PLANETTE

Leveraging CESM2 data for machine learning based S2S forecasting

Anisha Pal
Scientific Machine Learning Engineer
Planette AI

NCAR CESM Workshop 2025
11 June 2025
Boulder, CO, USA

Planette AI Team


Anisha Pal


Dr. Hansi Singh


Dr. Kalai Ramea


Dr. Kyle Heyblom


Dr. Aodhan J-Sweeney

AI Foundation Models

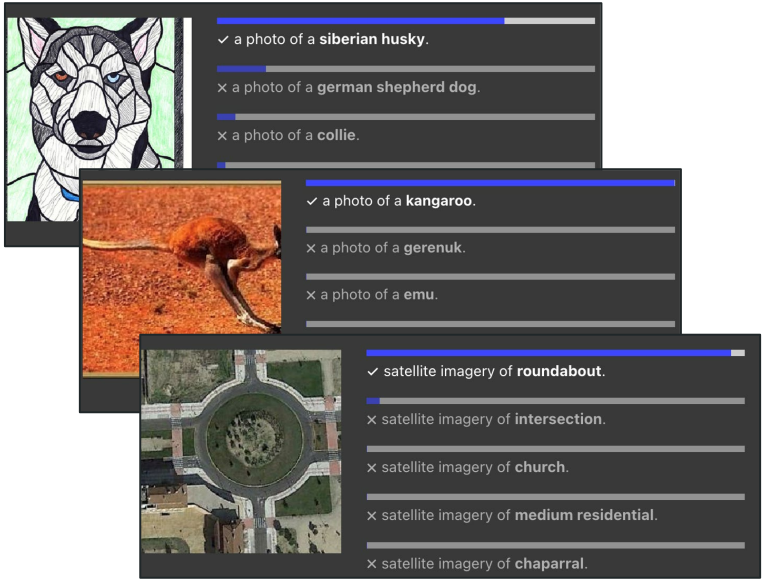
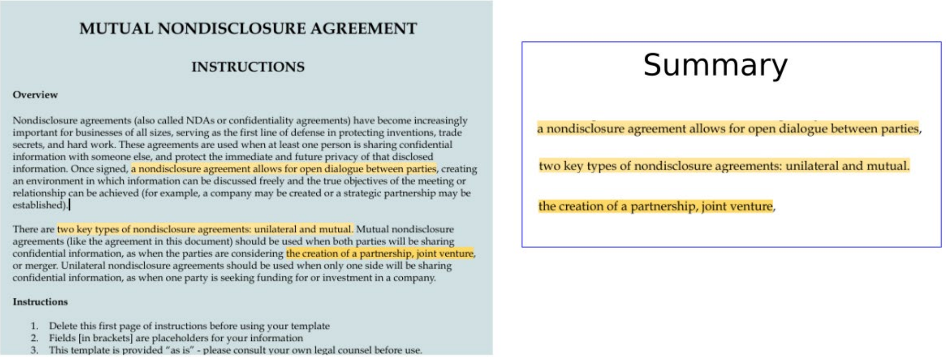


Image Classification [1]

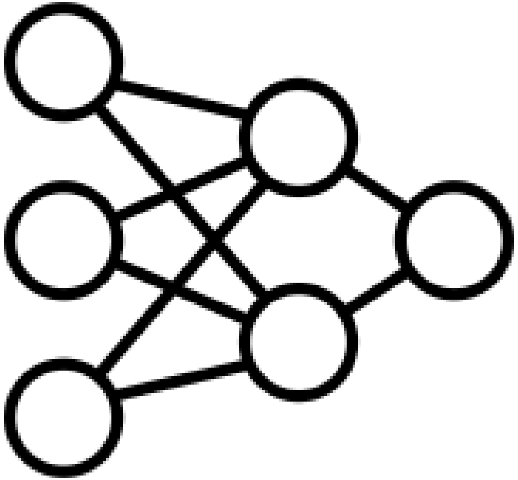
Image Segmentation [2]



Text Summarization [3]

Pretrain Broad ⟶ Finetune Precise

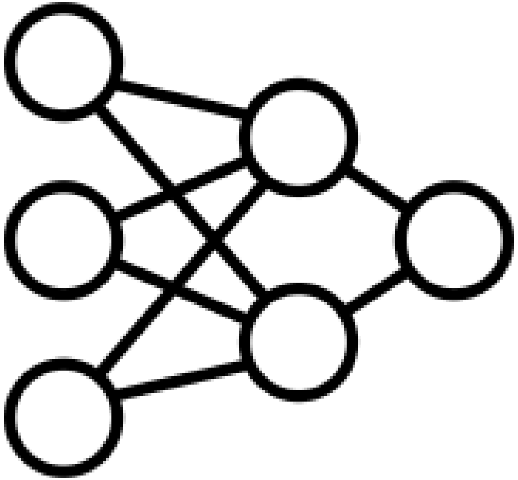
Data ~ 5M datapoints



Simple Objective/
Same as Finetuning
Objective

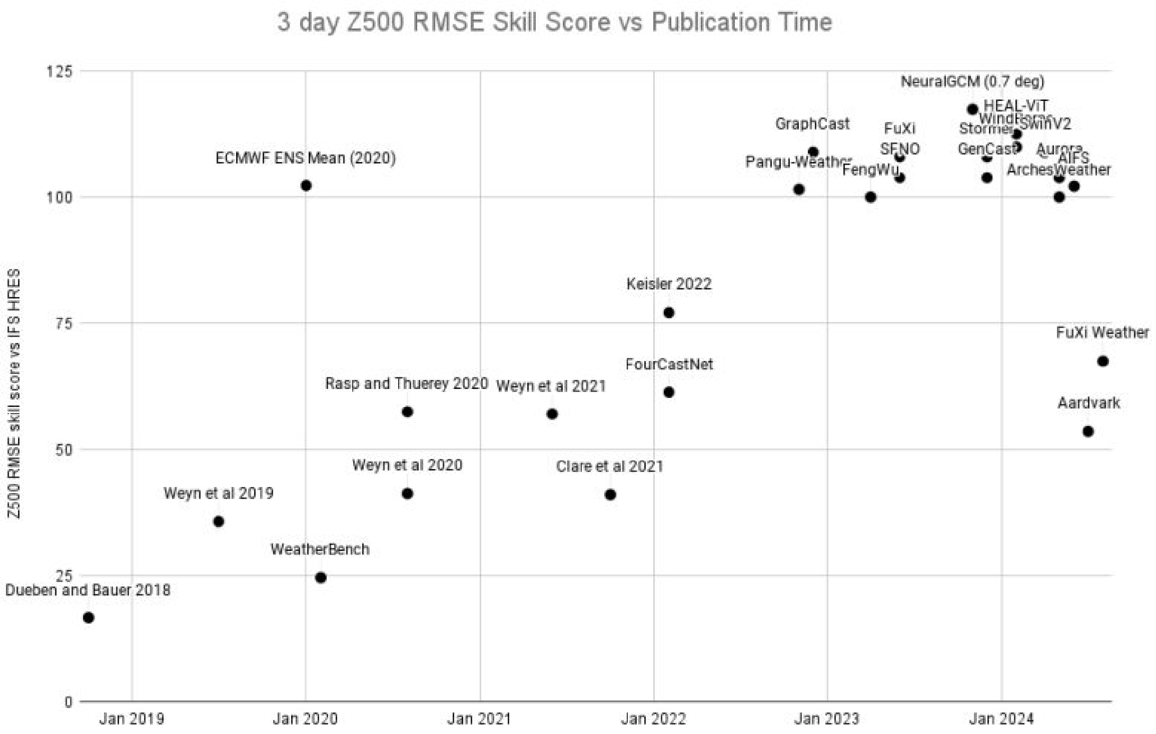
Pretrain Broad → Finetune Precise

Task Specific Data ~ 500



Specific Task

AI Based Weather Modelling



AI Weather model skill [4]

