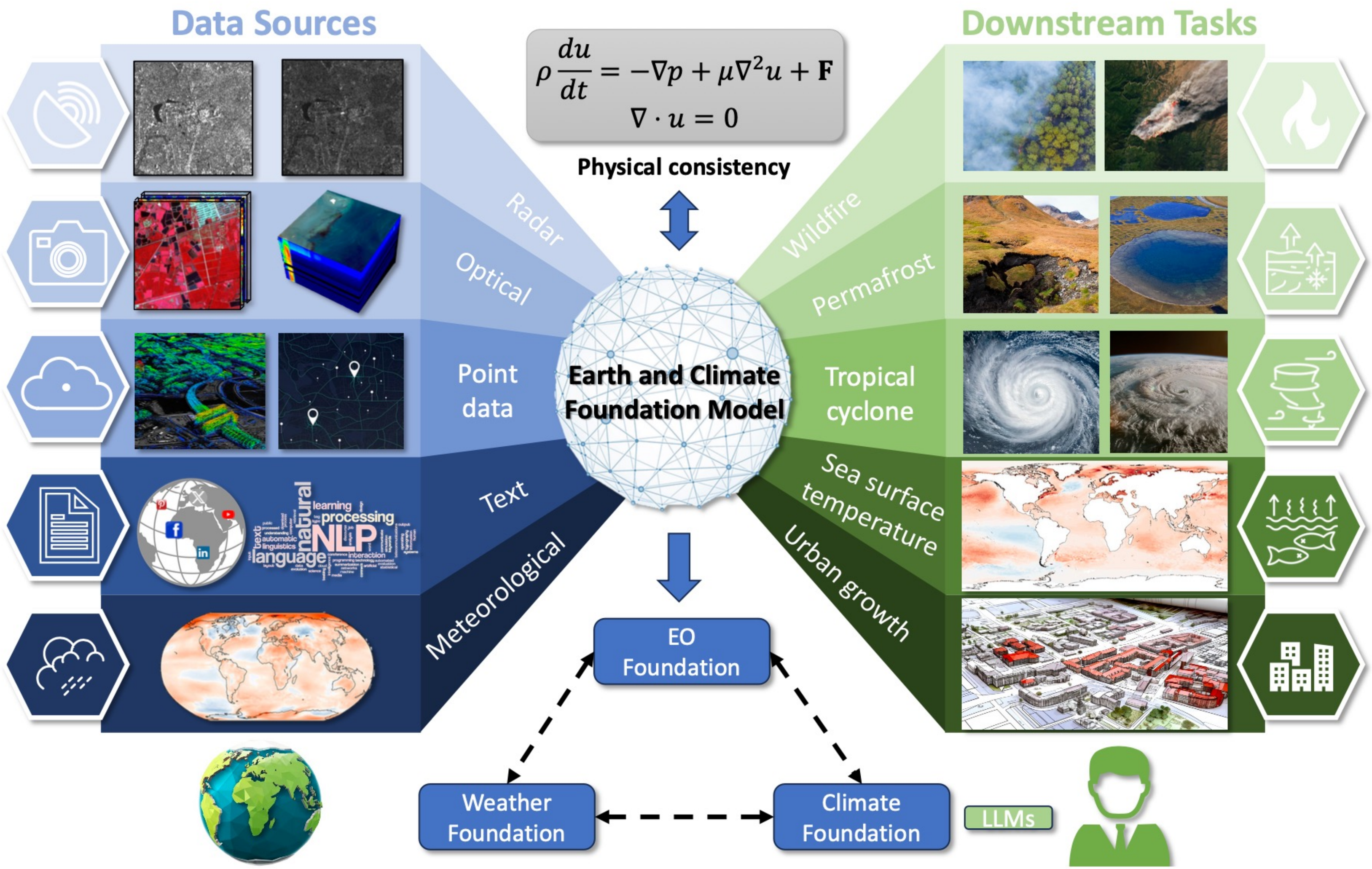
AI Foundation Models for Whole Atmosphere Climate: Development and Applications

Wenjun Dong[1,2]
Colleagues: David C. Fritts[1,2], Thoms Lund[2], Alan Liu[1], Han-Li Liu[3], Jonathan Snively[2], and Jia Yue[4]

[1]Center for Space and Atmospheric Research (CSAR), Embry-Riddle Aeronautical University, Daytona Beach, USA
[2]Global Atmospheric Technologies and Sciences (GATS), Boulder, USA
[3]High Altitude Observatory (HAO), National Center for Atmospheric Research (NCAR)), Boulder, USA
[4]NASA Goddard Flight Center

June 5, 2024
CCMC Workshop, Maryland on June 3-7, 2024

AI Foundation Model: Any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks.



Source: Zhu et al., 2024

Earth Foundation Model

Type	Definition
Earth Observation	Process and analyze data collected from Earth observation systems, such as satellites, radios, and ground-based sensors.
Weather	Predict short- to medium-term weather conditions based on atmospheric data. These models use meteorological data, including historical weather records and real-time observations, to forecast weather patterns and extreme weather events, providing essential information for weather prediction and nowcasting.
Climate	Aim at projecting long-term climate trends and changes based on historical and simulated climate data. These models integrate various data sources, including climate model outputs and historical climate records, to predict future climate scenarios and inform policy and planning for climate change mitigation and adaptation.

Recent Weather and Climate Foundation Models

Model Architecture	Model	Year	Affiliation	Temporal Resolution	Time Span	Spatial Resolution	Model Size (# params)	Cluster Size	Training Time
GNN	GraphCast	2022.12	Google	6 hrs	10 days	0.25°	36.7M	32 TPU v4	28 days
Transformer	FourCastNet	2022.02	NVIDIA	6 hrs	8.3 days	0.25°	433M–2.19B	64 A100	0.7 days
	Pangu	2022.11	Huawei	1 hr	7 days	0.25°	256M	192 V100	15 days
	ClimaX	2023.01	Microsoft	6 hrs	30 days	0.25°	108–111M	80 V100	3 days
	FengWu	2023.04	Shanghai AI	6 hrs	14 days	0.25°	427M	32 A100	17 days
	W-MAE	2023.04	Chengdu	6 hrs	16 days	0.25°	96.5M	8 A800	9.3 days
	FuXi	2023.06	Fudan	6 hrs	15 days	0.25°	1.56B	8 A100	1.5 days
	Stormer	2023.12	UCLA	6 hrs	10 days	1.4°	100–600M	128 A100	1 day
	FengWu-4DVar	2023.12	Shanghai AI	6 hrs	10 days	0.25°	427M	32 A100	17 days
	FengWu-GHR	2024.01	Shanghai AI	6 hrs	10 days	0.09°	4B	16 A100	3 days
	CaFA	2024.05	CMU	6 hrs	7 days	1.5°	200 M	4 A6000	7 days
	AUORA	2024.05	Microsoft	6 hrs	10 days	0.25°	1.3B	32 A100	17 days
Diffusion	SEEDs	2023.06	Google	N/A	16 days	2°	114M	16 TPU v4	0.75 days
	GenCast	2023.12	Google	12 hrs	15 days	1°	60M	32 TPU v4	5 days
Differential GCM	NeuralGCM	2023.11	Google	3.75–12 mins	10 days	0.7–2.8°	115–311M	128 TPU v5e	10 days

Note: Please see the last page for the related publication list

Recent Weather and Climate Foundation Models

2022

FourCastNet — *First AI Weather Foundation Model*

Pangu — *A 3D transformer equipped with earth-specific priors, showing the importance of embedding earth location in machine learning models*

GraphCast — *ML models that are now competitive with traditional weather forecasting methods*

2023

ClimaX — *ML can be trained using heterogeneous datasets spanning different variables, spatial-temporal coverage, and physical groundings*

FengWu — *Pushes skillful global medium-range weather forecasts beyond 10 days lead time.*

W-MAE — *Demonstrates the feasibility and potential of implementing pre-training techniques in weather forecasting*

FuXi — *Uses cascade machine learning to extend skillful global medium-range weather forecasts beyond 15 days*

SEEDs — *Provides an alternative to current operational weather forecasting paradigms, allowing computational resources saved by the statistical emulator to increase the resolution of physics-based model*

NeuralGCM — *The first GCM combining a differentiable solver for atmospheric dynamics with machine learning components.*

GenCast — *A probabilistic weather model with greater skill and speed than the top operational medium-range weather forecasts.*

2024

FengWu-GHR — *The first AI model running at 0.09° horizontal resolution, coarse-resolution models could be fine-tuned to achieve high resolution.*

FengWu-4DVar — *Integrates data assimilation with AI models*

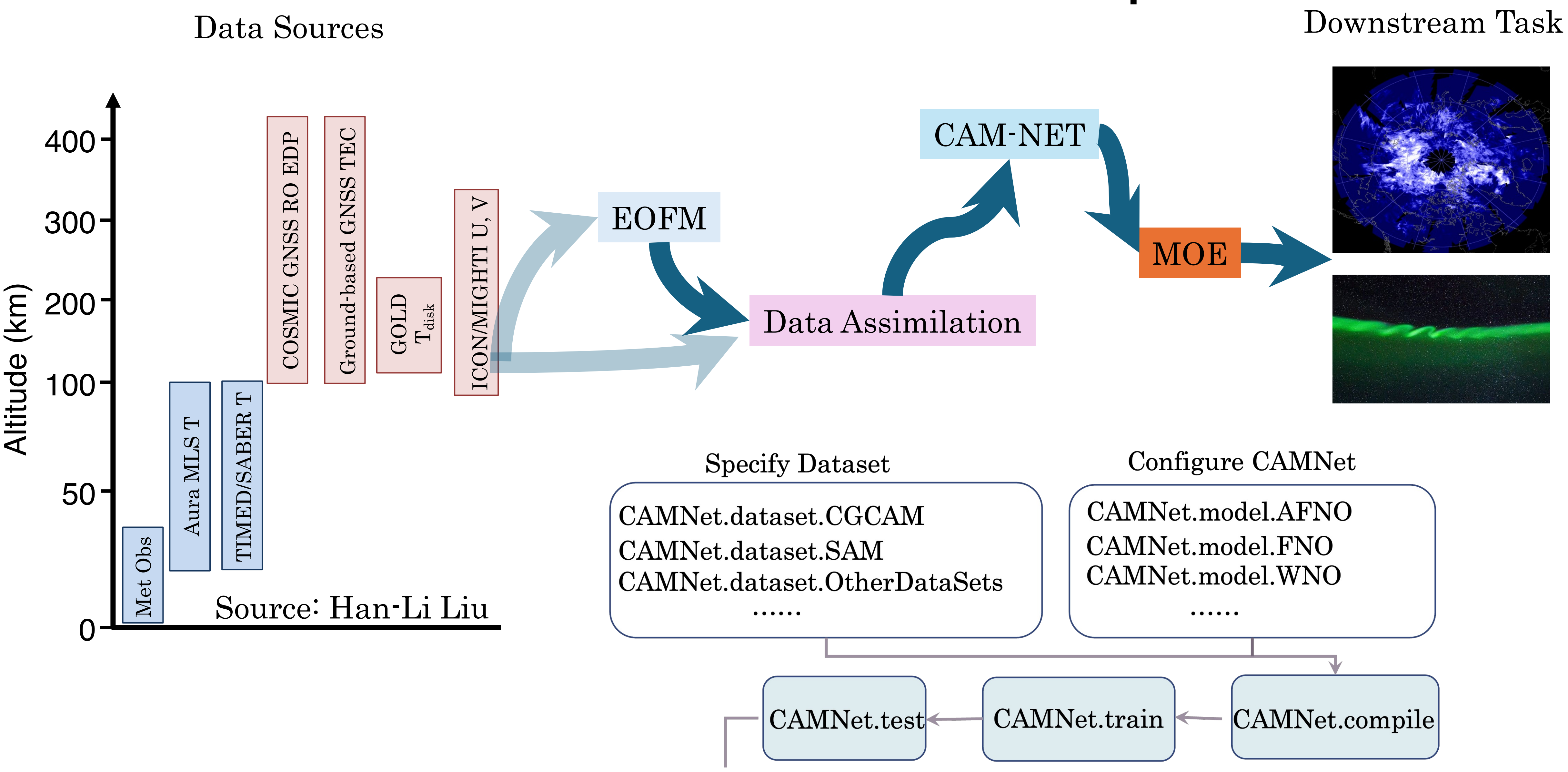
CaFA — *A factorized-attention-based model tailored for spherical geometries, reducing the quadratic complexity of attention to linear*

AUORA — *Can be pre-trained on heterogeneous datasets and fine-tuned for specific tasks.*

Key Features:

- ➢ AI models have demonstrated the potential to surpass traditional numerical physical models in accuracy while achieving several orders of magnitude increases in speed.
- ➢ AI models have been shown to successfully integrate with data assimilation techniques.
- ➢ AI models can handle heterogeneous datasets, including observations, simulations, and reanalysis datasets.

Unified Framework for Whole Atmosphere Model



Get Started with CAM-NET Using Just a Few Lines of Code

```
import os
import torch
import camnet
```

Step 1: Import CAM-NET package

```
torch.cuda.set_device(0)
device = torch.device("cuda")

data_path = "/home/datasets/CGCAM.mat"
```

Step 2: Load training datasets

```
train_loader, test_loader = camnet.datasets.cgcam(data_path, batch_size = 10, T_in = 10, T_out = 40,
type = "1e-3", sub = 1 )

ep = 50; o = 20; m = 12
```

Step 3: Configure model architecture and architecture parameters

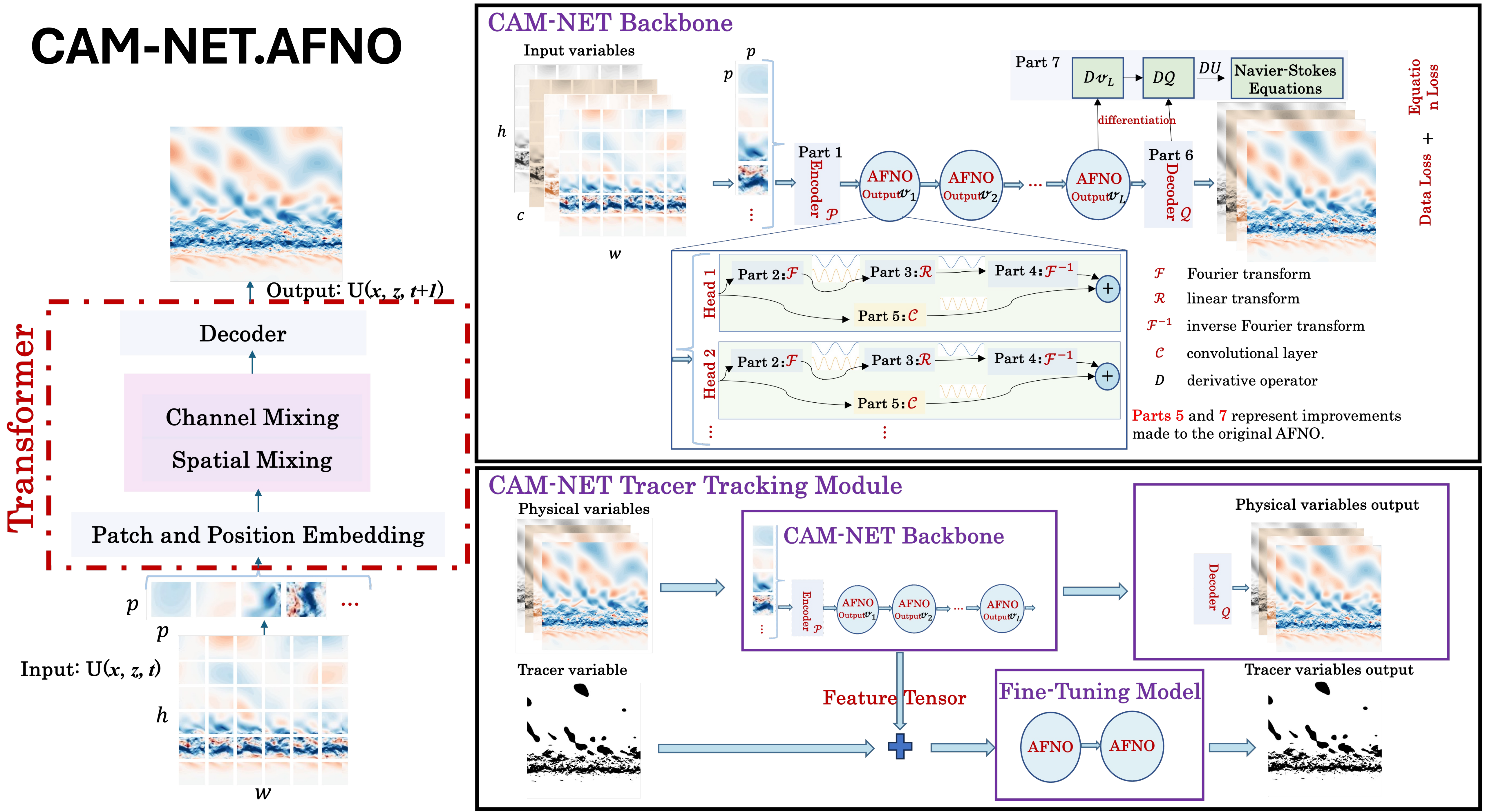
```
camnet_model = camnet.model.fourier(backbone = "FNO2d", o = o, m = m, t_in = 10, device = device)
camnet_model.compile()
```

Step 4: Configure training parameters and start training

```
camnet_model.opt_init("Adam", lr = 0.001, step_size=100, gamma=0.5)
camnet_model.train(epochs=ep, trainloader = train_loader, evalloader = test_loader)
```

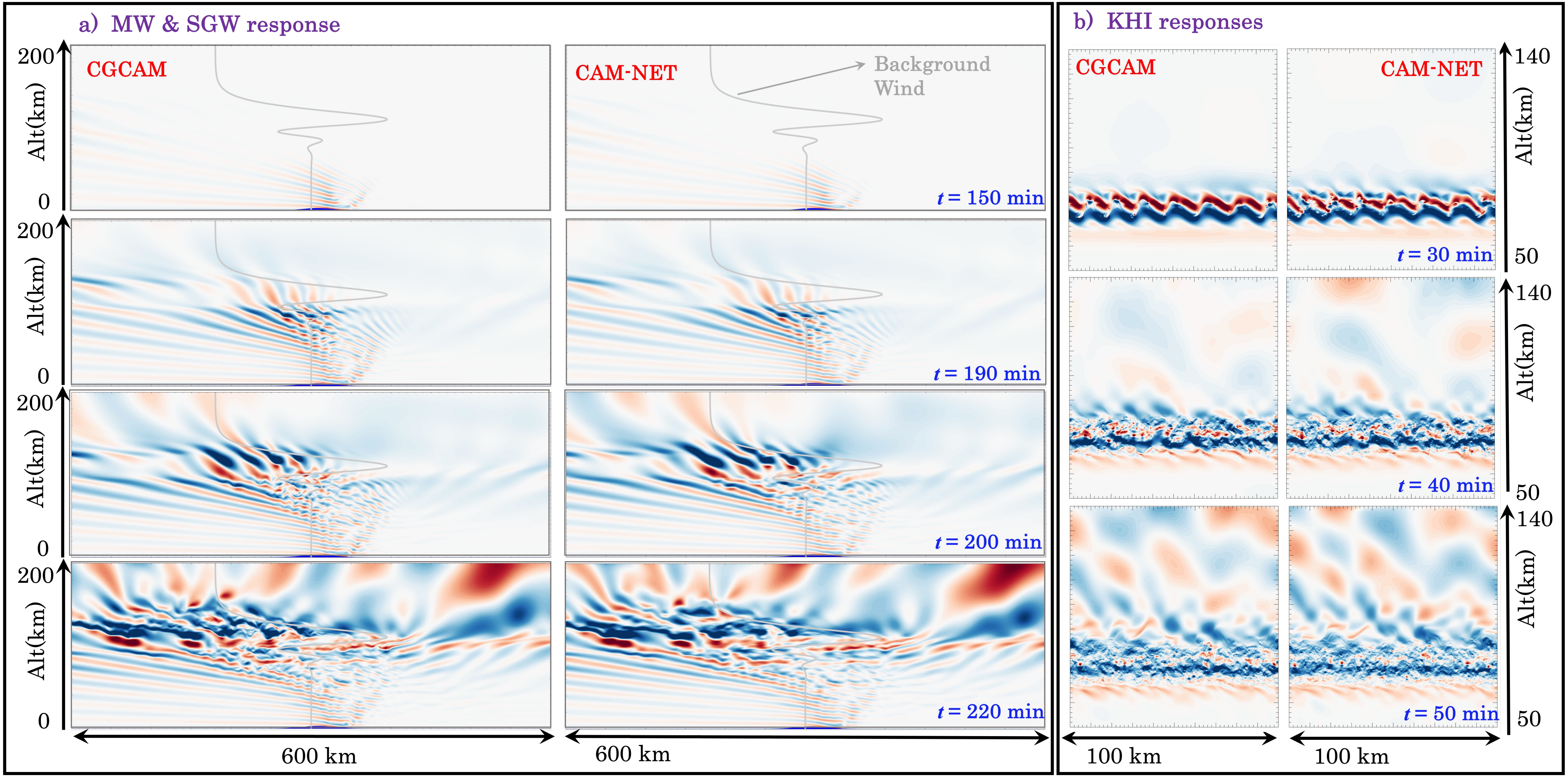
```
time_error = camnet_model.test(test_loader, path = fig_path, is_save = True, is_plot = True)
```

Step 5: Test Outputs

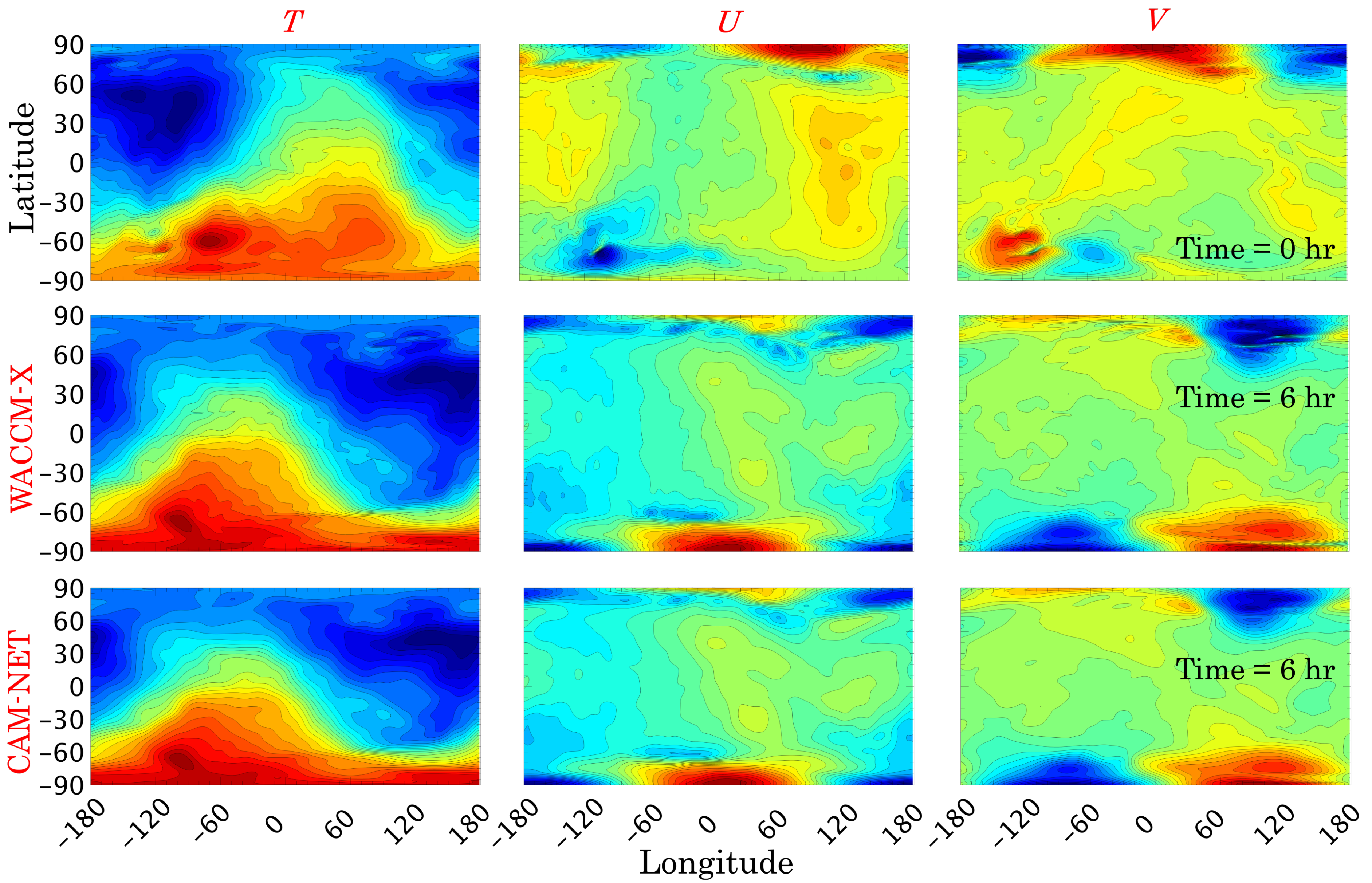


CAM-NET.AFNO
Output: U(x, z, t+1)
Decoder
Channel Mixing
Spatial Mixing
Patch and Position Embedding
Transformer
Input: U(x, z, t)
CAM-NET Backbone
Input variables
Part 7
Navier-Stokes Equations
differentiation
Part 1
Encoder
AFNO
Part 6
Decoder
Equation Loss
+
Data Loss
Head 1
Head 2
Part 2: ℱ
Part 3: ℛ
Part 4: ℱ⁻¹
Part 5: 𝒞
ℱ Fourier transform
ℛ linear transform
ℱ⁻¹ inverse Fourier transform
𝒞 convolutional layer
D derivative operator
Parts 5 and 7 represent improvements made to the original AFNO.
CAM-NET Tracer Tracking Module
Physical variables
Physical variables output
Tracer variable
Feature Tensor
Fine-Tuning Model
Tracer variables output

Example: Highly-Nonlinear Wave Dynamics



Example: WACCM-X and CAM-NET for Surface Variables: A Demo



Future Development

- **Train CAM-NET on WACCM-X (+DART) Simulation Data**

  Use simulation data from WACCM-X(+DART) to train the CAM-NET model.

- **Develop Mixture-of-Experts for Downstream Tasks**

  Implement Mixture-of-Experts models using techniques such as LORA (Low-Rank Adaptation), feature extraction, or other fine-tuning methods to enhance performance on specific downstream tasks.

- **Integrate JEDI+ AI Models for Data Assimilation**

  Develop and integrate JEDI with AI models to enable advanced data assimilation features in AI foundation models.

Weather and Climate Foundation Model Paper List

1. Pathak et al., (2022), FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, https://arxiv.org/abs/2202.11214.
2. Lam et al., (2022), GraphCast: Learning skillful medium-range global weather forecasting, https://arxiv.org/abs/2212.12794.
3. Bi et al., (2023), Accurate medium-range global weather forecasting with 3D neural networks, https://www.nature.com/articles/s41586-023-06185-3.
4. Chen et al., (2023), FengWu: Pushing the Skillful Global Medium-range Weather Forecast beyond 10 Days Lead, https://arxiv.org/abs/2304.02948.
5. Chen et al., (2023), FuXi: a cascade machine learning forecasting system for 15-day global weather forecast, https://www.nature.com/articles/s41612-023-00512-1.
6. Man et al.(2023), W-MAE: Pre-trained weather model with masked autoencoder for multi-variable weather forecasting, https://arxiv.org/abs/2304.08754.
7. Kochkov et al. (2023), Neural General Circulation Models for Weather and Climate, https://arxiv.org/abs/2311.07222.
8. