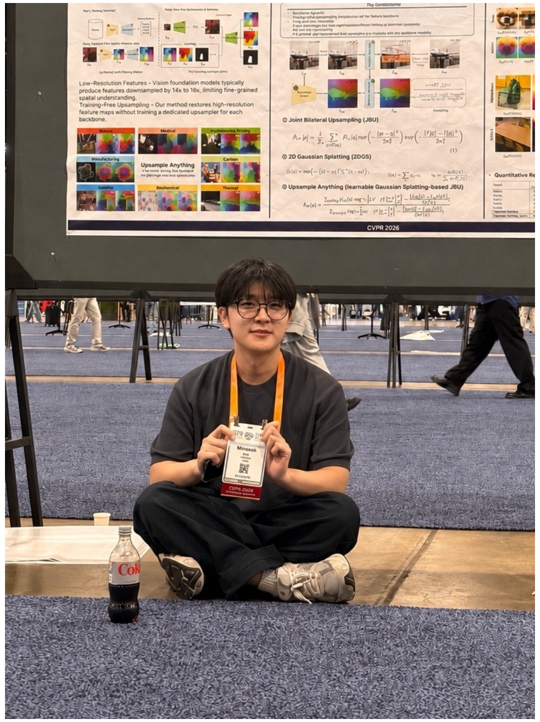
Domain-Adaptive Data Assimilation for Global AI Weather Forecasting

Minseok Seo[1], Noah Brenowitz[2], Doyi Kim[1], Hyesook Lee[1], Changick Kim [1]

2026.09.16

National Institute of Meteorological Sciences

About Me



Minseok Seo (서민석)
Ph.D student at KAIST

Research topic

Test-time Optimization
Domain Adaptation

Research Experiences (Weather Forecasting)

- **NVIDIA (INTERN)** , Research Scientist,, London, UK (July 2026 - Sep 2026)
- **SI-Analytics (military service)** , Research Scientist , Daejeon, Korea (Mar 2022 - Sep 2025)

Publication

Upsample Anything: A Simple and Hard to Beat Baseline for Feature Upsampling

Minseok Seo, Mark Hamilton, Changick Kim (KAIST, Microsoft, MIT)

CVPR 2026 (Compute Gold Star / Transparency Champion)

RainODE: Continuous-Time Precipitation Forecasting with Latent Neural ODEs

Yeeun Seong*, Doyi Kim*, **Minseok Seo**, and Changick Kim (KAIST)

ECCV 2026 (Project Lead)

Station2Radar: query-conditioned gaussian splatting for precipitation field

Doyi Kim, Minseok Seo, Changick Kim (KAIST)

ICLR 2026 (Project Lead)

Data-driven Precipitation Nowcasting Using Satellite Imagery

Young-Jae Park, Doyi Kim, **Minseok Seo**, Hae-Gon Jeon, Yeji Choi

AAAI 2025 (ORAL) (Project Lead)

Probabilistic Weather Forecasting with Deterministic Guidance-based Diffusion Model

Donggeun Yoon* , **Minseok Seo*** , Doyi Kim, Yeji Choi, and Donghyeon Cho

ECCV 2024 (Project Lead)

Long-Term Typhoon Trajectory Prediction: A Physics-Conditioned Approach Without Reanalysis Data

Young-Jae Park*, **Minseok Seo***, Doyi Kim, Hyeri Kim, Sanghoon Choi, Beomkyu Choi, Jeongwon Ryu, Sohee Son, Hae-Gon Jeon, Yeji Choi

ICLR 2024 (Spotlight)

Weather4cast at neurips 2022: Super-resolution rain movie prediction under spatio-temporal shifts

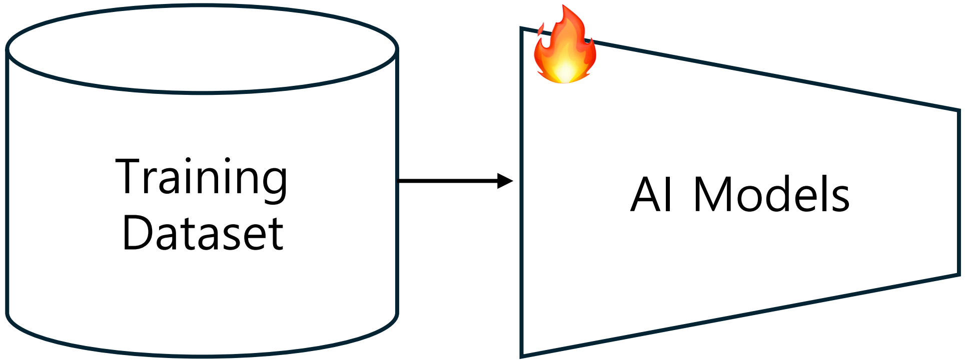
Weather4Forecast (Winner) (Project Lead)

NeurIPS 2022 competition track

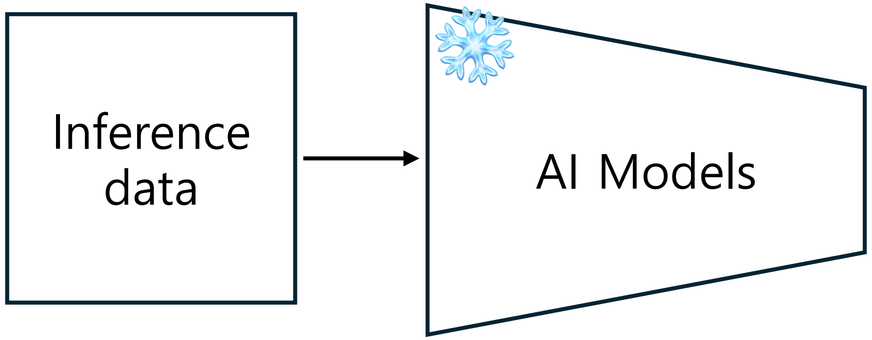
Google Microsoft NVIDIA

What Is Test-Time Optimization?

