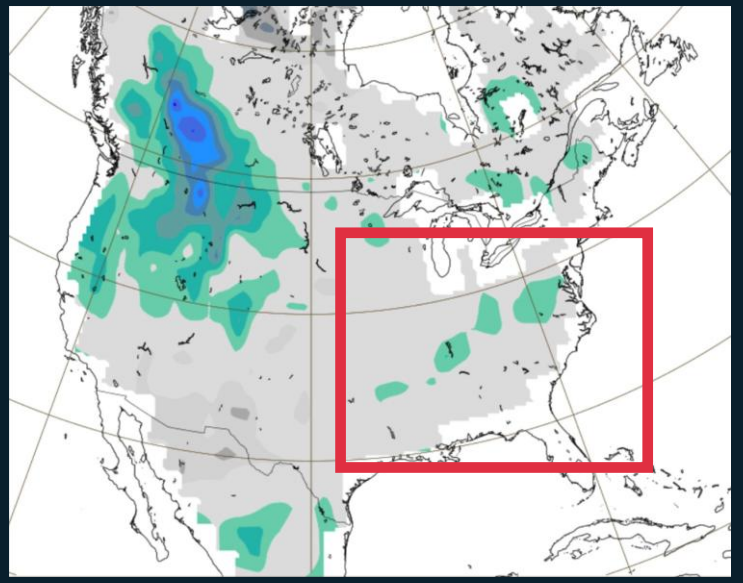
January 2026 Snowstorm

Mon 19 Jan 2026 - Sun 25 Jan 2026 — Quintile interval : >= 80%

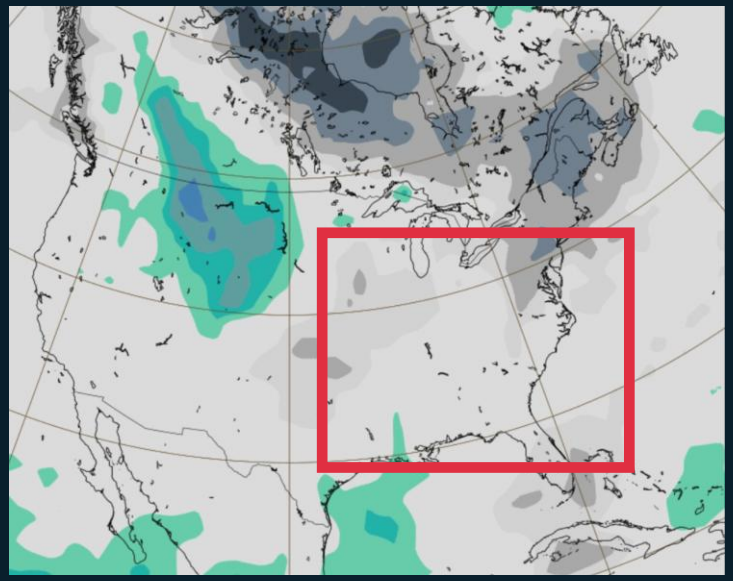
Observed



MicroDuet



AIFSgaia



Accumulated precipitation quintile probabilities

0 3 6 9 12 15 25 30 35 40 45 50 55 60 65 70 75 80 85 90 0.95 100%

Source: www.ecmwf.int

江苏省气候中心
Jiangsu Climate Center

AI Weather Quest ECMWF

Presentation by team LP

Second / third best ranked-team of the DJF Period for variable-averaged, period-aggregated scores for 1st / 2nd forecast windows

ECMWF

Destination Earth

WORLD METEOROLOGICAL ORGANIZATION

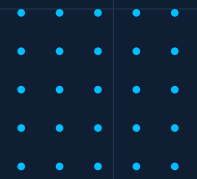
AI WEATHER QUEST

LPM Model for DJF Period Performance

Team | LP Team

Presenter | Lu Peng

Affiliation | Jiangsu Climate Center



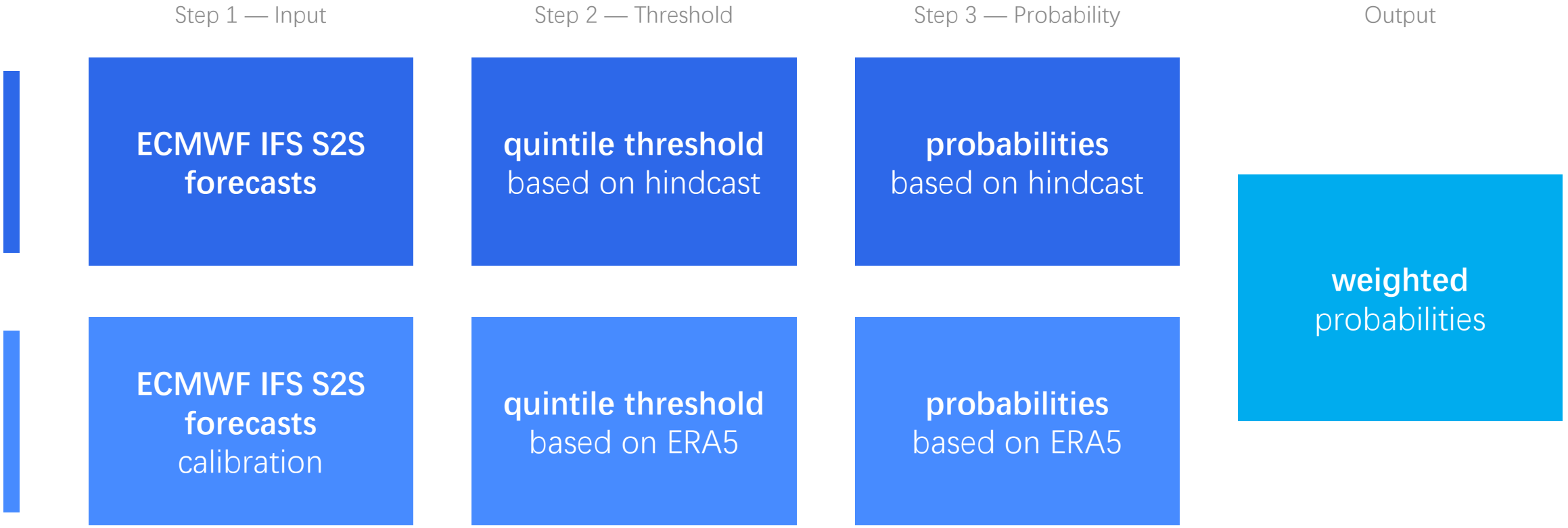
- LP Team: Lu Peng
- Acknowledgements
- Jiangsu Climate Center
- ECMWF AWQ organizers(Olga and Joshua)
- SpeedyWeather.jl (Milan and Maximilian)

江苏省气候中心
Jiangsu Climate Center

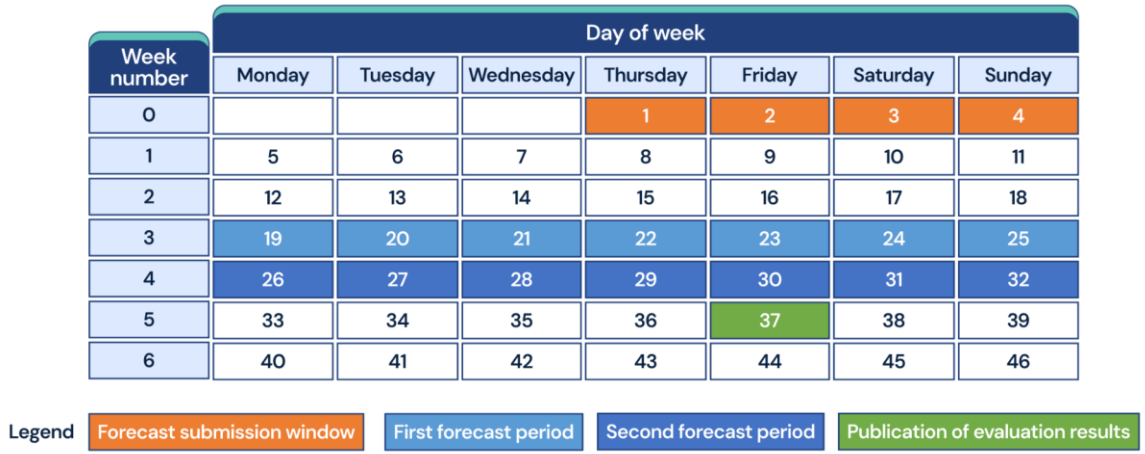
AI Weather Quest ECMWF

SpeedyWeather.jl

Forecast Pipeline



post-processing in pipeline B



Week number	Day of week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
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

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

Recent Systematic Bias (each grid)

Bias = mean(forecast_0 - obs)

