



TianXing-S2S: A data-driven probabilistic subseasonal to seasonal forecasting model based on multi-sphere coupling

Prof. Bin Mu

School of Computer Science and Technology, [Tongji University](#)

2026.05.15 [NanJing, China](#)



**TianXing-S2S: A data-driven
probabilistic subseasonal to
seasonal forecasting model
based on multi-sphere coupling**



1 Introduction

2 TianXing-S2S

3 Experiments

4 Discussions

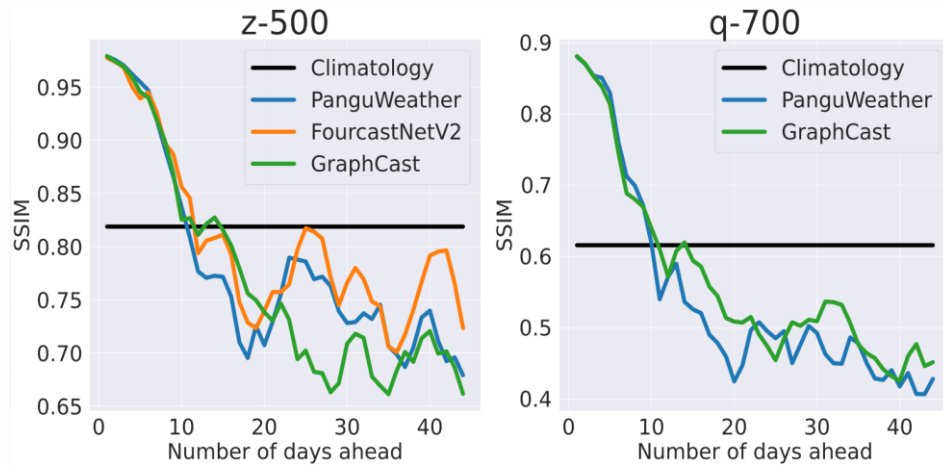


Awesome AI Weather Forecast Models

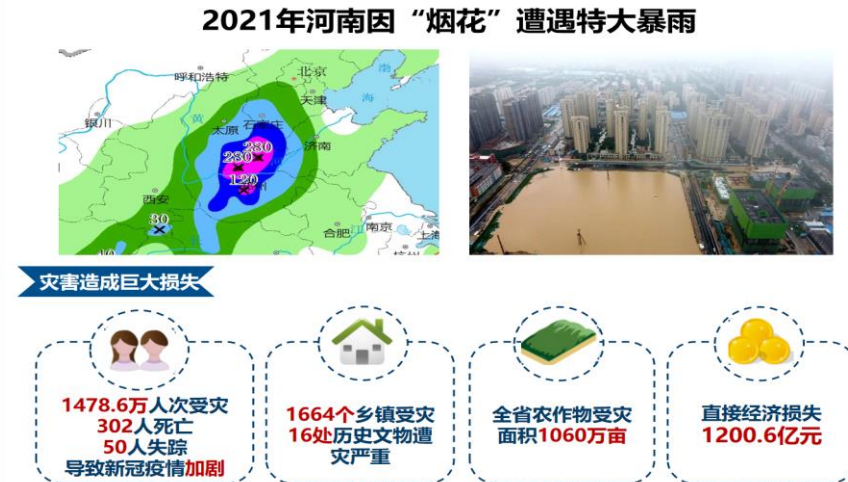
Name	institution	res	method	time
FourCastNet	NVIDIA	0.25°	FNO-based	2022.02
Pangu-weather	Huawei		Transformer-based	2022.11
FengWu	Shanghai AI Lab		Transformer-based	2023.04
FuXi	Fudan university		Transformer-based	2023.06
TianXing	Tongji university		Transformer-based	2024.07
AIFS	ECMWF		Transformer + GNN-based	2024.08
GraphCast + GenCast	Google DeepMind		GNN	2024.12
Aardvark	University of Cambridge		Transformer + CNN-based	2025.03
Aurora	Microsoft		Transformer-based	2025.05

- In recent years, AI models have been continuously iterated and updated, which has **profoundly promoted** the upgrading of modeling paradigms in the field of Earth sciences.

➤ However, current AI-based weather forecast models also face **challenges**



- Forecast skill rapidly declines **beyond 10 days** (sub-seasonal scale)



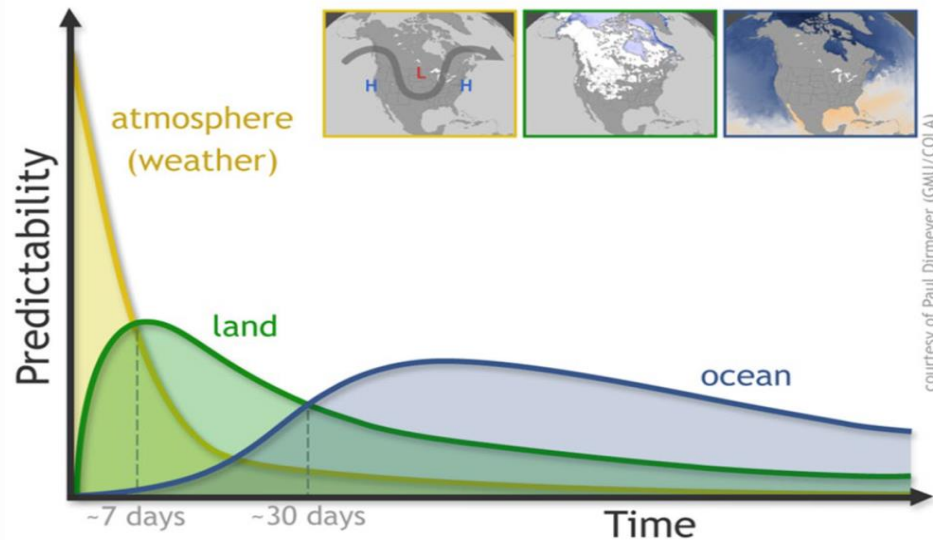
- The forecast results are difficult to meet the needs of **early-stage disaster warnings**

➤ The predictive ability of AI models in **S2S time scales** still needs to be improved.



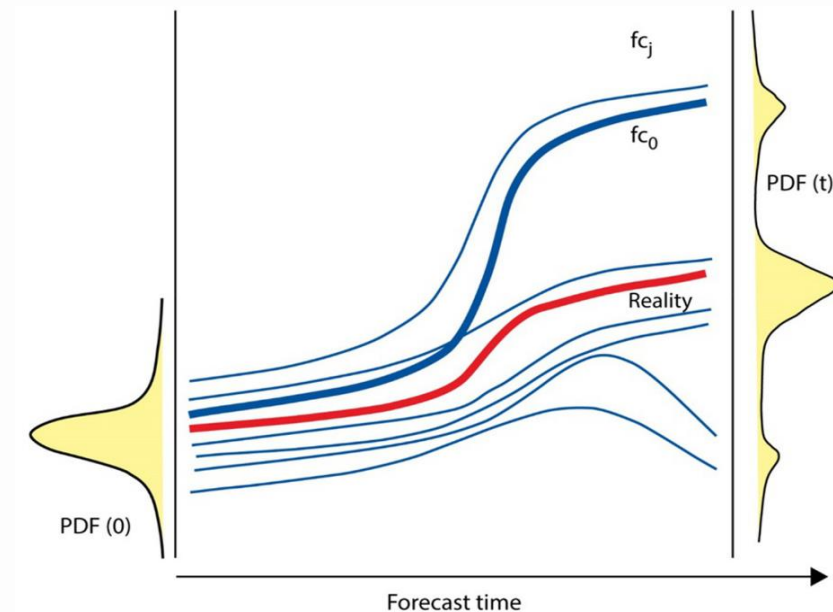
➤ From "AI Weather Forecast" to "AI Subseason to Seasonal Forecast"

Multi-shpere Coupling



- Need to incorporate variables from **multiple spheres**; And reasonably characterize it to reduce the noise interference caused by the increase in data volume

Ensemble Forecasts

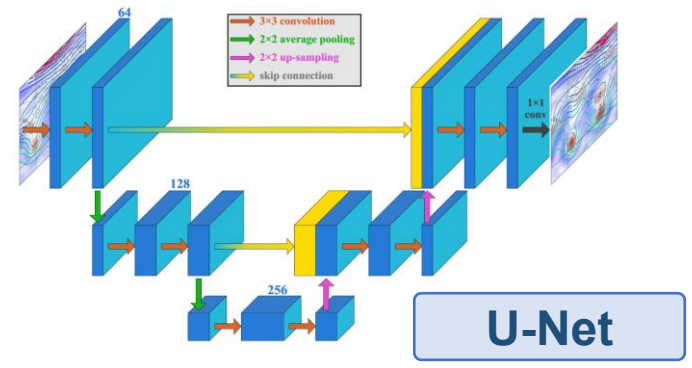


- Effective **uncertainty modeling** methods are needed, though the optimization method of AI models tends to produce deterministic and convergent forecast results

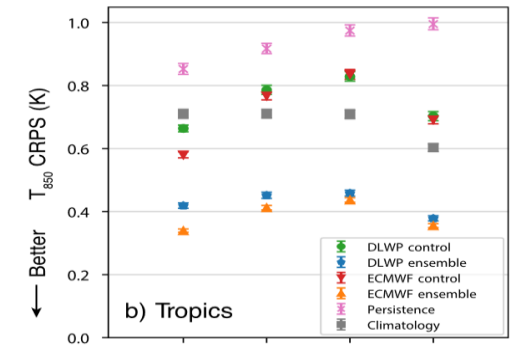
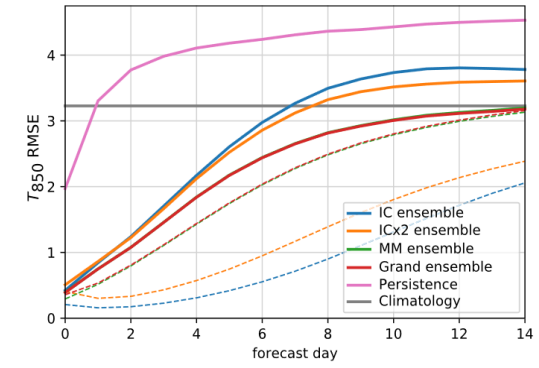


1.Introduction

U-Net

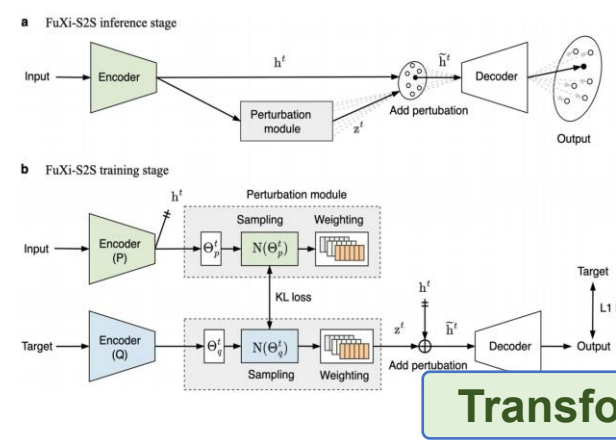


Weyn J A, Durran D R, Caruana R, et al. Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models[J]. *Journal of Advances in Modeling Earth Systems*, 2021, 13(7).



