AI-Driven Satellite Data Assimilation

Wei Han（韩 威）
hanwei@cma.gov.cn
Xiaoze Xu, Zeting Li and Yonghui Li

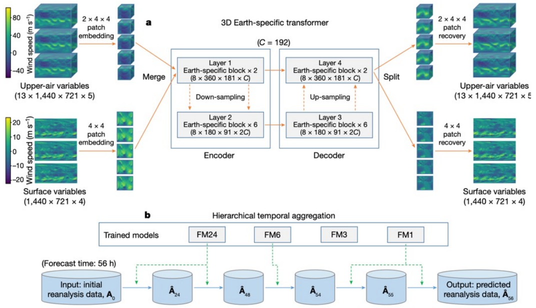
8th - 14th May 2025 at Resort Rio, Goa, India

CONTENT

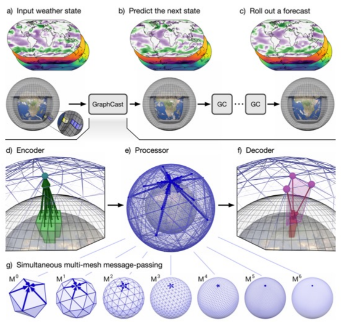
Background

Deterministic Forecasting

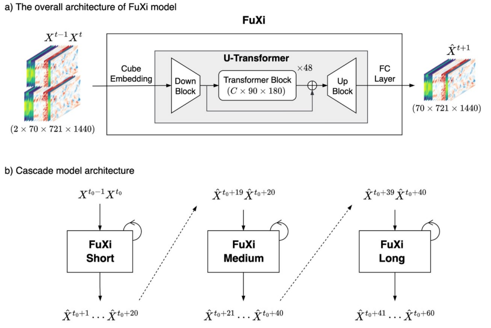
PanGu (Bi et al., 2023)



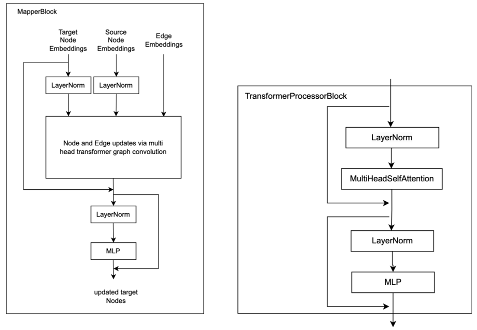
GraphCast (lam et al., 2023)



FuXi (Chen et al., 2023)

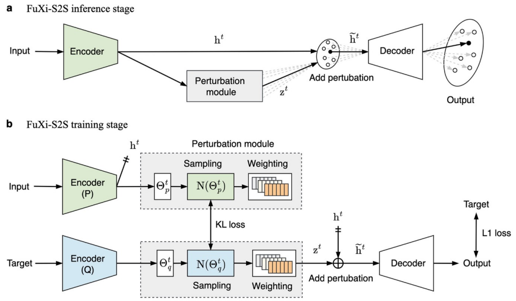


AIFS (Lang et al., 2023)

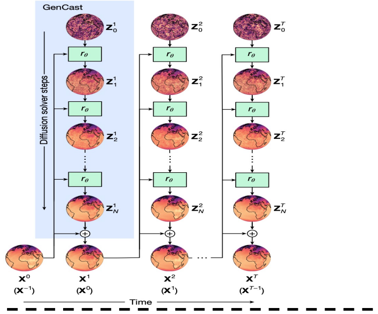


Ensemble Forecast

FuXi-S2S (Chen et al., 2024)

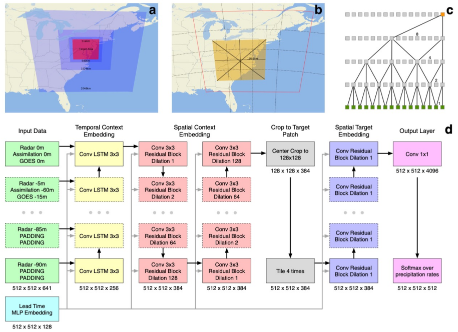


GenCast (price et al., 2025)

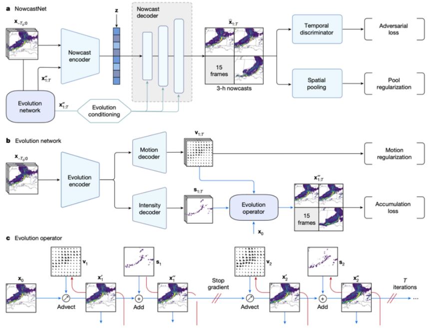


Precipitation Forecast

MetNet-2 (Espeholt et al., 2023)



NowcastNet (Zhang et al., 2023)

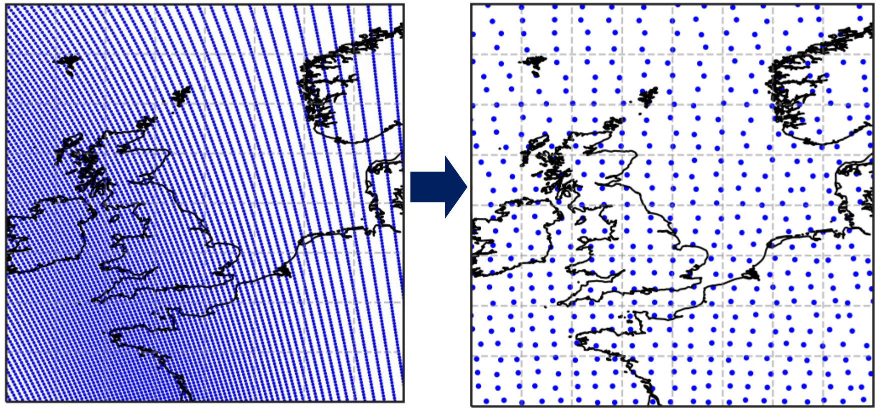


Deep learning-based weather forecast models surpass NWP models

Both NWP and DL-based weather forecasting models rely on the initial fields

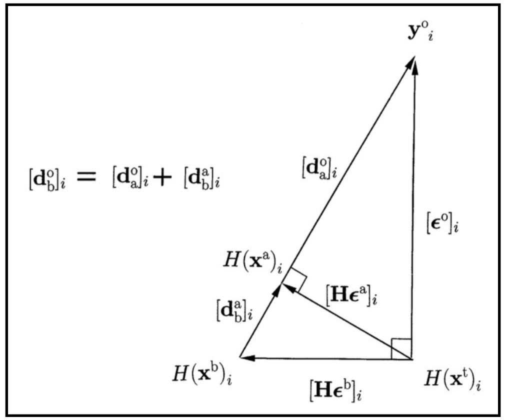
Background

Thinning/Selection



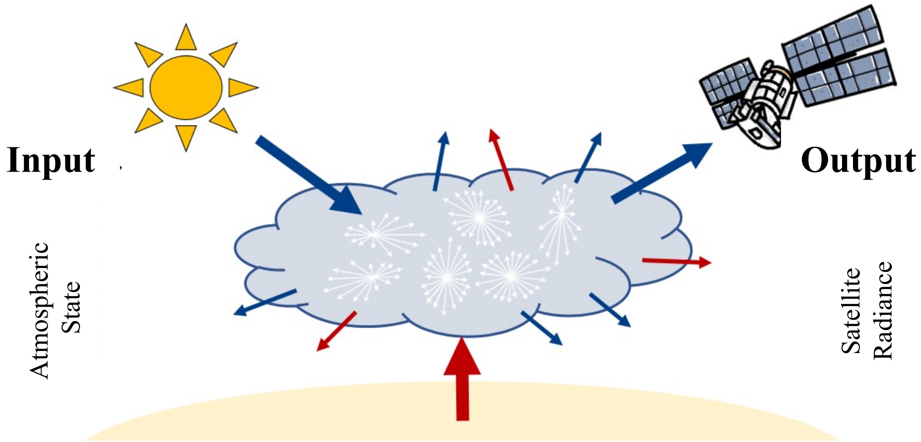
Duncan et al., 2024

Observation Errors



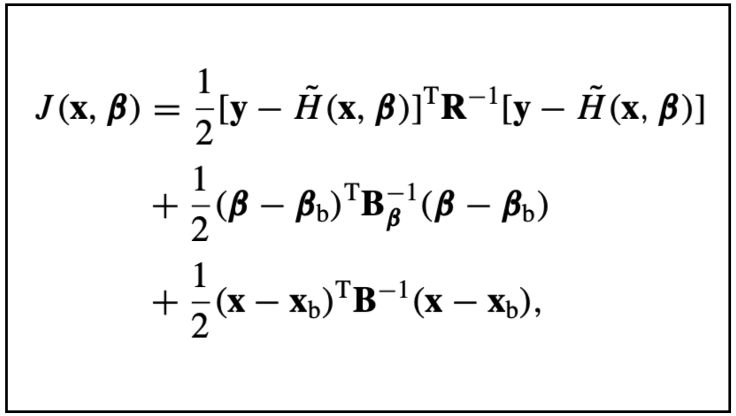
Desroziers et al., 2005

Observation Operator



e.g., RTTOV

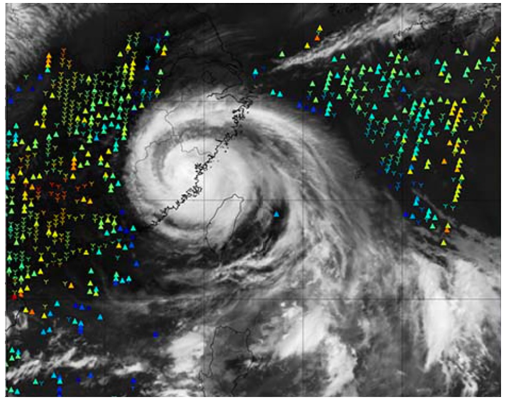
Bias Correction



$$J(\mathbf{x}, \boldsymbol{\beta}) = \frac{1}{2}[\mathbf{y} - \tilde{H}(\mathbf{x}, \boldsymbol{\beta})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y} - \tilde{H}(\mathbf{x}, \boldsymbol{\beta})] + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}_{\mathrm{b}})^{\mathrm{T}}\mathbf{B}_{\boldsymbol{\beta}}^{-1}(\boldsymbol{\beta} - \boldsymbol{\beta}_{\mathrm{b}}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_{\mathrm{b}}),$$

Auligne´ et al., 2007

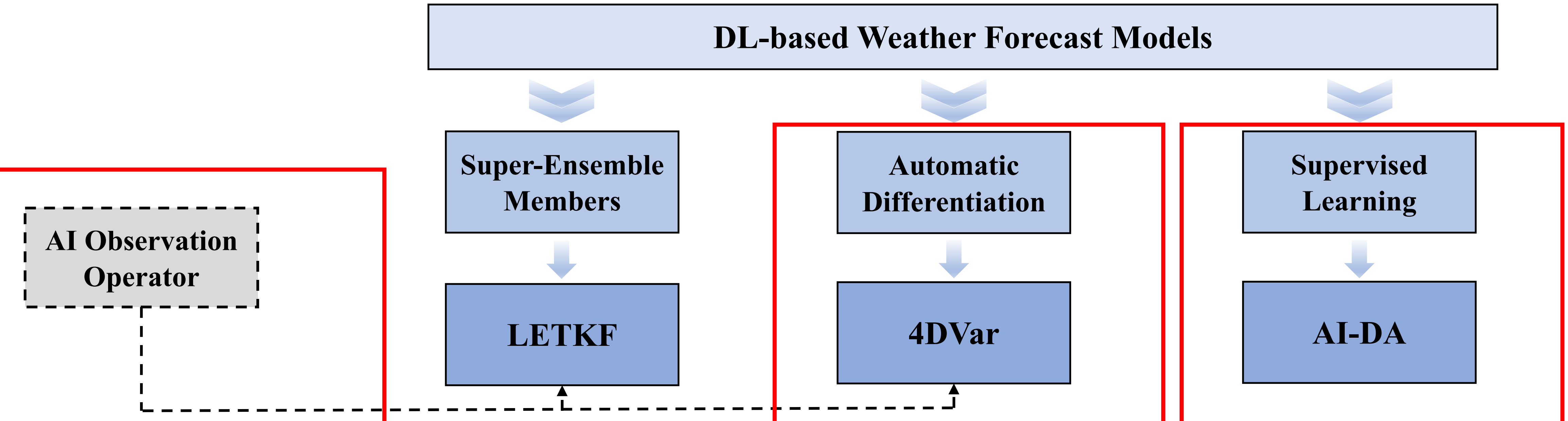
Cloud Detection (Partial Obervation)