Forecast_corrected = forecast - Bias

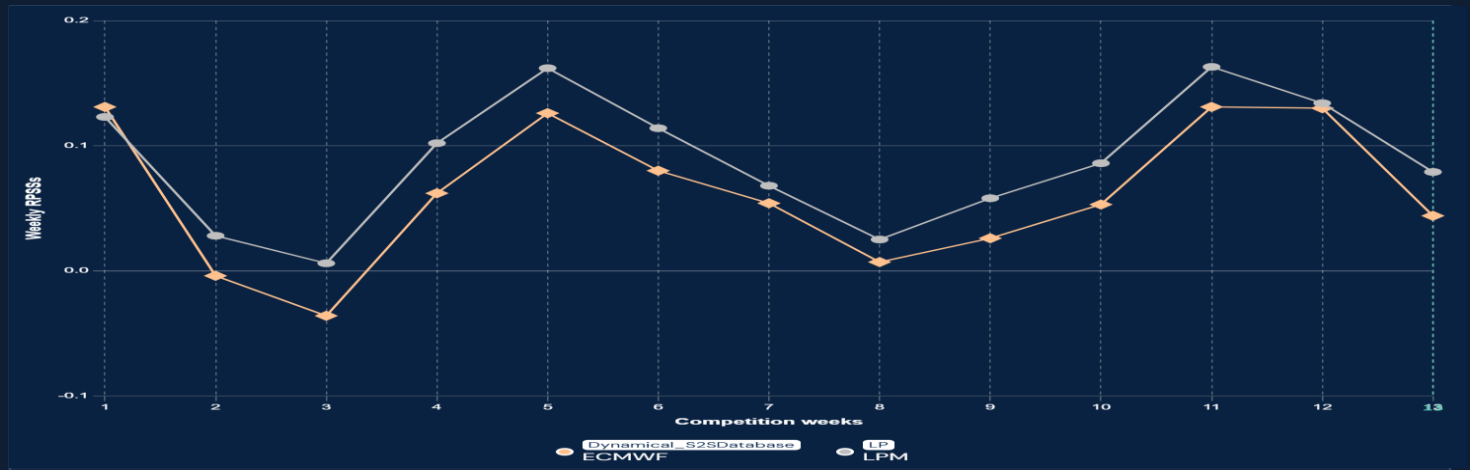
post-processing	Tas	MSLP	Pr
windows	Week [-4,-1]	Week [-3,-1]	No post-processing

TAS Performance: Model Comparison (RPSS)

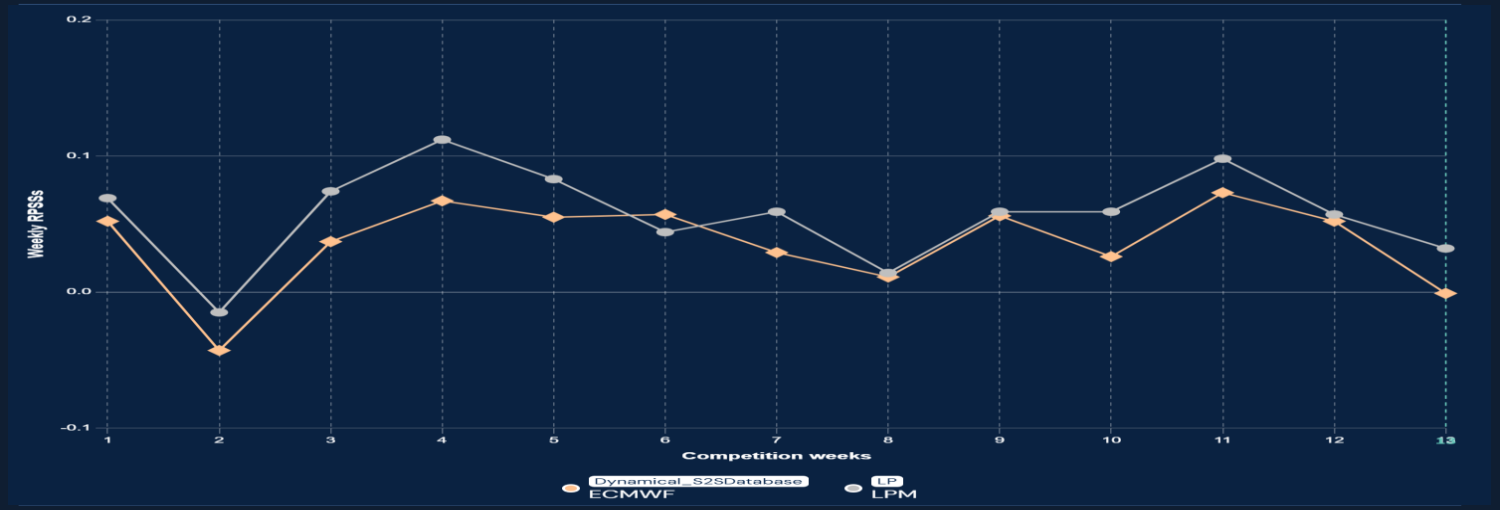
	First Window	Second Window
LPM	0.089	0.057
ECMWF IFS S2S	0.062	0.036

- LPM consistently outperforms ECMWF IFS S2S in both windows.

Weekly RPSS - First Window



Weekly RPSS - Second Window

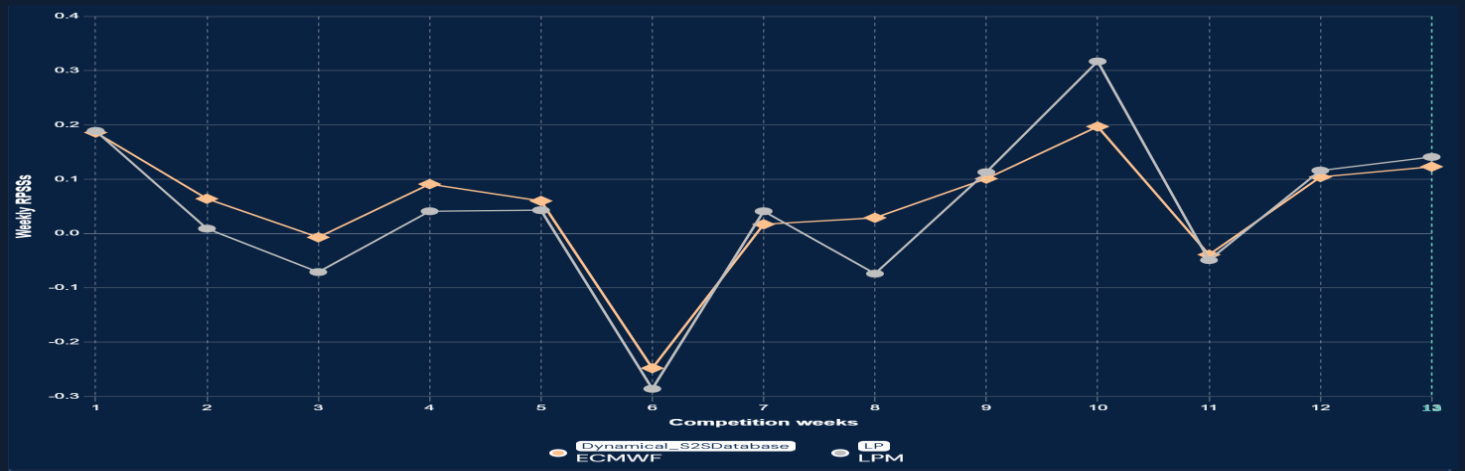


MSLP Performance: Model Comparison (RPSS)

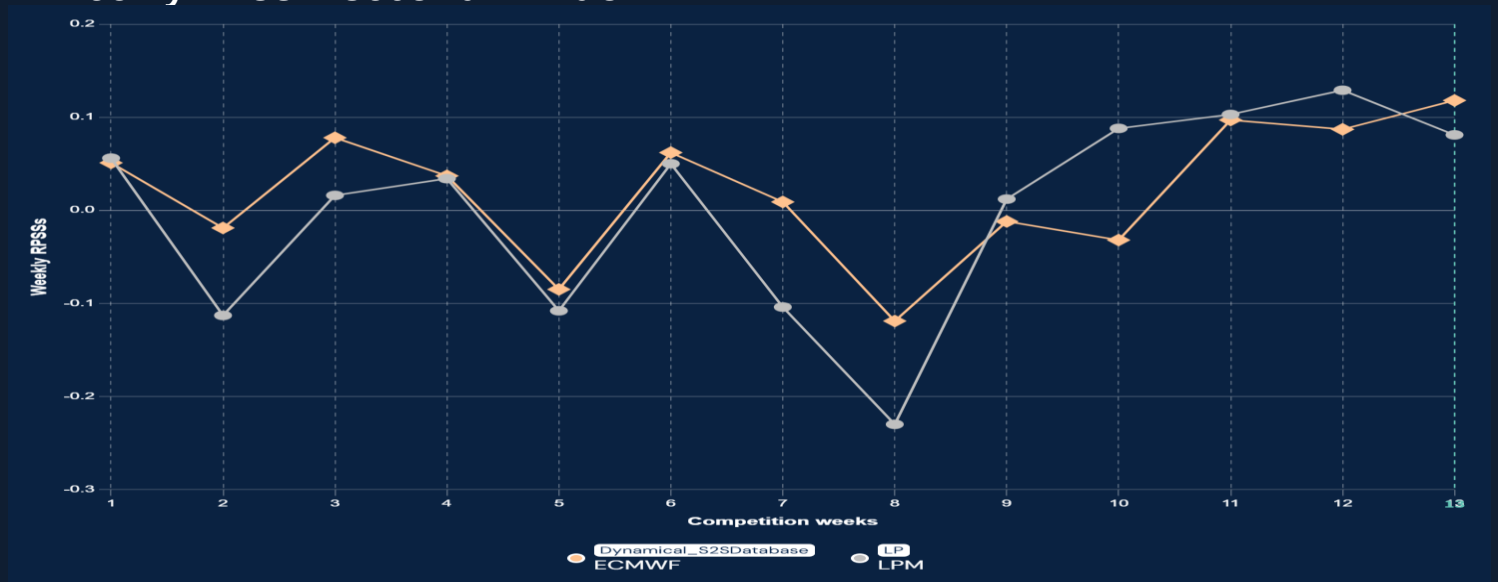
	First Window	Second Window
LPM	0.047	0.001
ECMWF IFS S2S	0.056	0.020

- ECMWF IFS S2S model outperforms LPM in both forecast windows.