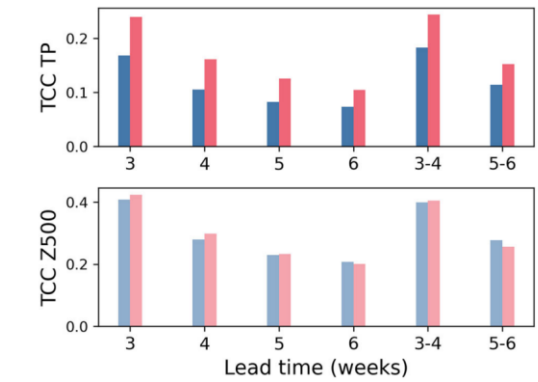
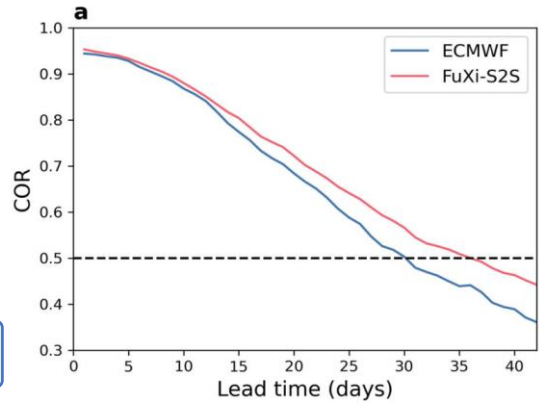
First Attempt

FuXi-S2S



Transformer

Chen L, Zhong X, Li H, et al. A machine learning model that outperforms conventional global subseasonal forecast models[J]. *Nature Communications*, 2024, 15(1): 6425.



Outperform NWP

➤ The work of scholars has fully demonstrated the feasibility of AI models in S2S prediction problems, prompting us to attempt to unify multi-sphere customized coupling and effective representation of uncertainty. → **TianXing-S2S**



Predictands in TianXing-S2S

Sphere	Level	Variable	Abbreviation
Atmosphere	Pressure levels at 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 850, 925, and 1000 hPa	specific humidity	Q
		temperature	T
		u component of wind	U
		v component of wind	V
		geopotential	Z
	Single levels	2-meter temperature	T2M
		outgoing longwave radiation	OLR
		total precipitation	TP
		mean sea level pressure	MSLP
		10-meter u wind component	U10
		10-meter v wind component	V10
Ocean	Single levels	sea surface temperature	SST
		sea-ice cover	SIC
Land	Depth levels at 0-7, 7-28, and 28-100 cm	soil temperature	ST
		soil moisture	SM
Interface flux	Single levels	mean surface latent heat flux	LHF
		mean surface sensible heat flux	SHF

Multiple spheres:

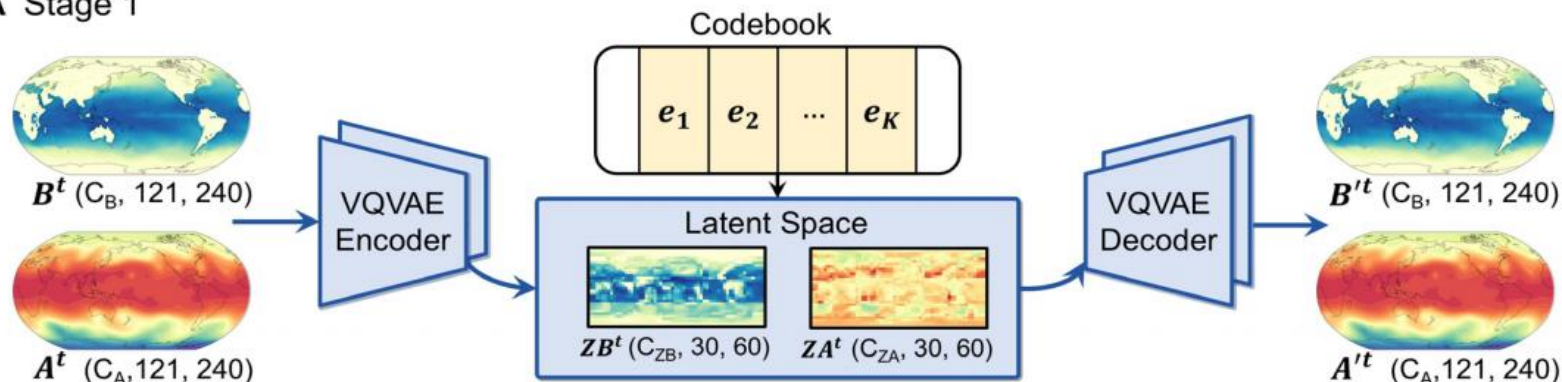
TianXing-S2S forecasts 81 key variables spanning four Earth system spheres—atmosphere, ocean, land, and interface flux.

Resolution:

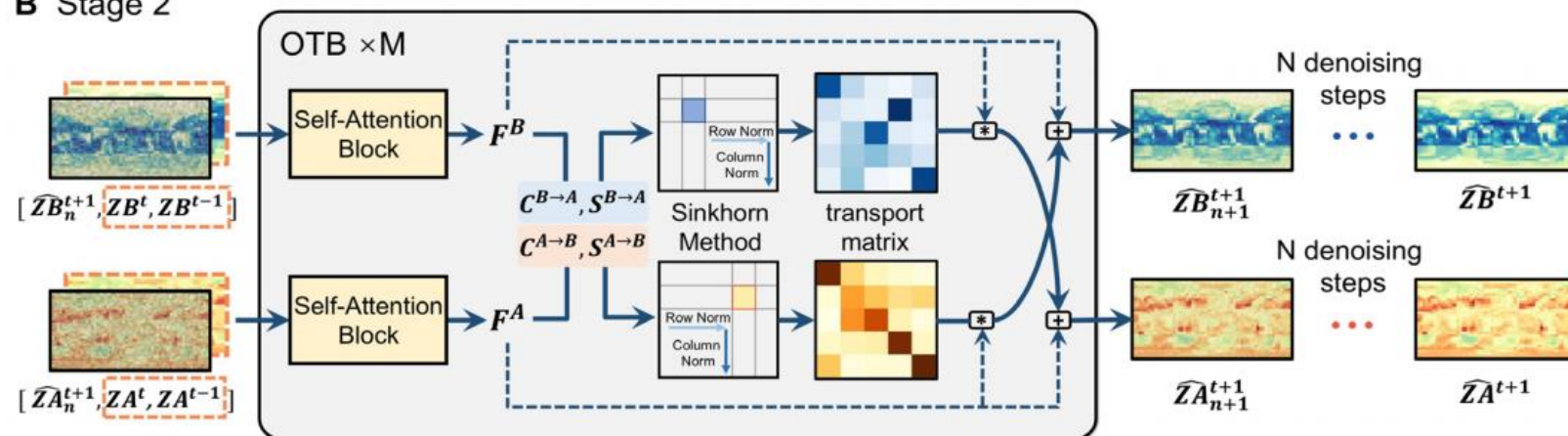
This comprehensive multi-sphere coverage enables physically consistent coupling between spheres, supporting global daily-mean ensemble predictions at 1.5° resolution.

Model structure of TianXing-S2S

A Stage 1



B Stage 2



Condition (dashed box) Skip Connection (dashed arrow) $[\cdot]$ Channel Concatenation \otimes/\oplus Element-wise Multiplication / Addition

Stage 1

Group pre-training **VQ-VAE autoencoders** for multi-sphere variables

Stage 2

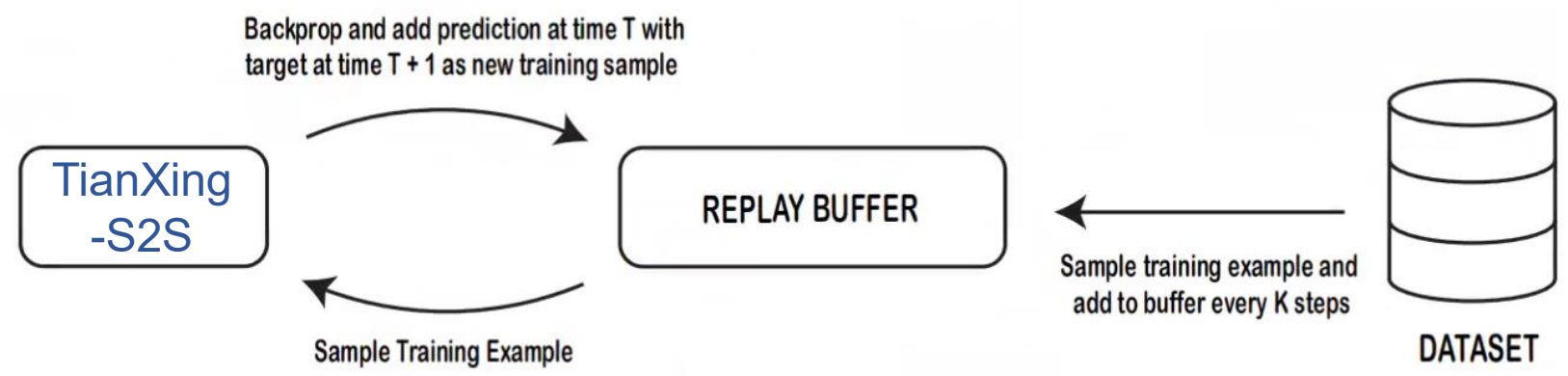
Train Transformer-based **diffusion model** to represent uncertainty based on the encoded latent space features

OTB

Based on the **optimal transfer concept**, additional normalization constraints are added to transfer only **the most valuable physical information** between atmospheric variables and multi-sphere boundary condition variables



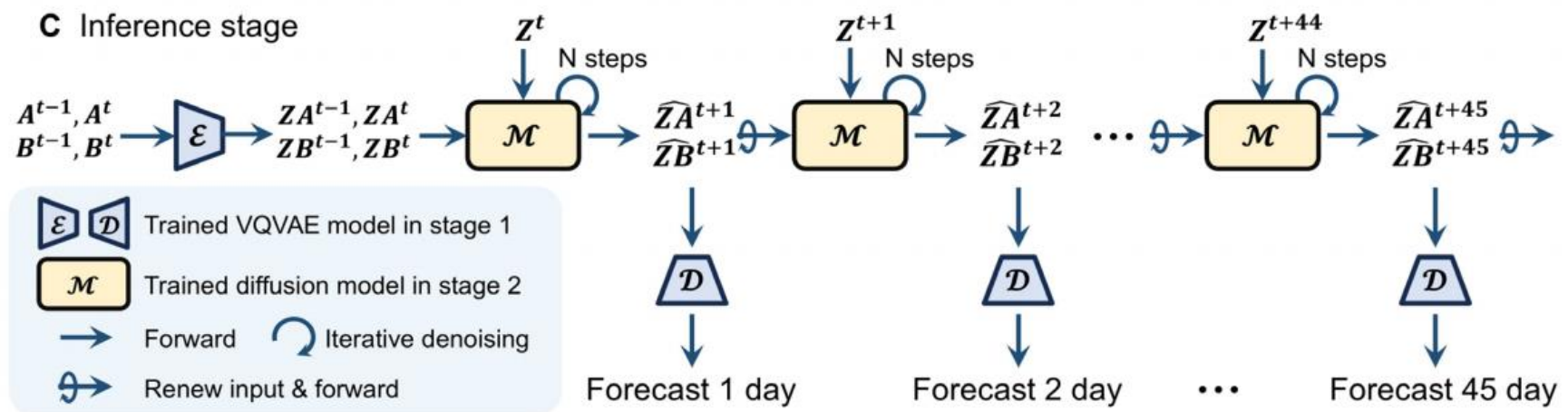
Training strategy of TianXing-S2S



Bodnar C, Bruinsma W P, Lucic A, et al. A foundation model for the Earth system[J]. Nature, 2025.