Data Source - ERA5

Temporal Scale - Hourly

Multiple Downstream Tasks

AI Based S2S Forecasting - Challenges

TASK COMPLEXITY ↑

DATA AVAILABILITY ↓

ERROR BUILDUP ↑

AI based **S2S** forecasting is **difficult** compared to AI based weather forecasting

Increased **task complexity** and additional constraints

Earth System Model Data to the rescue !

ERA5

- 1980 (1940) - 2024
- 527 months
- 1 ensemble member

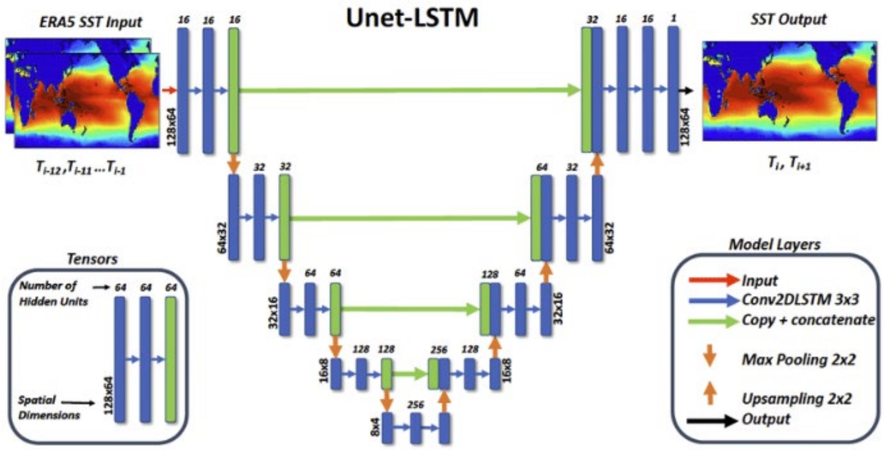
Total Data Points - **527**

CESM2

- 1850 - 2100
- ~3000 months
- 100 ensemble members

Total Data Points - **300,000** !

Our Model

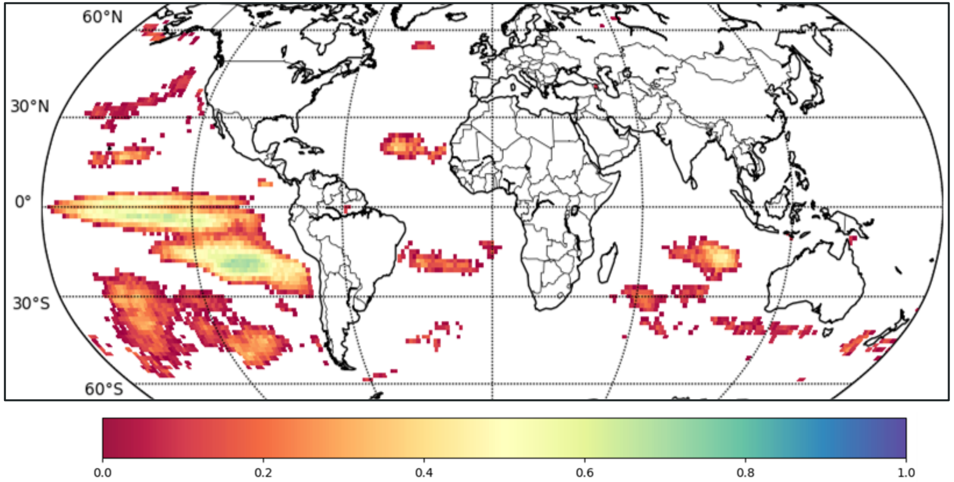


UnetLSTM architecture [5]

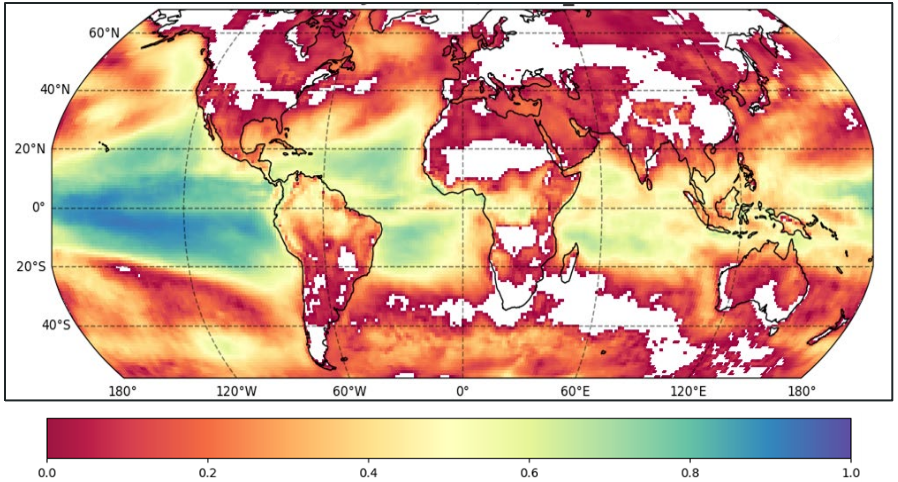
- **Unet-LSTM** - Spatiotemporal
- **Quantile Regression**
- Training Data - **CESM; 1850 - 1922**
- Test Data - **ERA5; 2008 - 2024**
- Grid Resolution - **0.9° x 1.25 °**
- Grid - **60°S to 60°N** , 180°W to 180°E
- Input - **11** variables
- Output - **T2M, Precip**
- Lead Times - **{0,1,2,3} months**
- **Metric -** Anomaly R2 ↑