Li et al., 2019

The data processing is complex, and many useful observations are discarded

AI in Satellite Data Assimilation



DL-based Weather Forecast Models

Super-Ensemble Members

LETKF

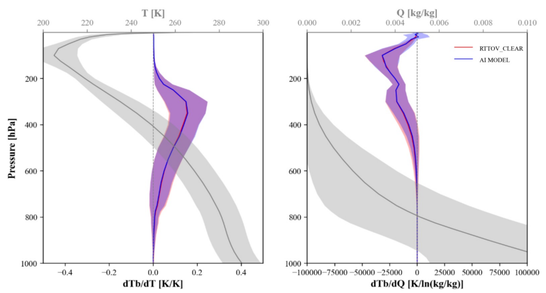
Automatic Differentiation

4DVar

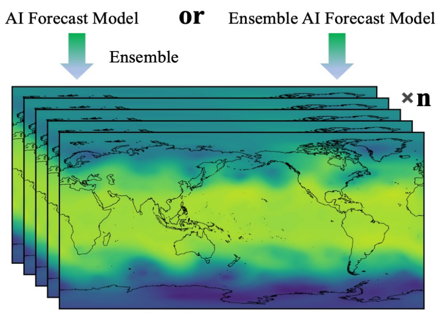
Supervised Learning

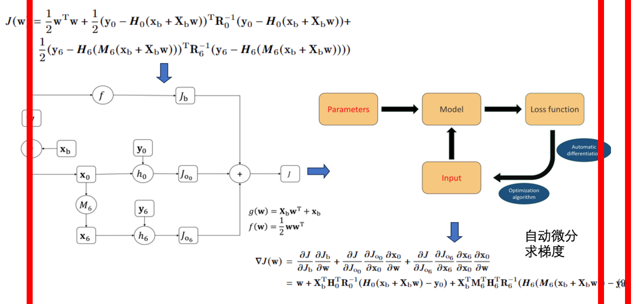
AI-DA

AI Observation Operator

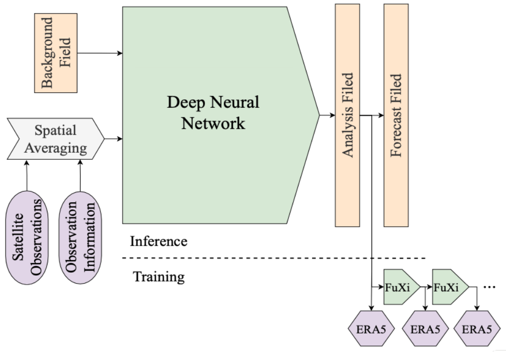


All-sky AI _RTM



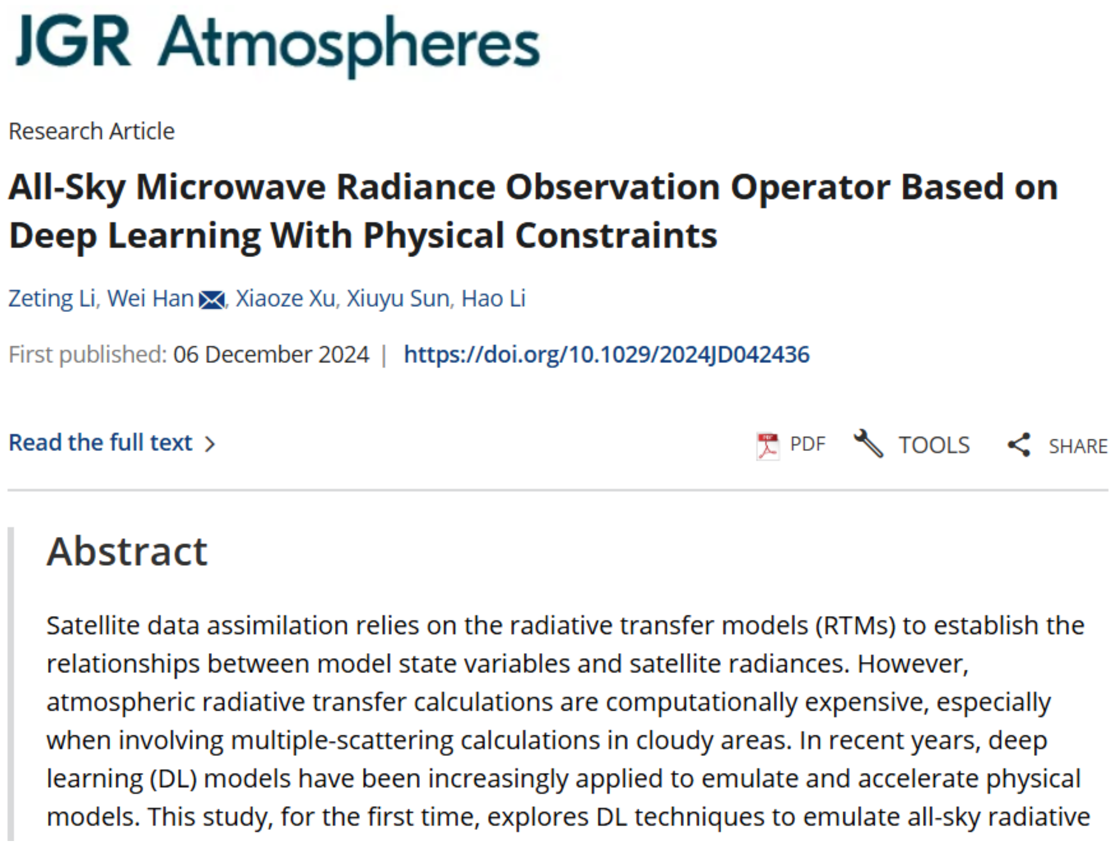


AI-4DVar



AI-DA

All-sky AI-RTM



JGR Atmospheres

Research Article

All-Sky Microwave Radiance Observation Operator Based on Deep Learning With Physical Constraints

Zeting Li, Wei Han ✉, Xiaoze Xu, Xiuyu Sun, Hao Li

First published: 06 December 2024 | https://doi.org/10.1029/2024JD042436

Read the full text ›

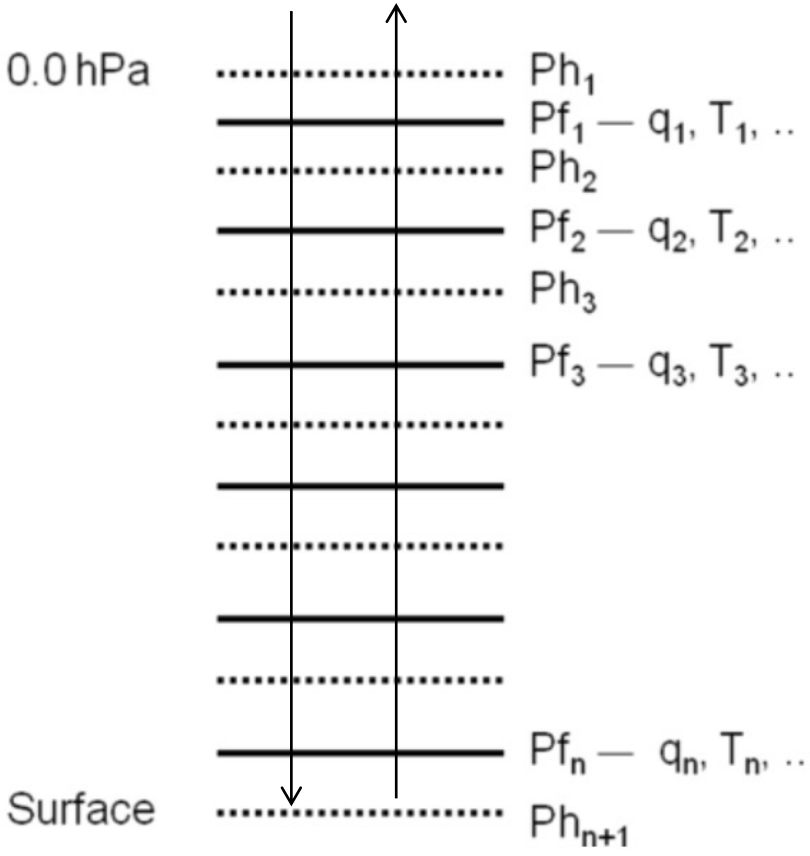
PDF TOOLS SHARE

Abstract

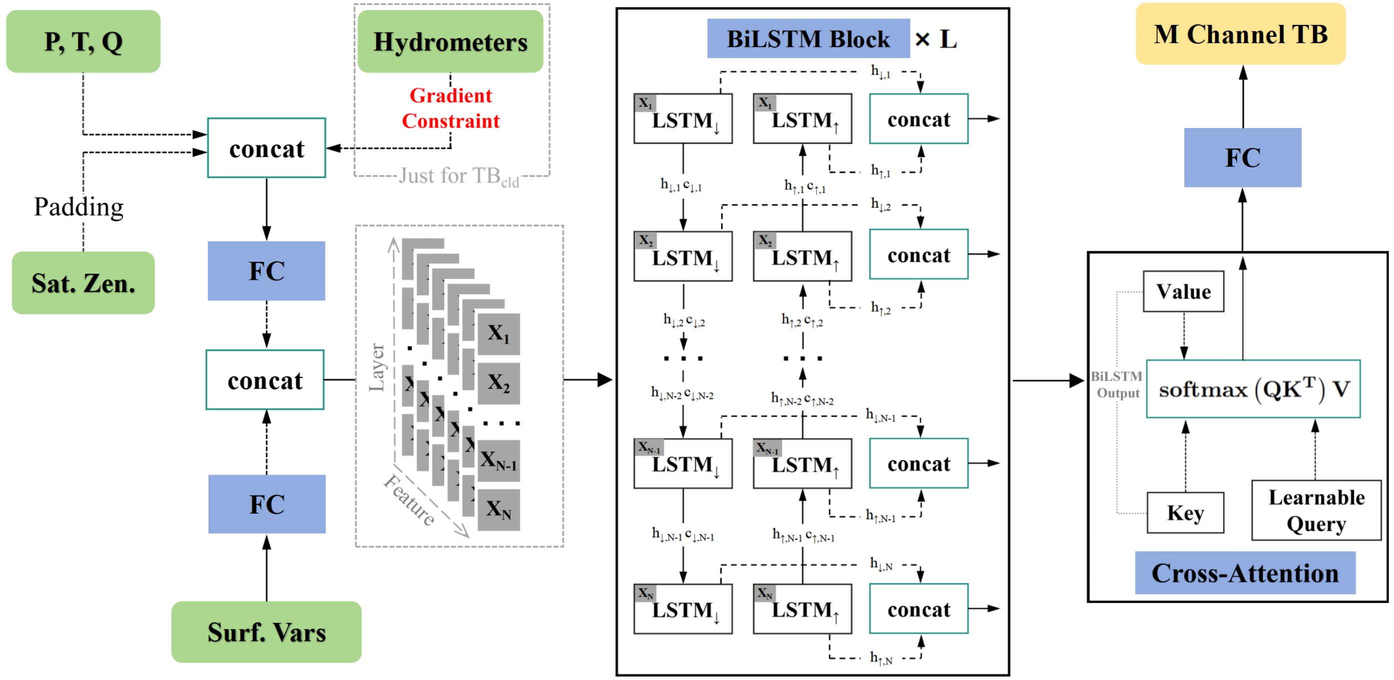
Satellite data assimilation relies on the radiative transfer models (RTMs) to establish the relationships between model state variables and satellite radiances. However, atmospheric radiative transfer calculations are computationally expensive, especially when involving multiple-scattering calculations in cloudy areas. In recent years, deep learning (DL) models have been increasingly applied to emulate and accelerate physical models. This study, for the first time, explores DL techniques to emulate all-sky radiative

All-sky AI-RTM

Radiative Transfer Calculation



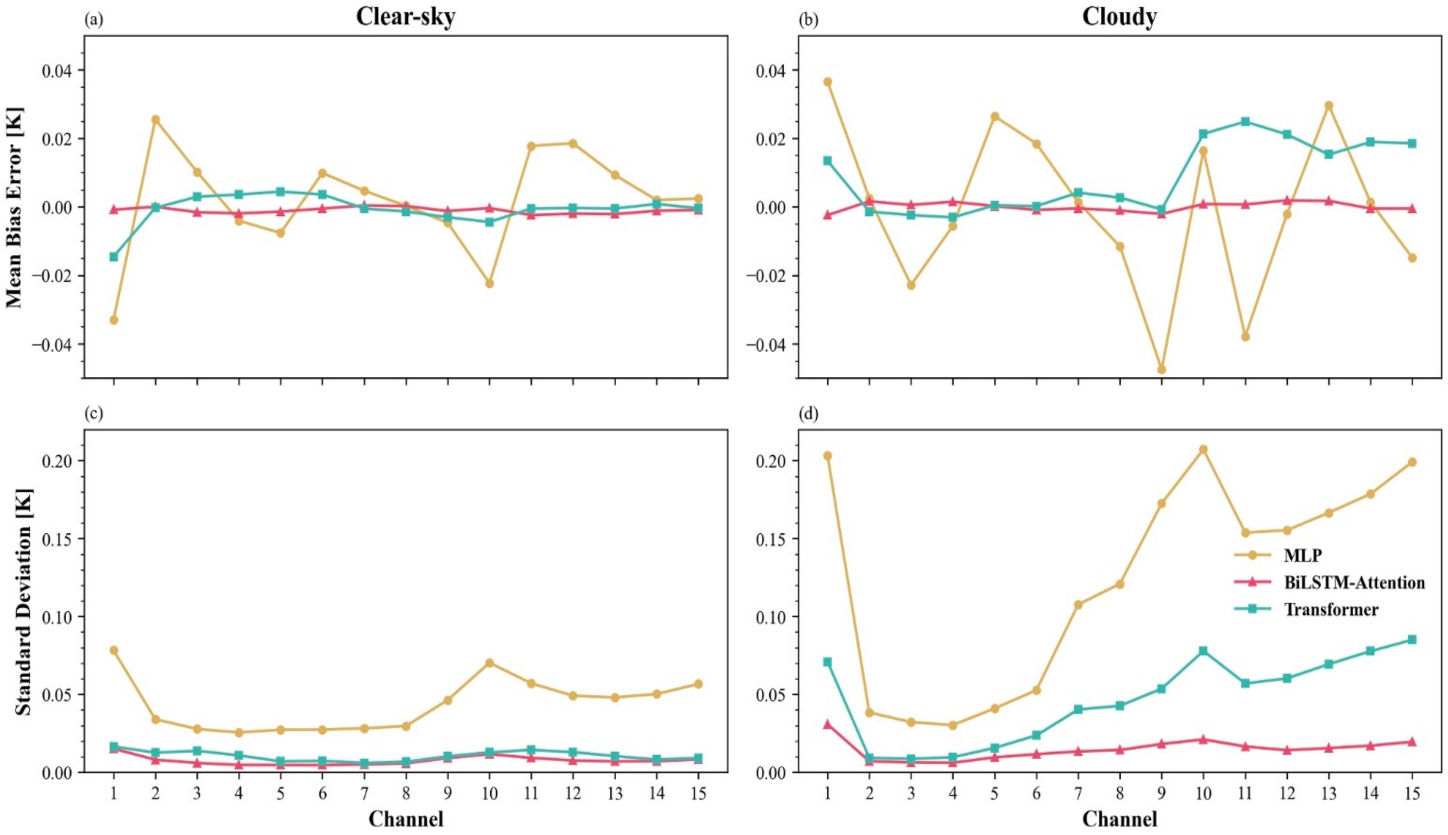
Architecture of AI-RTM



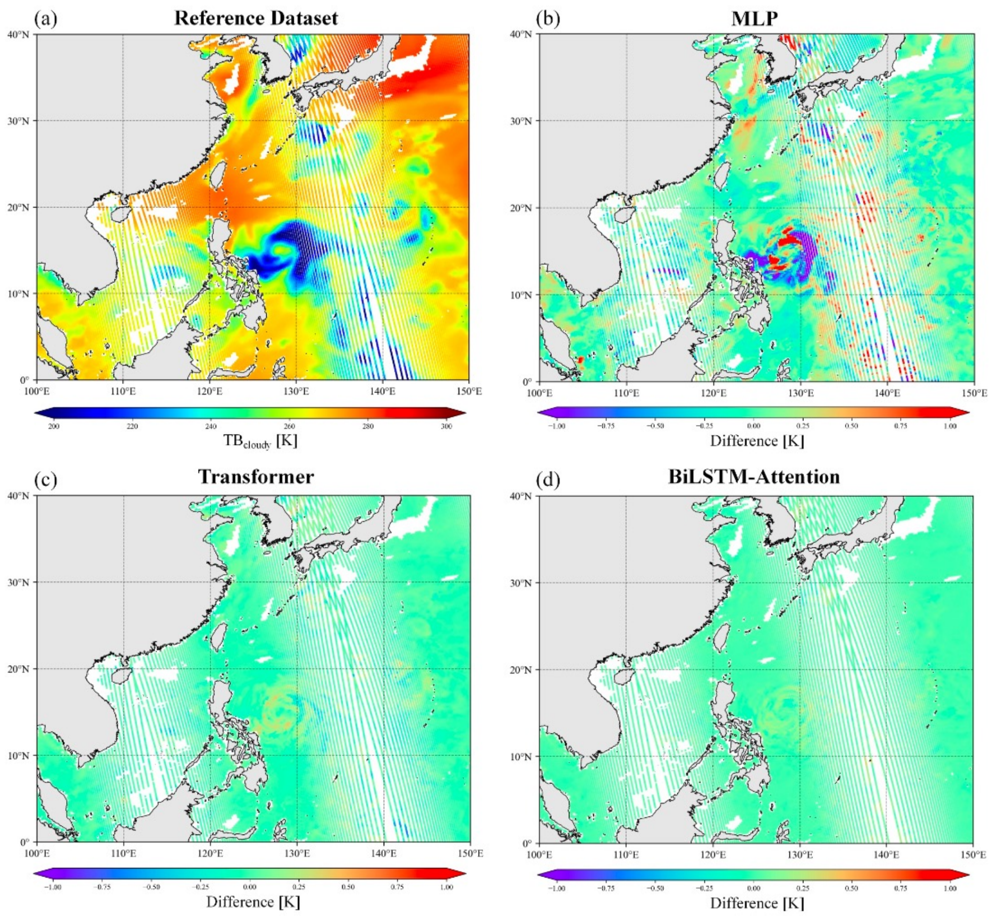
- LSTM captures the relationships between layers
- The learnable Q matrix captures the sensitivity of channels to different layers.

All-sky AI-RTM

Simulation errors of each channel

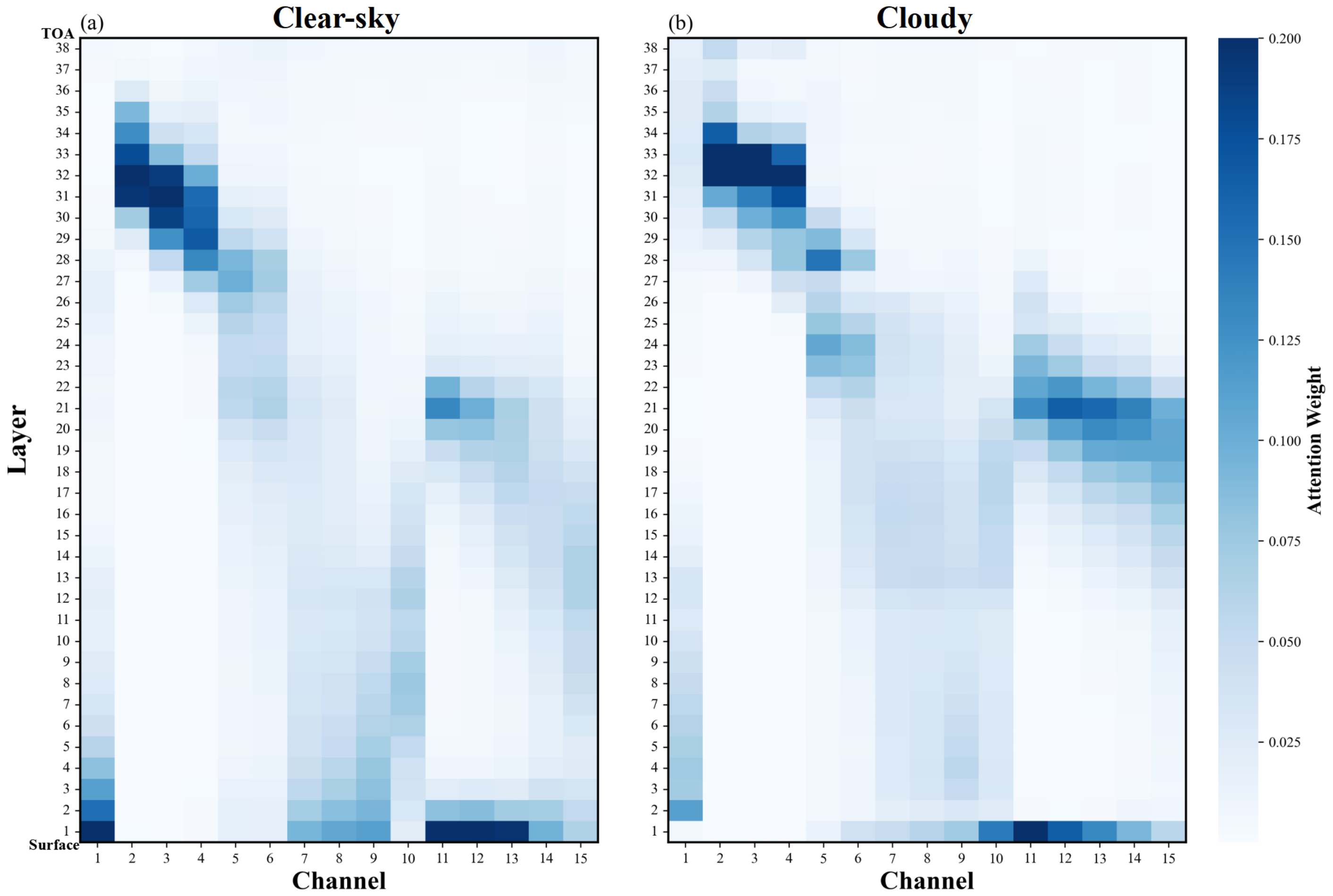


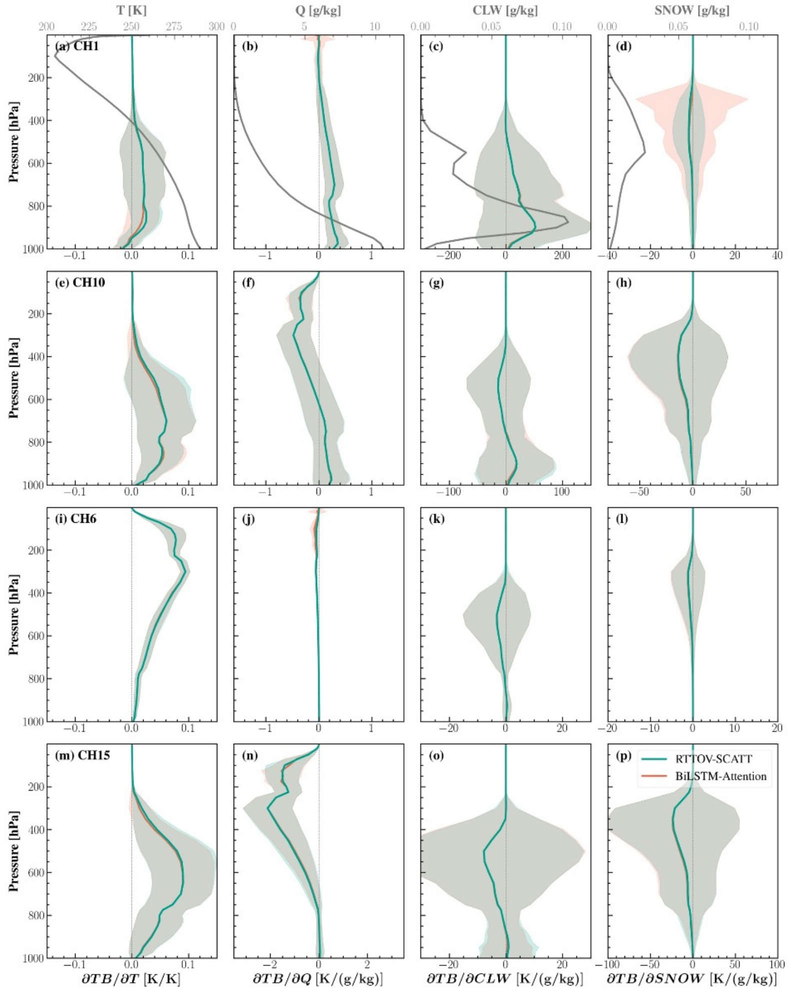
Spatial distribution of simulation errors



The computational efficiency has improved by 20 times, with over 99.5% of the differences within 0.01K.

All-sky AI-RTM





- The Attention weights derived from the AI-RTM resembled the weighting functions of MWHS-2 channels

- The Jacobians derived from the AI-RTM are stable and accurate

AI-RTM (Considering Uncertainty)

JGR Machine Learning and Computation

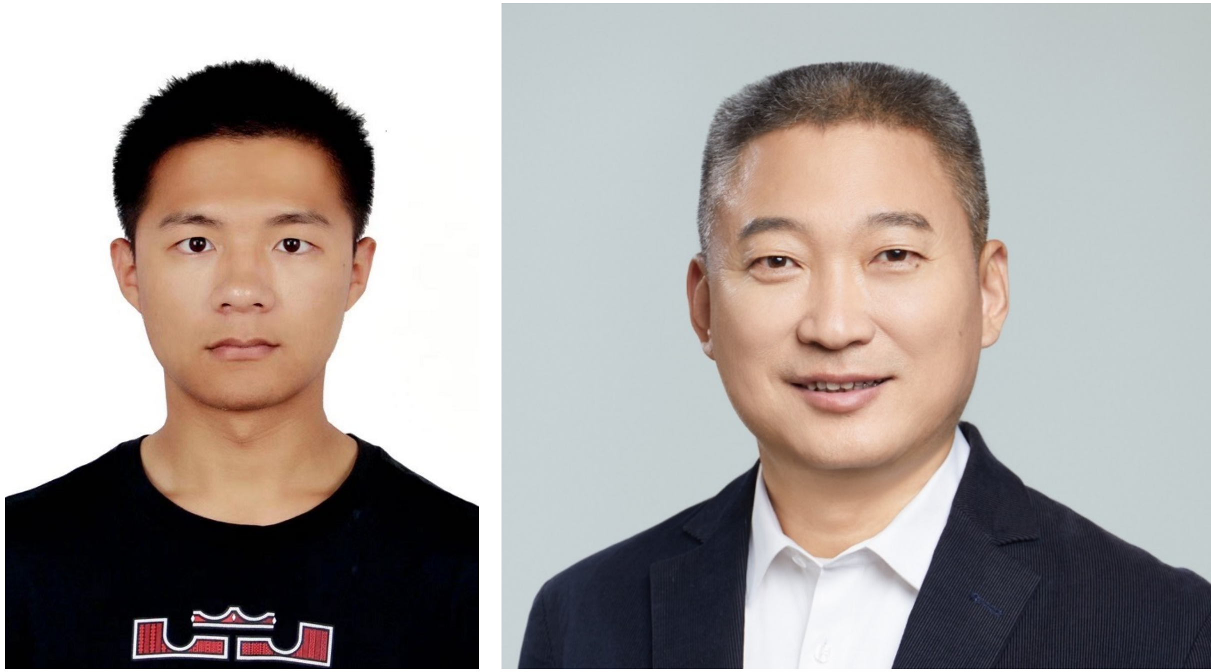
Research Article | Open Access

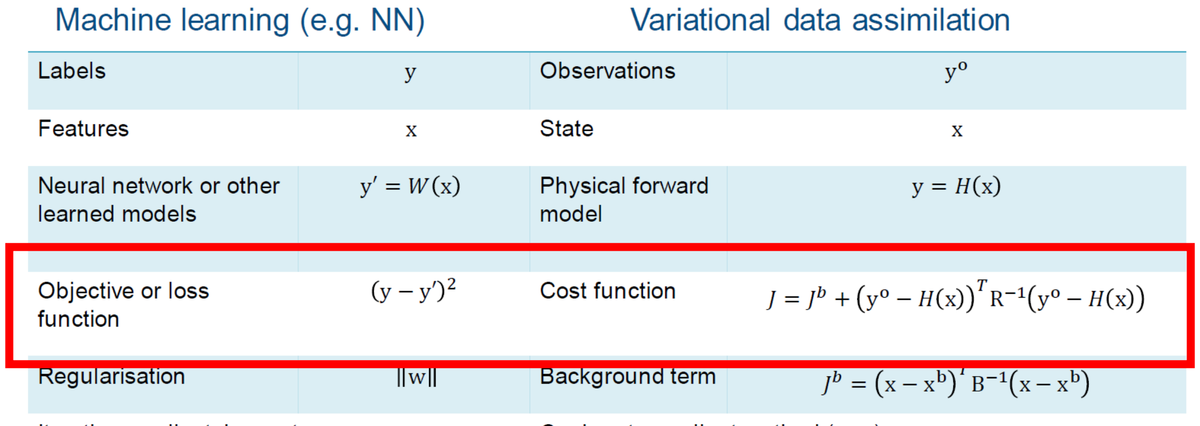
A Machine Learning-Based Observation Operator for FY-4B GIIRS Brightness Temperatures Considering the Uncertainty of Label Data

Yonghui Li, Wei Han, Wansuo Duan, Zeting Li, Hao Li

First published: 19 March 2025 | https://doi.org/10.1029/2024JH000449

SECTIONS PDF TOOLS SHARE



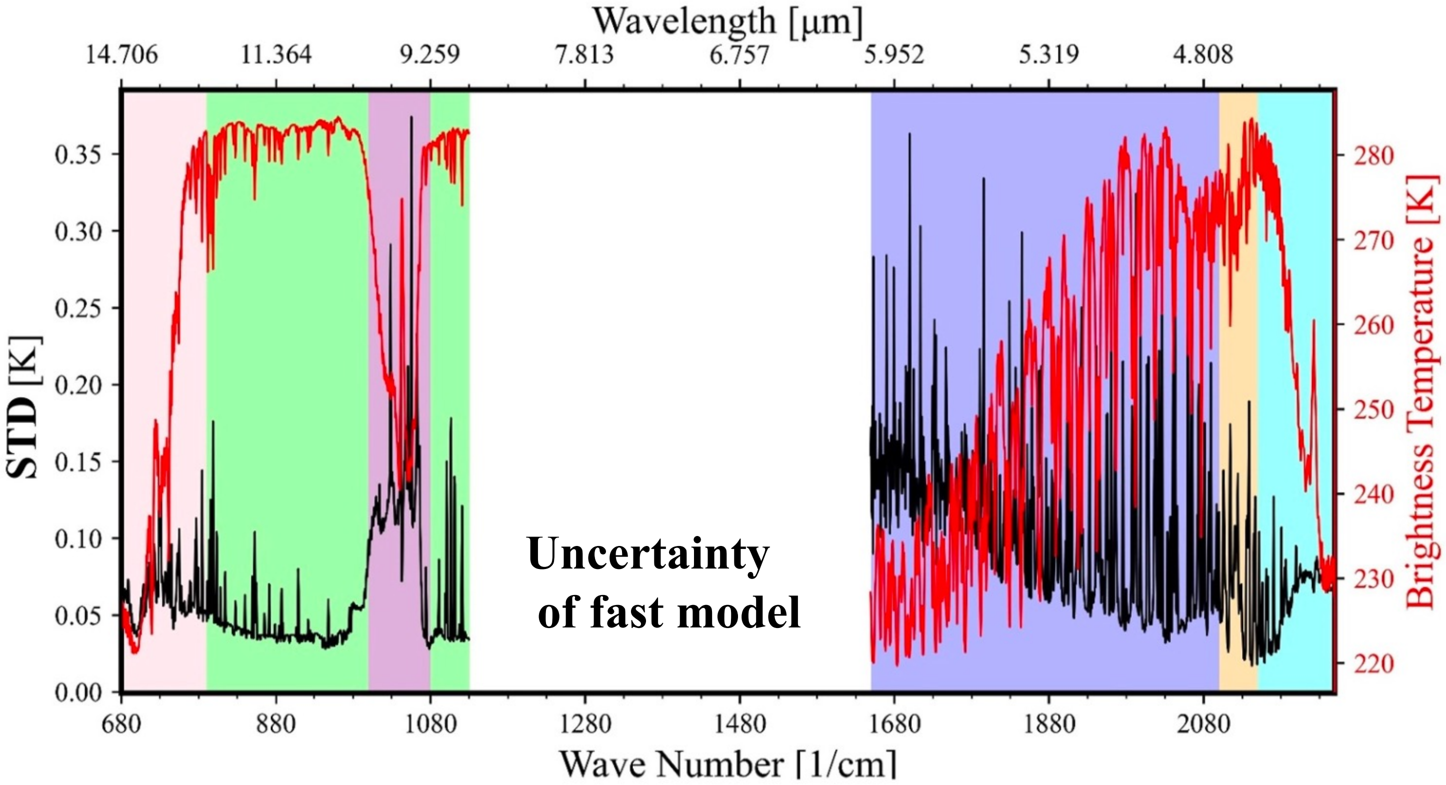


Machine learning (e.g. NN)		Variational data assimilation	
Labels	y	Observations	y^o
Features	x	State	x
Neural network or other learned models	$y' = W(x)$	Physical forward model	$y = H(x)$
Objective or loss function	$(y - y')^2$	Cost function	$J = J^b + (y^o - H(x))^T R^{-1}(y^o - H(x))$
Regularisation	$\|w\|$	Background term	$J^b = (x - x^b)^T B^{-1}(x - x^b)$

Use DA methodology to improve ML training

Alan Geer, 2022

AI-RTM (Considering Uncertainty)



The STD, used to quantify uncertainty, of the statistics. The black line represents the square root of the diagonal elements of the Σ matrix, the red line represents the BT

- The loss function, which considers the simulation uncertainty of each channel, is used to optimize the weighted loss for each channel.

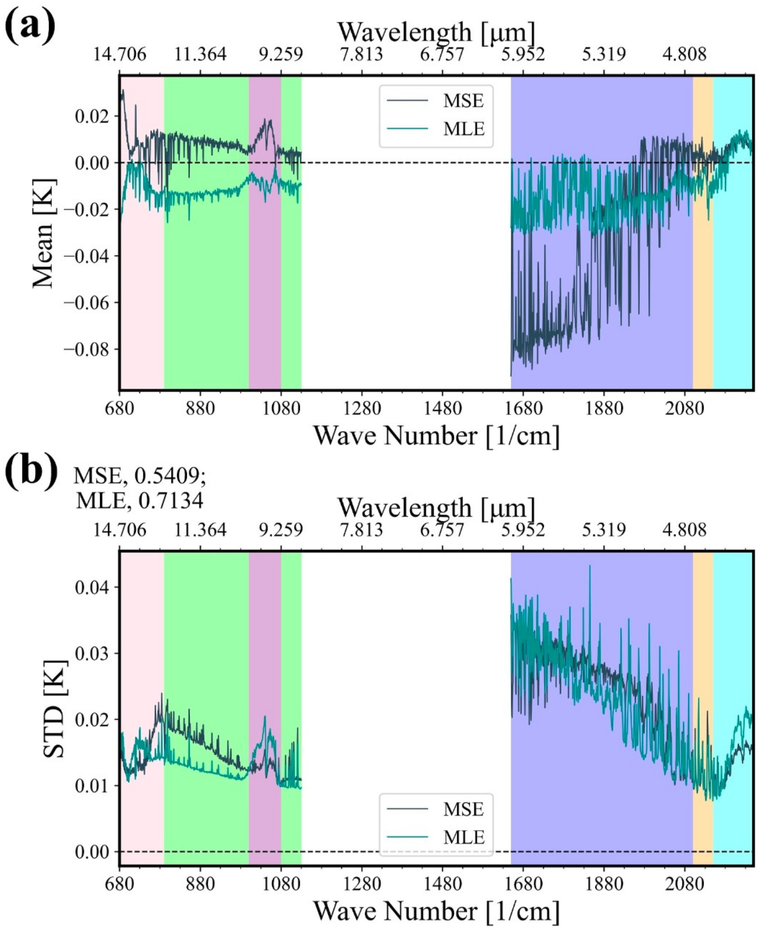
$$L(\theta;y,x) = \frac{1}{\sqrt{2\pi|\Sigma|}}\exp\left(-\frac{(y-\hat{y}(\theta,x))^T\Sigma^{-1}(y-\hat{y}(\theta,x))}{2}\right)$$

$$\Downarrow$$

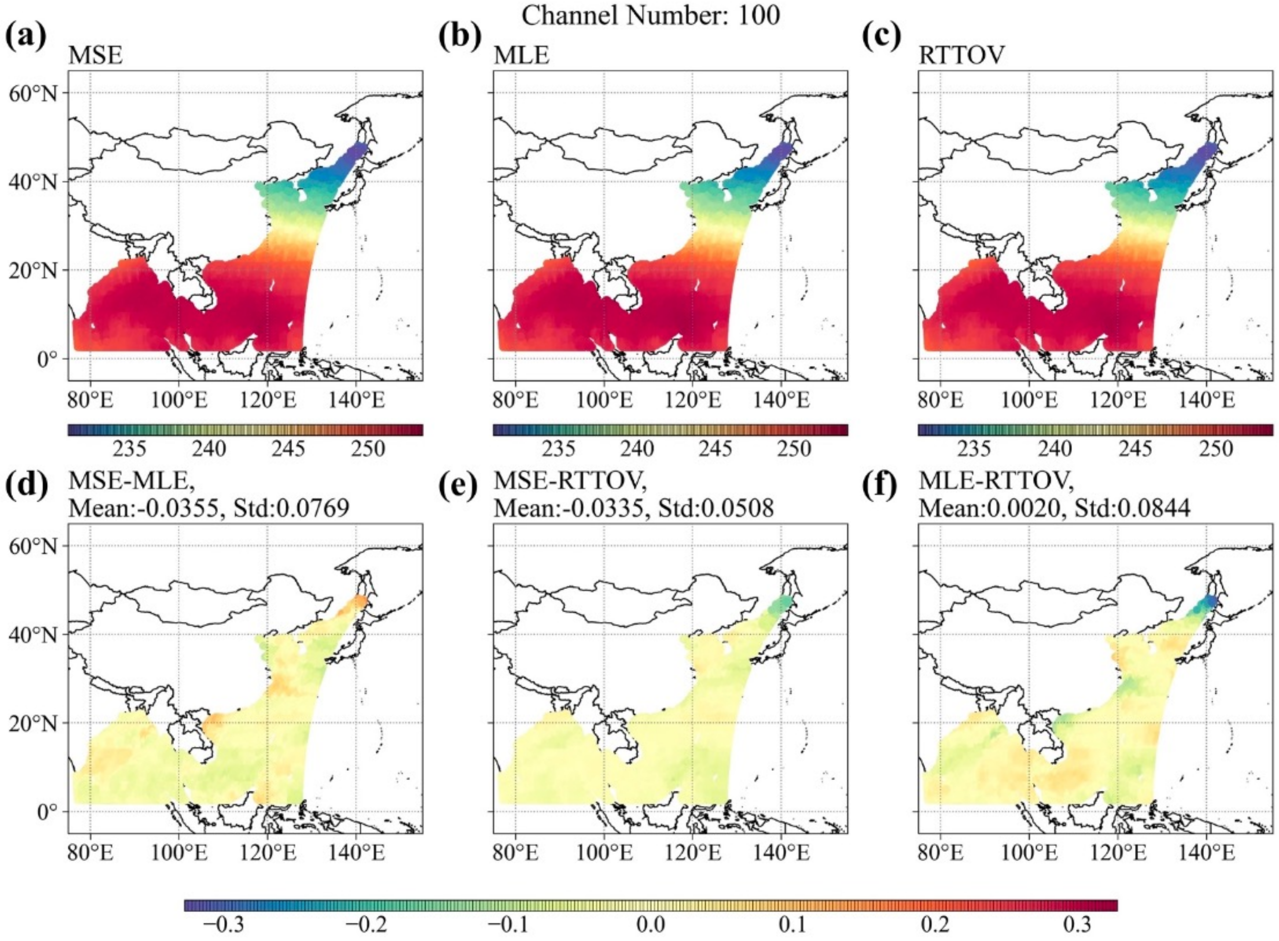
$$J_{MLE}(\theta) = (y-\hat{y}(\theta,x))^T\Sigma^{-1}(y-\hat{y}(\theta,x))$$

AI-RTM (Considering Uncertainty)

Simulation errors of each channel

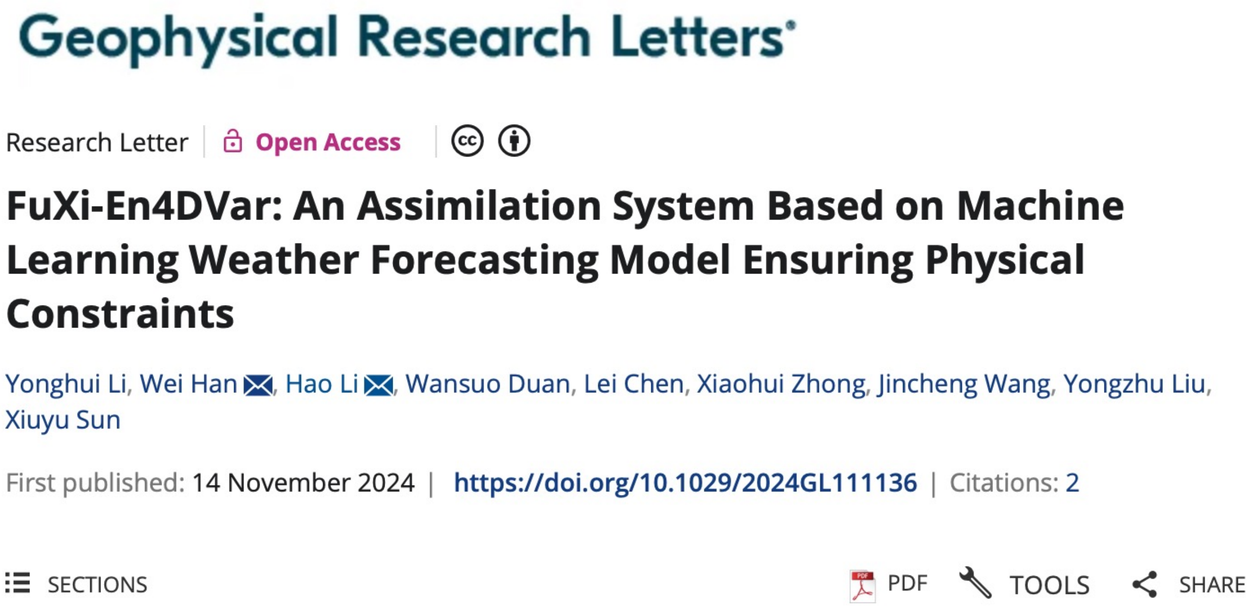


Spatial distribution of simulation errors



Compared to the MSE loss, the model using the MLE(weighted) loss shows smaller simulation errors.

FuXi-En4DVar



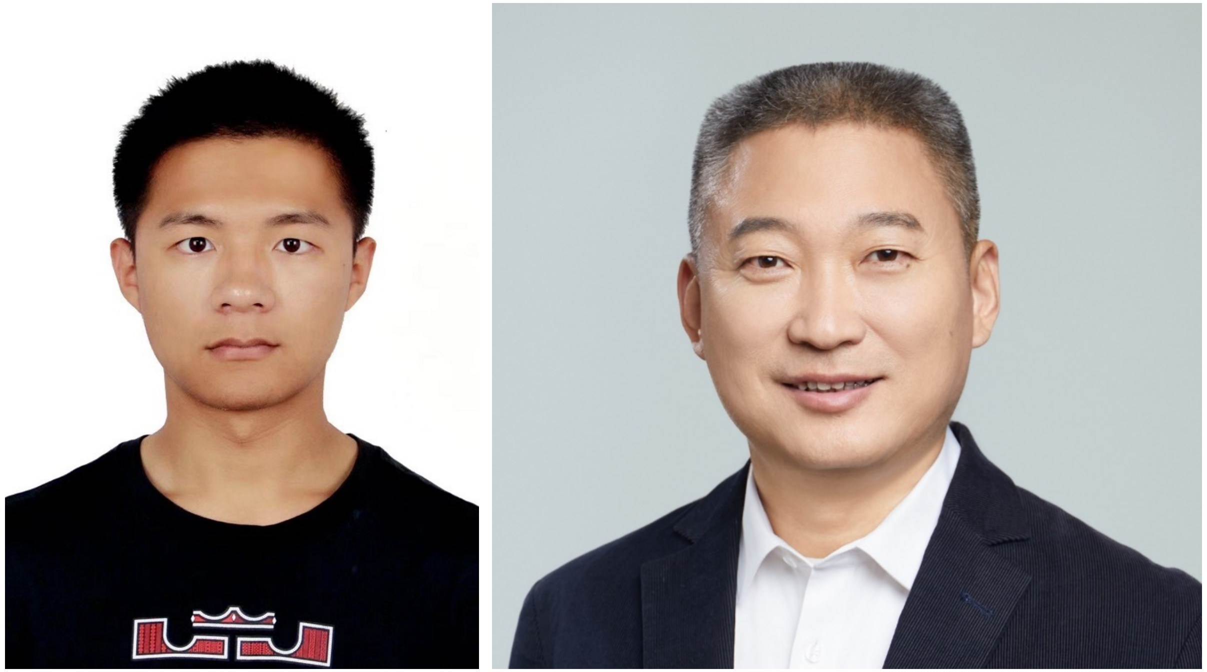
Geophysical Research Letters

Research Letter | Open Access

FuXi-En4DVar: An Assimilation System Based on Machine Learning Weather Forecasting Model Ensuring Physical Constraints

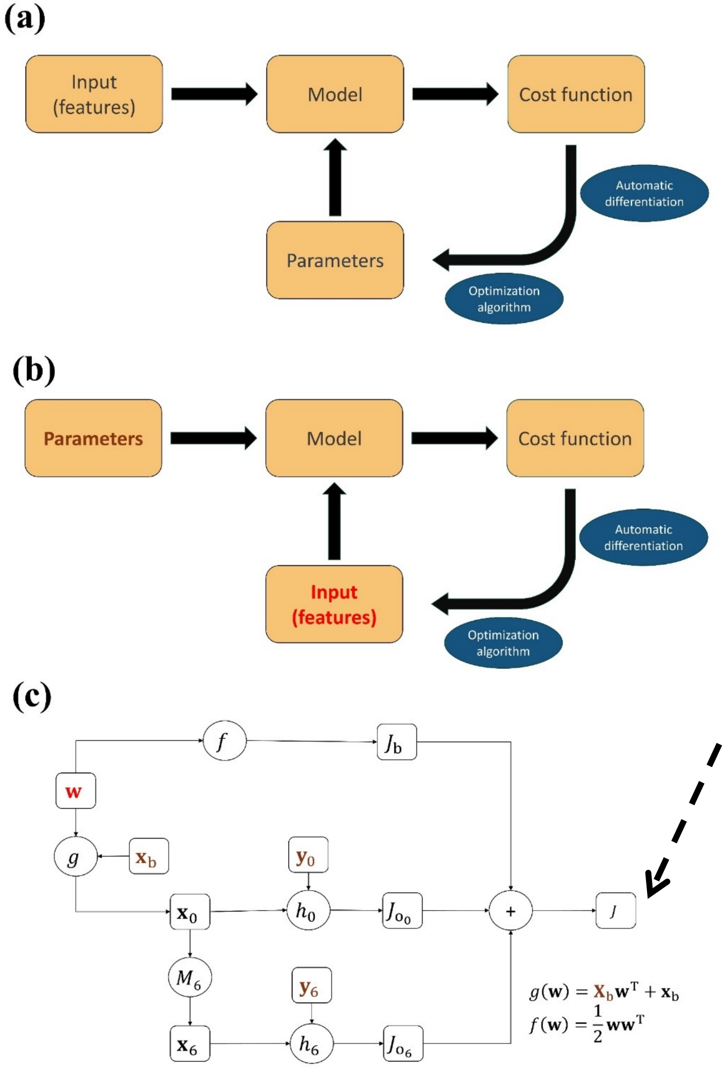
Yonghui Li, Wei Han ✉, Hao Li ✉, Wansuo Duan, Lei Chen, Xiaohui Zhong, Jincheng Wang, Yongzhu Liu, Xiuyu Sun

First published: 14 November 2024 | https://doi.org/10.1029/2024GL111136 | Citations: 2



FuXi-En4DVar

Architecture of FuXi-En4DVar

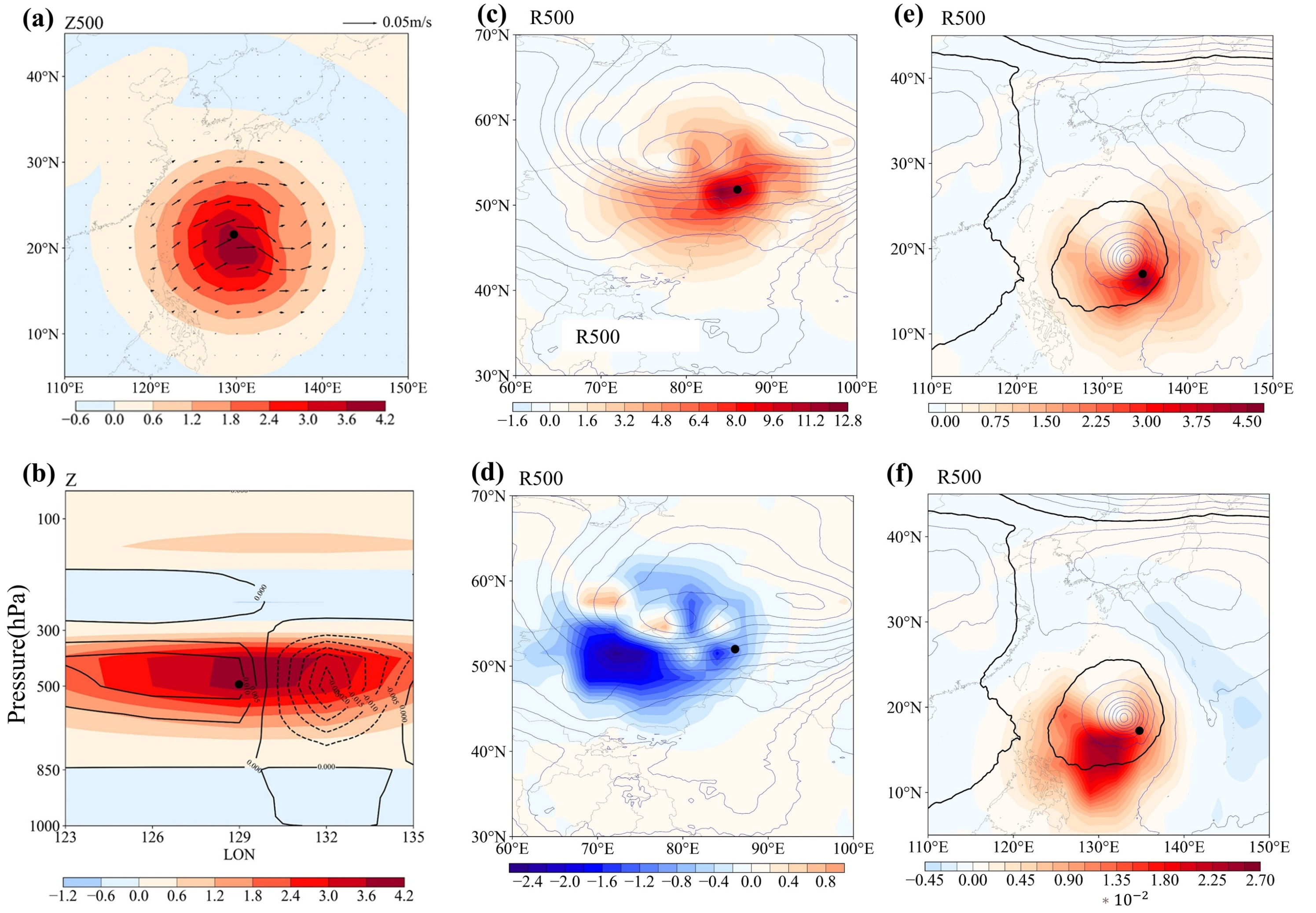


$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\mathbf{y}_0 - \boldsymbol{H}_0(\mathbf{x}_b + \mathbf{X}_b\mathbf{w}))^T\mathbf{R}_0^{-1}(\mathbf{y}_0 - \boldsymbol{H}_0(\mathbf{x}_b + \mathbf{X}_b\mathbf{w})) + \frac{1}{2}(\mathbf{y}_6 - \boldsymbol{H}_6(\boldsymbol{M}_6(\mathbf{x}_b + \mathbf{X}_b\mathbf{w})))^T\mathbf{R}_6^{-1}(\mathbf{y}_6 - \boldsymbol{H}_6(\boldsymbol{M}_6(\mathbf{x}_b + \mathbf{X}_b\mathbf{w}))))$$

- Based on the automatic differentiation property of the AI model, gradients are obtained without relying on traditional tangent-linear and adjoint models.

- Based on the fast inference capability of the AI model, a large number of ensemble members are generated to estimate the background error covariance matrix with flow-dependent features

FuXi-En4DVar

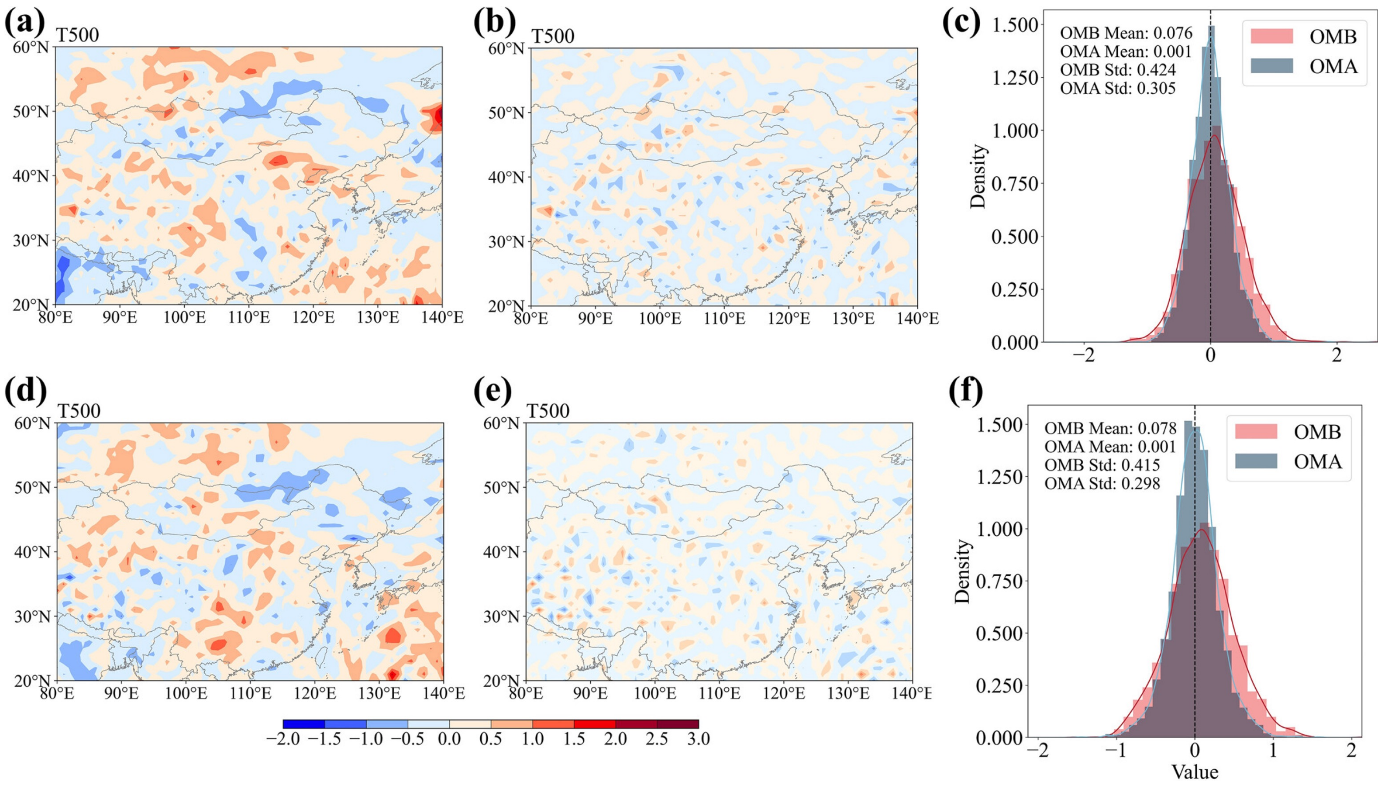


Single-observation test

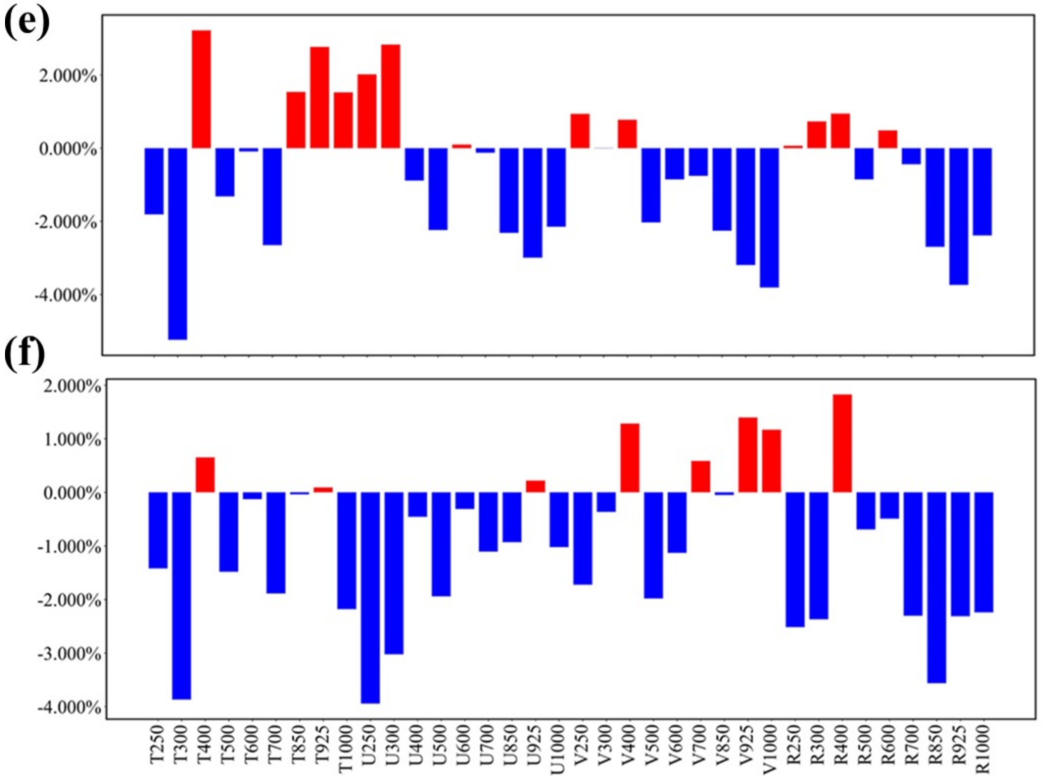
- FuXi-En4DVar assimilation system effectively adjusts the analysis increment within the constraints of physical balance.

FuXi-En4DVar

OMB and OMA

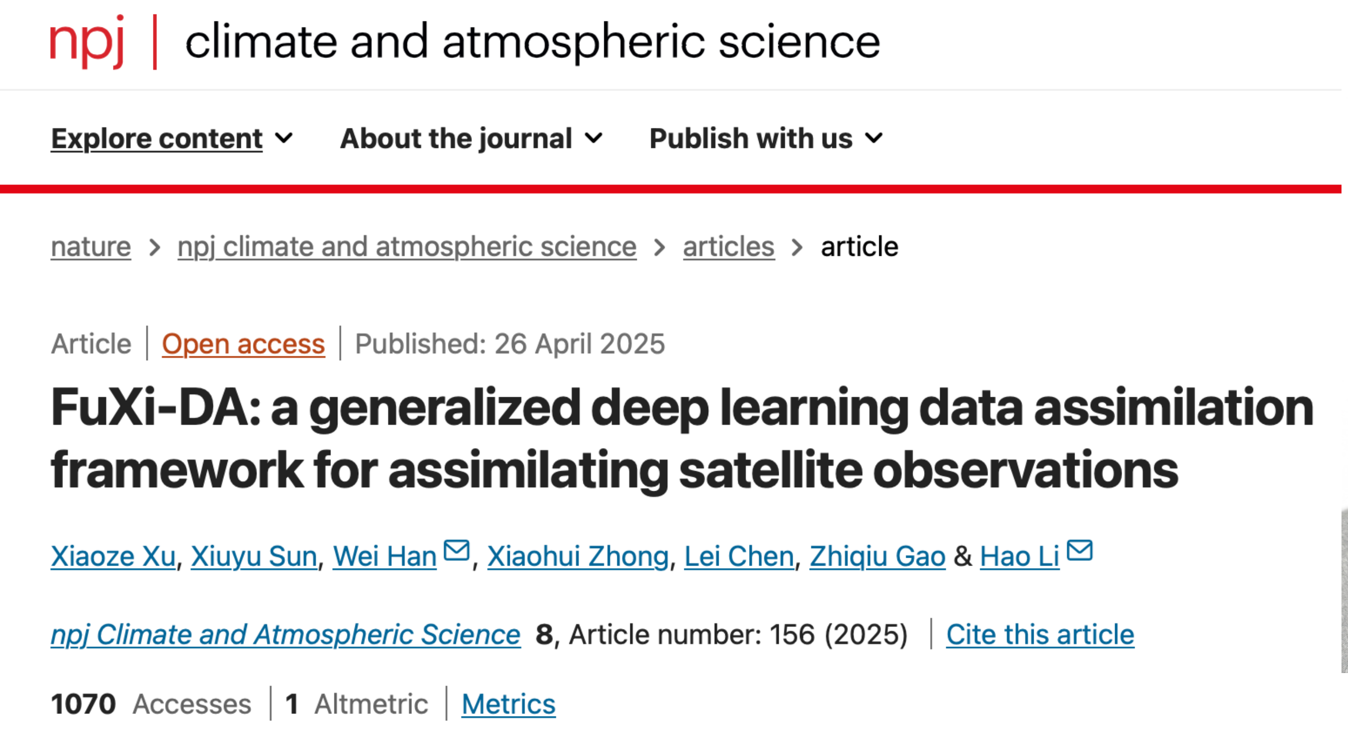


RMSE reduction rate after assimilating simulated observations



Analysis error is reduced after assimilating simulated observations.

FuXi-DA: End-To-End Assimilation



npj | climate and atmospheric science

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › npj climate and atmospheric science › articles › article

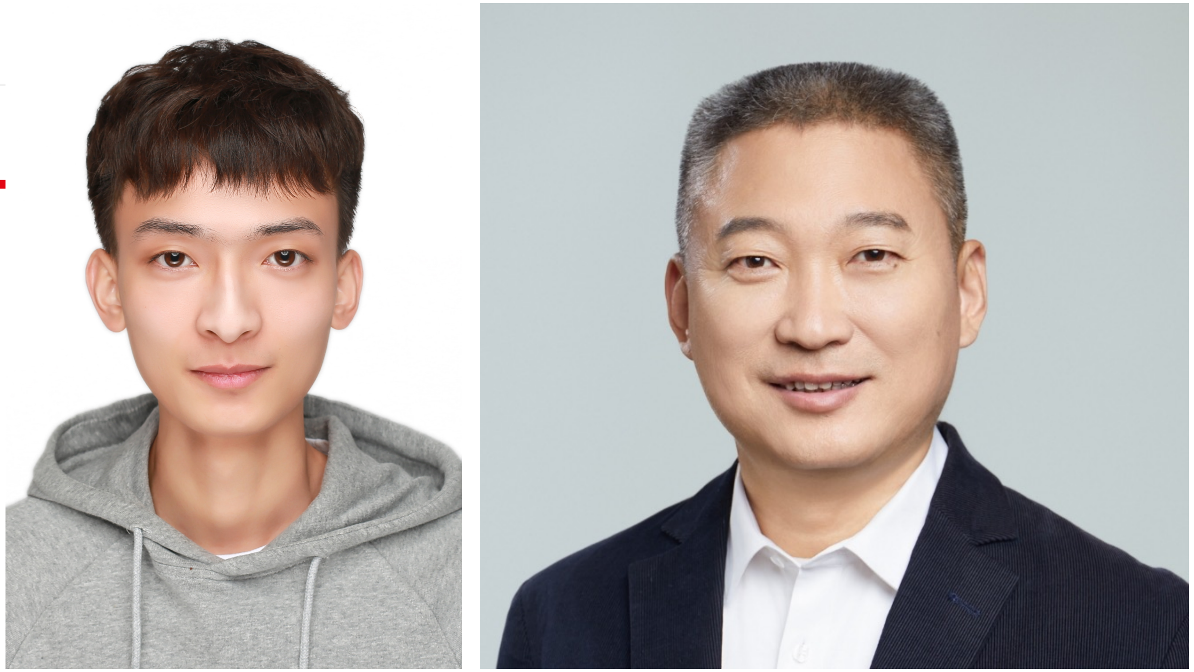
Article | Open access | Published: 26 April 2025

FuXi-DA: a generalized deep learning data assimilation framework for assimilating satellite observations

Xiaoze Xu, Xiuyu Sun, Wei Han ✉, Xiaohui Zhong, Lei Chen, Zhiqiu Gao & Hao Li ✉

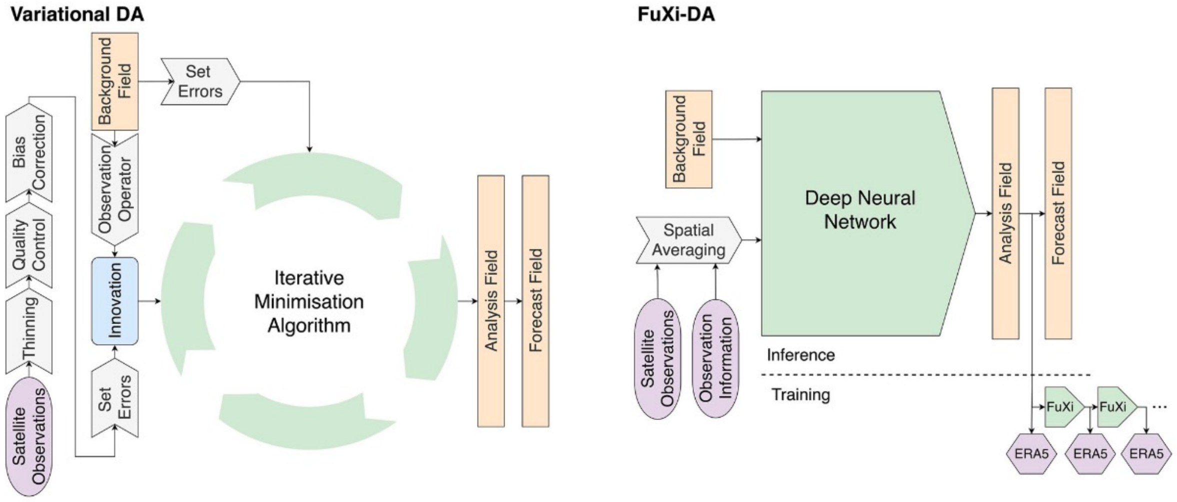
npj Climate and Atmospheric Science **8**, Article number: 156 (2025) | Cite this article

1070 Accesses | **1** Altmetric | Metrics



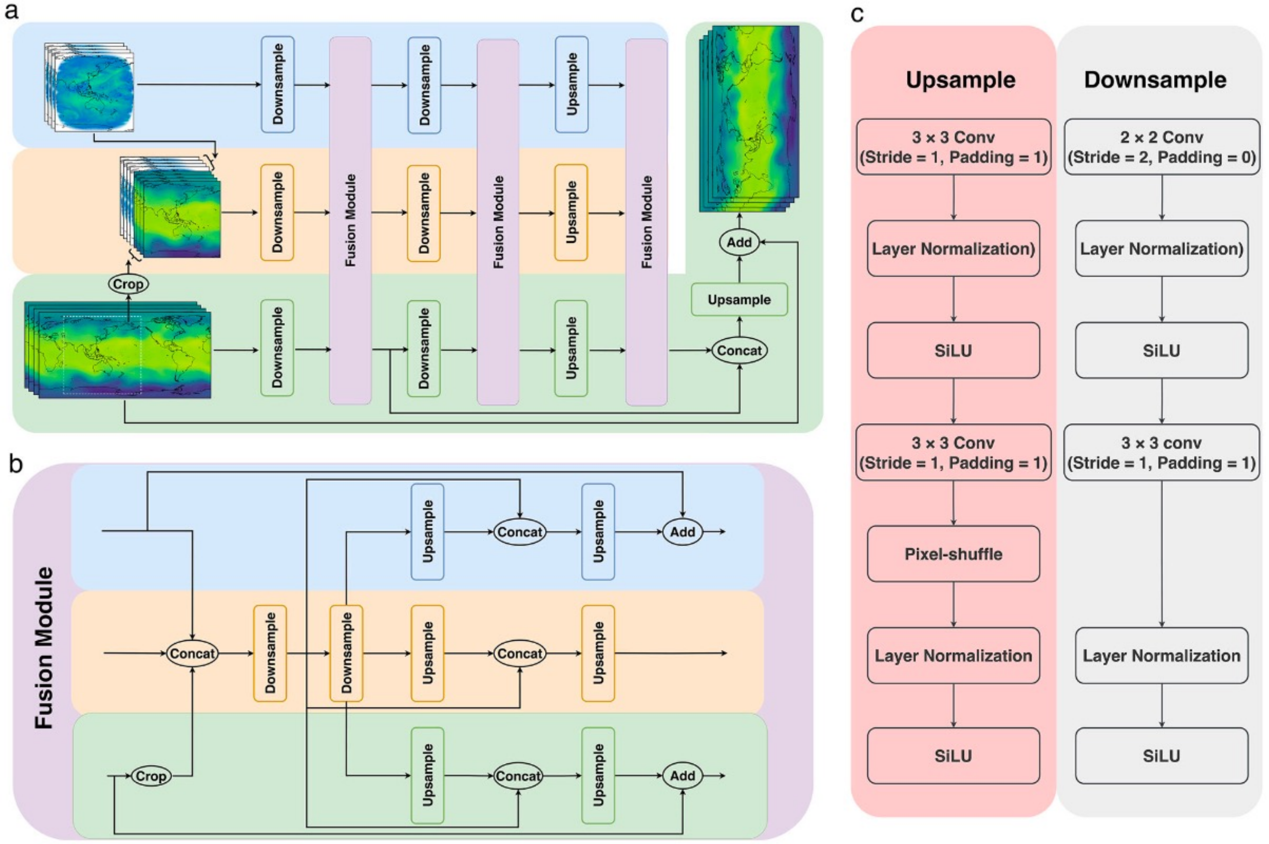
FuXi-DA : End-To-End Assimilation

The comparison between the variational data assimilation and Fuxi-DA



- **No thinning** (grid-based super observation)
- **No observation operators** (feature space interaction)
- **No error estimation** (automatic weight learning)
- **No bias correction** (spatial-temporal encoding)
- **Reduced computational cost** (fast inference)
- ➢ **Joint training of assimilation and forecasting** models to improve long-term forecast performance

Architecture of FuXi-DA model

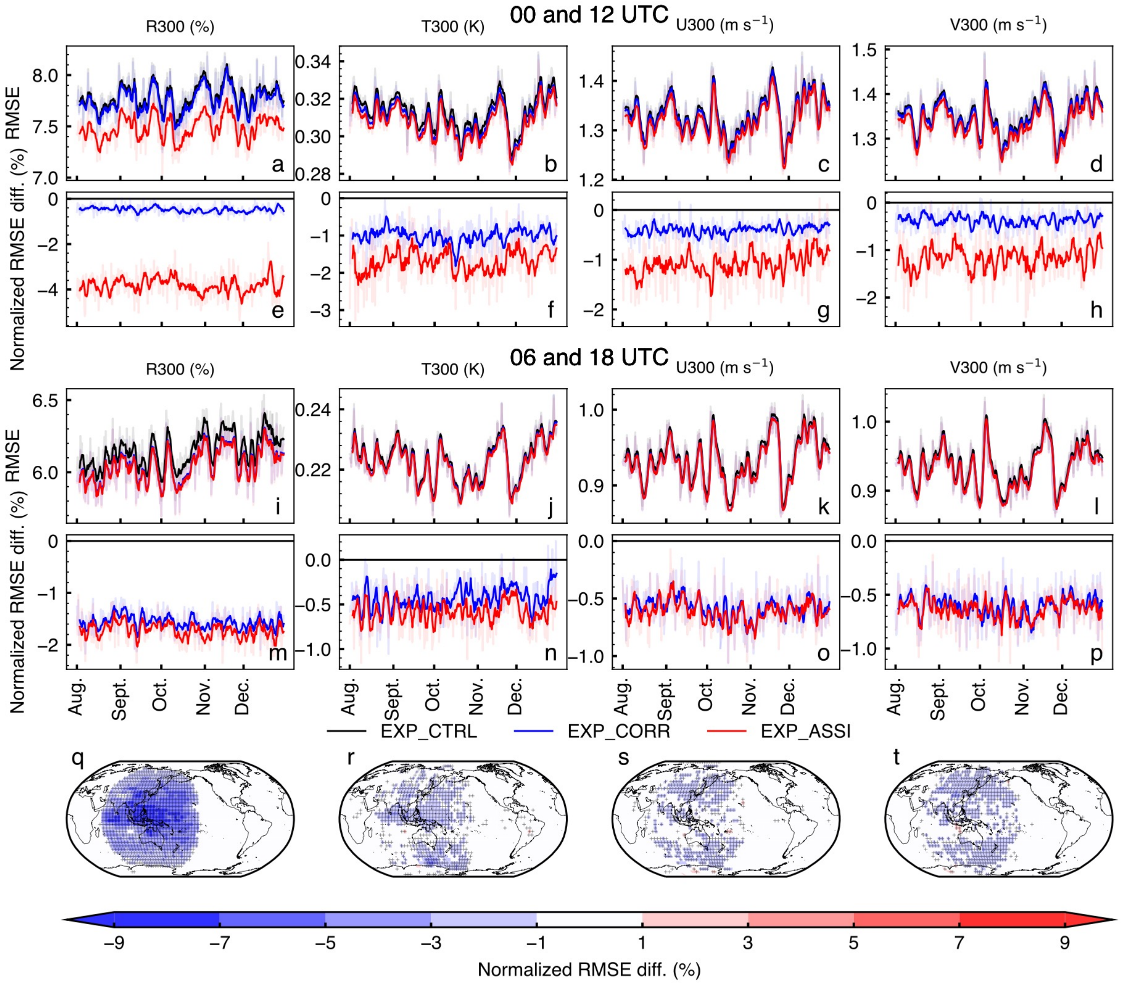


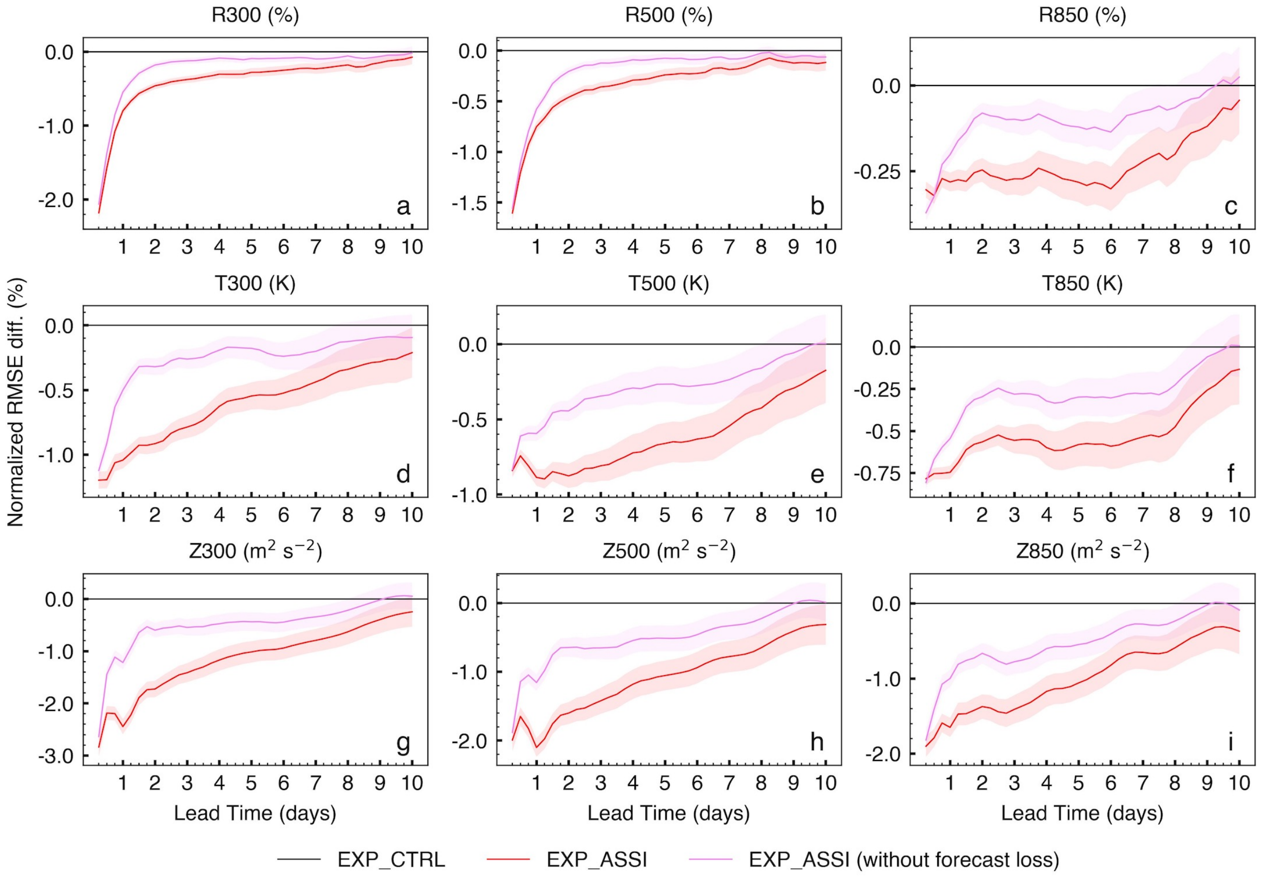
DA process is treated as an incremental learning process

$$x^a = x^b + BH^T(HBH^T + R)^{-1}[y^o - H(x^b)]$$

Operator Learning

FuXi-DA

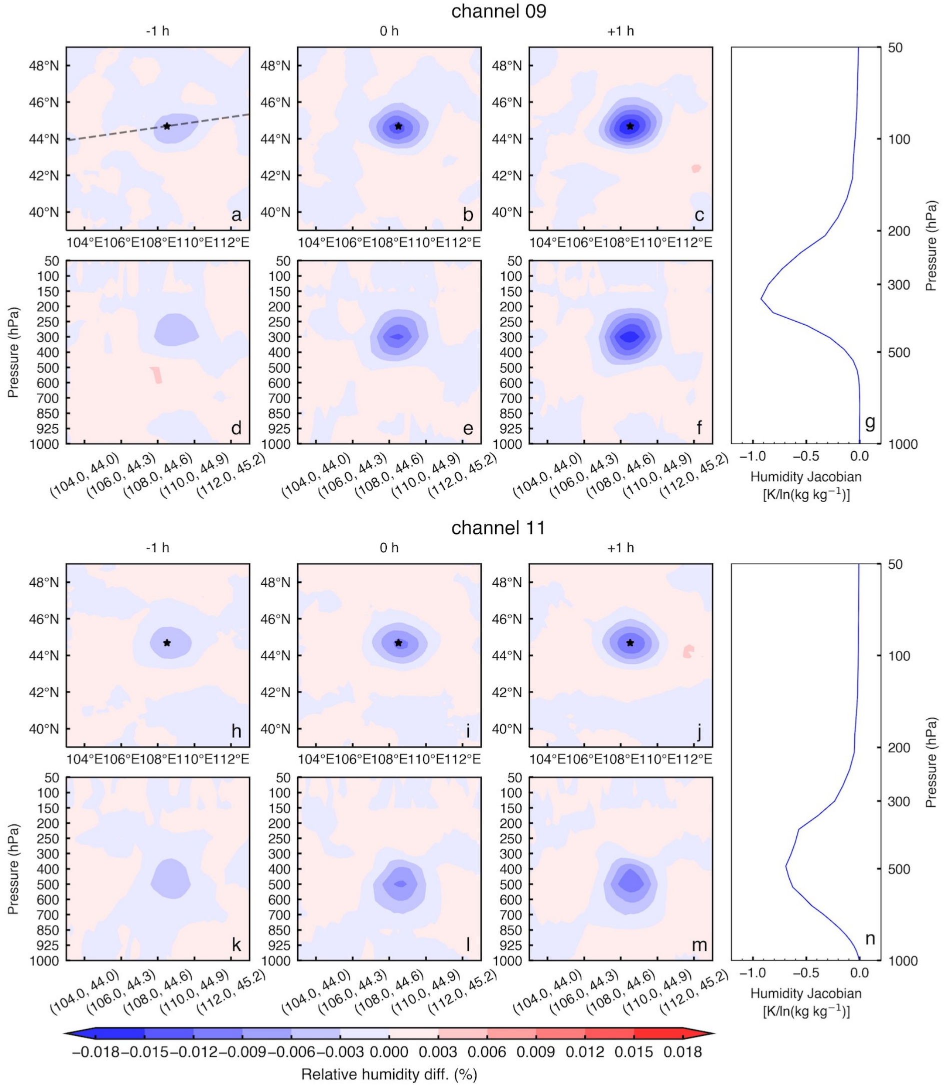




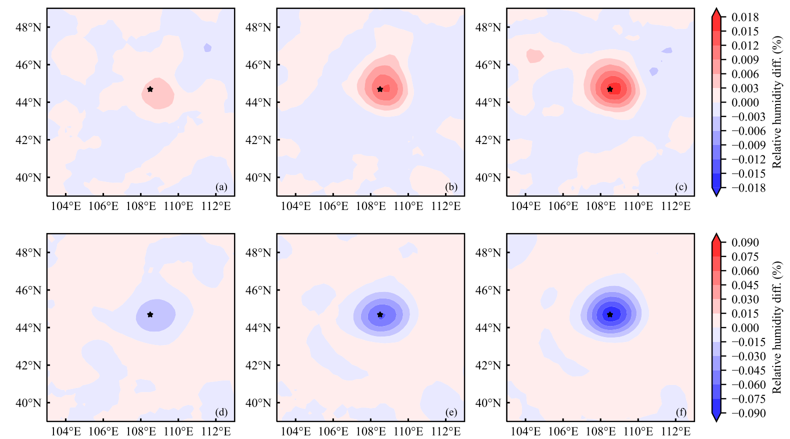
The analysis (left) and forecast (right) field errors are reduced After assimilating FY-4B/AGRI data

FuXi-DA

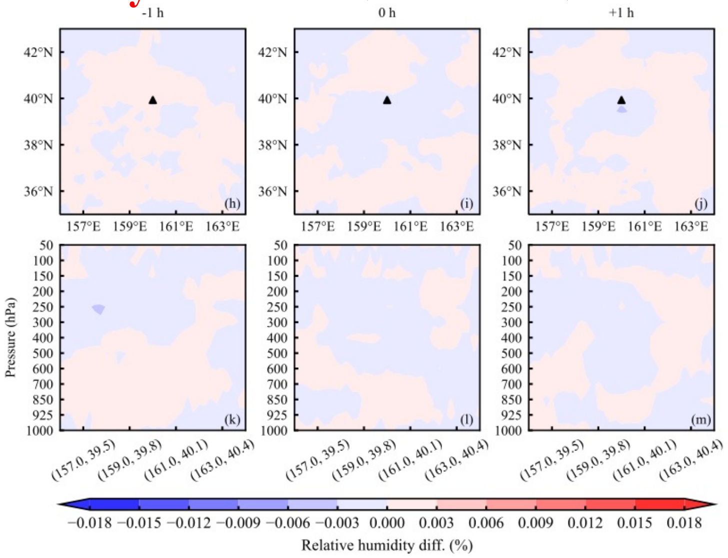
Analysis increment of a 1K perturbation (Top: Channel 9; Bottom: Channel 11)



Analysis increment of -1K and 5K perturbation **clear** condition (Channel 9)



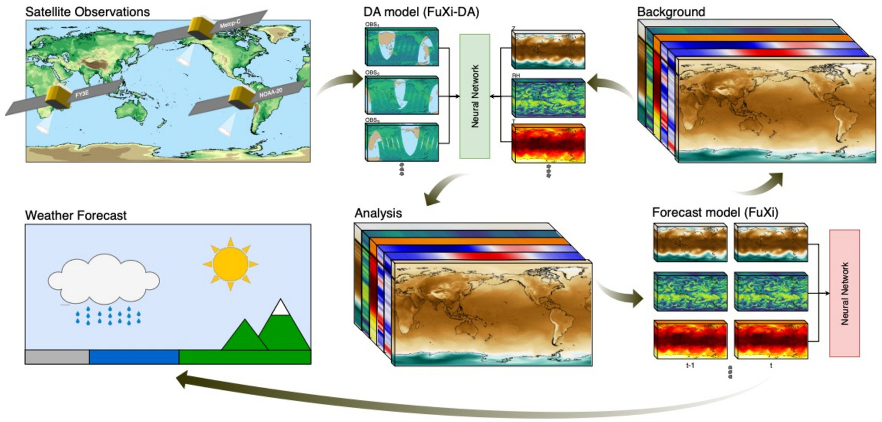
Analysis increment of 1K perturbation under **cloudy** condition（Channel 9）



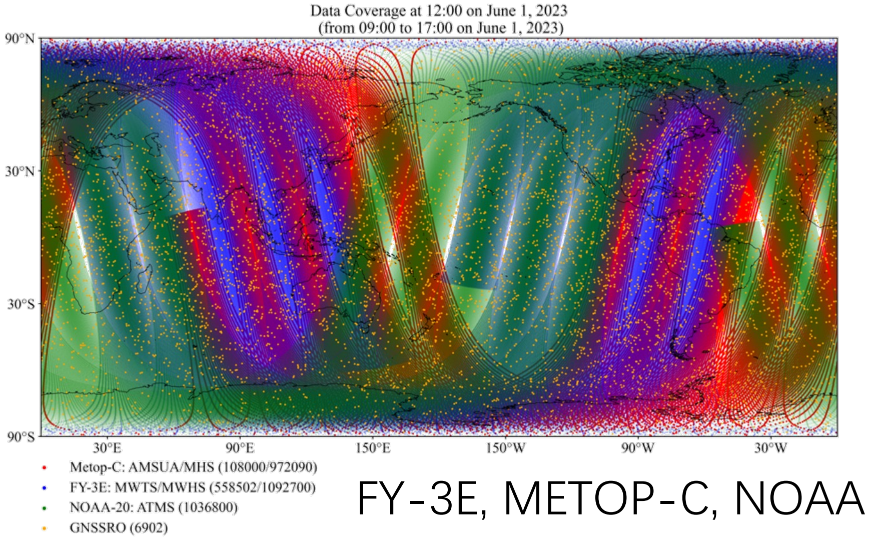
- Demonstrates consistency with prior knowledge of atmospheric physics and the ability to automatically distinguish between cloudy and clear-sky observations.

FuXi-Weather

Schematic of the FuXi Weather System



Observation distribution

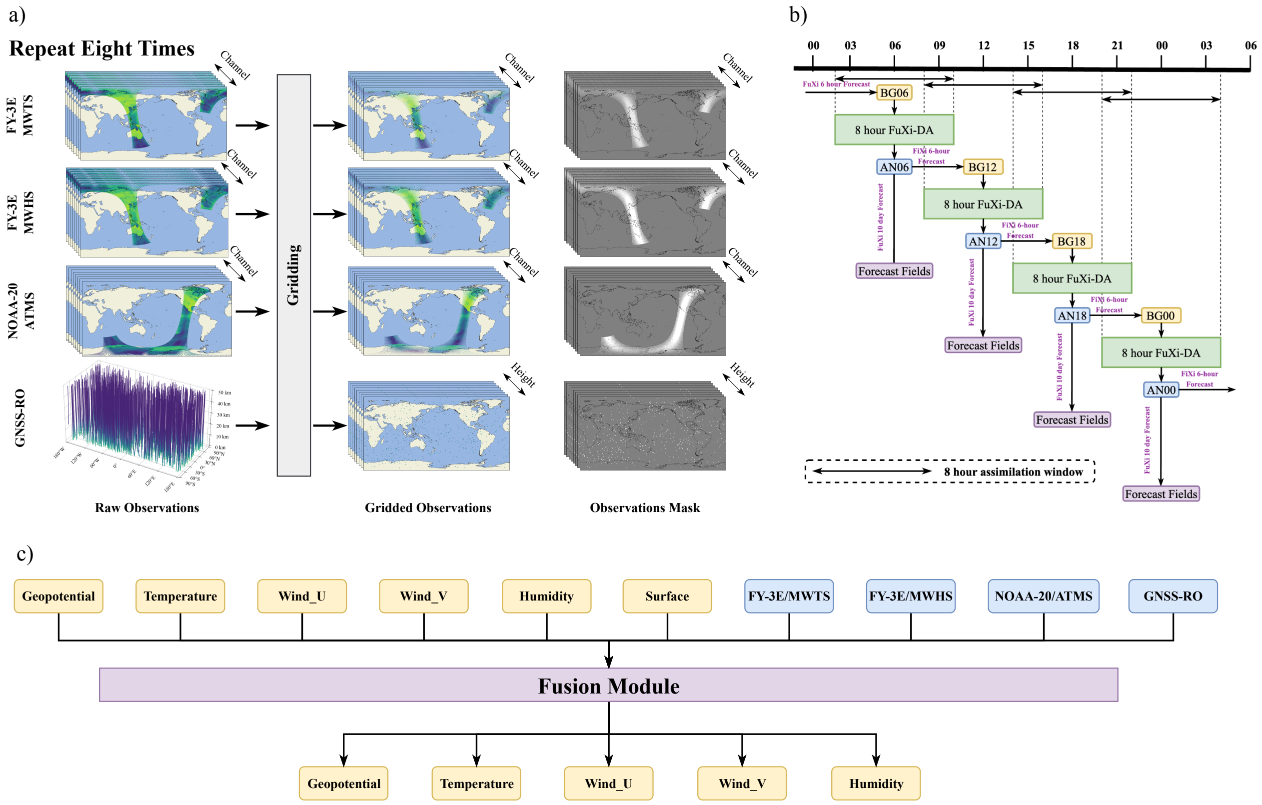


FY-3E, METOP-C, NOAA-20

All grids, surfaces, and channels

MW+RO

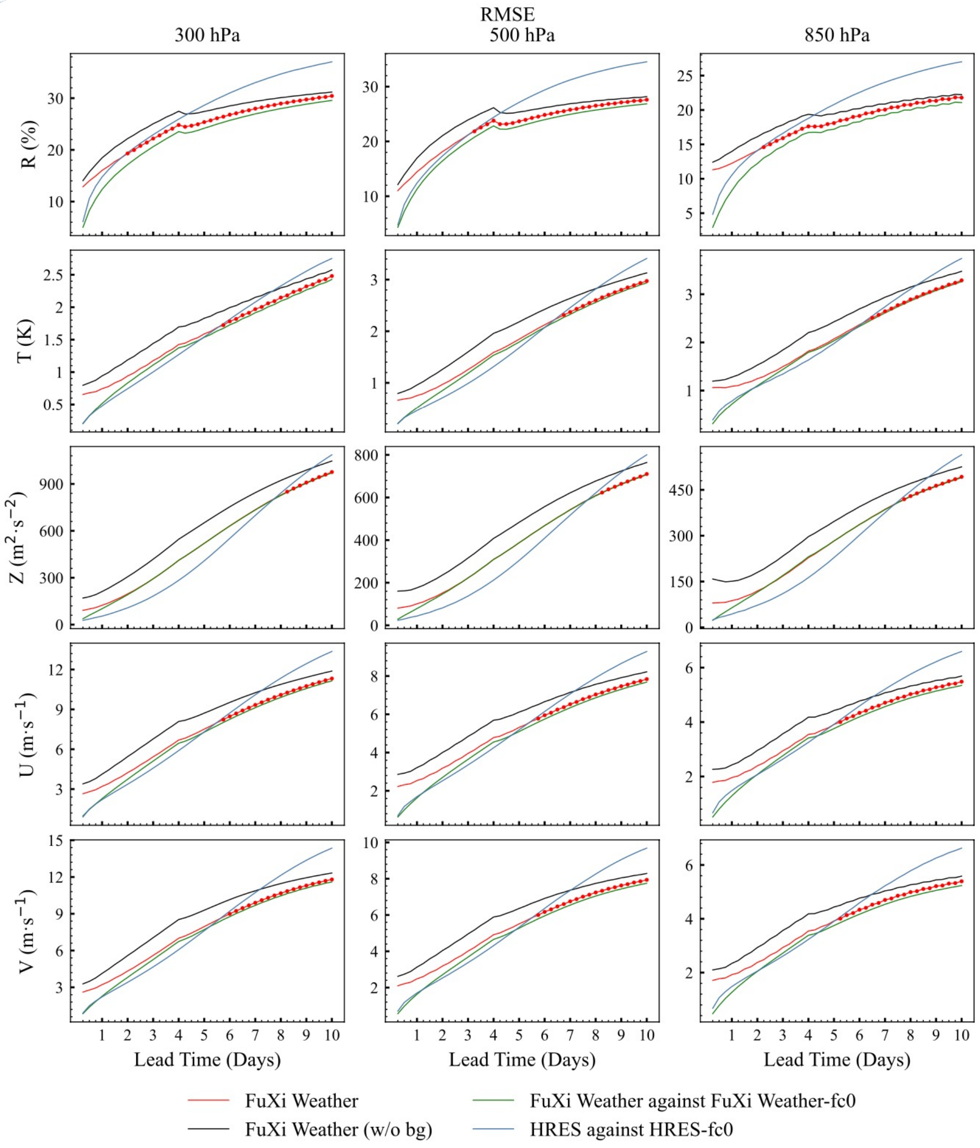
Observation processing (a), cycle flow (b), and assimilation model's input/output (c)

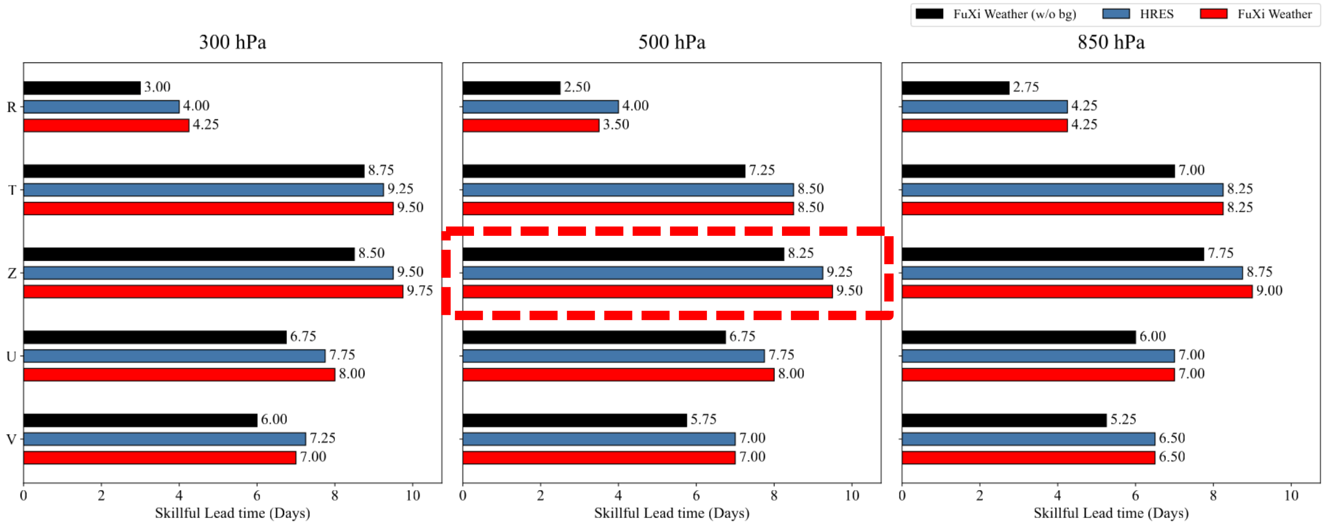


- Adopt an online learning strategy to dynamically update the weather forecasting system
- Achieve end-to-end joint optimization of assimilation and forecasting

arXiv:2408.05472

FuXi-Weather



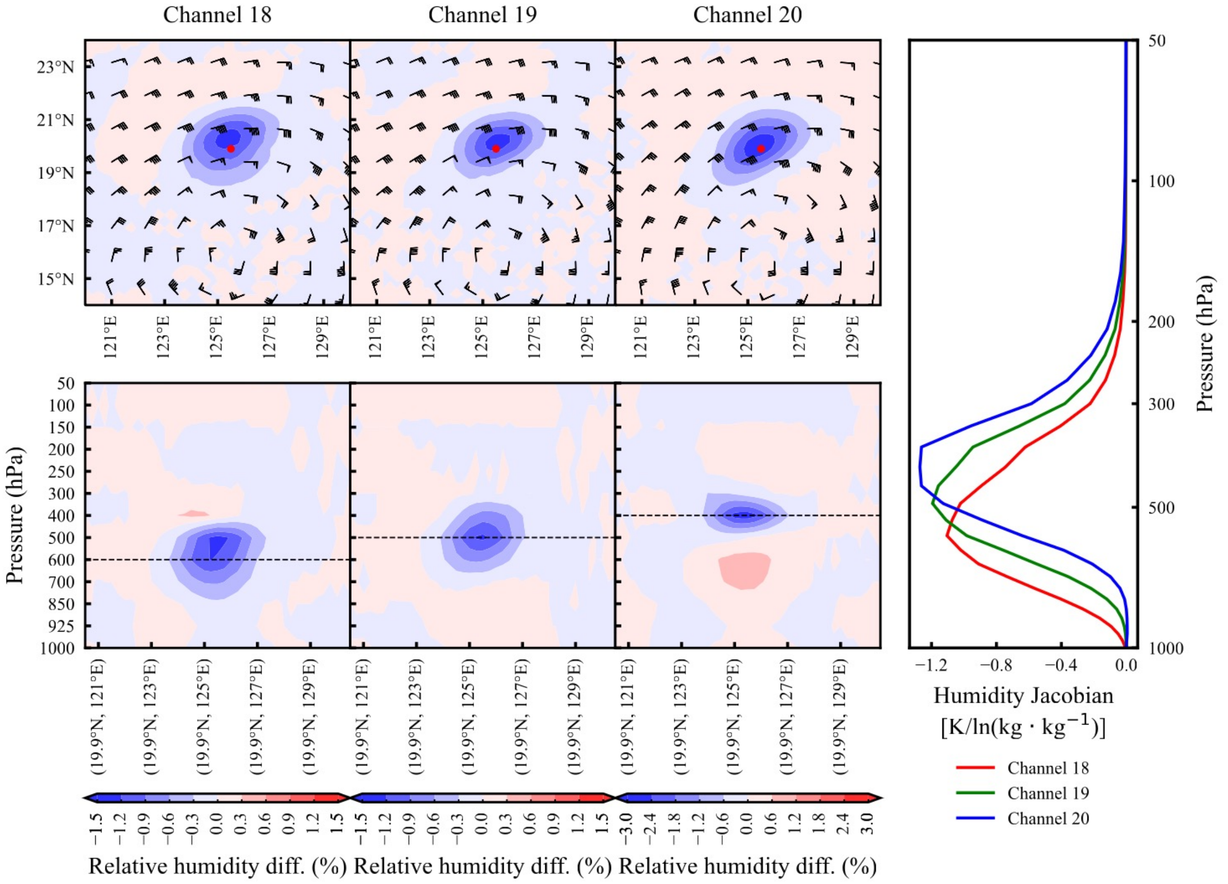


- The forecast accuracy is comparable to HRES, and the later-stage forecast is better than HRES (right, red dots)

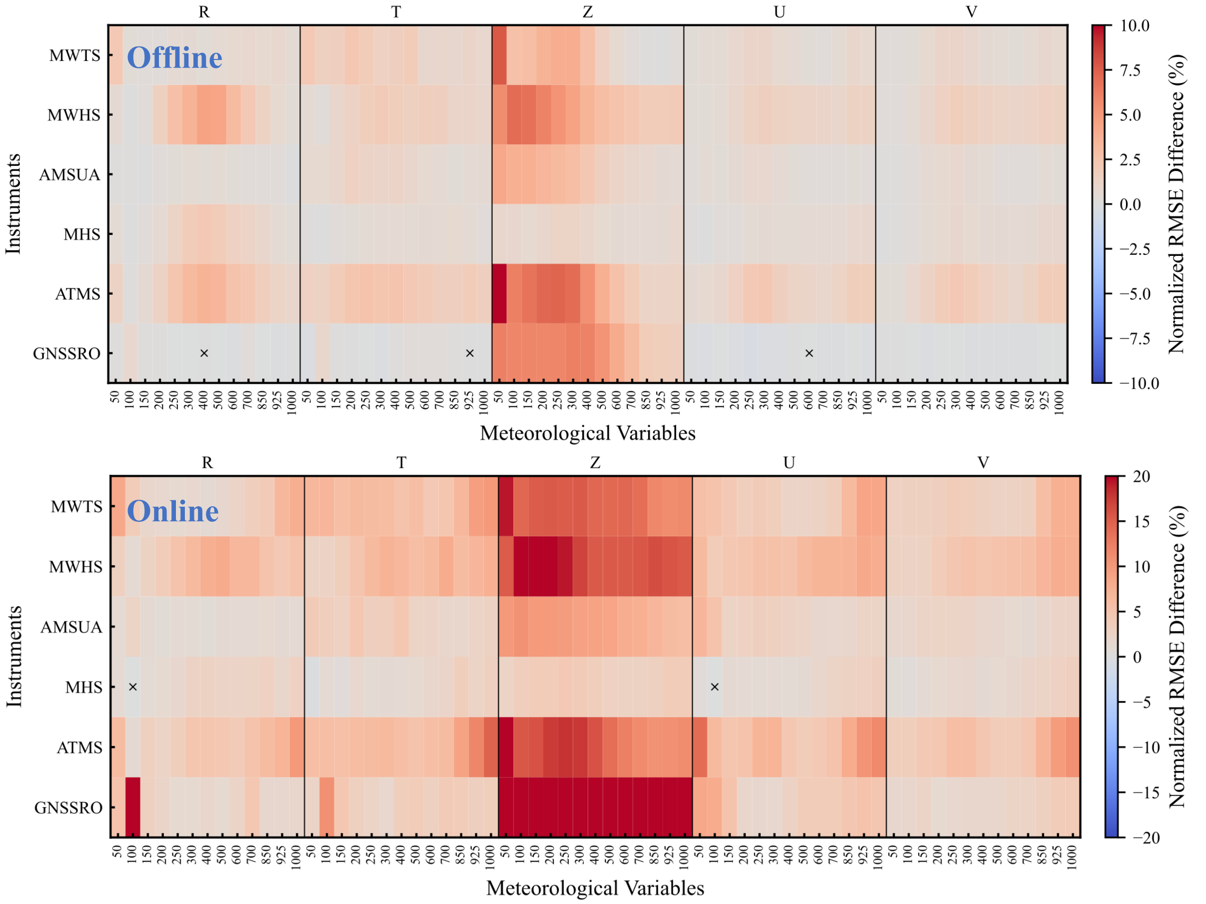
- Among the 15 evaluated variables, 7 have skillful forecast lead times exceeding HRES, with the skillful lead time for Z500 extended from 9.25 days in ECMWF HRES to 9.50 days (left)

FuXi-Weather

Observation Perturbation Test



Data Denial Experiment



Assimilation model maintains physical consistency.

Conclusions

- **AI-RTM** significantly improves computational efficiency and achieves high accuracy in both forward and Jacobian calculations, making it suitable for future assimilation application research.
- The **FuXi-En4DVar** assimilation system can generate flow-dependent background error covariance matrices, while avoiding the use of the adjoint model, thereby simplifying the solution of the cost function.
- **FuXi-DA** effectively assimilates satellite observations, simplifies the observation data processing workflow, and reduces the consumption of computational resources.
- **FuXi-Weather** achieves forecast accuracy comparable to HRES, with smaller forecast errors in the long-term forecasts.

References

① Li, Y., **Han, W.***, Duan, W., Li, Z., & Li, H.(2025). A **machine learning-based observation operator** for FY-4B **GIIRS** brightness temperatures considering **the uncertainty of label data**. *Journal of Geophysical Research: Machine Learning and Computation*, 2, e2024JH000449. https://doi.org/10.1029/2024JH000449

② Xu Xiaoze , Xiuyu Sun, **Wei Han***, Xiaohui Zhong, Lei Chen, Zhiqiu Gao,Hao Li. **Fu Xi-DA: a generalized deep learning data assimilation framework for assimilating satellite observation**[J]. *npj Climate and Atmospheric Science*, 2025, 8(1): 156. https://doi.org/10.1038/s41612-025-01039-3.

③ Li, Z., **Han, W.***, Xu, X., Sun, X., Li, H., 2024. **All-Sky Microwave Radiance Observation Operator** Based on Deep Learning With Physical Constraints. *Journal of Geophysical Research: Atmospheres* 129, e2024JD042436. https://doi.org/10.1029/2024JD042436

④ Li, Y., **Han, W.***, Li, H., Duan, W., Chen,L., Zhong, X., et al. (2024). **Fuxi-En4DVar**:An assimilation system based on machine learning weather forecasting model ensuring physical constraints. *Geophysical Research Letters*, 51, e2024GL111136. https://doi.org/10.1029/2024GL111136

⑤ Sun, X., Zhong, X., Xu, X., Huang, Y., Li, H.*, Neelin, J.D.*, Chen, D., Feng, J., **Han, W.***, Wu, L., Qi, Y.*, 2024. **FuXi Weather**: A data-to-forecast machine learning system for global weather. https://doi.org/10.48550/arXiv.2408.05472

⑥ Allen, A., Markou, S., Tebbutt, W. *et al.* **End-to-end data-driven weather prediction**. *Nature* (2025). https://doi.org/10.1038/s41586-025-08897-0