Weekly RPSS - First Window



Weekly RPSS - Second Window

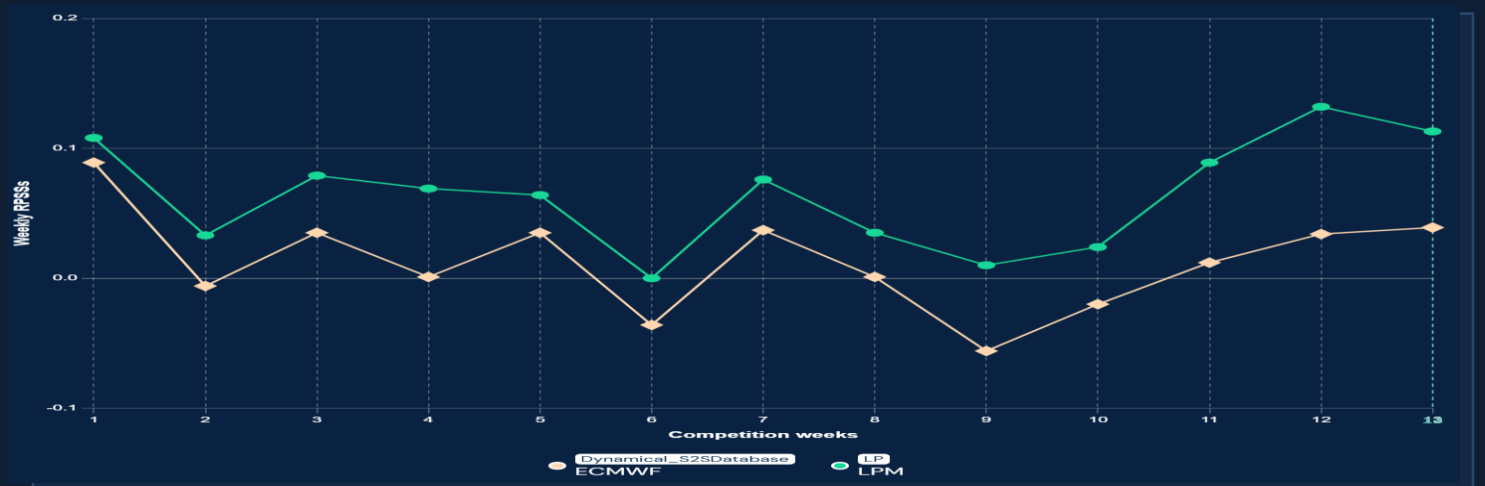


PR Performance: Model Comparison (RPSS)

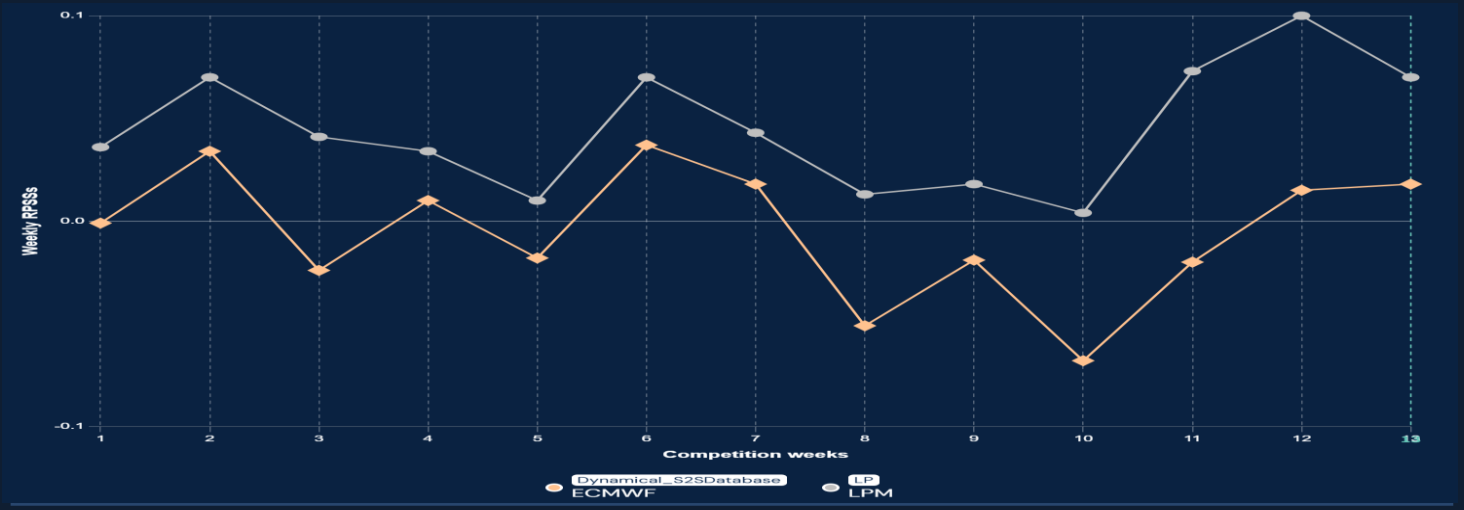
	First Window	Second Window
LPM	0.065	0.046
ECMWF IFS S2S	0.013	-0.005

- LPM clear outperforms ECMWF IFS S2S for rain.
- Simple method: no training, no extra parameters.
- Not SOTA But new baseline