Effect of Training Data on ERA5 Test set

T2M Anomalies



Trained on ERA5



Trained on CESM2

CESM data helps the model learn important large scale features such as ENSO

Lead - 0 month

Increasing Ensemble Members

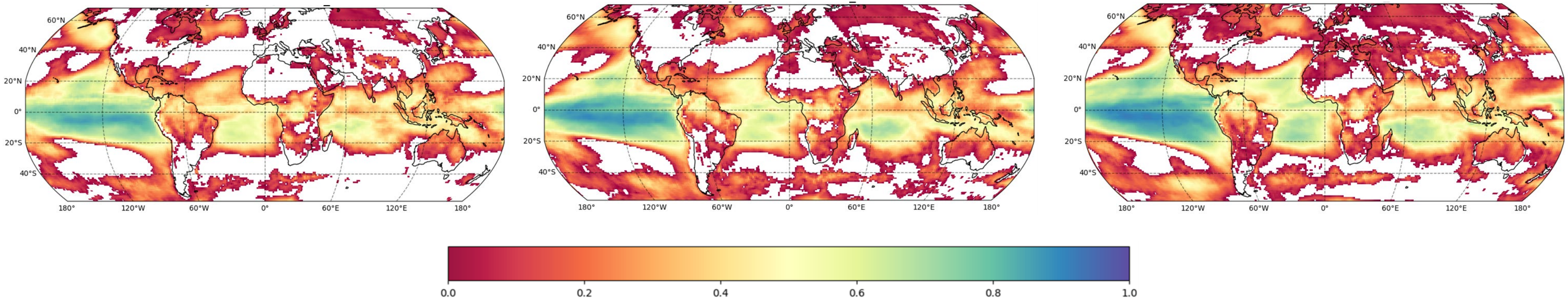
T2M Anomalies

Ensemble Members →

10

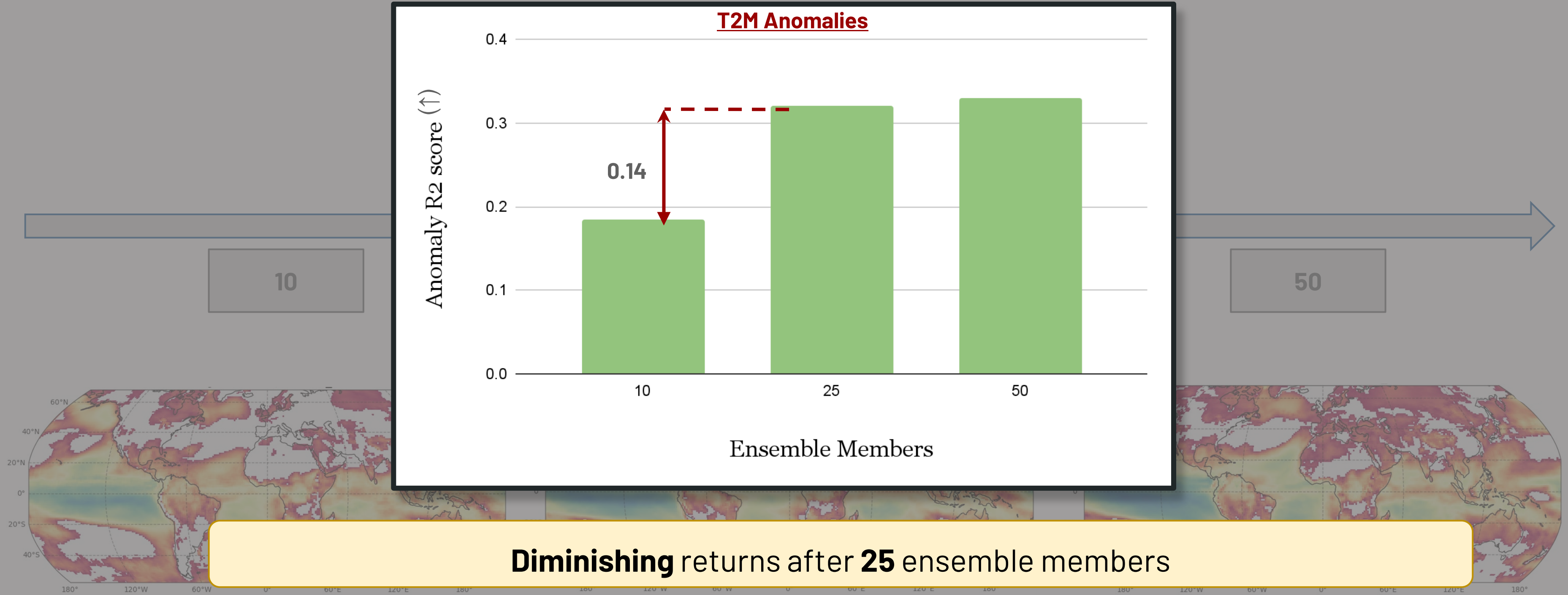
25

50



Lead - 0 month

Increasing Ensemble Members

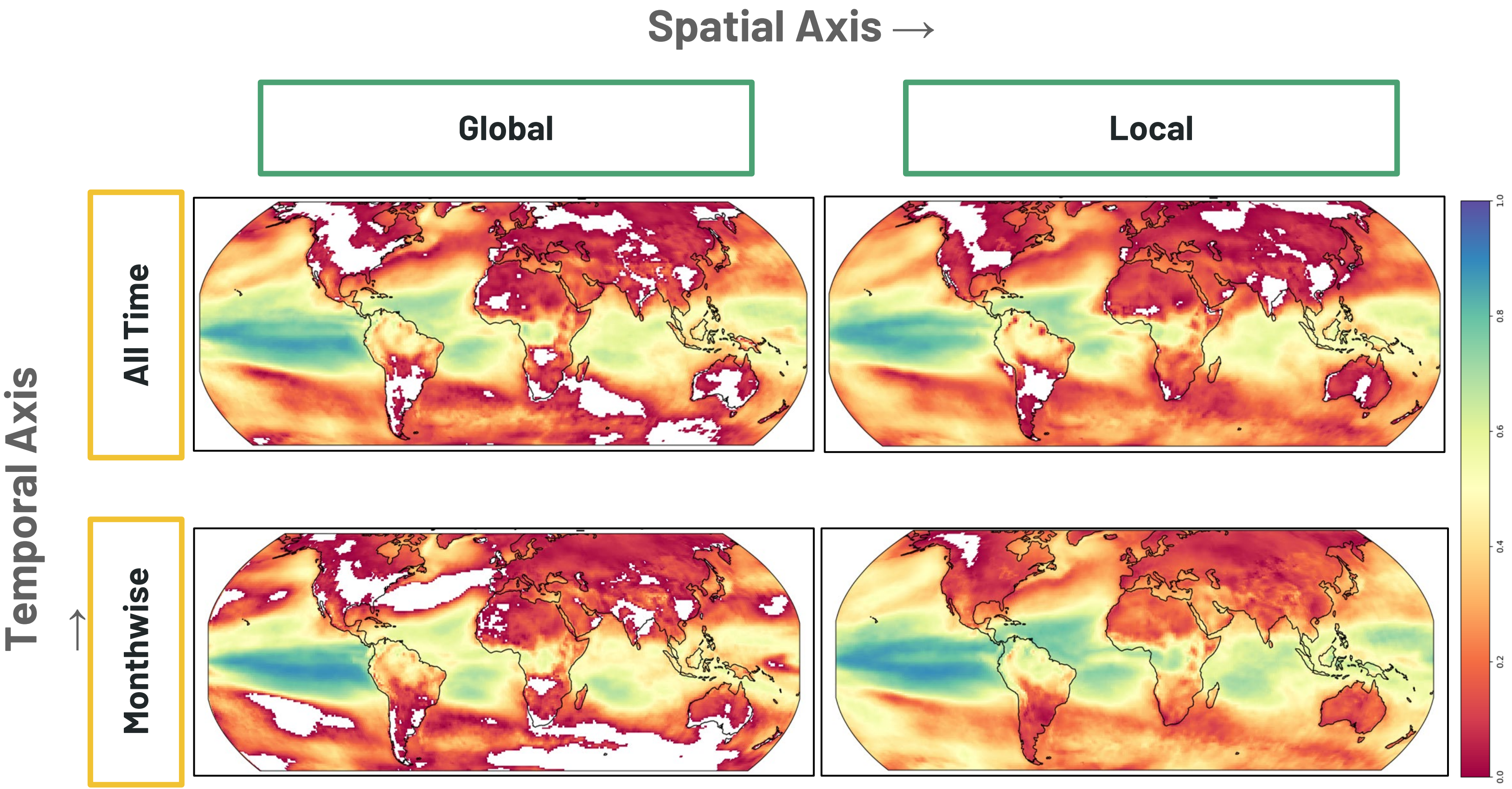


Diminishing returns after **25** ensemble members

Additional ensemble members contribute **minimally** to **unlearned large scale** variability