Pre-training
Single step training, fitting the noise of a certain denoising step

Fine-tuning Training:
Gradually increase the prediction time, allowing the model to train on its own predictions, and use a replay buffer strategy to improve training efficiency.

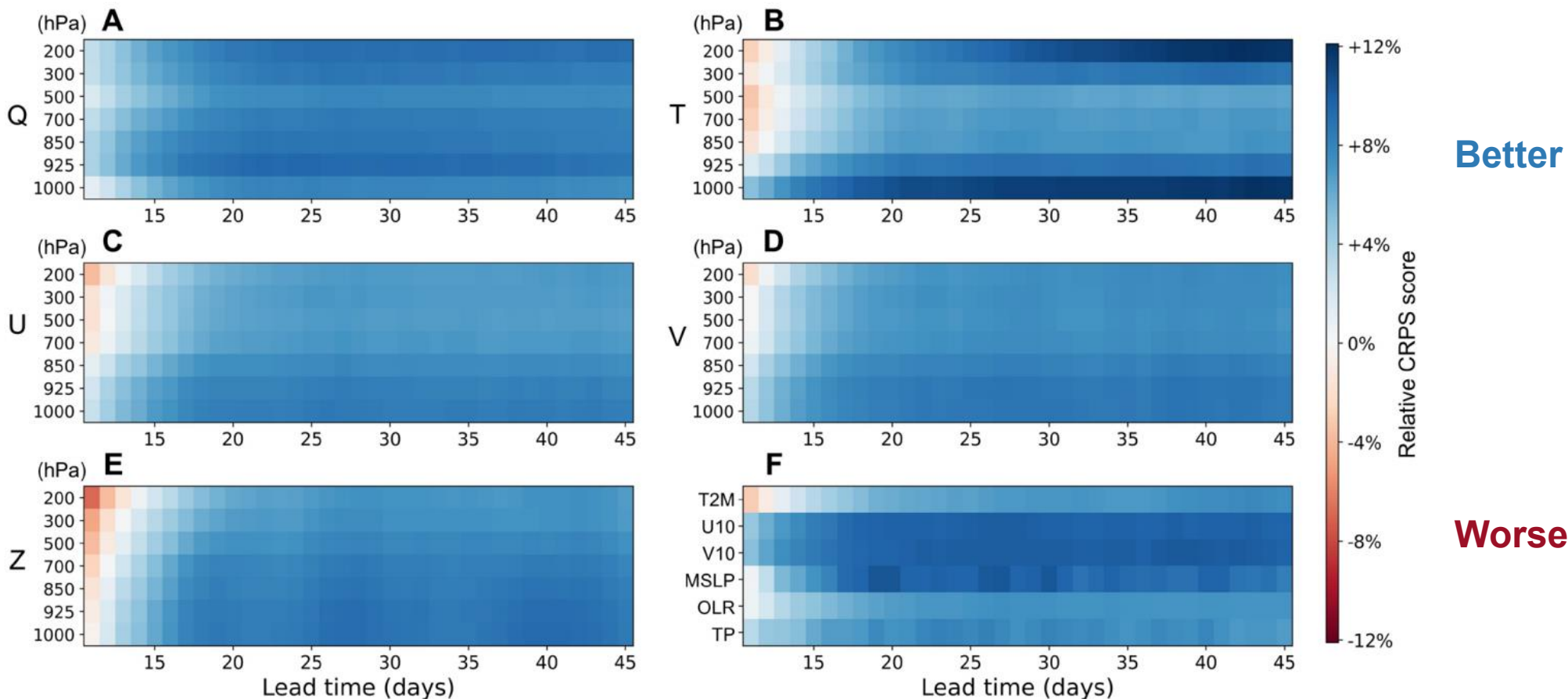


Inference stage:
Autoregressive Forecasting
1 day... 45 days... 180 days



Continuous ranked probability score (CRPS)

v.s. ECMWF S2S

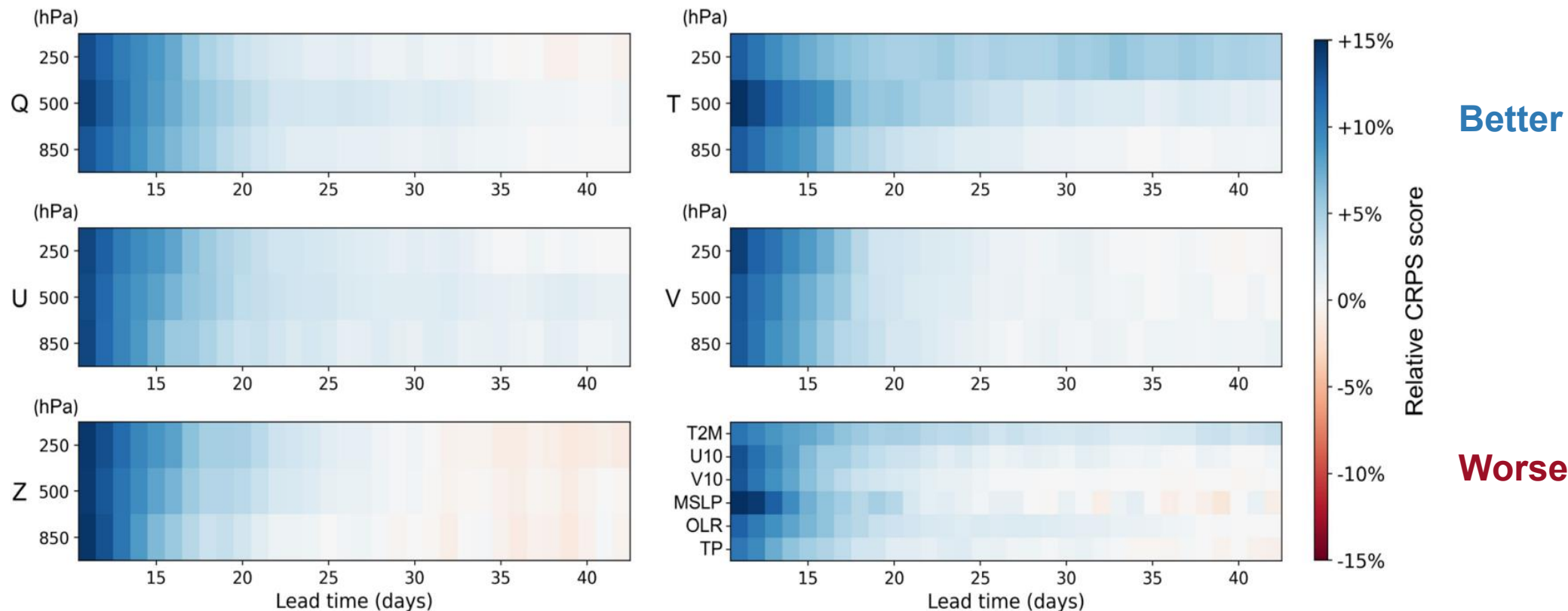


- On many key variables, TianXing-S2S demonstrates **superior** predictive power compared to ECMWF S2S.



Continuous ranked probability score (CRPS)

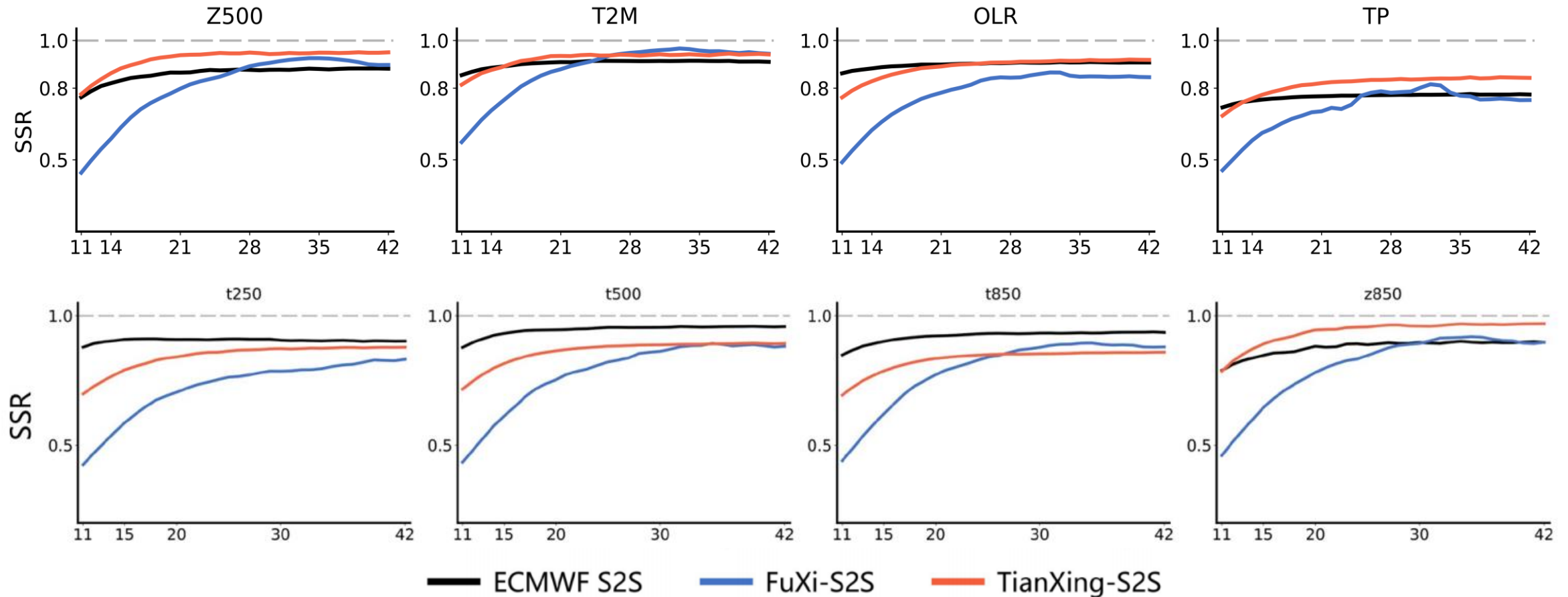
v.s. FuXi-S2S



- Compared to the state-of-the-art (SOTA) FuXi-S2S model, TianXing-S2S achieves a **comparable** level and is even **superior** on some variables.



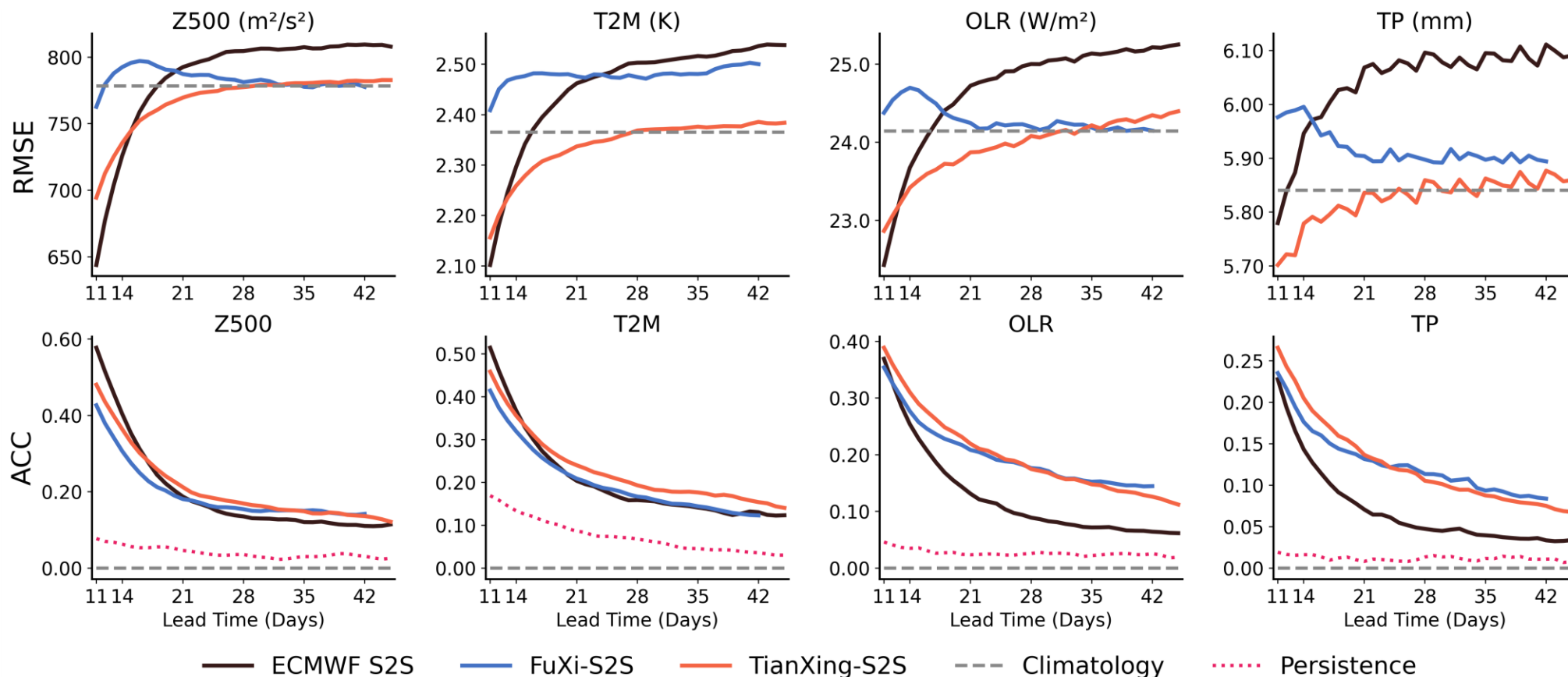
Spread-skill ratio (SSR)



- The SSR forecast skill of TianXing-S2S is generally **better** than that of FuXi-S2S and comparable to that of ECMWF S2S, verifying the **reliability** of TianXing-S2S ensemble forecasts.



Root mean square error (RMSE) & Anomaly correlation coefficient (ACC)



- Compared to the ECMWF S2S/FuXi-S2S, the TianXing-S2S demonstrates **superior** RMSE/ACC performance across key variables such as Z500/T2M/OLR/TP.

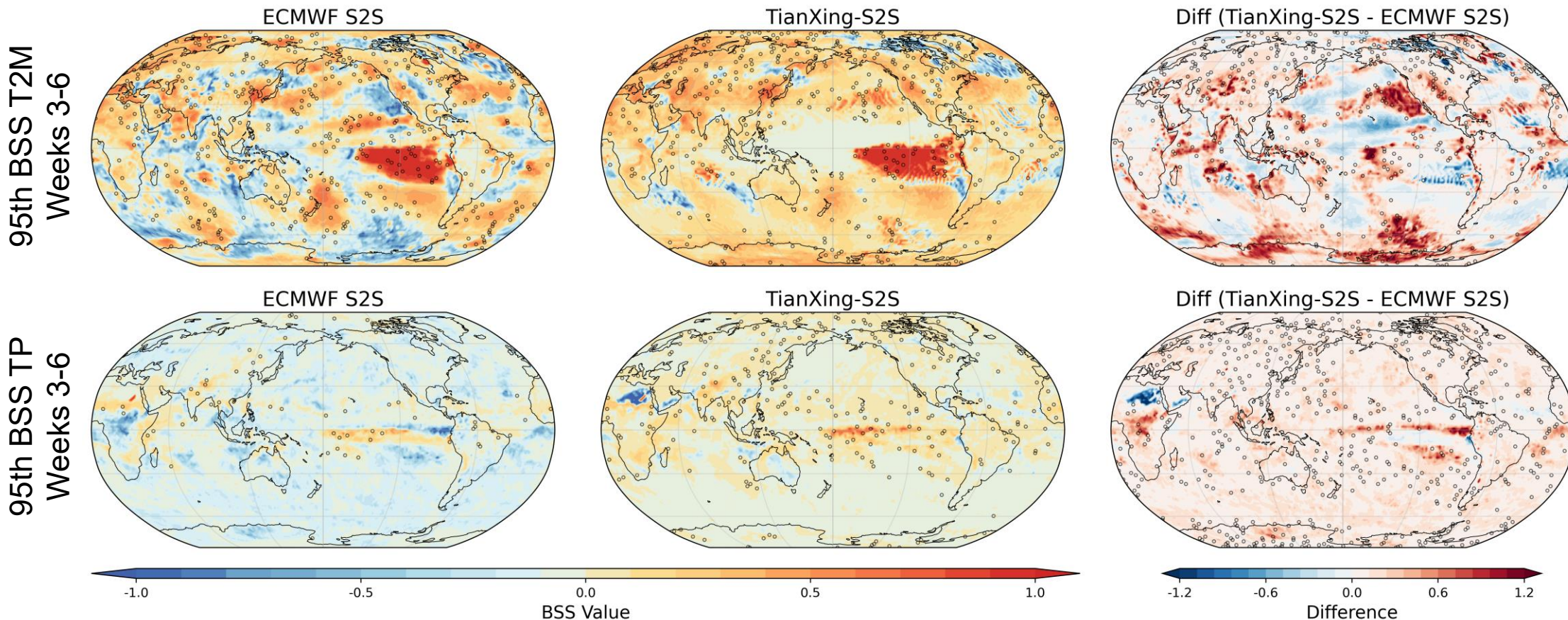


3.Experiments

14

Brier skill score (BSS)

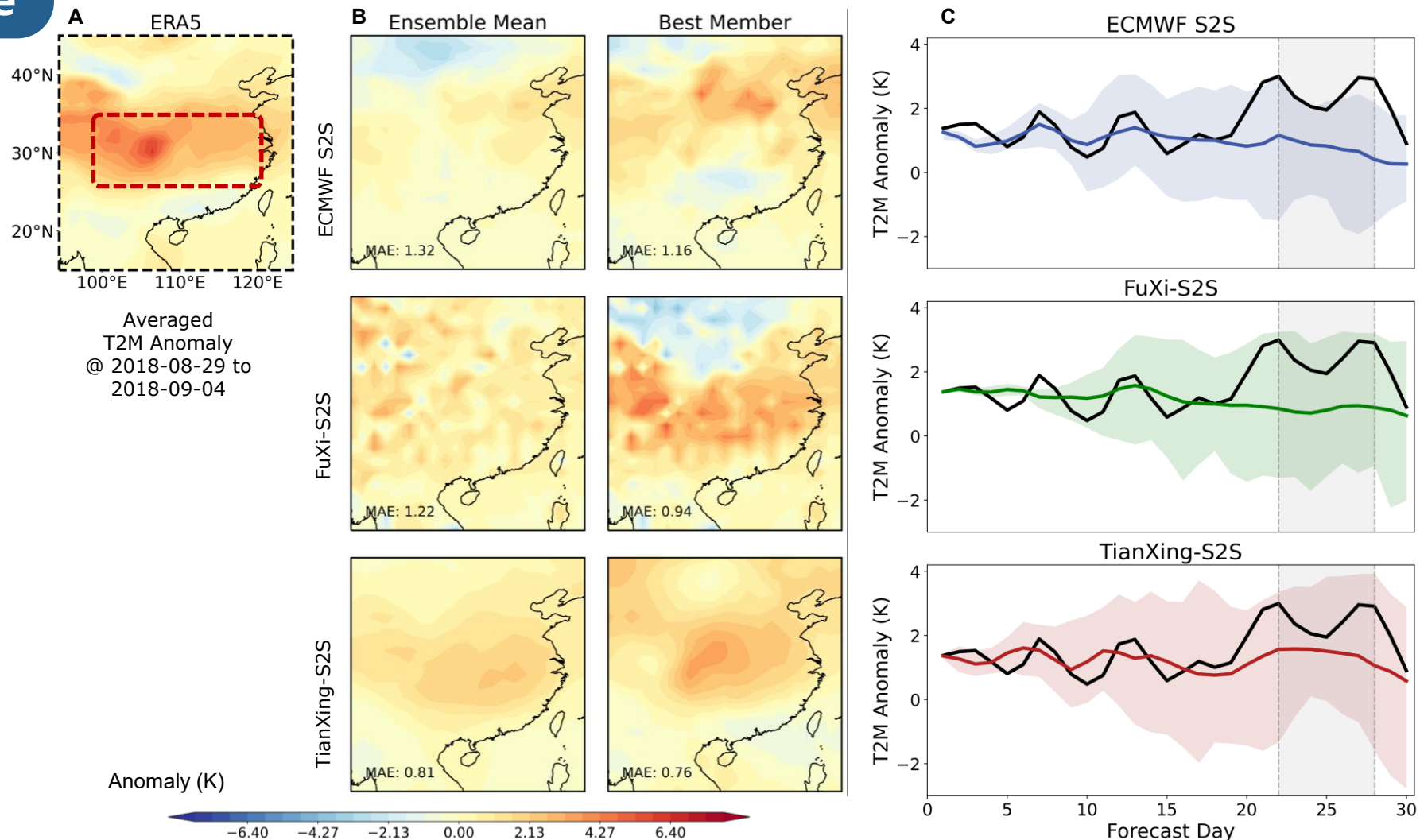
BSS = 1: optimal; BSS > 0: positive skill; BSS < 0: negative skill



- TianXing-S2S **outperforms** the ECMWF S2S ensemble forecast return system overall in extreme BSS scores on key T2M/TP variables.



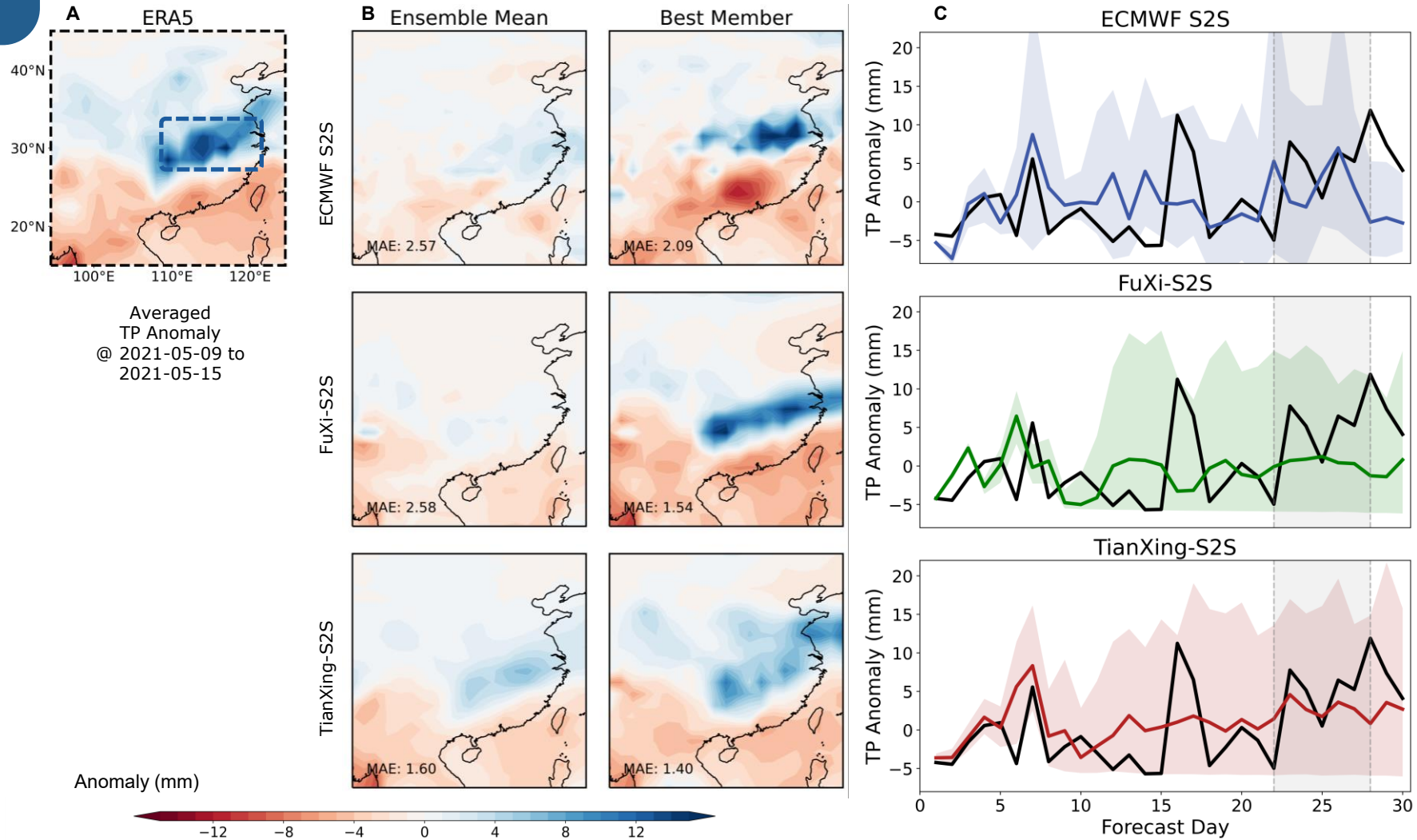
Heat wave



➤ For the 2018 heat wave, TianXing-S2S yields the **lowest** MAE (0.81 / 0.76); its ensemble **envelope brackets** the late-period surge that ECMWF and FuXi-S2S miss.



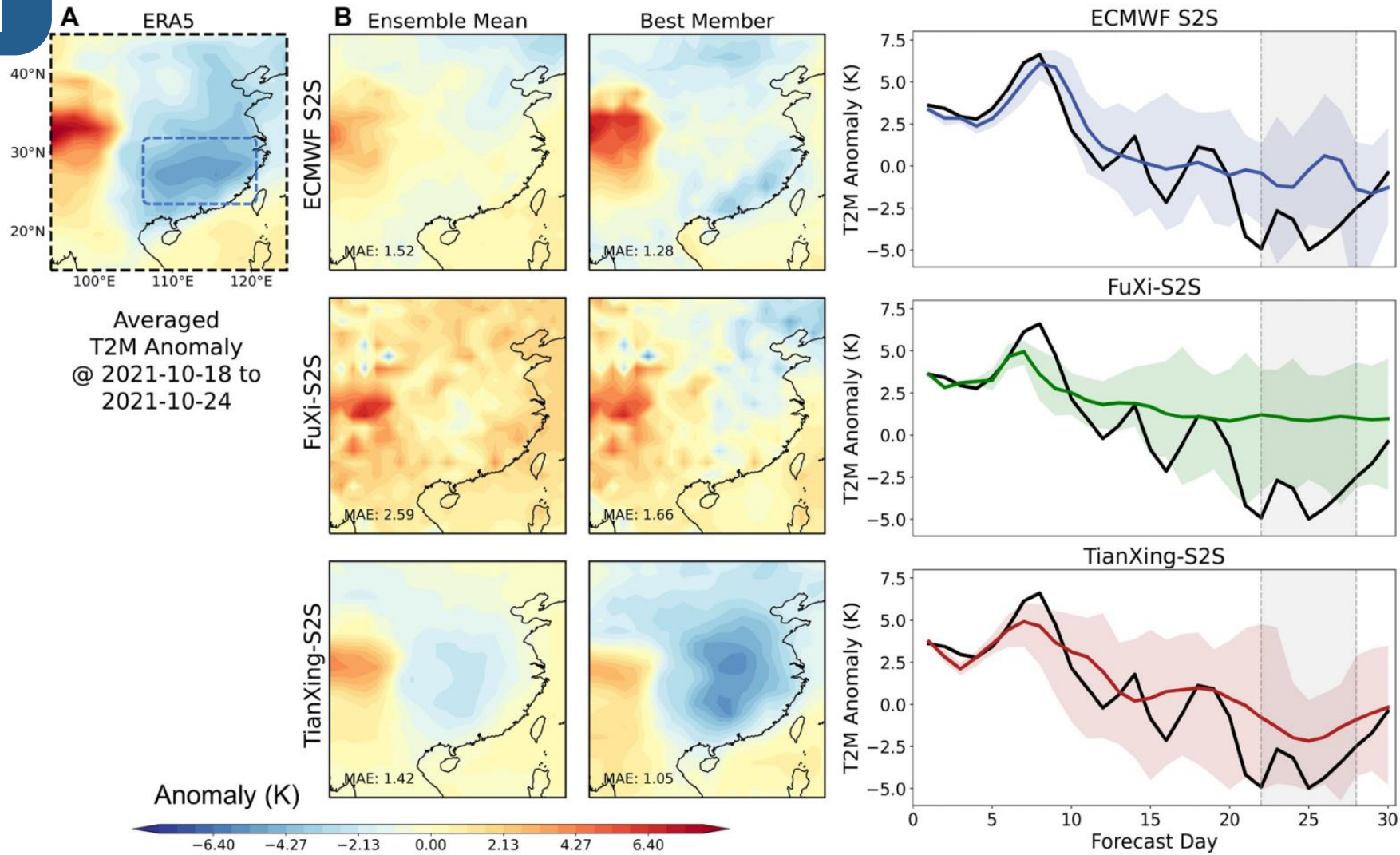
Meiyu



- For the 2021 Meiyu event, only TianXing-S2S's ensemble mean reproduces the **wet rainband along the Yangtze**, achieving the lowest MAE (1.60 / 1.40).



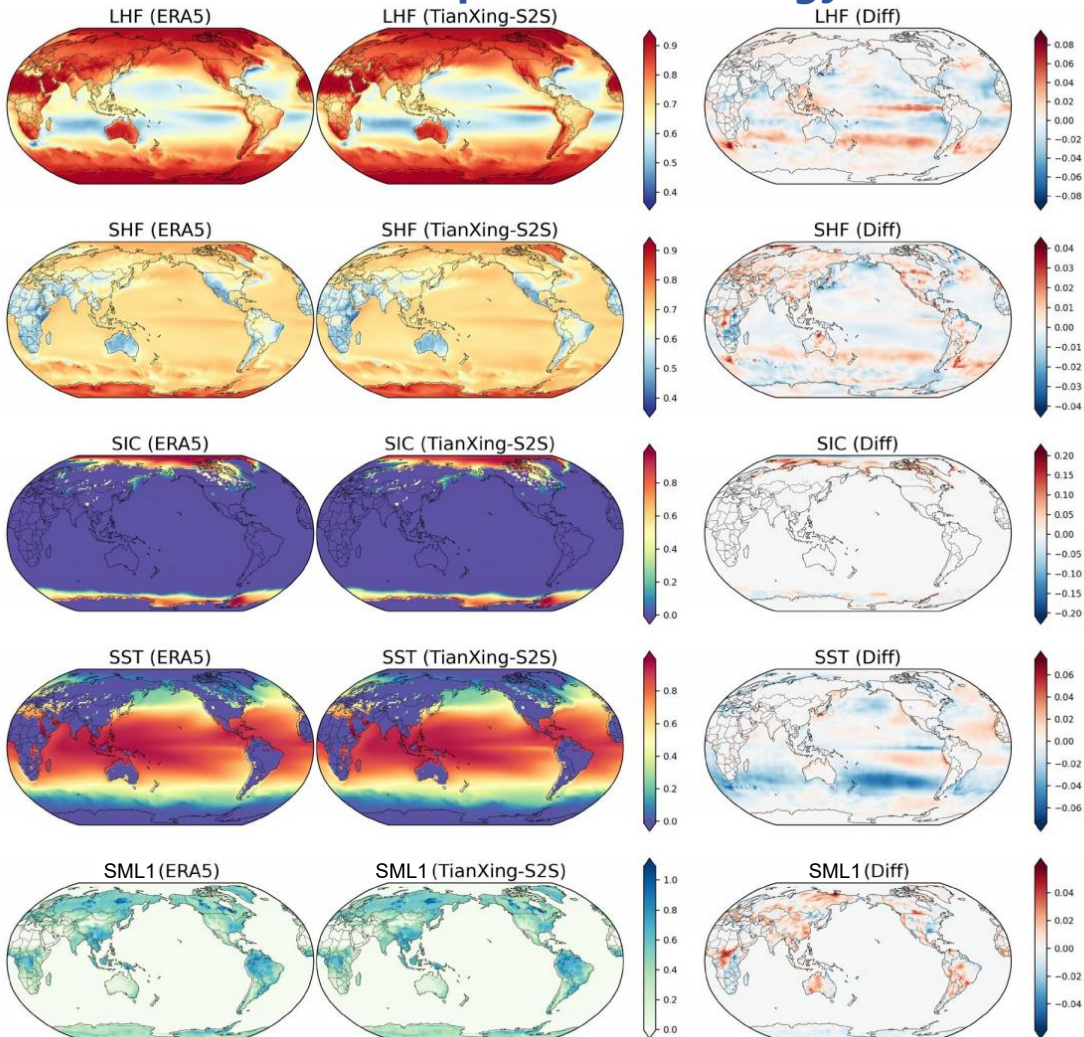
Cold spell



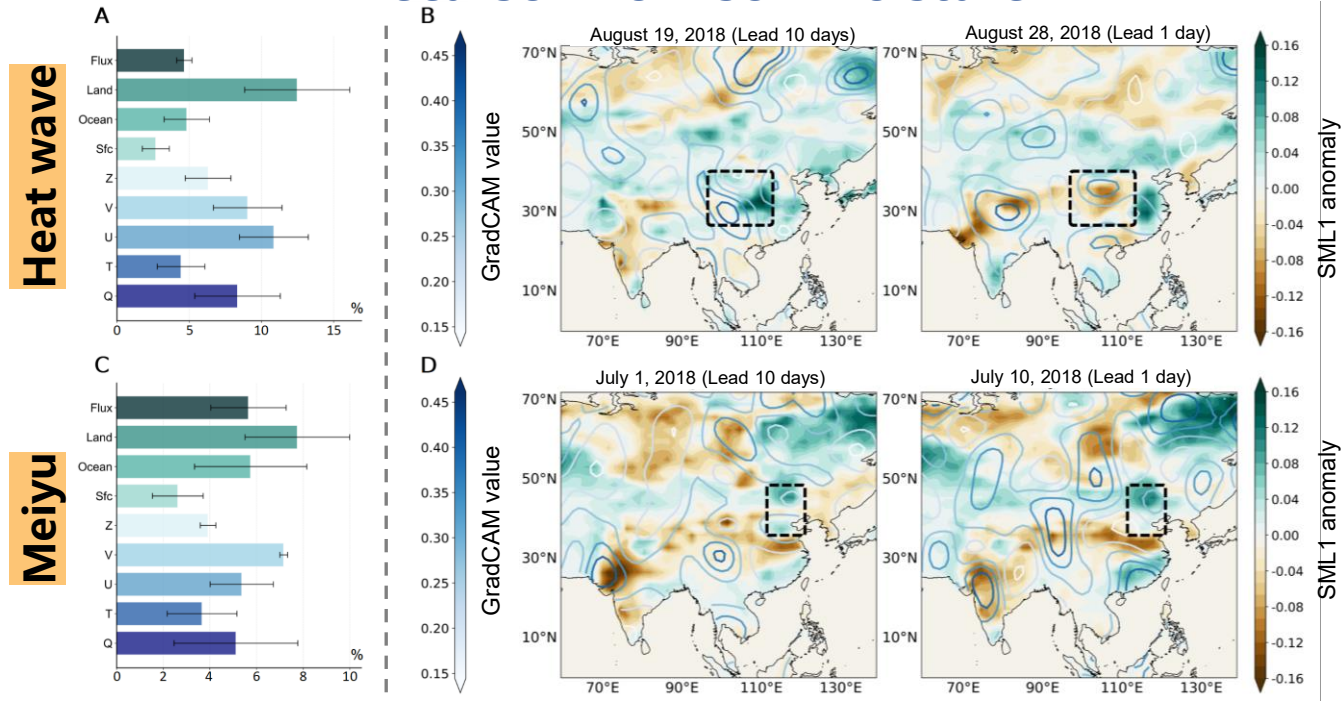
- For the 2021 cold spell, TianXing-S2S best captures the **cold core over southeastern China** (MAE 1.42 / 1.05), while FuXi-S2S markedly underestimates the anomaly.

Multi-sphere attributions

Multi-sphere climatology



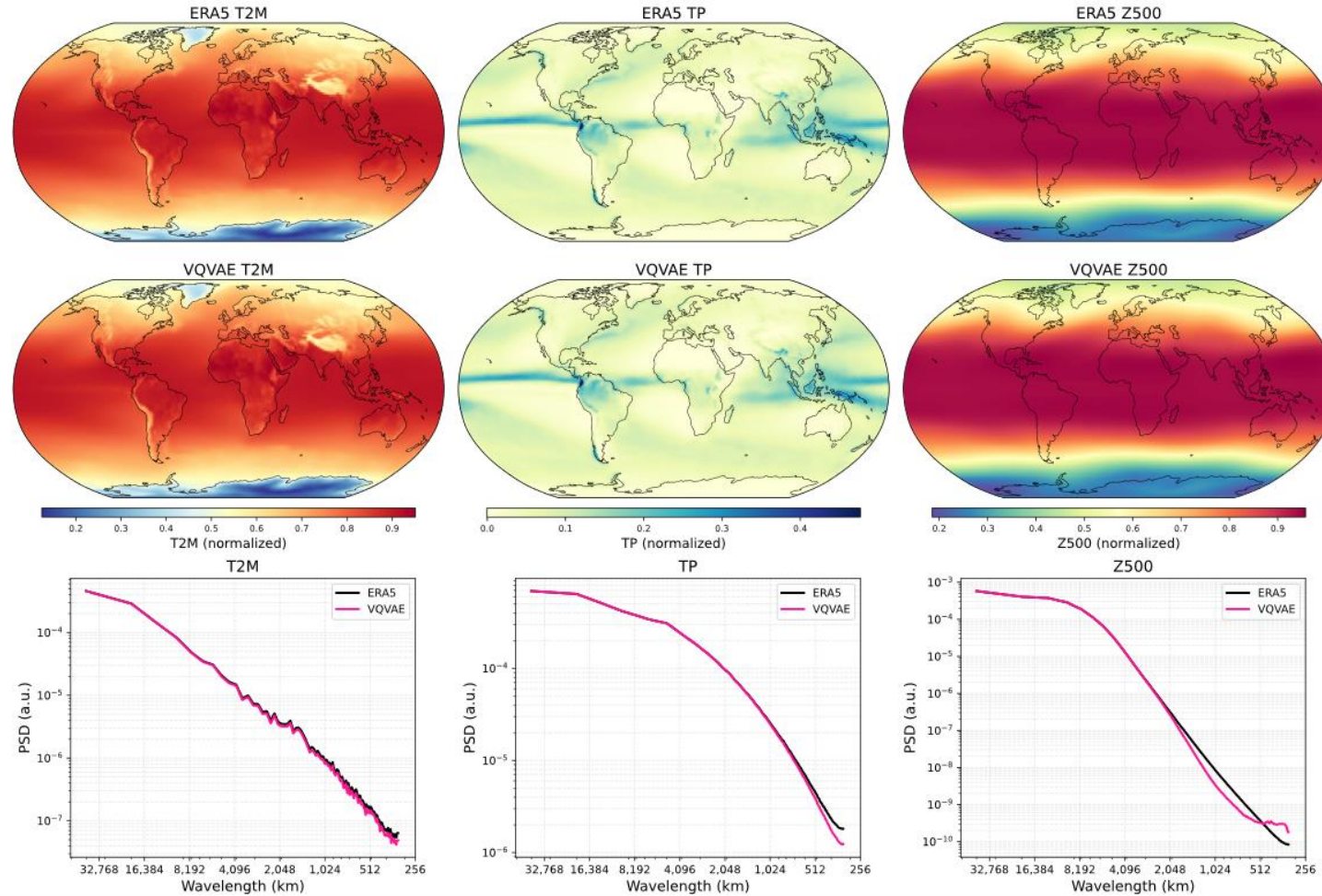
Precursor from soil moisture



➤ By combining the Perturbation Importance Method (PIM) and the GradCAM method, TianXing-S2S identified **soil sphere variables** as important precursor signals.



VQ-VAE accelerates computational efficiency

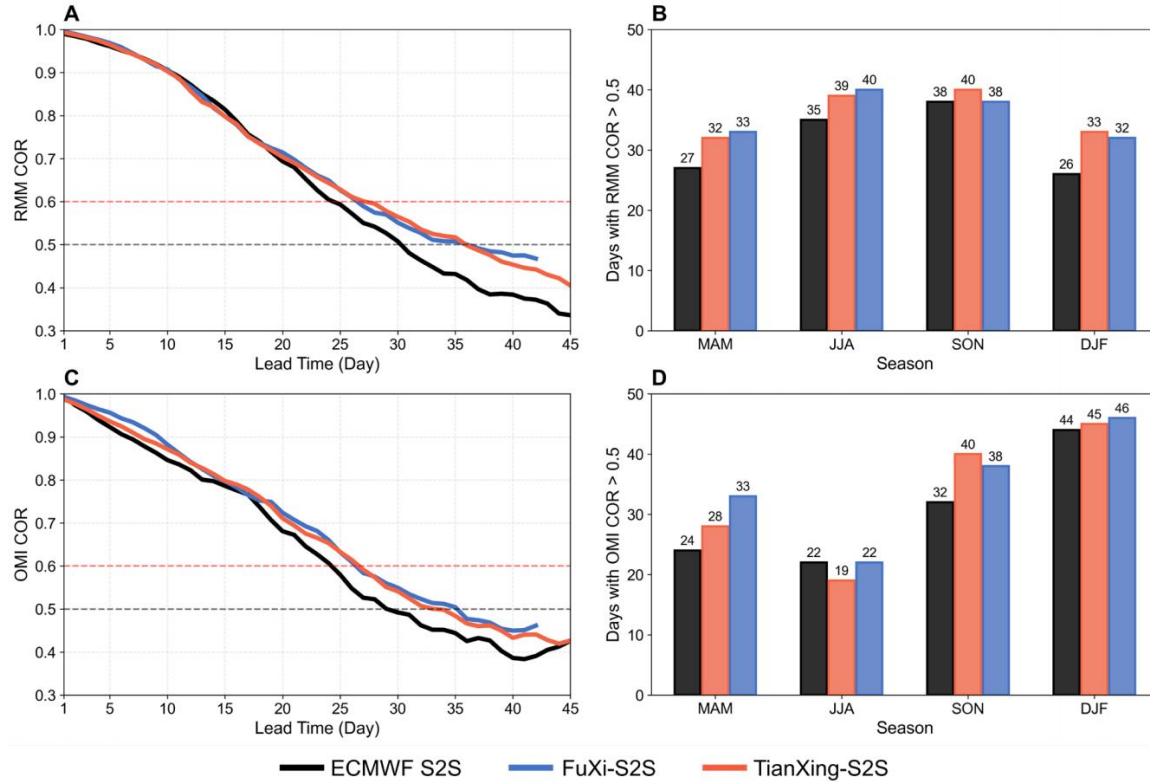


- With $\approx \times 20$ reduced data volume, highly consistent reconstruction of climate state and power spectrum, VQ-VAE further **improves the efficiency** of AI models.

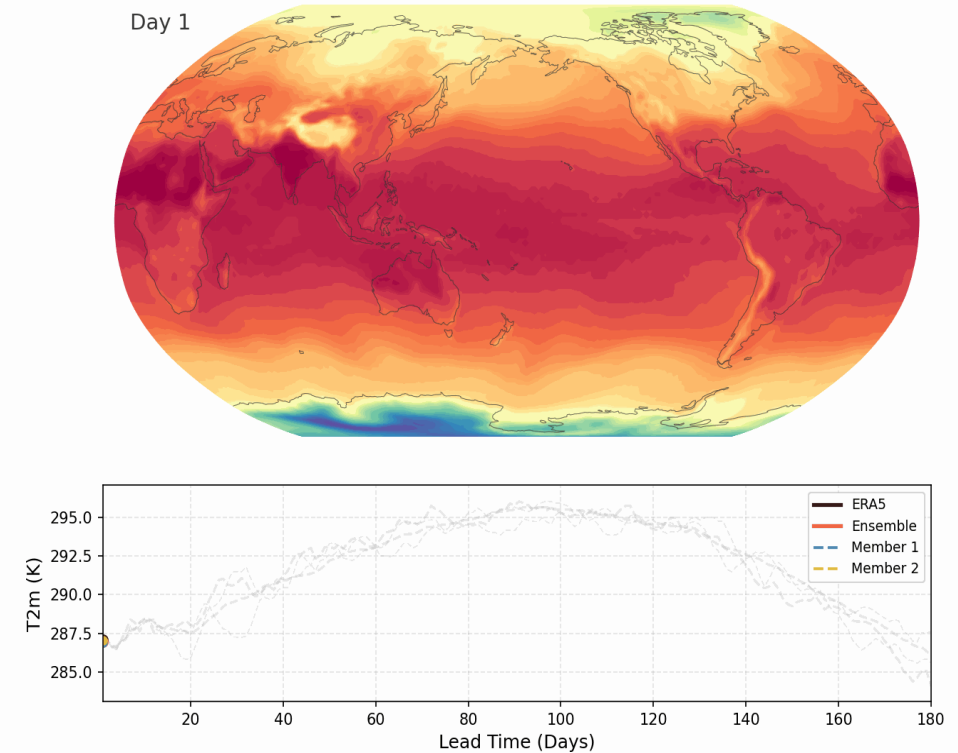


Seasonal forecast outlook

MJO simulations



180-day long term forecasts (T2M)



- TianXing-S2S has high forecasting skill for MJO and can reproduce **stable modes** and **seasonal cycles** in long-term forecasts.



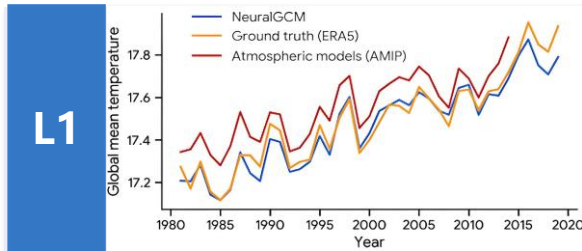
Mu B, Chen Y, Yuan S, et al. **Skillful Subseasonal-to-Seasonal Forecasting of Extreme Events with a Multi-Sphere Coupled Probabilistic Model**[J]. arXiv preprint arXiv:2512.12545, 2025.

- **1. Multi-sphere + Diffusion Modeling:** TianXing-S2S integrates 81 multi-sphere variables via sphere-specific **VQVAEs** and a latent **diffusion** framework, with an **Optimal Transport Block** coupling atmospheric and boundary dynamics.
- **2. Comprehensive Performance:** TianXing-S2S **outperforms** ECMWF S2S and FuXi-S2S in ensemble mean, spread, and tail coverage, with **notable gains** on heat waves, cold spells, and precipitations.
- **3. Soil Moisture as Precursor Signal:** TianXing-S2S identifies **soil moisture** as a key precursor, with captured **land-atmosphere memory** underpinning its enhanced skill on heat waves at extended lead times.

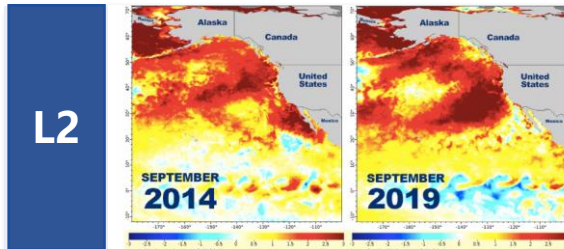


- Setting higher requirements: from "predictable" to "explainable and traceable"

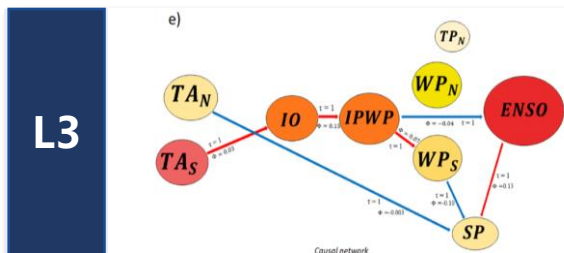
From the surface to the core,
meticulous attribution



- Can we accurately predict future evolution?



- Which regions are most affected by disturbances, and which physical processes are most critical?



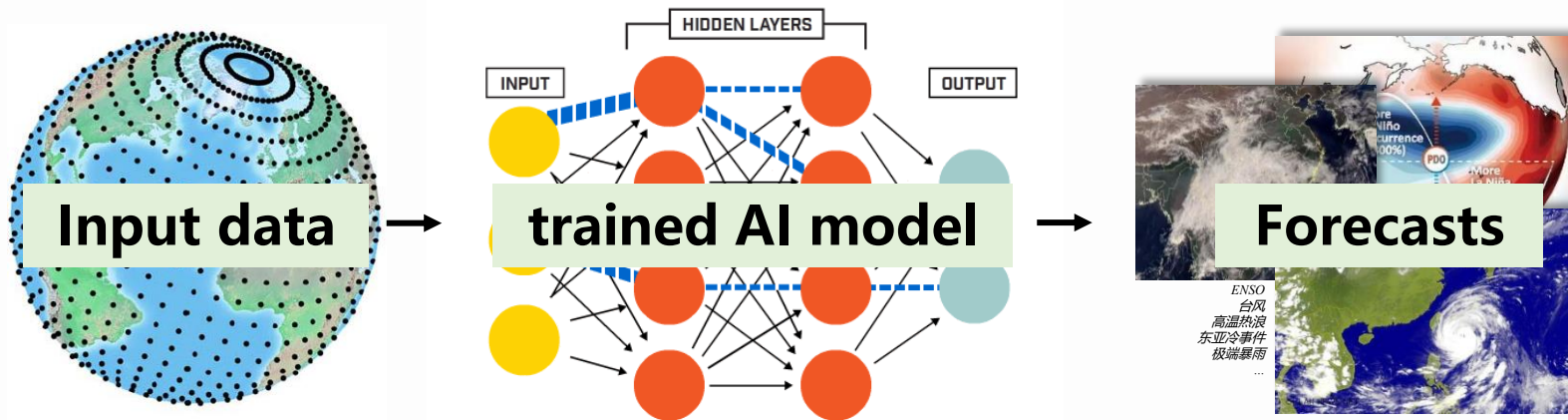
- What are the physical mechanisms underlying the evolution of signals or disturbances? What is the causal chain?

From cause to effect,
driving modeling and cognition

Research on the predictability of AI models
should be put on the agenda.

- Predictability research focuses on **clarifying the causes and mechanisms of forecast uncertainty**, and then providing methods and approaches to **reduce** forecast uncertainty.

Predictability Research Paradigm



Type I

- Assumes the forecast model is "perfect"
- Investigates forecast errors arising from **initial-condition** errors

Type II

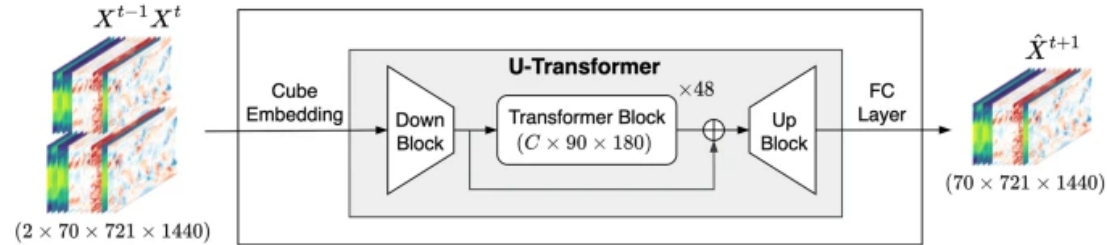
- Assumes the initial conditions are "perfect"
- Investigates forecast errors arising from **model** errors

➤ Weather Events: Typhoon Track

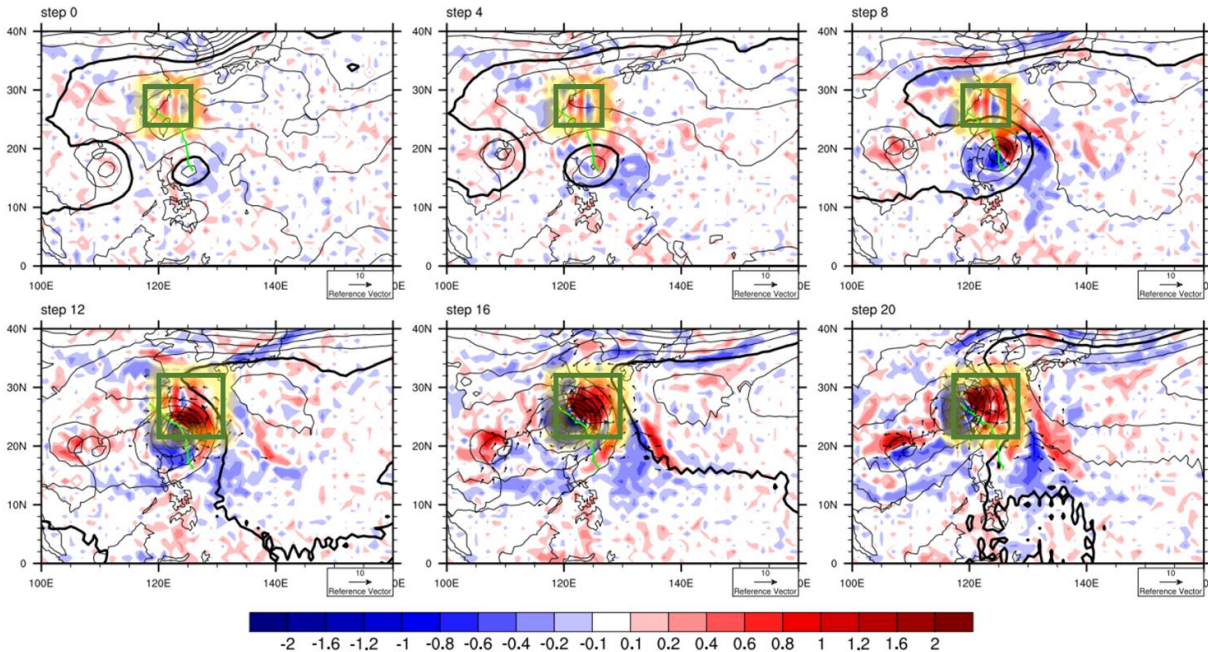
a) The overall architecture of FuXi model

FuXi model

Chen et al., 2023, npj CAS



The optimal growing initial error (OGIE) that causes the largest Typhoon track uncertainty identified in FuXi



Taking Typhoon Gaemi (2024) as an example

- Shades represent the **potential vorticity anomalies**.
- The contours represent the geopotential height at 500hPa of control forecast (**Z500**).
- The thick black line marks the Western Pacific Subtropical High of control forecast (**WPSH**).
- The green line represents the typhoon track of control forecast.

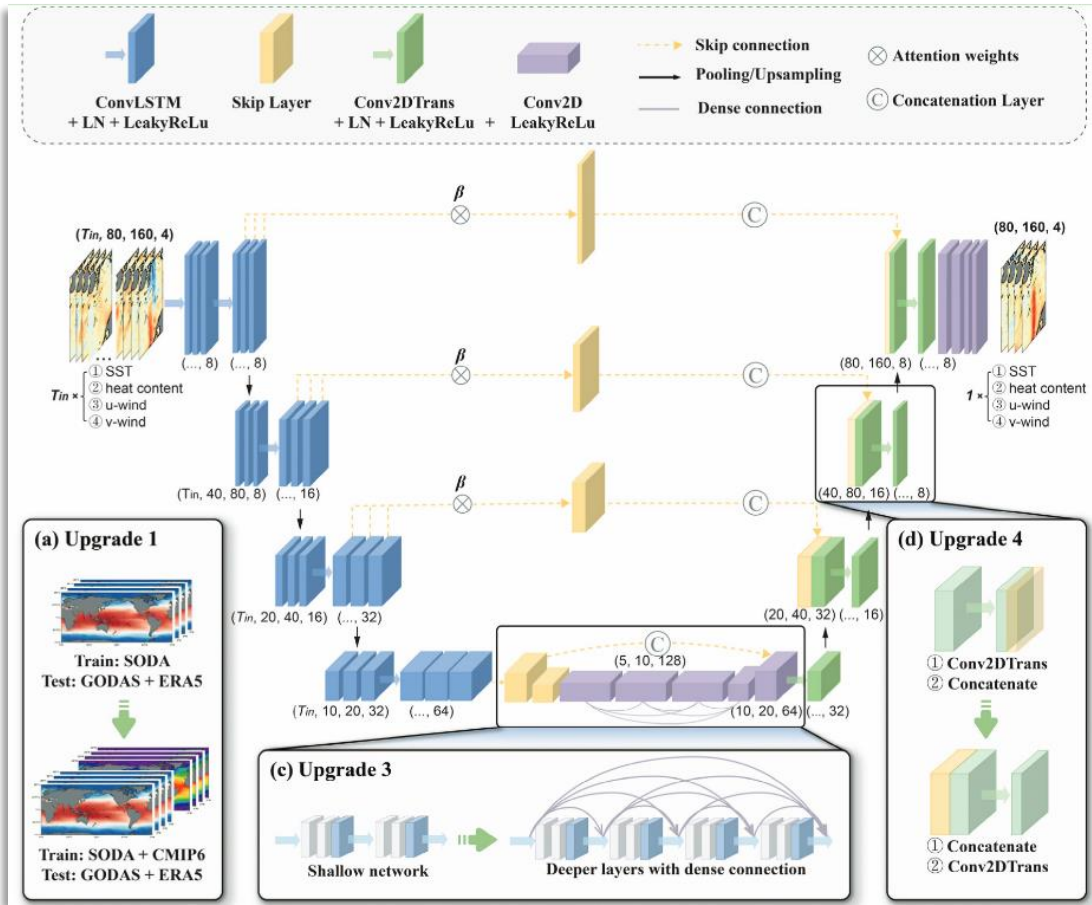
Illustrations

- Typhoons tend to **move along the boundary of WPSH**.
- Our identified OGIE induces positive potential vorticity anomalies at the boundary of WPSH, **significantly weakening WPSH and causing it to retreat eastward**.
- The Typhoon then moves northward along the retreating boundary of WPSH, leading to the largest track forecast errors.

Qin et al., 2026, QJRMS

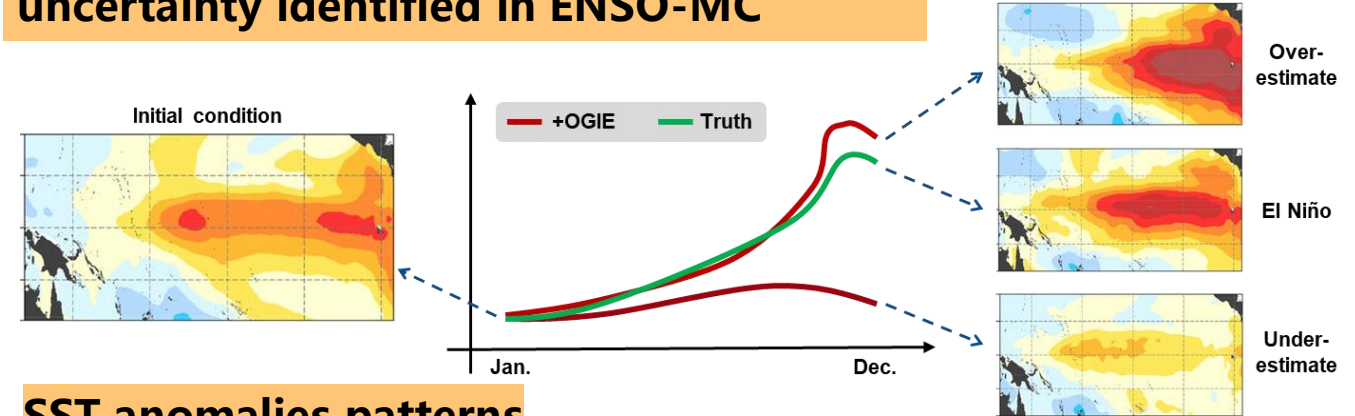
Climate Events: Intensity of Two Types of El Niño

ENSO-MC model

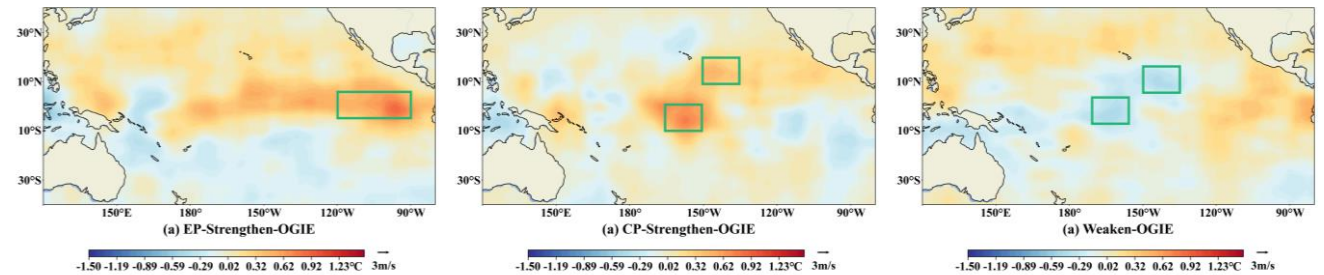


Mu et al., 2022, GMD; Qin et al., 2024, QJRM

The OGIEs that cause the largest intensity uncertainty identified in ENSO-MC



SST anomalies patterns

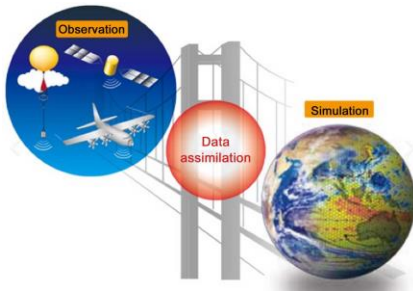


Illustrations

- Our identified OGIEs can significantly lead to **overestimation or underestimation** of two types of El Niño events.
- We superimpose the identified OGIEs into the numerical model (GFDL CM2p1), and they can induce **error evolutions very similar to those in AI models**

- We have identified such initial errors that **can significantly affect forecast accuracy**
- These initial errors can be reduced through a variety of means, thereby **enhancing forecast skill**

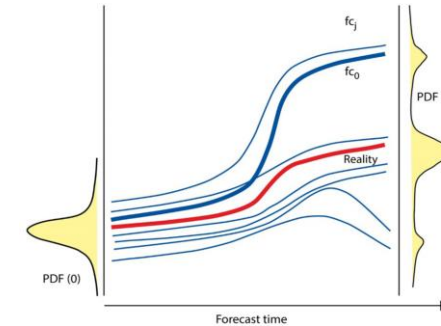
e.g., **Data Assimilation**



Targeted Observation



Ensemble Forecast



etc.

More importantly

- By tracing the evolution of such initial errors in AI models, **certain latent mechanisms hidden within BIG DATA can be uncovered**

TianXing-S2S has been invited to participate in the 2nd Artificial Intelligence Meteorological Model Forecast Demonstration Project (AIM-FDP 2.0)

Thanks for your attention!

 arXiv paper: <https://doi.org/10.48550/arXiv.2512.12545>

Bin Mu

School of Computer Science and Technology, [Tongji University](#)

2026.05.15 NanJing, China