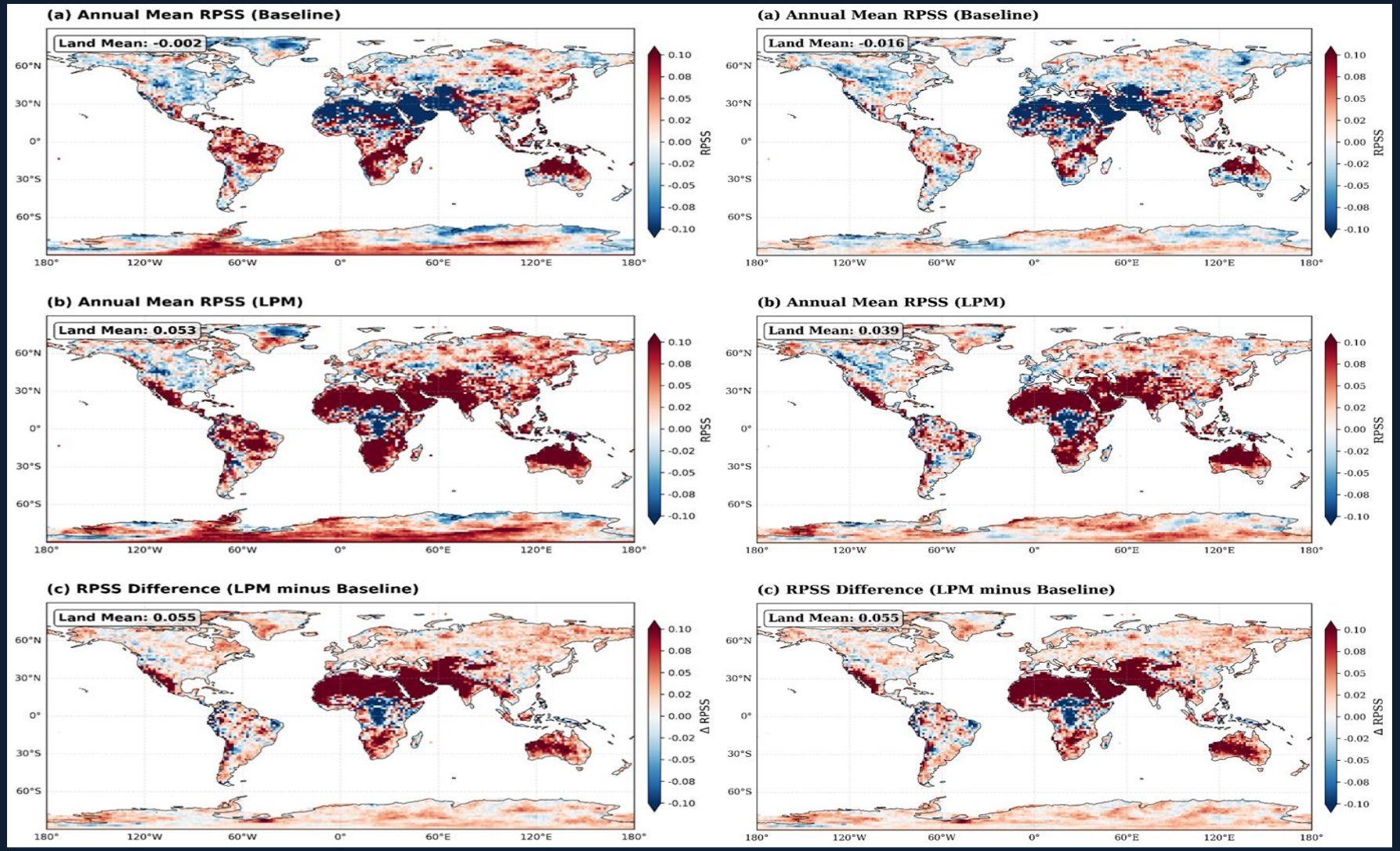
Weekly RPSS - First Window



Weekly RPSS - Second Window



Global Skill Improvement: LPM vs. Baseline



CORE RESULT

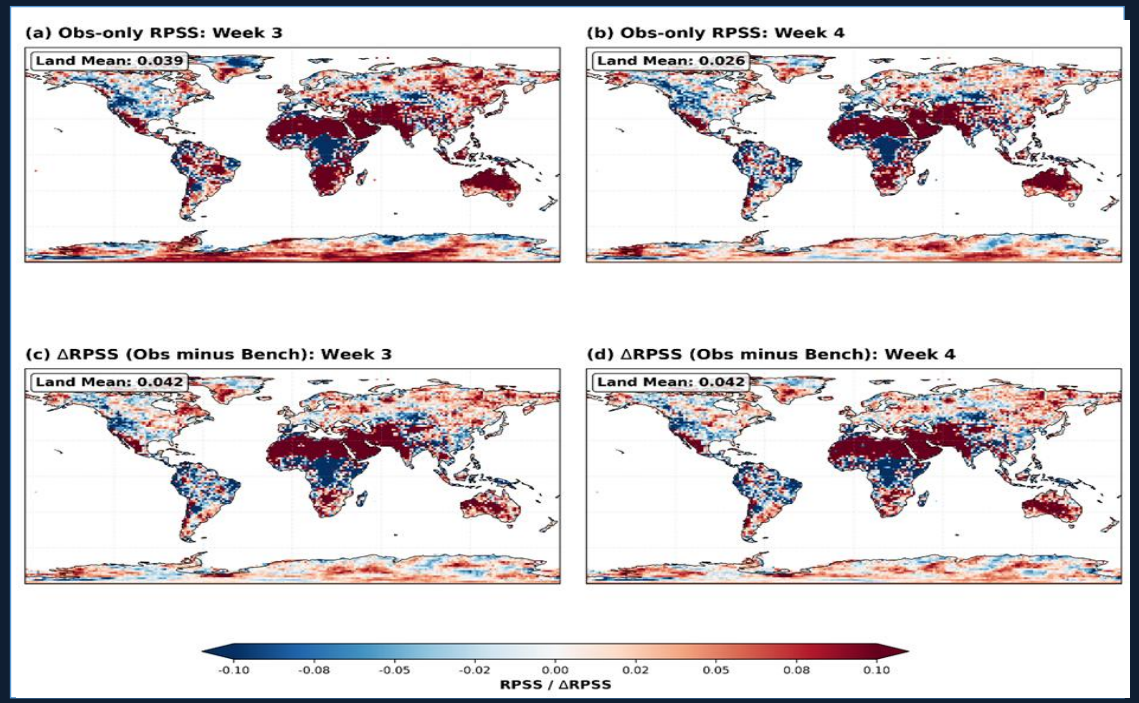
- LPM > Baseline (Pipeline A)
- Global Land Mean: +0.055

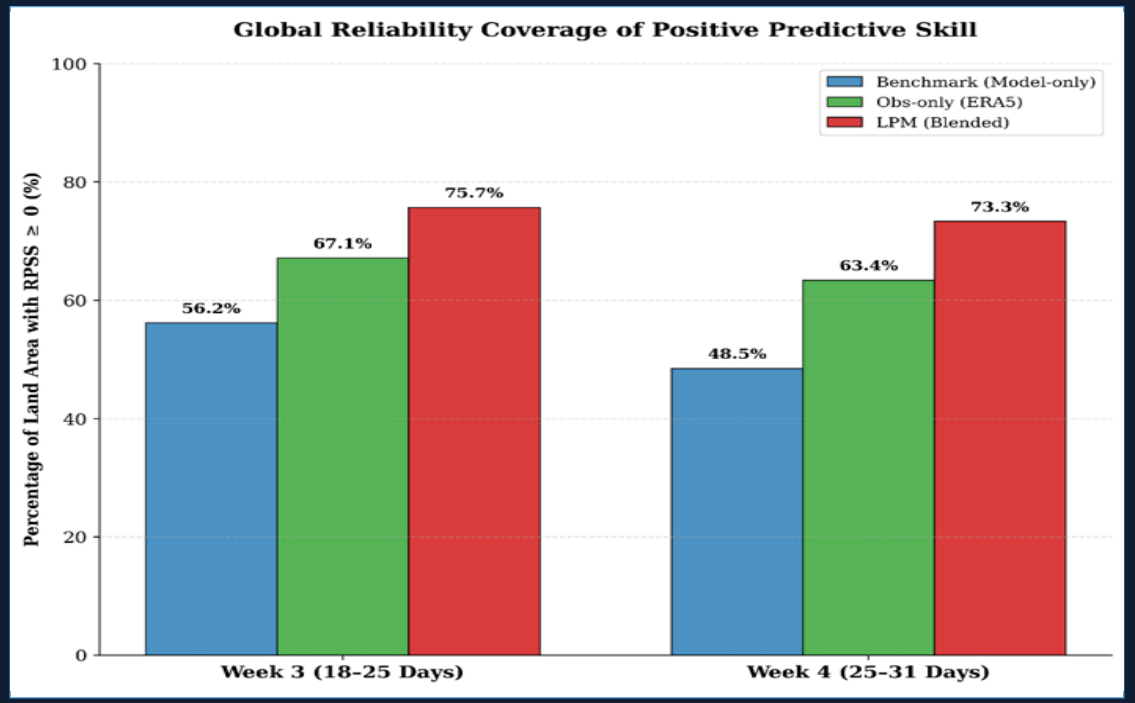
REGIONAL IMPROVEMENT

Large skill gain in:

- Tropical Africa
- Middle East & South Asia
- Australia

Pipeline A, B and A+B in PR





Pipeline B significantly outperforms Pipeline A (Baseline) for precipitation.

Blending A and B is superior to using Pipeline A or Pipeline B alone with the highest percentage of positive RPSS (RPSS > 0) globally

Future Work & Challenges

Stage I: Data-Driven AI Model
Pangu / GraphCast/ Fuxi / AIFS/ Fengshun

Stage II: Hybrid Model
NeuralGCM

Stage III: Differentiable Atmospheric Model
CliMA / SpeedyWeather (Enzyme.jl)

1. Time & Memory

- Small time steps (CFL)
- Many steps use too much memory.

2. Software Stability

- Enzyme.jl still developing
- Bugs in complex models

Note:

Work is still at an early stage
Not used in this competition

Thank You!

For questions or discussion,
please feel free to contact me by email:

670123489@qq.com

ECMWF

AI Weather Quest ECMWF

Presentation by team AIFS

Spotlighted as the team with the best ranked purely data-driven model, also scoring well on the featured Northern hemisphere extratropical cold spell case studies

ECMWF

Destination Earth

WORLD METEOROLOGICAL ORGANIZATION

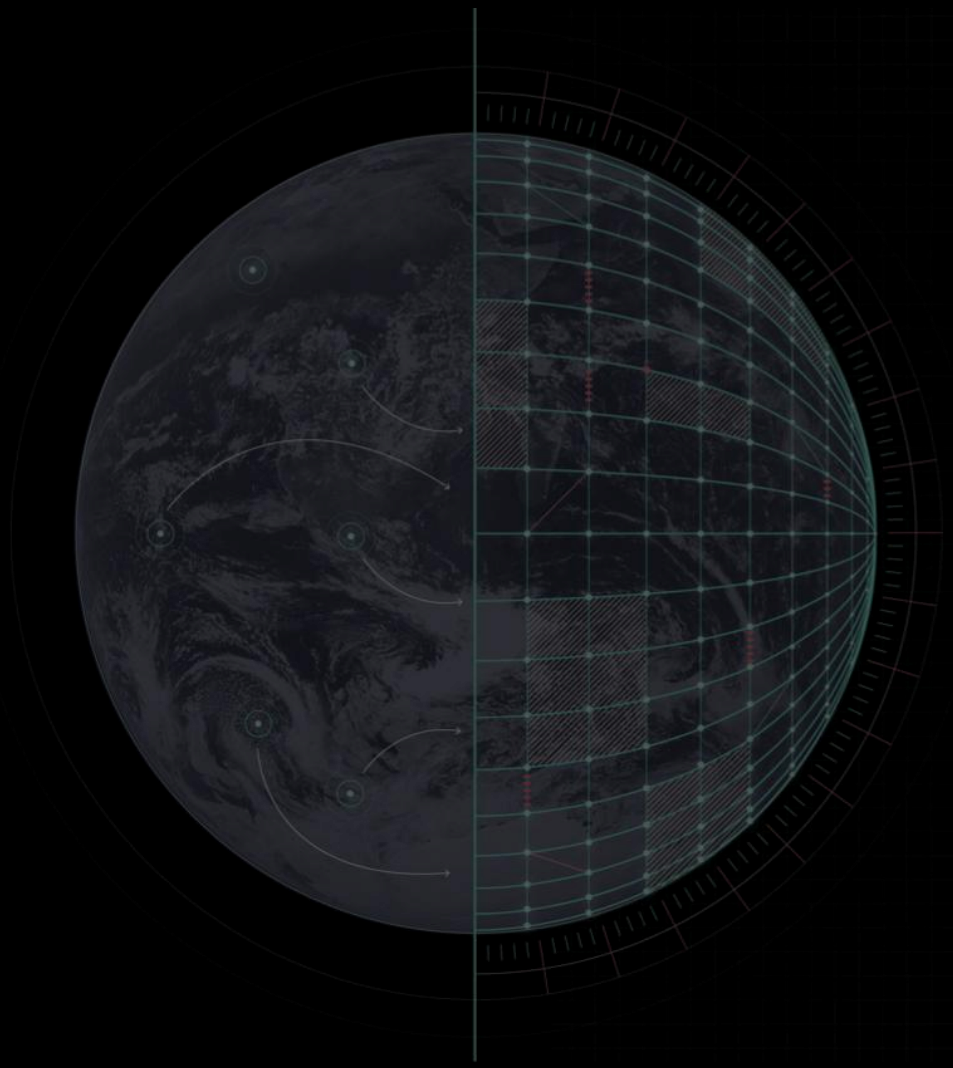
AI Weather Quest: DJF Webinar

AIFS-Team

Jakob Schloer

and many colleagues across ECMWF

jakob.schloer@ecmwf.int



ECMWF

Funded by the European Union

Destination Earth

AIFS-Team for the AI Weather Quest





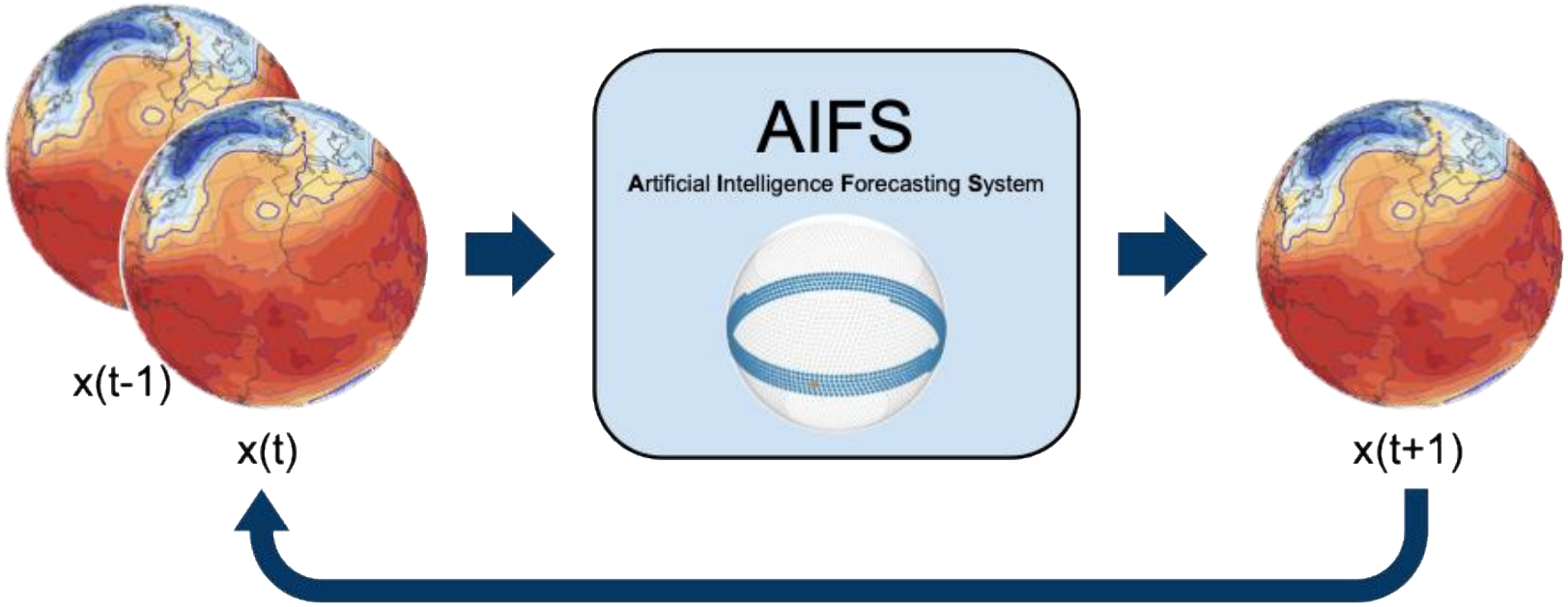
Sara Hahner, Lorenzo Zampieri, Rachel Furner, Mariana Clare, Gareth Jones




Chris Roberts, Simon Lang, Steffen Tietsche, Jakob Schloer, and many colleagues across ECMWF

AIFS: Data-driven Weather Forecasting Systems

Artificial Intelligence Forecasting System



$$\hat{x}(t+1) = x(t) + f_\Theta(x(t), x(t-1))$$

AIFS is a **fully data-driven model** trained on ERA5!

Lang et al. (2024a)

AIFS: Data-driven Weather Forecasting Systems

AIFS-CRPS architecture

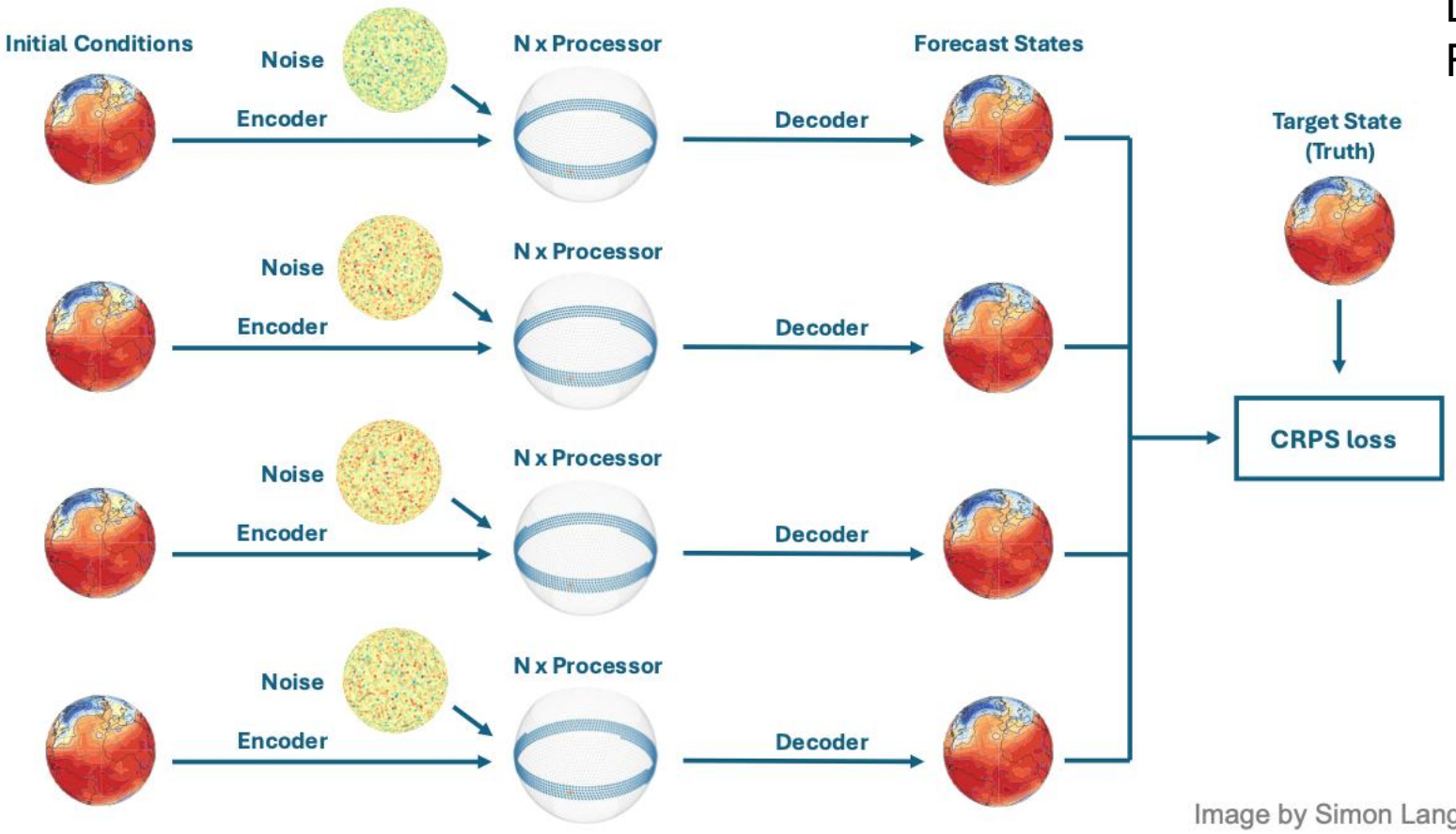
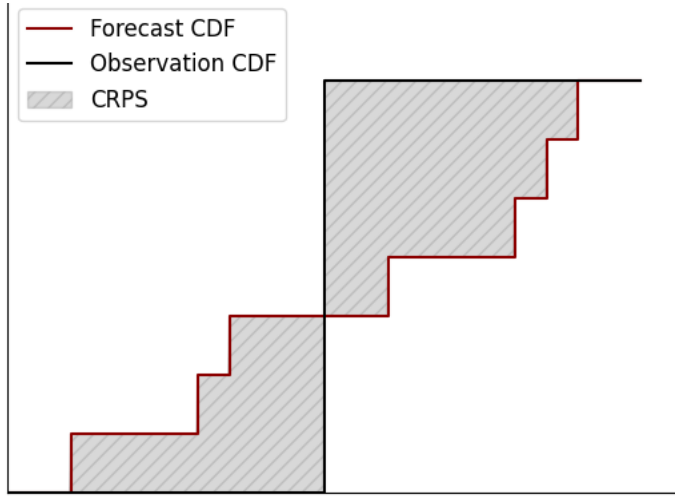


Image by Simon Lang

Loss-function (almost-fair Continuous Ranked Probability Score):

$$\text{afCRPS}_\alpha := \alpha\,\text{fCRPS} + (1-\alpha)\text{CRPS}$$



Artificial Intelligence Forecasting System

Our models

AIFS-hera

medium-range model

- Troposphere + Surface variables
- 6h time-step
- 12-step rollout-finetuning

Lang et al. AIFS-CRPS (2024b)

AIFS-hera

Model design choices

1. Data selection	• ERA5 at O96 (~1°) resolution • Global variables at surface and pressure levels
2. Learning objective	• Autoregressive forecasting with 6h time-step • 4 ensemble members
3. Loss functions	Almost-fair CRPS
4. Architecture	GraphTransformer with Encoder-Processor-Decoder structure
5. Training	• Pre-training: 1-step prediction • Fine-tuning: 2-step rollout + step-wise increasing rollout up to 12-steps

Artificial Intelligence Forecasting System

Our models

AIFS-hera

medium-range model

- Troposphere + Surface variables
- 6h time-step
- 12-step rollout-finetuning

Lang et al. AIFS-CRPS (2024b)

AIFS-gaia

longer time-step model

- Troposphere + surface variables + stratosphere
- 24h time-step
- 3-step rollout fine-tuning
- Revised variable weighting

AIFS-thalassa

model with surface ocean

- Troposphere + surface variables + surface ocean + sea ice
- 24h time-step
- 3-step rollout fine-tuning

All versions use 3 day rollout (time-step x number of steps)

AI Weather Quest

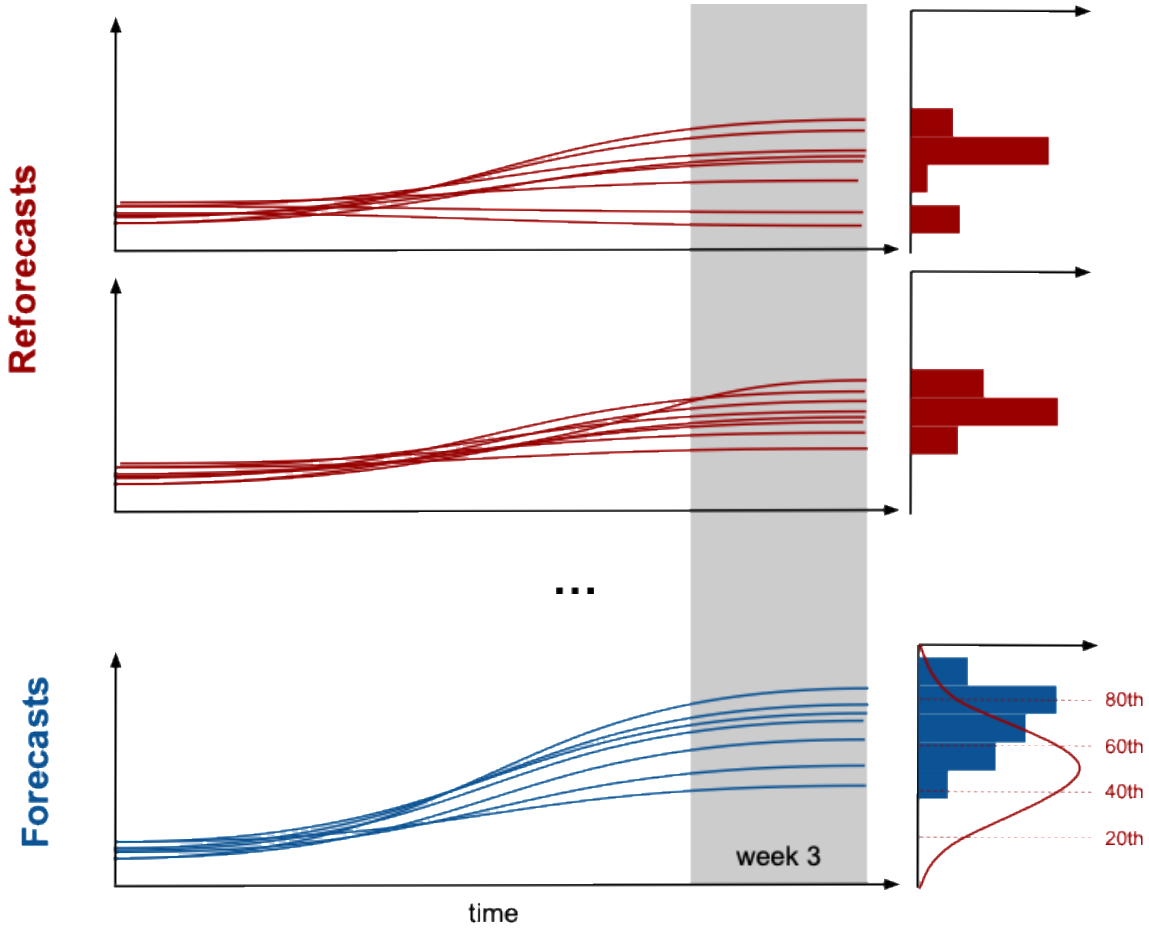
Forecast submission

Real-time forecast:

- 200-member ensemble
- Initial conditions from ERA5T
- Quintile probabilities for week 3 and 4

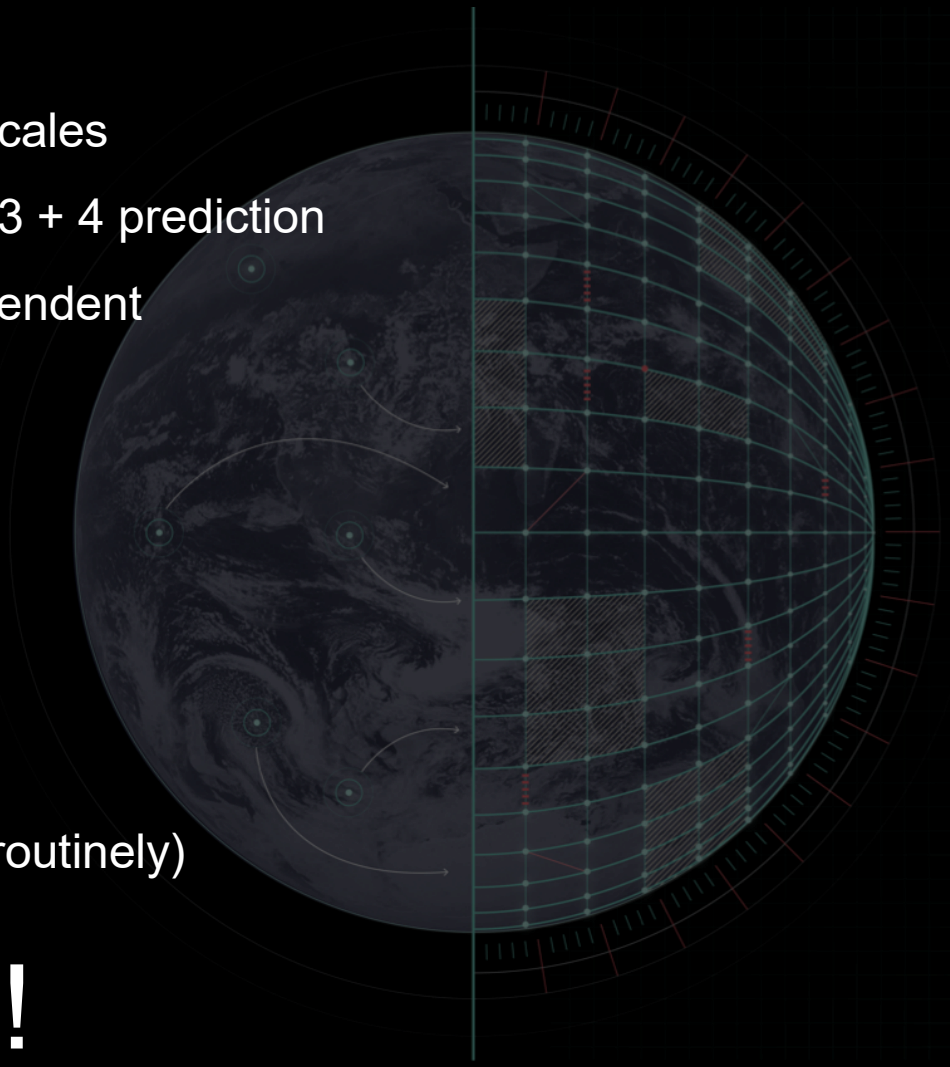
Reforecast:

- Reforecast period: past 20-years
- 10 members each date



Summary

- Fully data-driven models outperform IFS on sub-seasonal timescales
- Autoregressive models trained up to 3 days generalize to week 3 + 4 prediction
- Differences between model versions are small and variable dependent



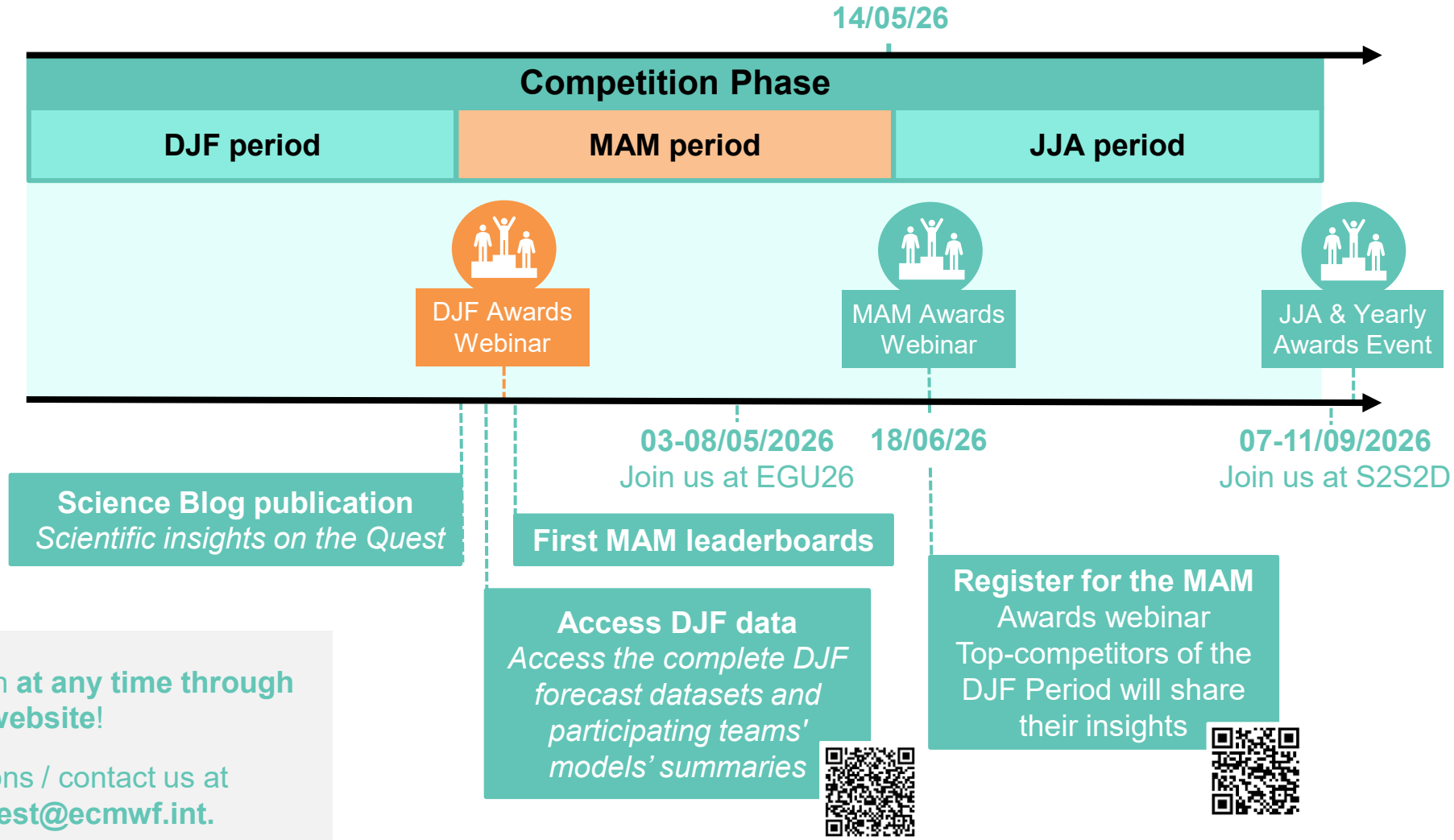
Next steps:

- Update on AIFS-gaia coming soon
- Post-processing of AIFS
- AIFS sub-seasonal system to be semi-operational in 2026 (run routinely)

Thank you!

jakob.schloer@ecmwf.int

Key milestones and actions



Join the competition **at any time through the website**!

Ask us questions / contact us at **aiweatherquest@ecmwf.int.**

Key milestones and actions

Call for Papers - International Journal of Climatology and Meteorological Applications

Advances in Machine Learning for Weather and Climate: Modelling, Forecasting, and Applications

The increasing use of Machine Learning (ML) to resolve research questions in both weather forecasting and climate modelling presents the potential to substantially enhance our understanding of the Earth system, quantify uncertainties in climate projections, and enhance the applicability of predictions.

This special issue is a collaboration between two RMetS journals, welcoming submissions from variety of disciplines and interests in Machine Learning.

Submission deadline: 31 December 2026

AI Weather Quest ECMWF

Thanks!

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

ECMWF

Destination Earth

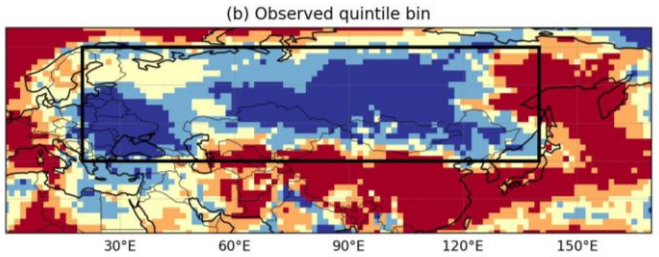
WORLD METEOROLOGICAL ORGANIZATION

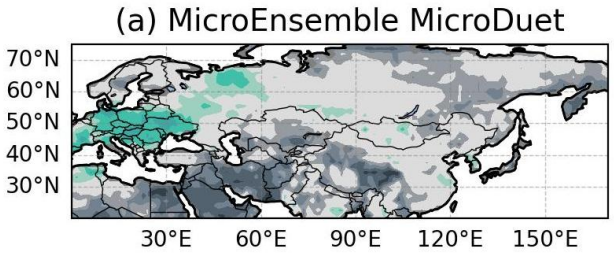
Additional slides

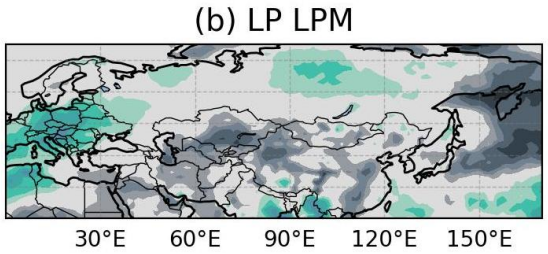
Northern hemisphere extratropical cold spell case studies

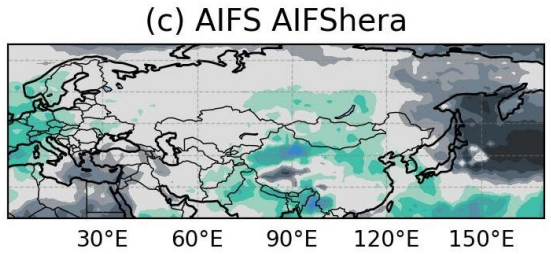
AI Weather Quest ECMWF

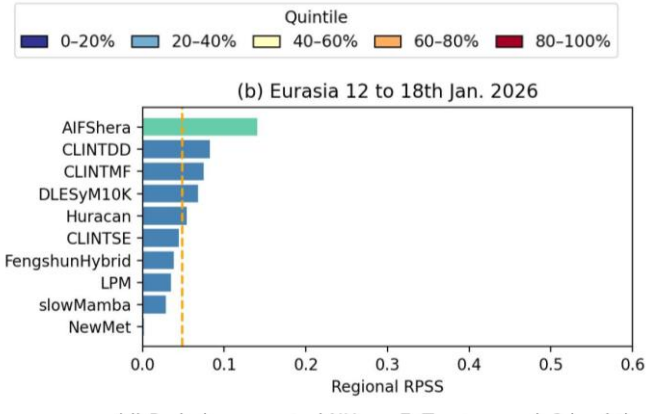
Examples of colder-than-normal quintile (< 20%) forecasts

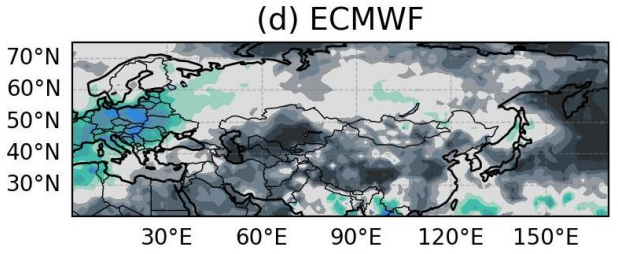


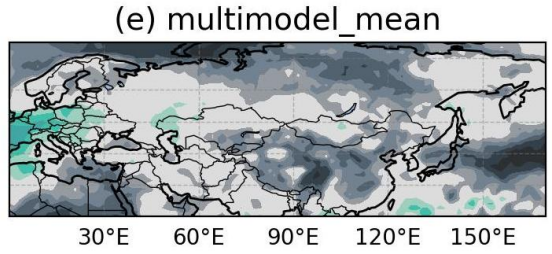


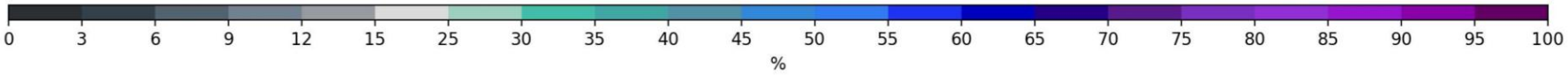












ECMWF

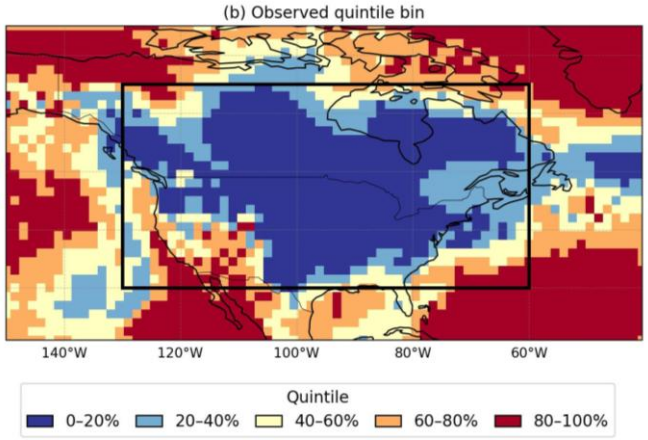
Destination Earth

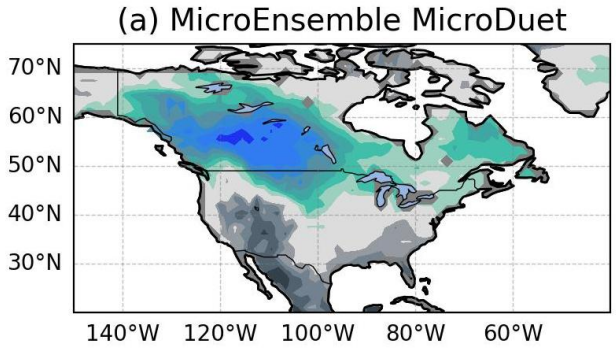
WORLD METEOROLOGICAL ORGANIZATION

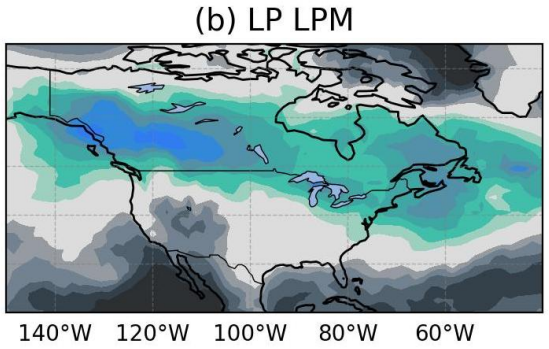
Northern hemisphere extratropical cold spell case studies

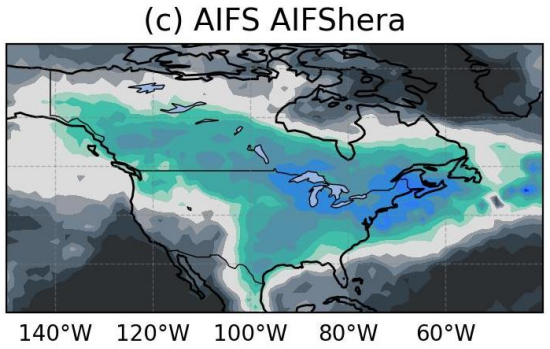
AI Weather Quest ECMWF

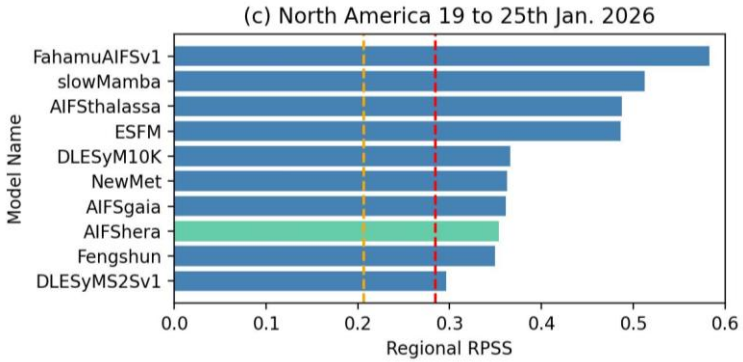
Examples of colder-than-normal quintile (< 20%) forecasts

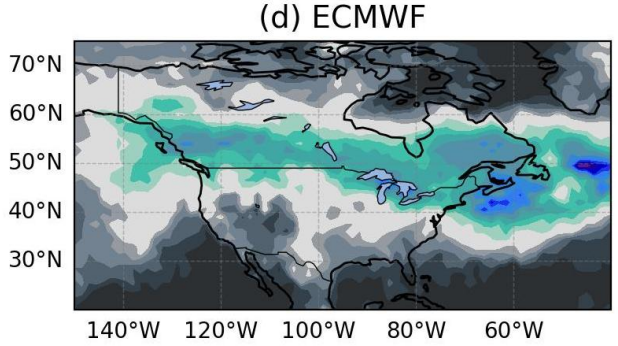


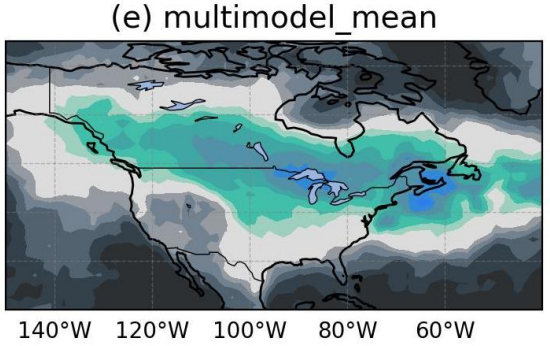


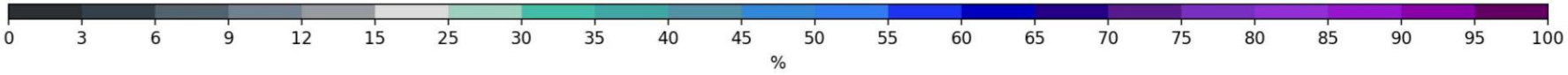












ECMWF

Destination Earth

WORLD METEOROLOGICAL ORGANIZATION