Capturing **additional patterns** from the noise requires a more **complex model** → **NIVA**

Effect of Data Normalization Strategies on **T2M Anomalies**



Lead - 0 month

Effect of Data Normalization Strategies on T2M

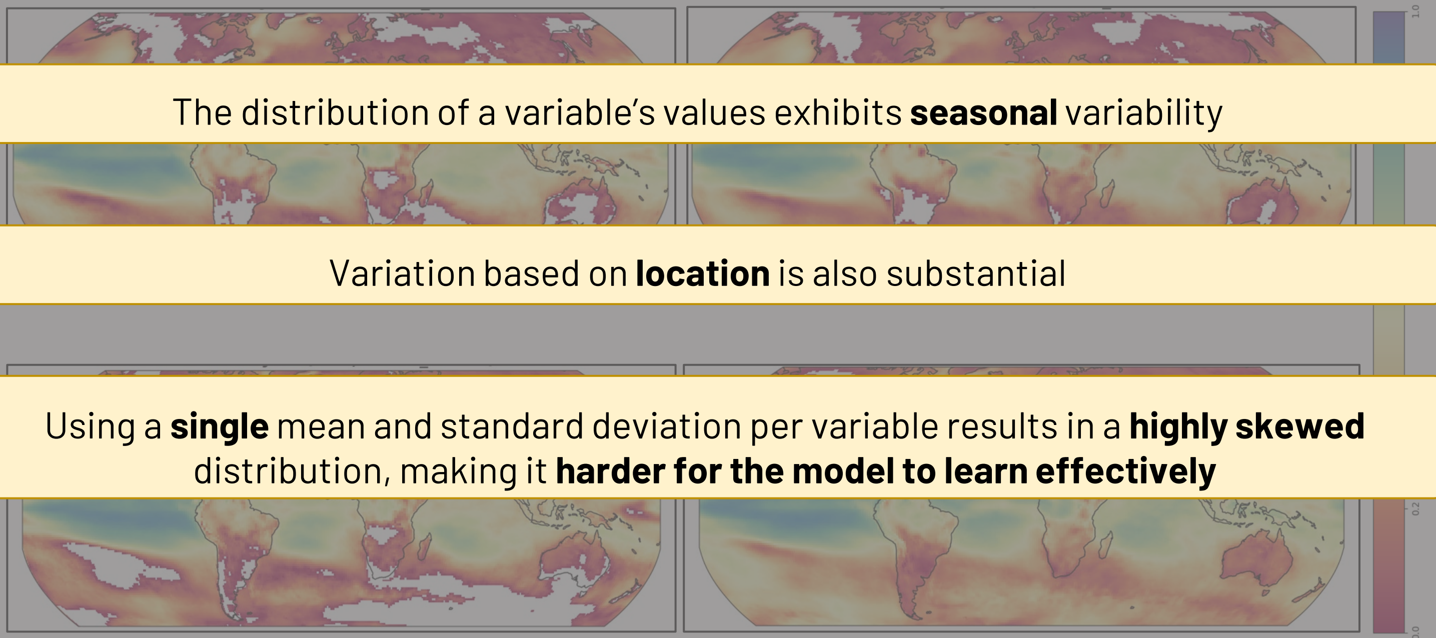
Spatial Axis →

Global | Local

Temporal Axis ↑

All Time

Monthwise



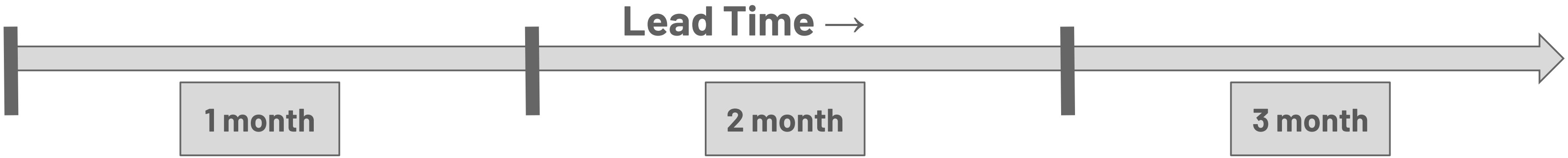
The distribution of a variable's values exhibits **seasonal** variability

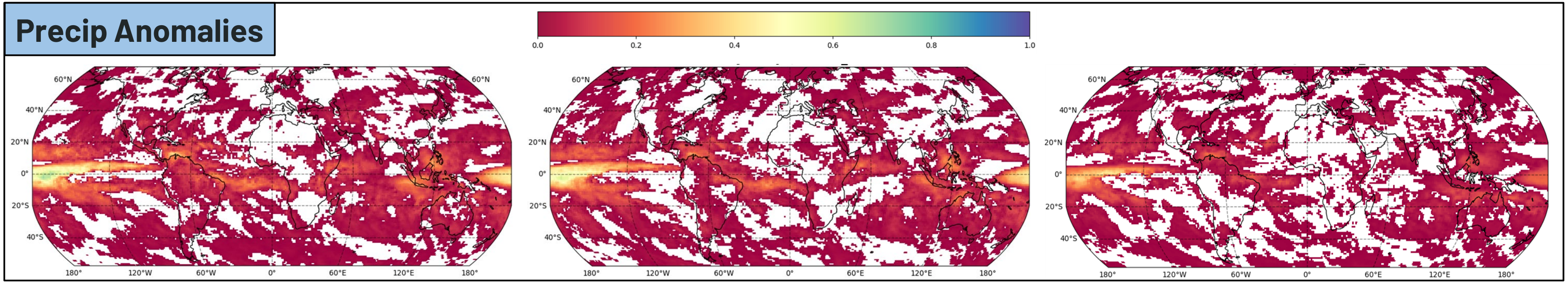
Variation based on **location** is also substantial

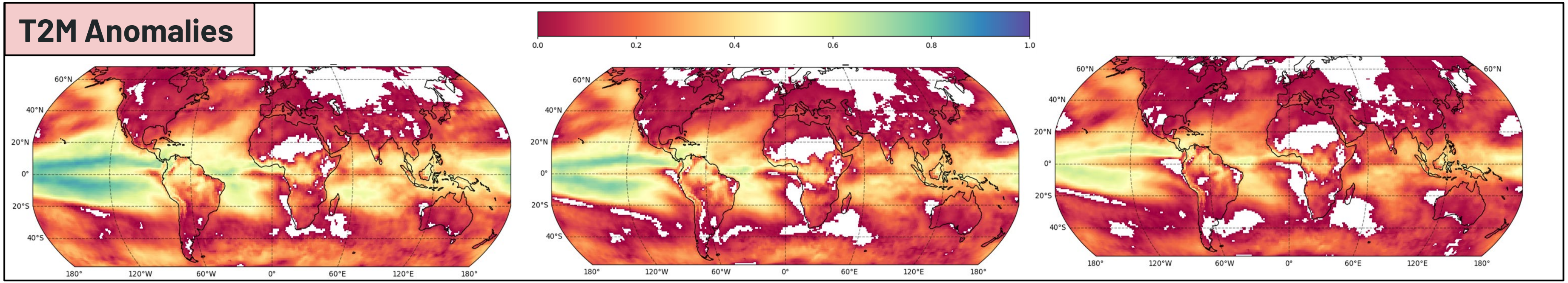
Using a **single** mean and standard deviation per variable results in a **highly skewed** distribution, making it **harder for the model to learn effectively**

Lead - 0 month

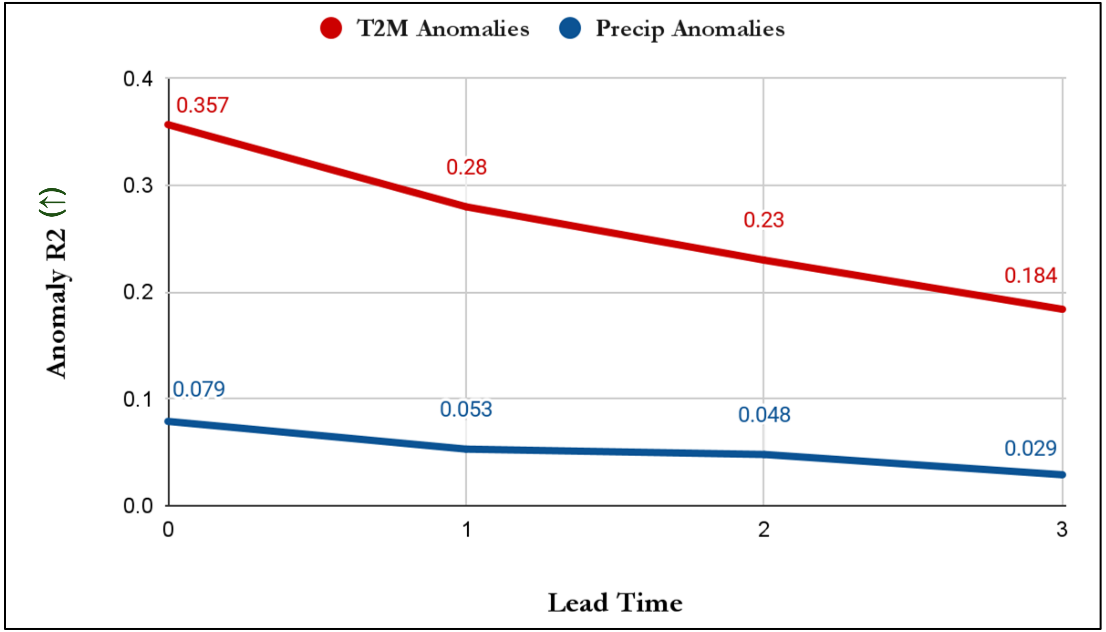
Different Lead Times







Different Lead Times



Model demonstrates greater skill in predicting **temperature** than precipitation

Predictive skill **declines with lead time**, but not as sharply as anticipated.

Key Takeaways

CESM2 data can help create the **next generation of AI foundation models** for S2S scale

Performance is very **sensitive the distribution of data** - picking the right normalization scheme is important

More data is beneficial but **model complexity** also needs to be scaled accordingly

References

[1] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., … Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *arXiv [Cs.CV]*. Retrieved from http://arxiv.org/abs/2103.00020

[2] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., … Girshick, R. (2023). Segment Anything. *arXiv [Cs.CV]*. Retrieved from http://arxiv.org/abs/2304.02643

[3] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv [Cs.CL]*. Retrieved from http://arxiv.org/abs/1810.04805

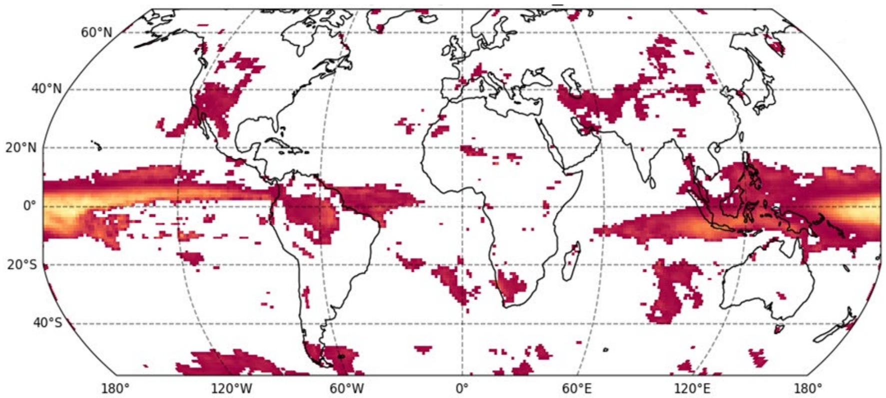
[4] Rasp, S. (2024, November 13). *Here is an updated version of my chart showing the progress of AI-weather. . . | Stephan Rasp | 22 comments [Online forum post]*.https://www.linkedin.com/posts/stephan-rasp-4154481aa_here-is-an-updated-version-of-my-chart-showing-activity-7262397352424992770-7TJU

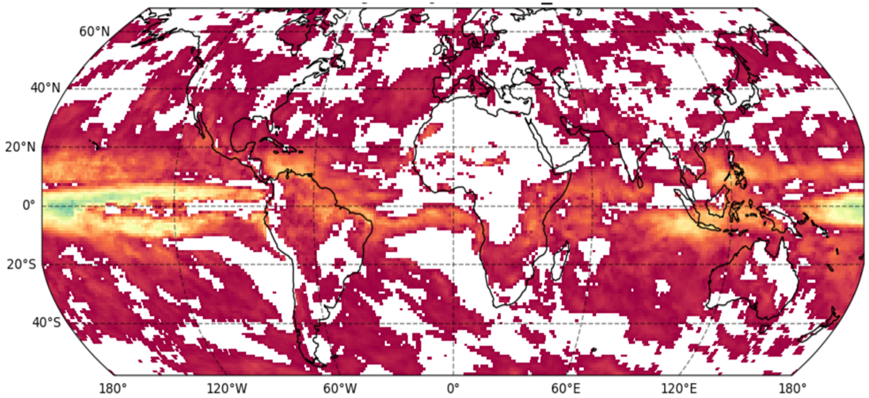
[5] Taylor J and Feng M (2022) A deep learning model for forecasting global monthly mean sea surface temperature anomalies. *Front. Clim.* 4:932932. doi: 10.3389/fclim.2022.932932

Thank You !
Any Questions ?

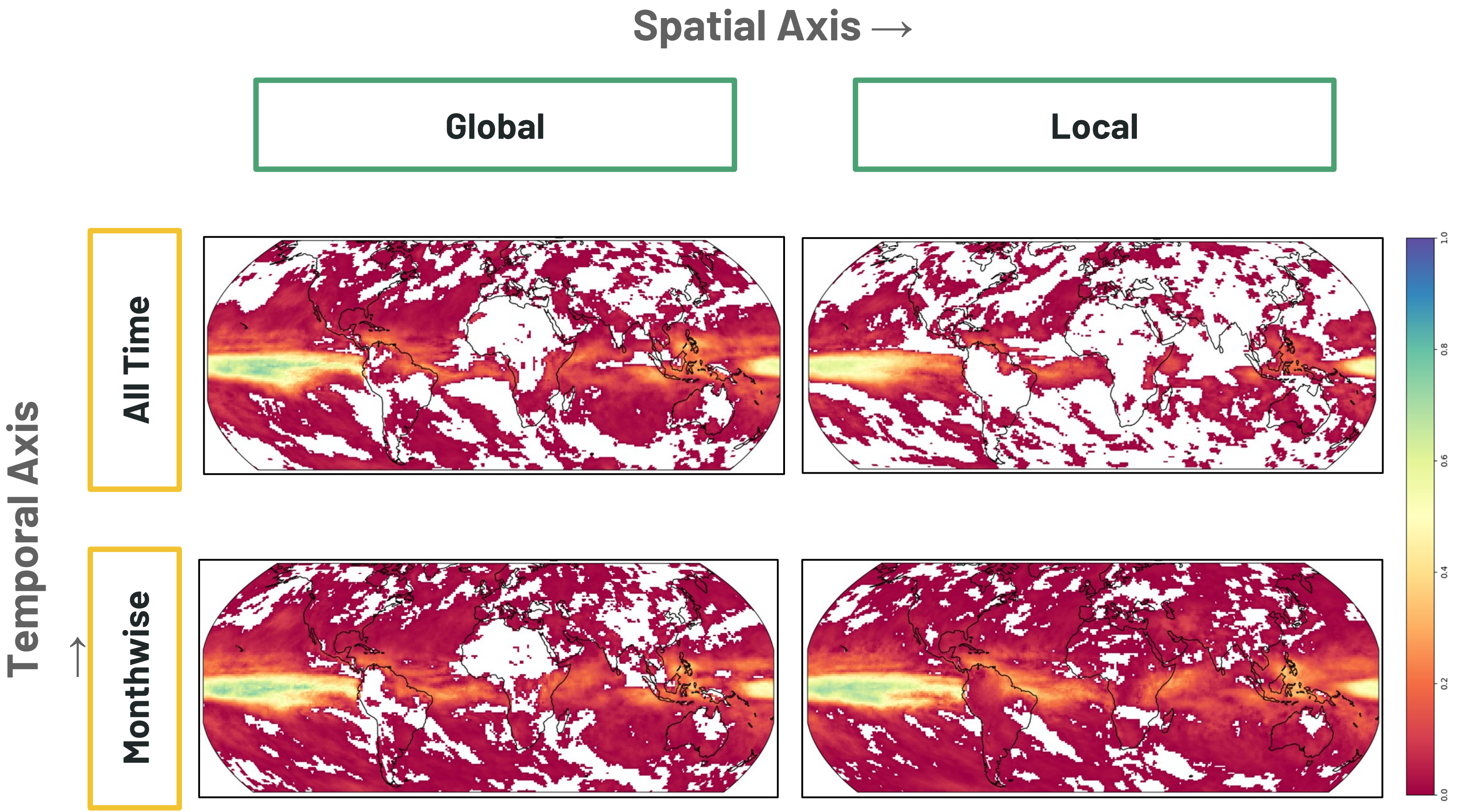
Extra Slides

Effect of Training Data





Effect of Data Normalization Strategies on **Precipitation**

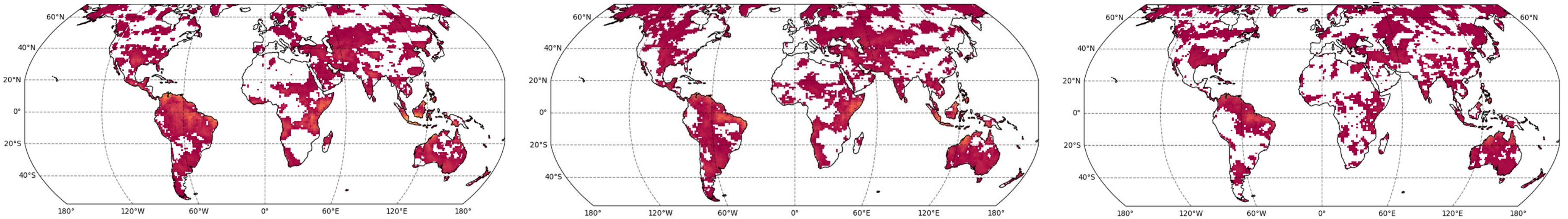


Zero Shot vs Fine Tuning **Precipitation (Land)**

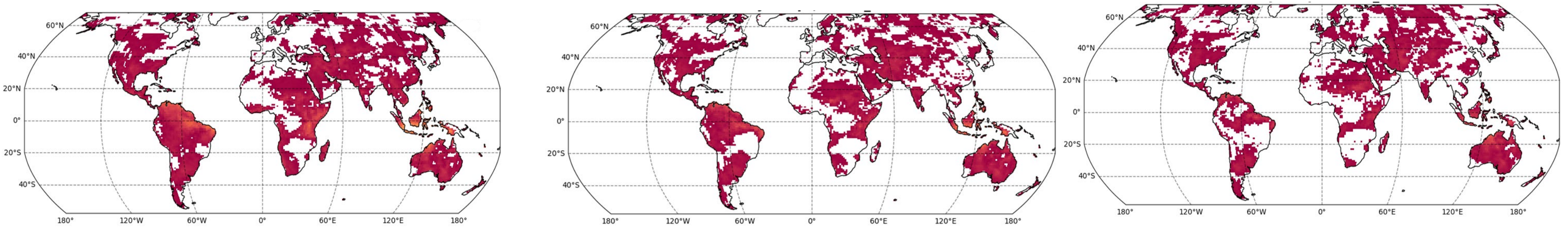
Lead Time →

1 month | **2 month** | **3 month**

Zero Shot



Fine Tuned



Zero Shot vs Fine Tuning T2M (Land)

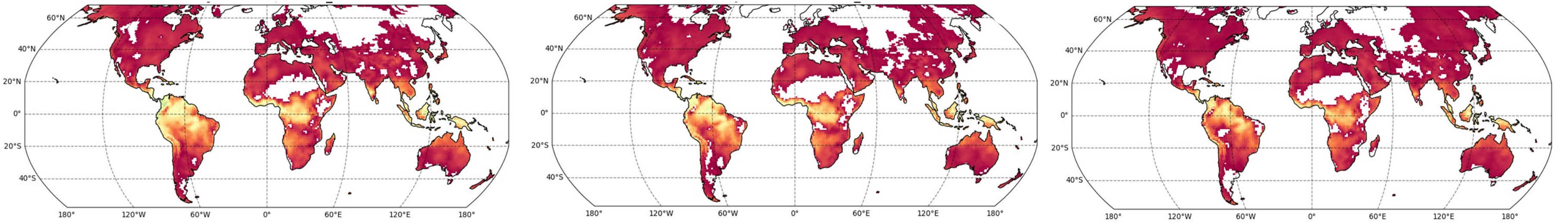
Lead Time →

1 month

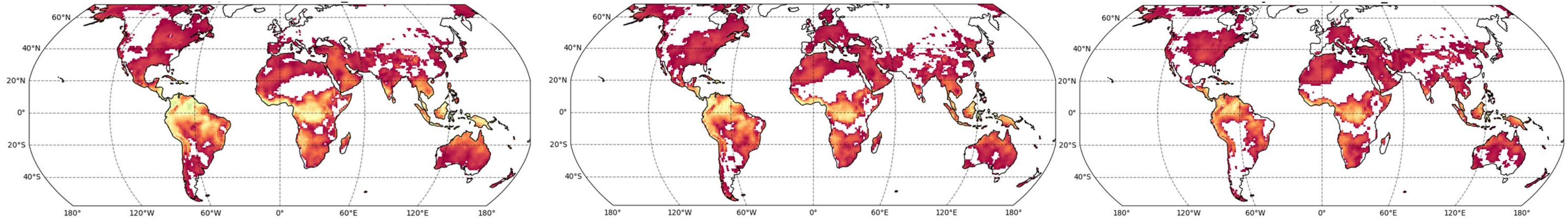
2 month

3 month

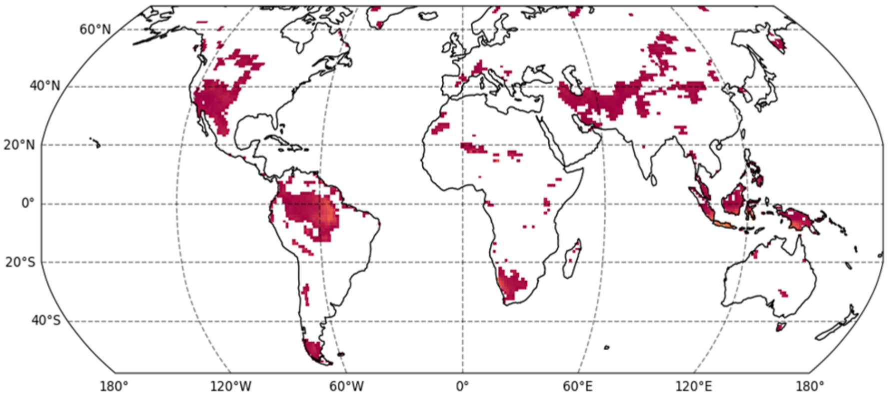
Zero Shot

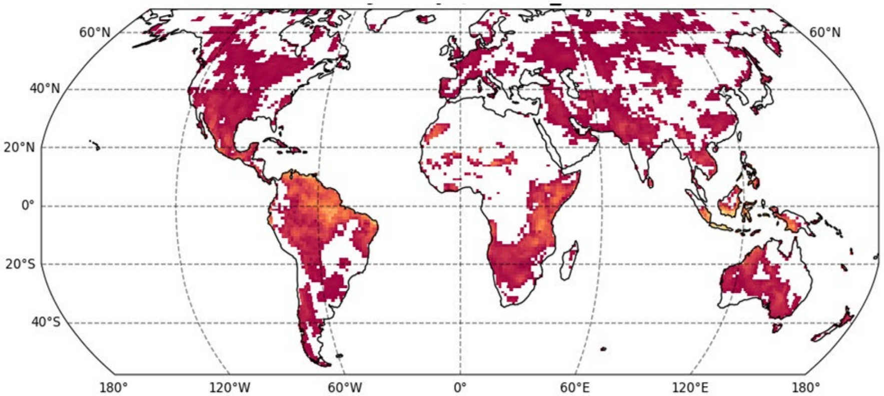


Fine Tuned

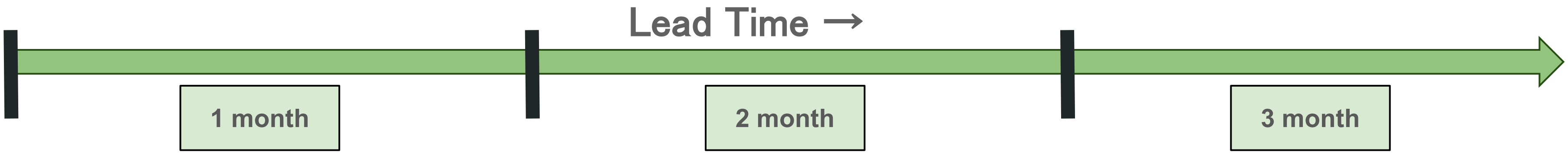


Effect of Training Data

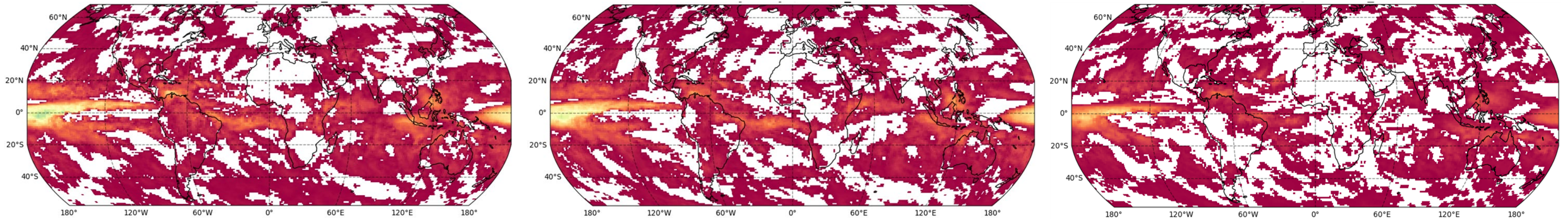




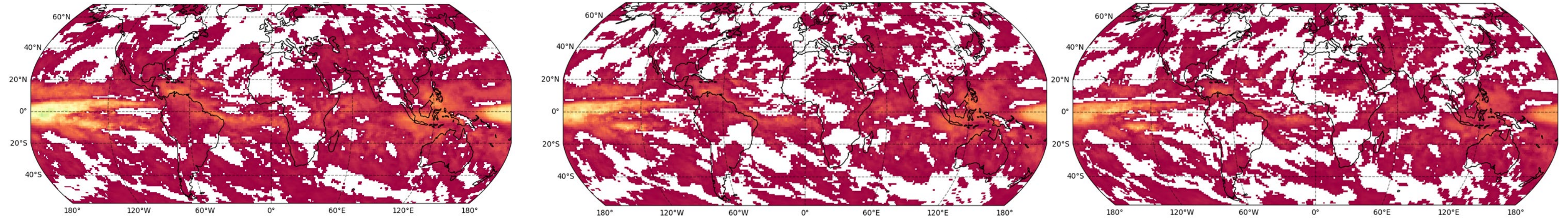
Zero Shot vs Fine Tuning Precipitation



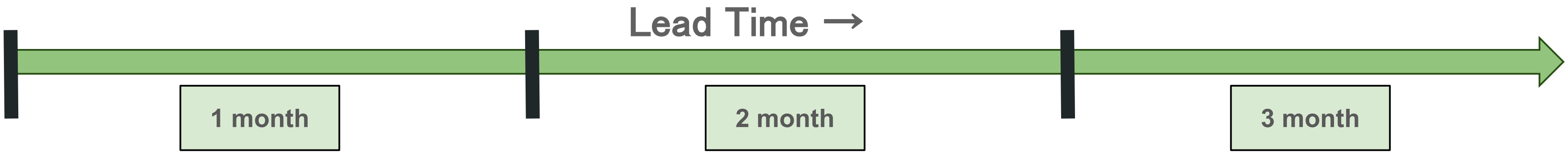
Zero Shot

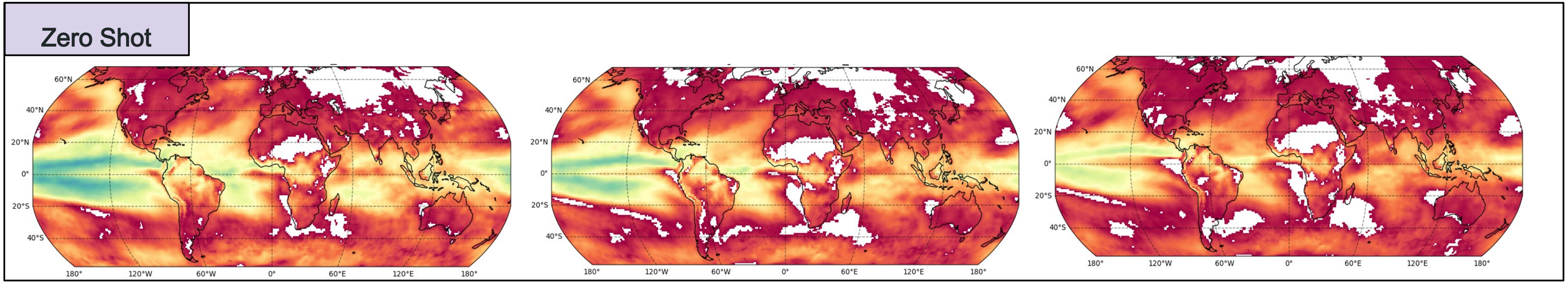


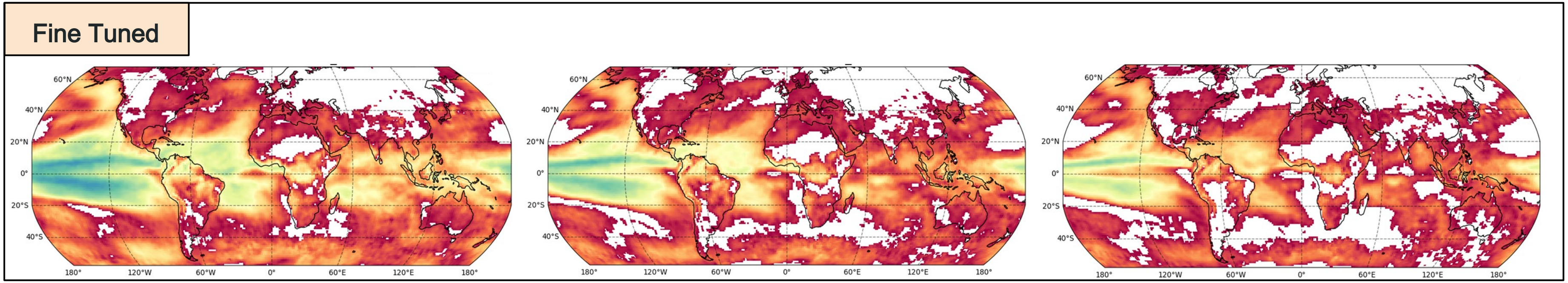
Fine Tuned



Zero Shot vs Fine Tuning T2M







Values

		Precip (All)			
1 month		2 month		3 month	
Zero Shot	Fine Tuned	Zero Shot	Fine Tuned	Zero Shot	Fine Tuned
0.053	0.054	0.048	0.049	0.029	0.036

		Precip (Land)			
1 month		2 month		3 month	
Zero Shot	Fine Tuned	Zero Shot	Fine Tuned	Zero Shot	Fine Tuned
0.043	0.0375	0.035	0.032	0.021	0.0311

Values

		T2M (All)			
1 month		2 month		3 month	
Zero Shot	Fine Tuned	Zero Shot	Fine Tuned	Zero Shot	Fine Tuned
0.28	0.3	0.23	0.24	0.184	0.2

		T2M (Land)			
1 month		2 month		3 month	
Zero Shot	Fine Tuned	Zero Shot	Fine Tuned	Zero Shot	Fine Tuned
0.16	0.18	0.13	0.15	0.116	0.12