Standard AI Model Workflow

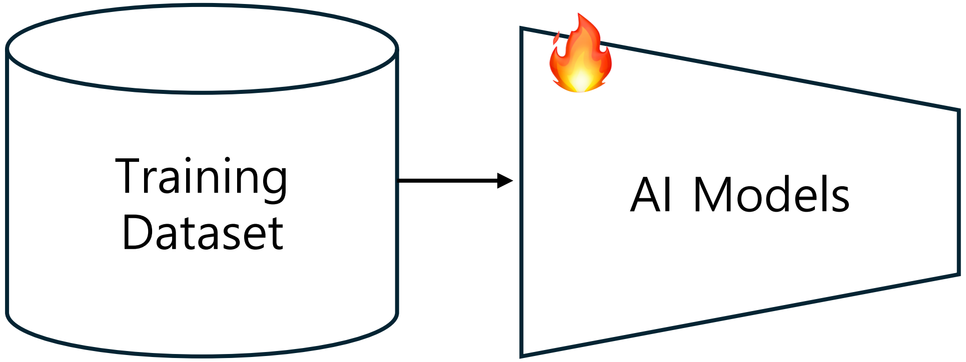


Phase 1. Training

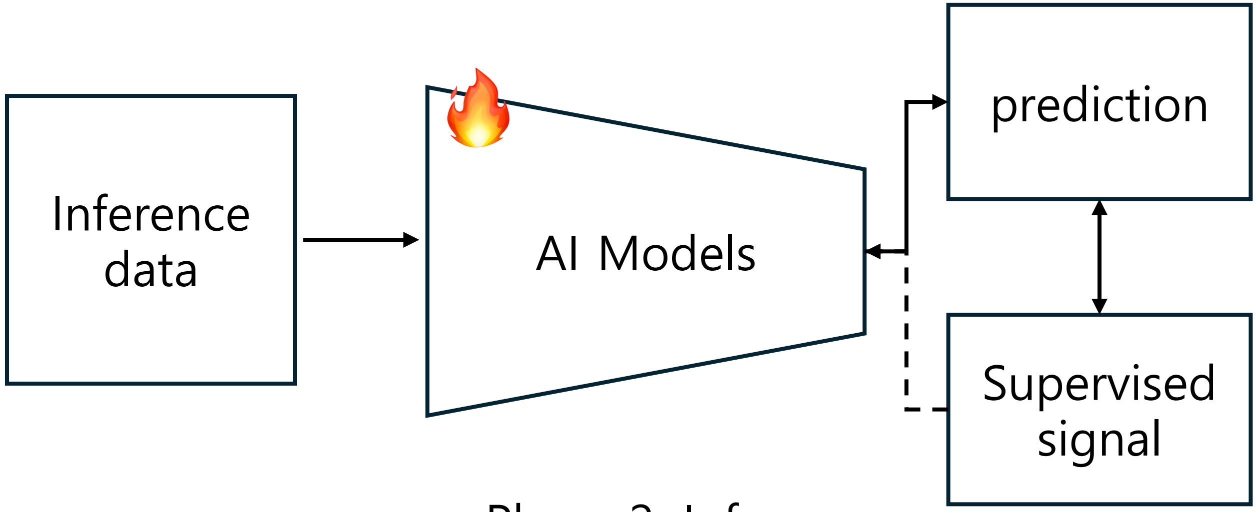


Phase 2. Inference

Test-time optimization Workflow

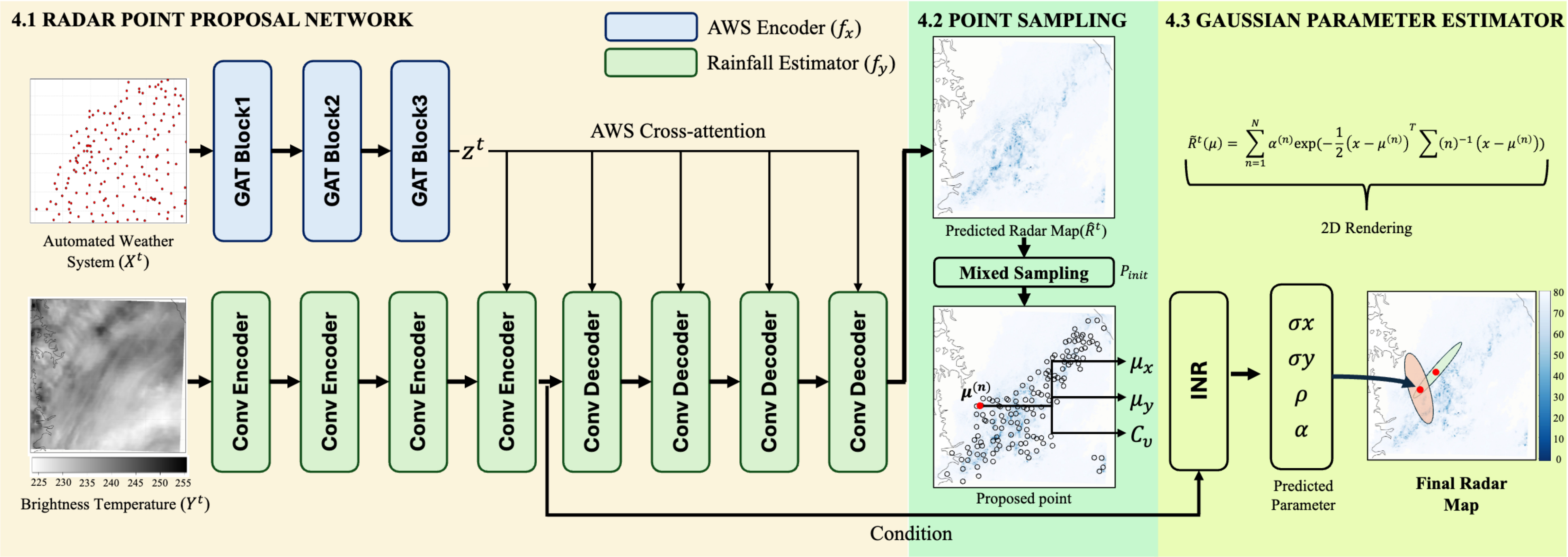


Phase 1. Training



Phase 2. Inference

Observation-Guided Test-Time Optimization: Station2Radar (1)



[1] Kim, Doyi, Minseok Seo, and Changick Kim. "Station2Radar: Query-Conditioned Gaussian Splatting for Precipitation Field." *International Conference on Learning Representations*. Vol. 2026. 2026.

Observation-Guided Test-Time Optimization: Station2Radar (2)

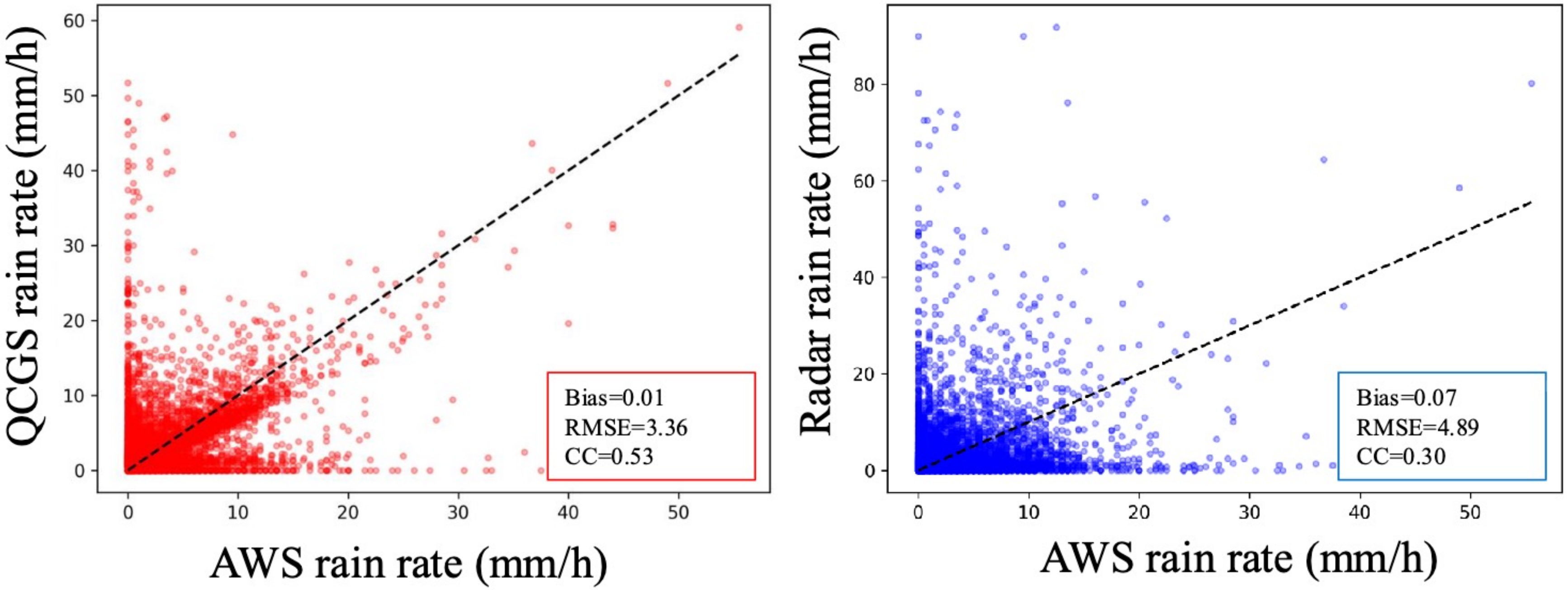
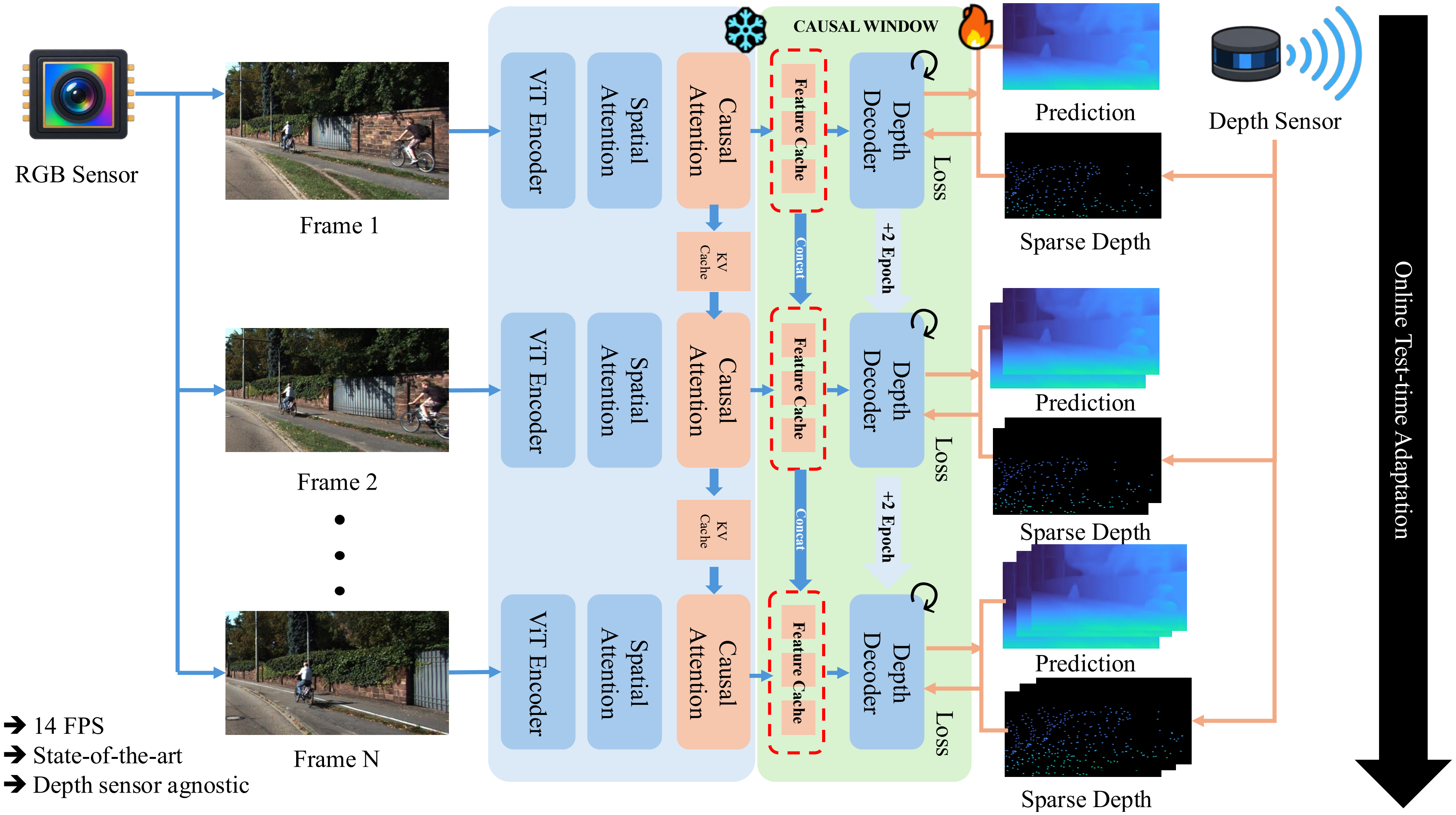


Figure 6: Bias, RMSE, and correlation coefficient (CC) of AWS rain rate compared with QCGS and radar. QCGS consistently achieves lower bias and RMSE and higher CC relative to radar, demonstrating closer agreement with gauge observations.

[1] Kim, Doyi, Minseok Seo, and Changick Kim. "Station2Radar: Query-Conditioned Gaussian Splatting for Precipitation Field." *International Conference on Learning Representations*. Vol. 2026. 2026.

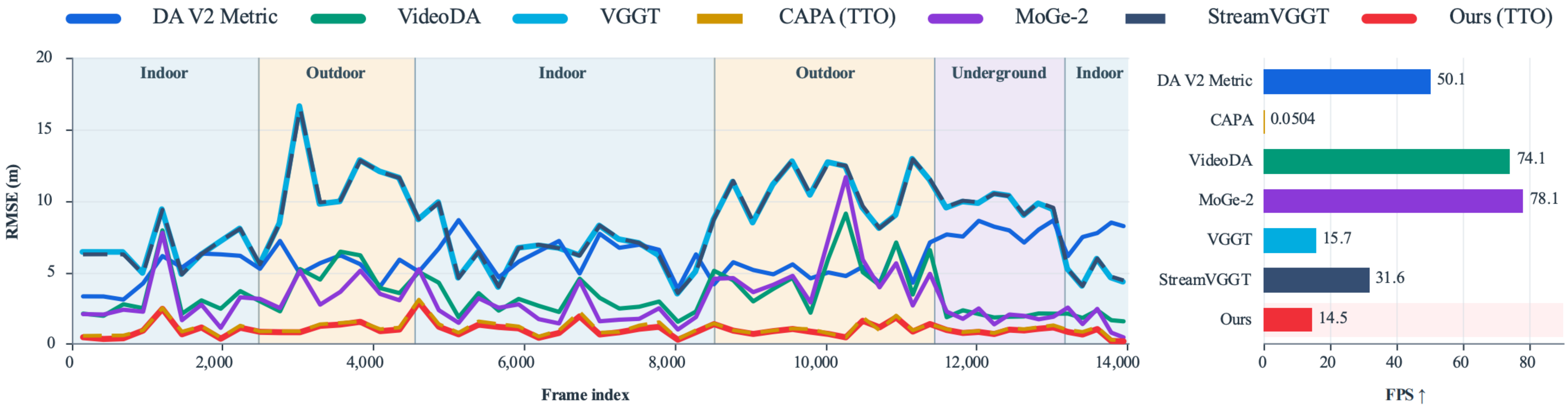
Observation-Guided Test-Time Optimization: StreamTTO (1)



- ➔ 14 FPS
- ➔ State-of-the-art
- ➔ Depth sensor agnostic

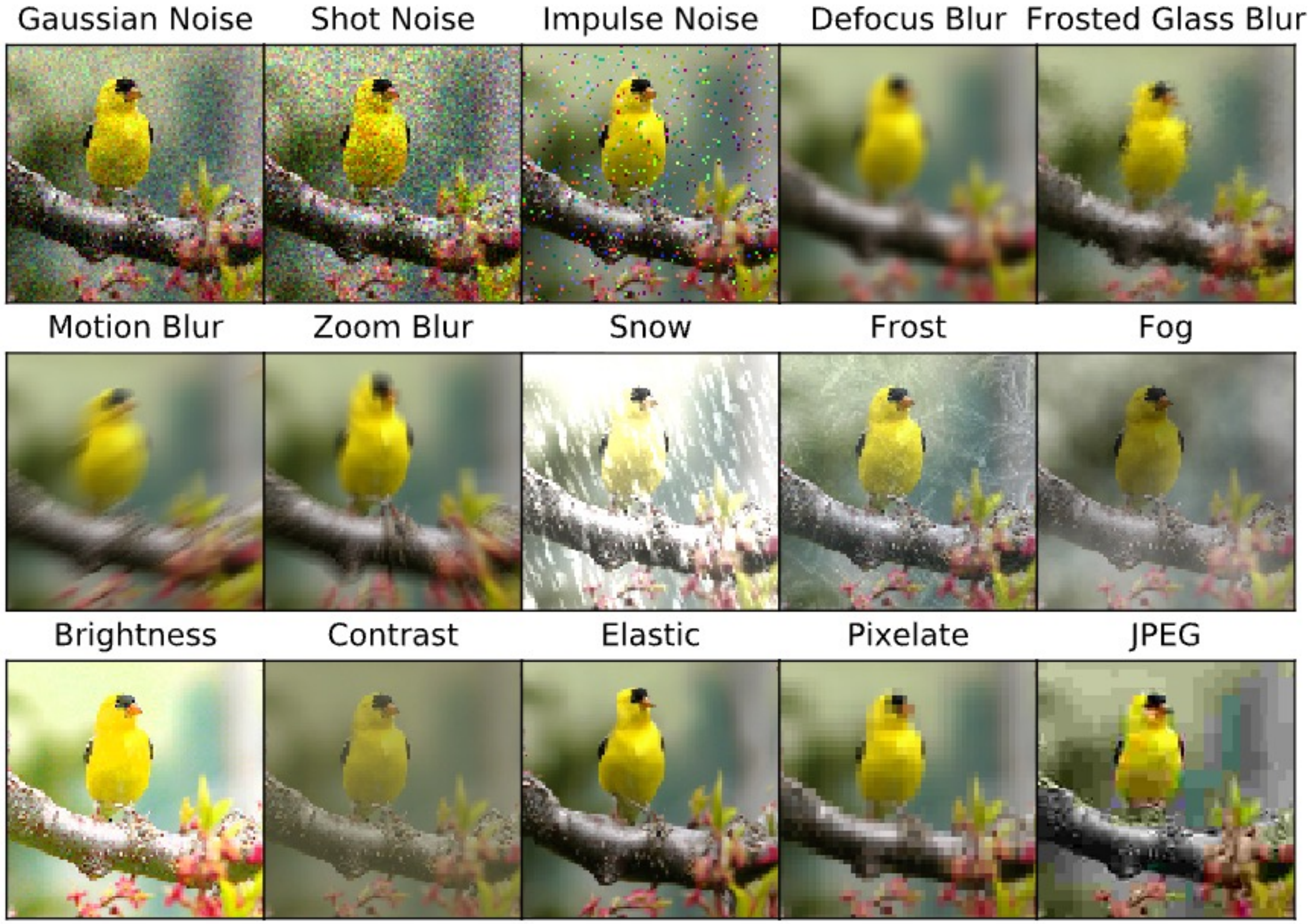
[2] Under review

Observation-Guided Test-Time Optimization : StreamTTO (2)



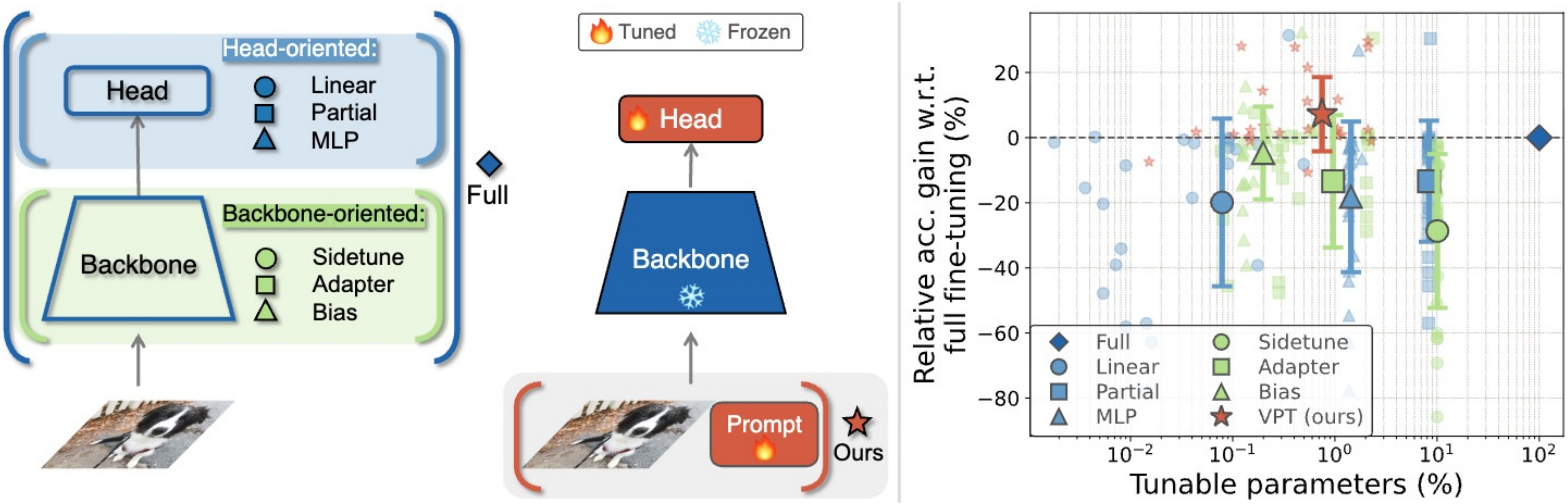
[2] Under review

Why Does TTO Work? Sensitivity to Distribution Shift (1)



[3] Su, Jiawei, Danilo Vasconcellos Vargas, and Kouichi Sakurai. "One pixel attack for fooling deep neural networks." *IEEE Transactions on Evolutionary Computation* 23.5 (2019): 828-841.
[4] Hendrycks, Dan, and Thomas Dietterich. "Benchmarking neural network robustness to common corruptions and perturbations." *arXiv preprint arXiv:1903.12261* (2019).

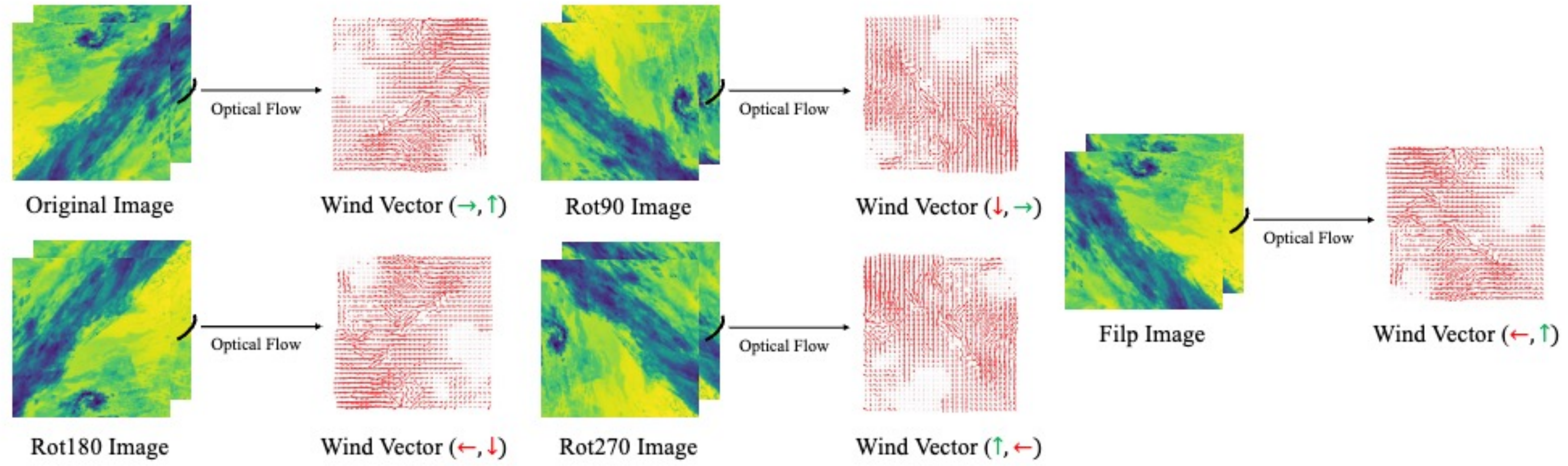
Why Does TTO Work? Sensitivity to Distribution Shift (2)



(a) Existing tuning protocols (b) Visual-Prompt Tuning (VPT) (c) Results on visual classification tasks

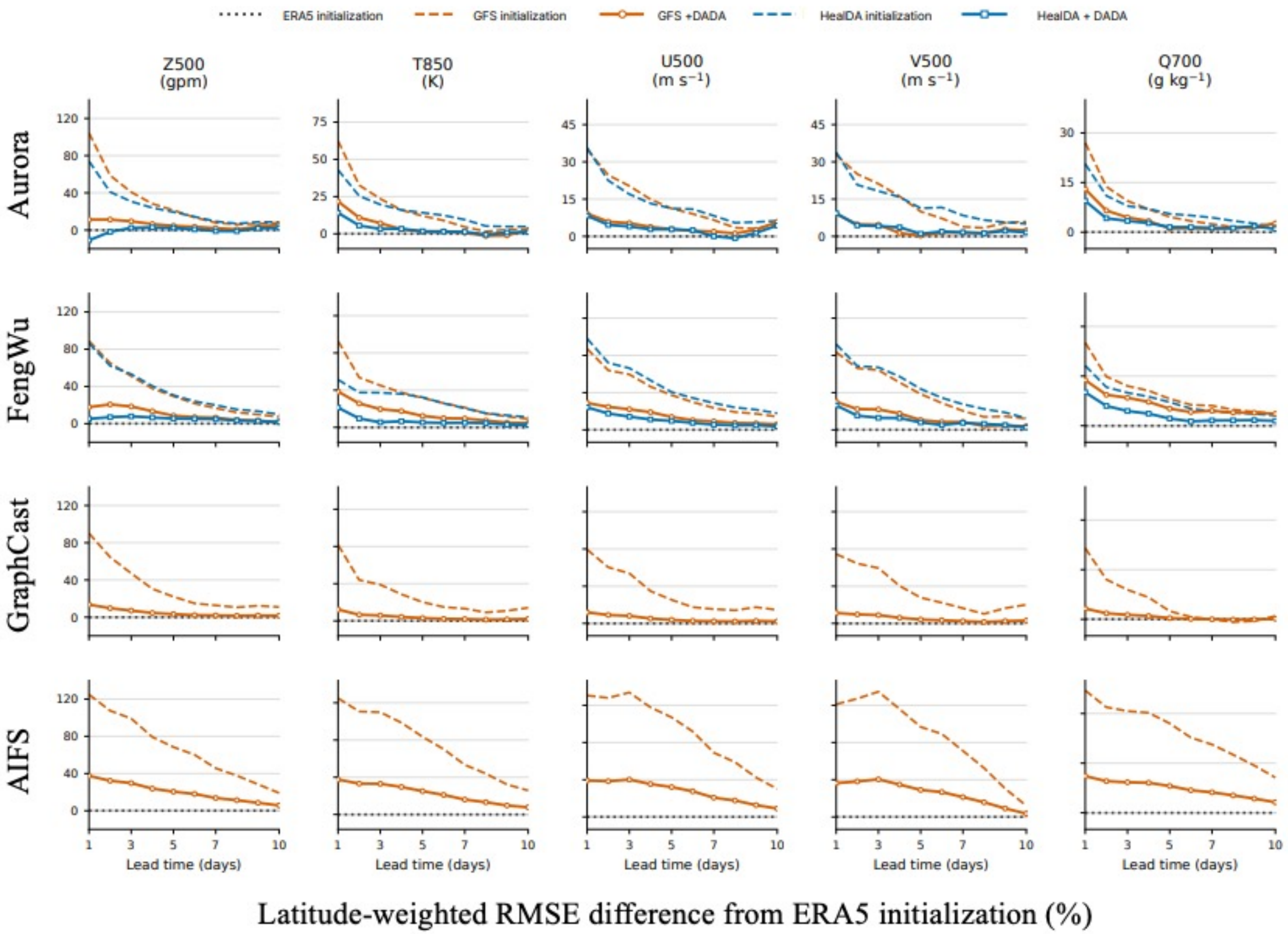
[5] Jia, Menglin, et al. "Visual prompt tuning." *European conference on computer vision*. Cham: Springer Nature Switzerland, 2022.

Why Does TTO Work? Sensitivity to Distribution Shift (3)

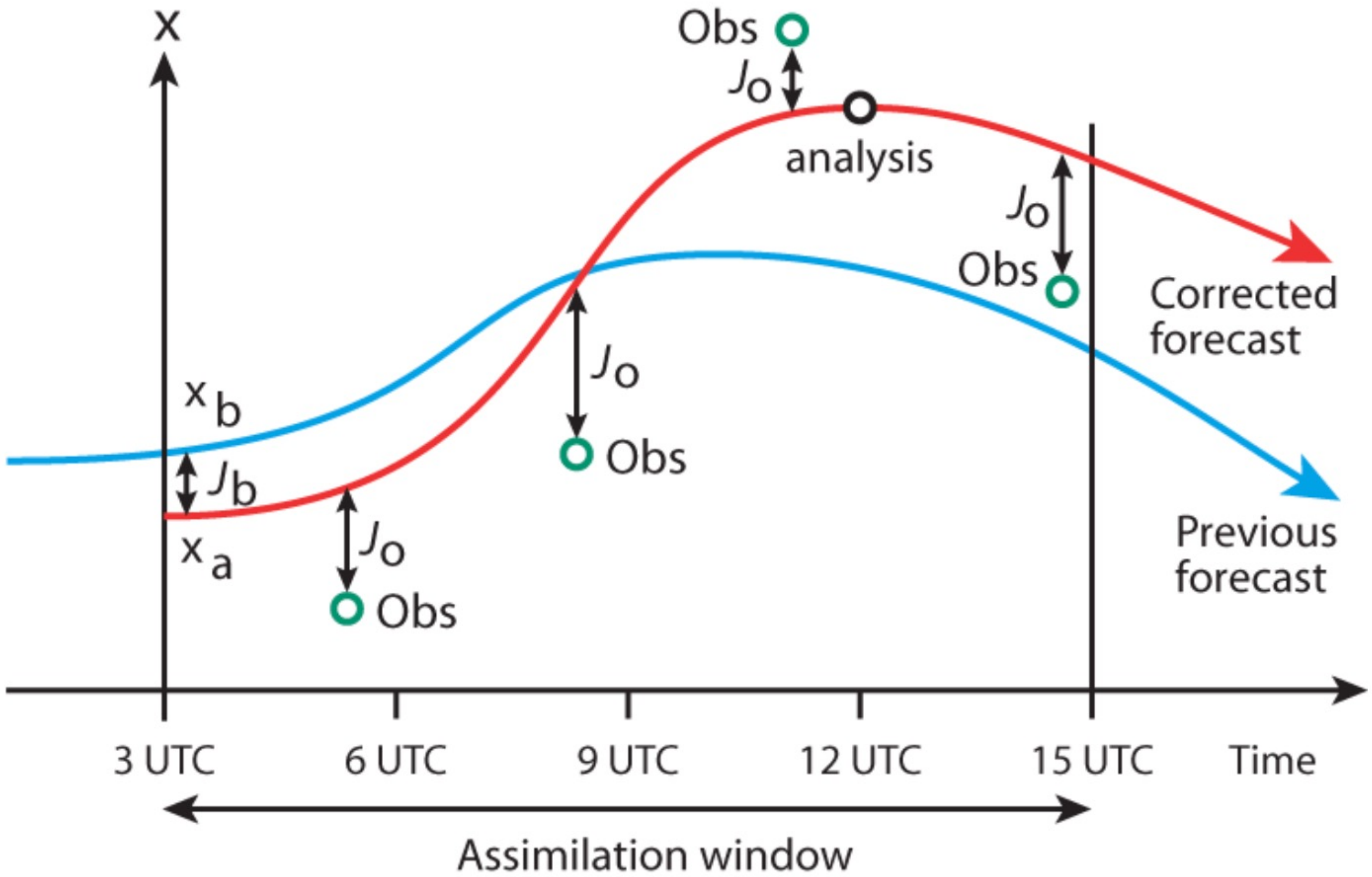


[6] Seo, Minseok, et al. "Domain generalization strategy to train classifiers robust to spatial-temporal shift." *arXiv preprint arXiv:2212.02968* (2022).

Distribution Shift in Weather Forecasting



Data Assimilation Naturally Addresses Distribution Shift



[6] https://www.ecmwf.int/en/about/media-centre/news/2017/20-years-4d-var-better-forecasts-through-better-use-observations

4D-Var: Optimizing the Initial State over Time

Analysis field

Backgroud field

Observation operator

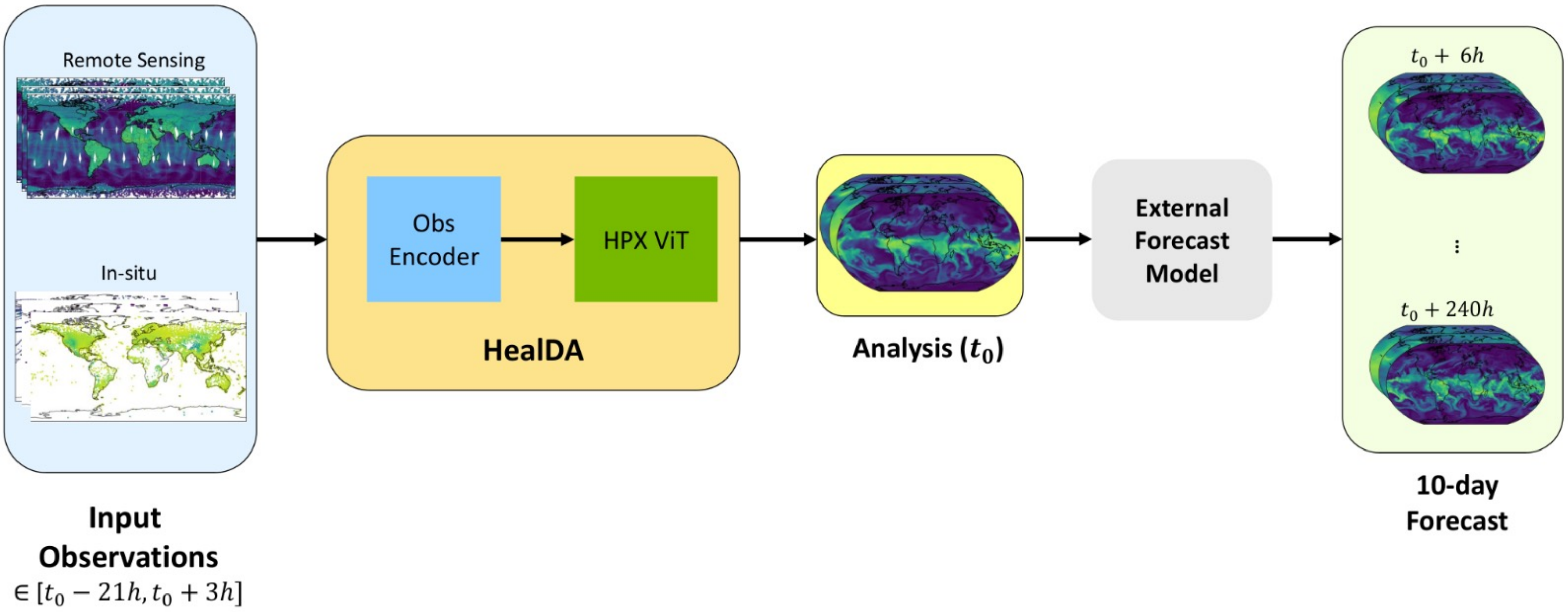
$$\mathbf{x}_0^* = \arg\min_{\mathbf{x}_0} \left[\underbrace{\|\mathbf{x}_0 - \mathbf{x}_b\|^2_{\mathbf{B}^{-1}}}_{\text{Background}} + \sum_k \underbrace{\|\mathcal{H}_k(\mathcal{M}_{0\to k}(\mathbf{x}_0)) - \mathbf{y}_k\|^2_{\mathbf{R}_k^{-1}}}_{\text{Observation}} \right]$$

background error covariance

observation error covariance

[6] https://www.ecmwf.int/en/about/media-centre/news/2017/20-years-4d-var-better-forecasts-through-better-use-observations

HealDA: Highlighting the importance of initial errors in end-to-end AI weather forecasts



[6] Gupta, Aayush, et al. "HealDA: Highlighting the importance of initial errors in end-to-end AI weather forecasts." arXiv preprint arXiv:2601.17636 (2026).

HealDA: Highlighting the importance of initial errors in end-to-end AI weather forecasts

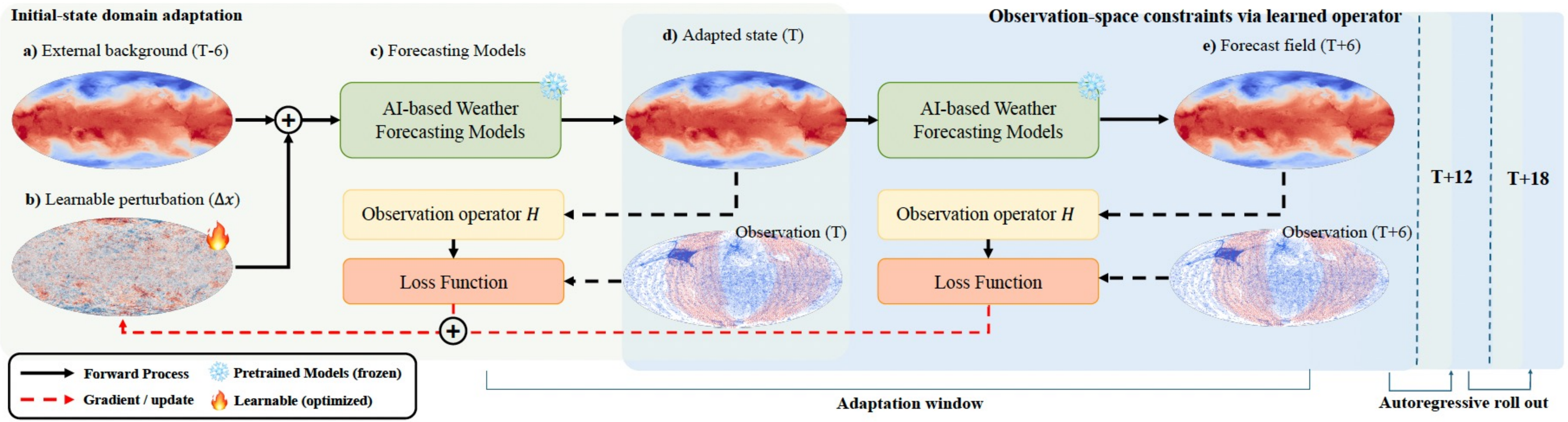
(a) RMSE Scores (vs ERA5 unless otherwise indicated)

Method	Validation Protocol	IC RMSE			RMSE at 1 day			RMSE at 5 days		
		Z500	T850	U500	Z500	T850	U500	Z500	T850	U500
ERA5 [32, Fig. 12] ‡	**~ 2018**	–	**0.34**	**0.74**	–	–	–	–	–	–
Aardvark [15] *	1°, 2018	62	0.94	2.3	93	1.0	2.8	450	2.2	5.9
FuXi Weather [18] *	0.25°, 2023–2024	80	1.0	2.2	100	1.1	2.5	390	2.1	5.3
XiChen [17] *	1.40625°, 2023	100	1.1	–	150	1.2	–	490	2.3	–
IFS ENS †	1°, 2022	26	0.71	1.11	47	0.77	1.46	280	1.62	4.22
HealDA-initialized FCN3 †	HPX64, 2022	47	0.70	1.68	67	0.78	1.98	325	1.76	4.71
HealDA-initialized Atlas †1	HPX64, 2022	47	0.70	1.68	66	0.81	1.94	308	1.73	4.55

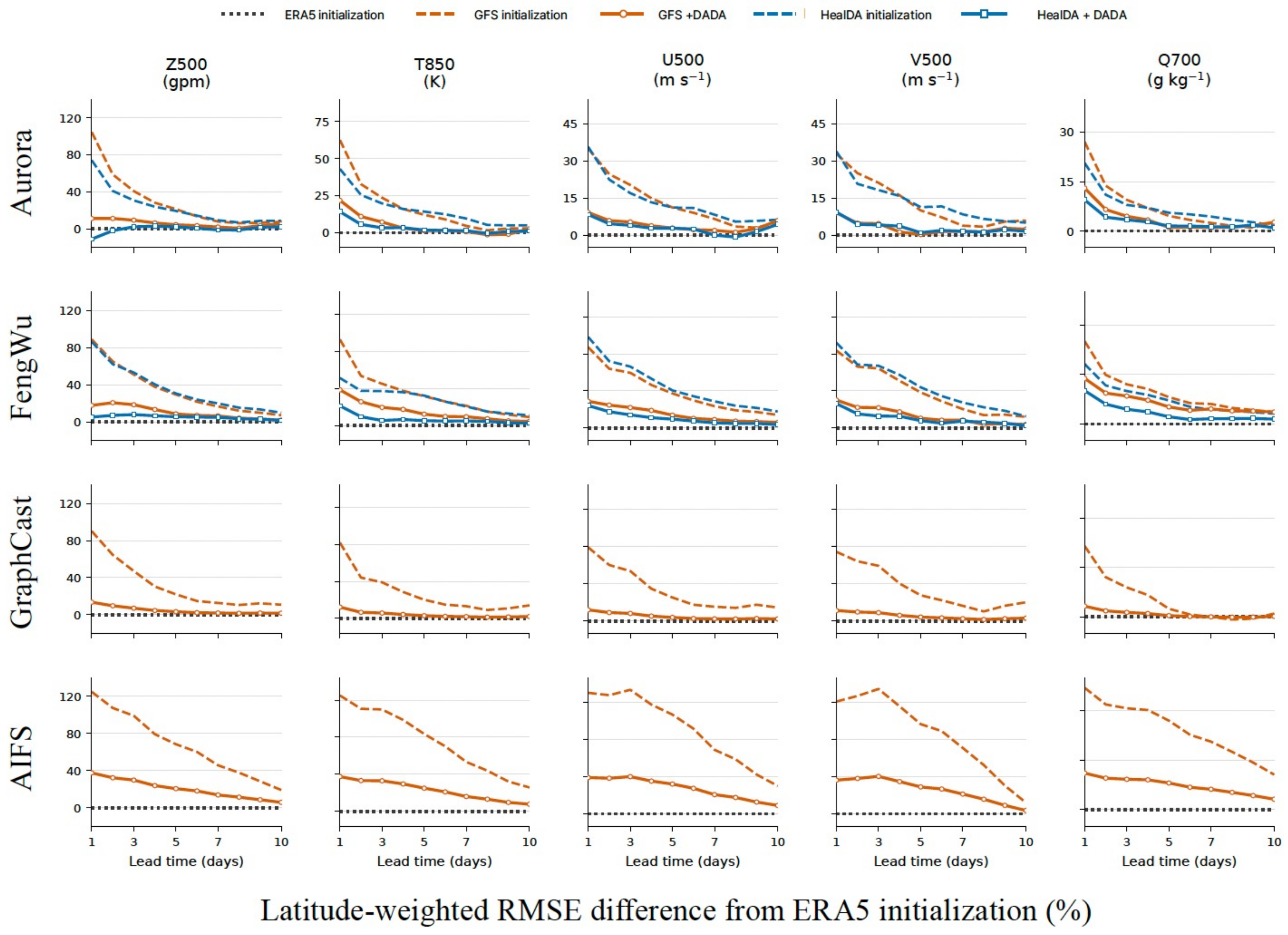
‡ Spread of DA ensemble reported. * Single-member forecast. Trained with multi-step loss functions. † Ensemble mean scores reported.

[6] Gupta, Aayush, et al. "HealDA: Highlighting the importance of initial errors in end-to-end AI weather forecasts." arXiv preprint arXiv:2601.17636 (2026).

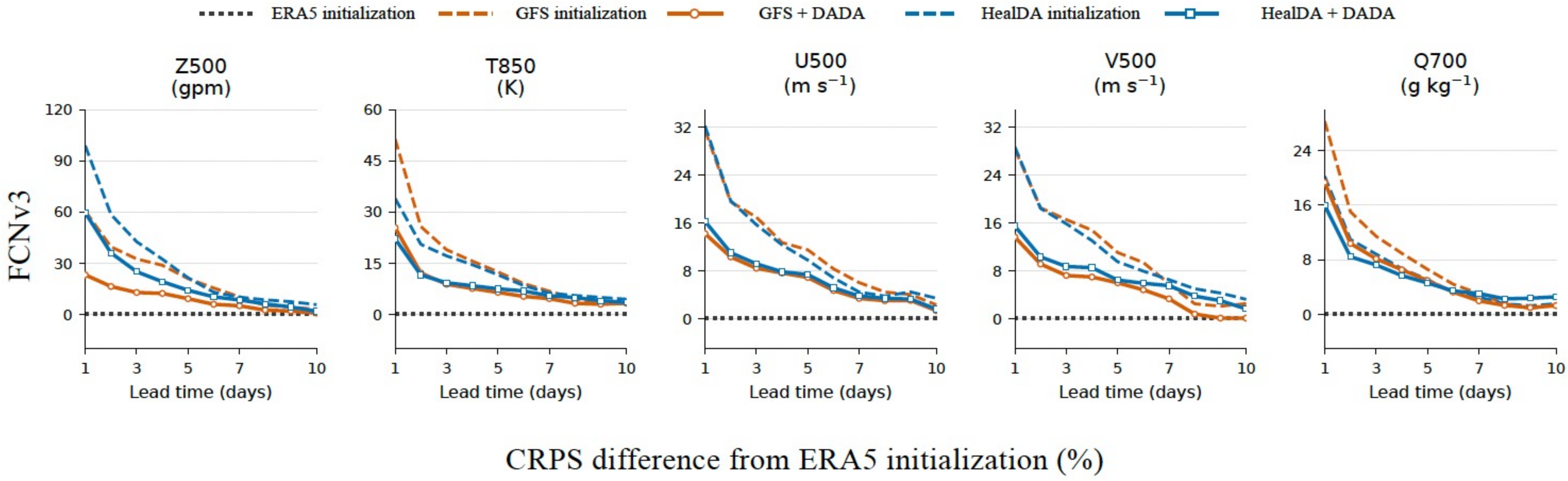
Domain-Adaptive Data Assimilation for Global AI Weather Forecasting



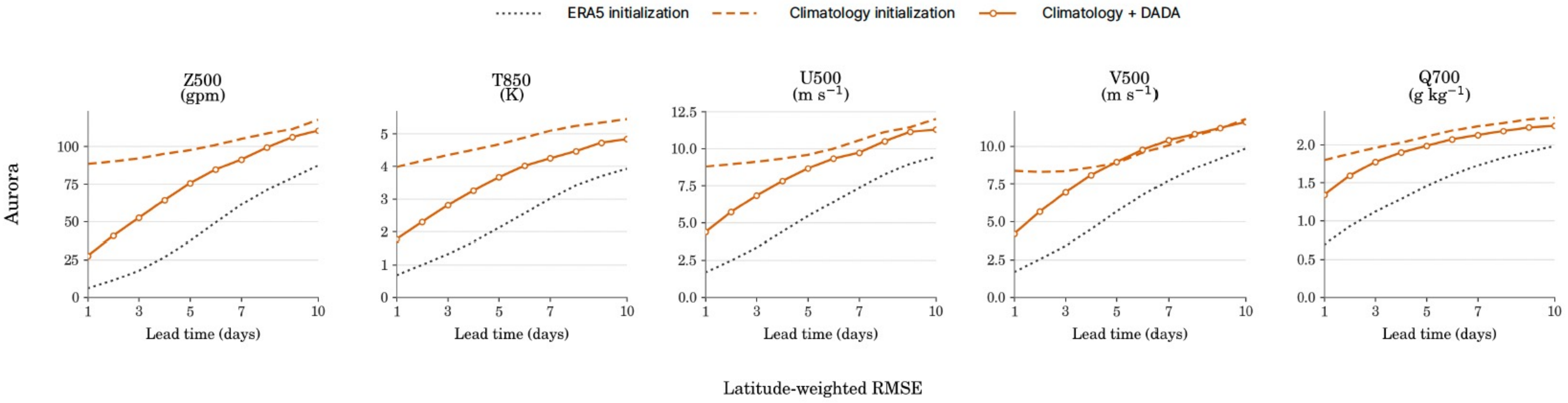
Adaptation across forecasting models and initial conditions



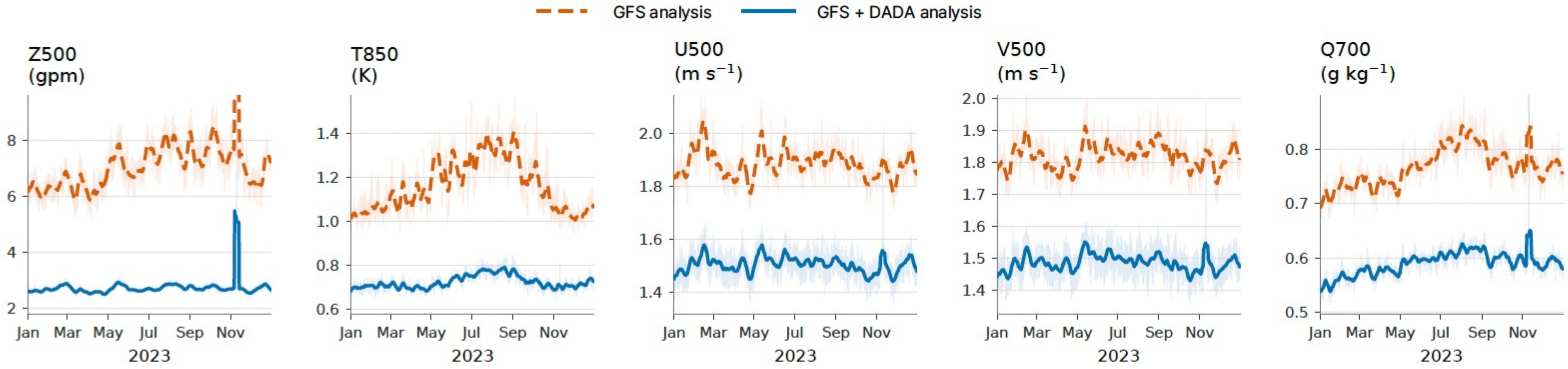
Probabilistic forecast skill with FCNv3



Assimilation from a climatological background



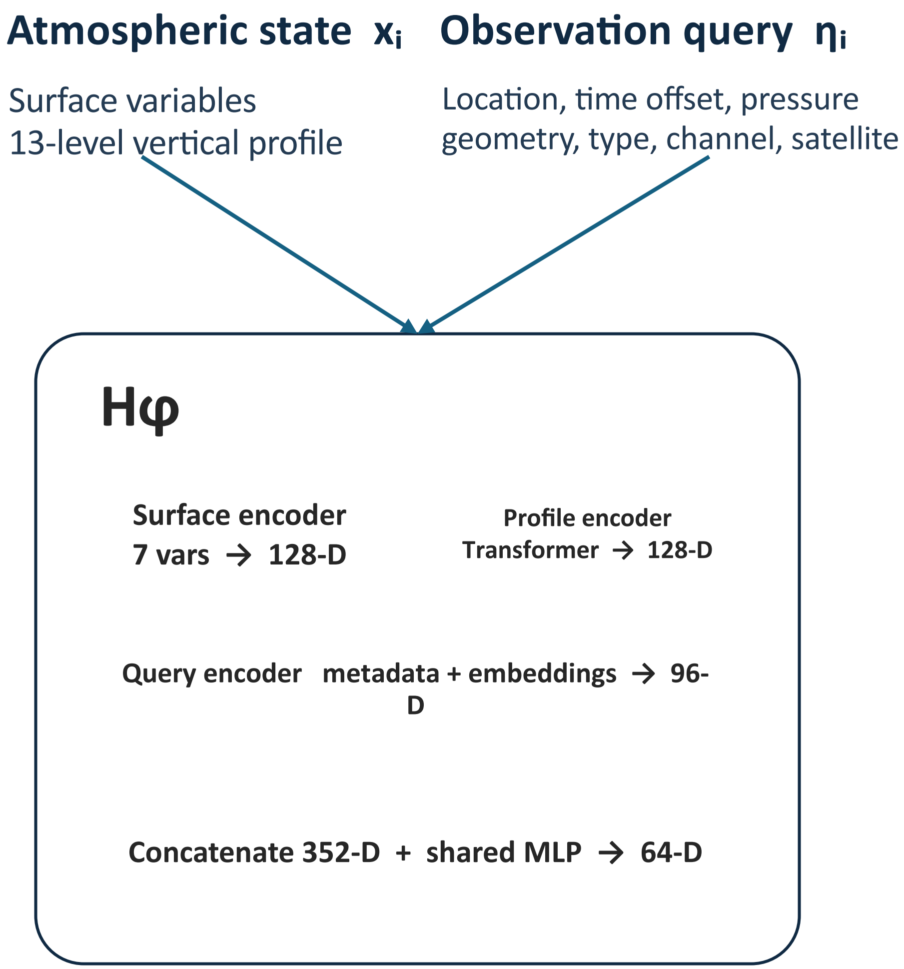
Adaptation of GFS analysis fields throughout 2023



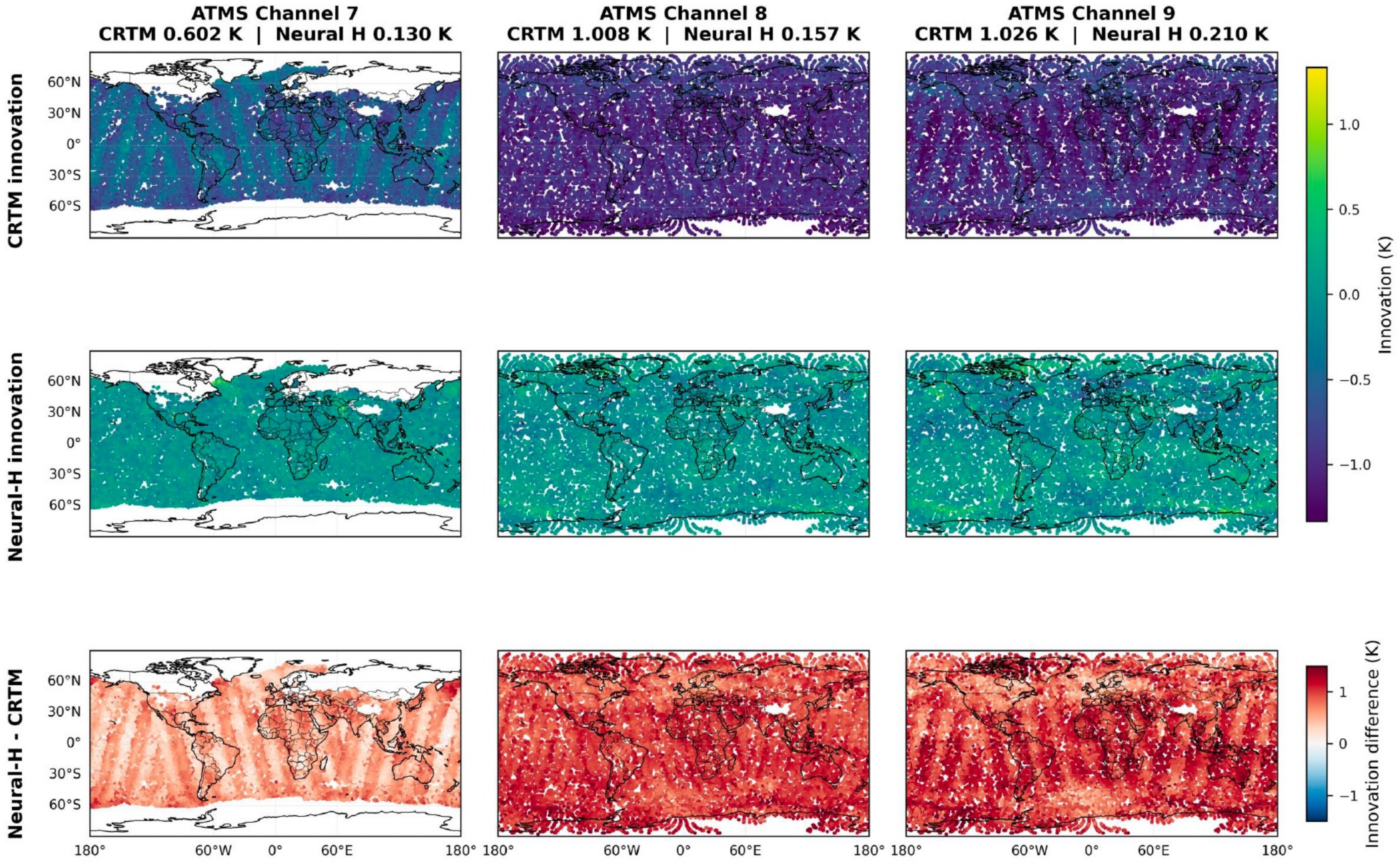
Learned observation operator H

Supplementary Table 1 Observation-space errors (K) for ATMS brightness temperatures. Centered SD is calculated after removing the channel-wise mean GSI/CRTM error.

Channel	H_ϕ RMSE	GSI/CRTM	
		Raw RMSE	Centered SD
ATMS 2	**2.476954**	3.155379	3.152329
ATMS 3	**1.958710**	2.625560	2.624673
ATMS 4	**1.164067**	1.959157	1.707803
ATMS 5	**0.460943**	1.074024	0.595001
ATMS 6	**0.192121**	0.408941	0.193317
ATMS 7	**0.137100**	0.607351	0.212642
ATMS 8	0.173212	0.992811	**0.140191**
ATMS 9	0.237705	1.009791	**0.202866**
ATMS 10	0.577272	0.716508	**0.210620**
ATMS 11	1.347094	0.692231	**0.220185**
ATMS 12	2.739497	0.793479	**0.220285**
ATMS 13	4.331110	1.065720	**0.311399**
ATMS 14	5.365727	1.181464	**0.450714**
ATMS 15	5.810463	1.011856	**0.699220**
ATMS 16	**3.129026**	3.432366	3.428995
ATMS 17	**1.964668**	2.423137	2.339071
ATMS 18	1.276355	1.653800	**1.229600**
ATMS 19	1.383326	1.398924	**1.170310**
ATMS 20	1.522965	1.392776	**1.199080**
ATMS 21	1.686363	1.348542	**1.252914**
ATMS 22	1.782852	1.314962	**1.314958**

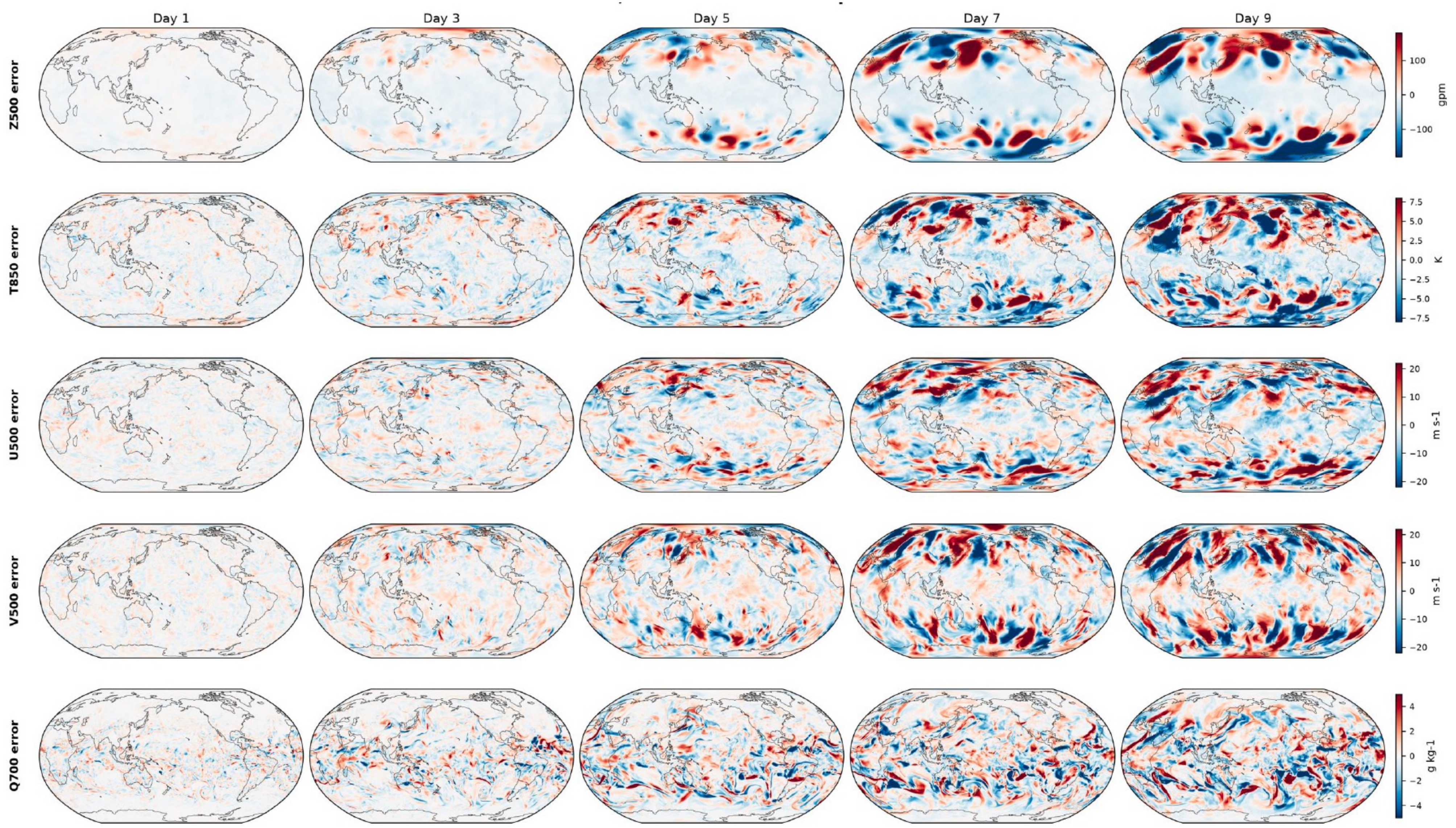


Learned observation operator H



16,384 matched FOVs per channel; no spatial interpolation

Visualization of Spatial Forecast Errors



Conclusion and Discussion

What We show

- DADA is the first framework to explicitly address the domain gap caused by changes in background states.
- DADA improves forecast performance with both AI-based and NWP-based backgrounds.
- DADA complements existing data assimilation systems because it still requires an informative background state.

Where the Field May Go

- Observation-to-prediction models such as AI-DOP avoid this particular background-state domain gap by design.
- If fully end-to-end observation-to-forecast learning becomes practical, DADA may no longer be necessary.
- However, the vast amount of accumulated gridded training data makes a complete transition difficult in the near term.

Practical Pathway

Observations → Data Assimilation → Background State → DADA → AI Forecast

Open to Collaboration

AI for weather and climate, and observation-driven research across domains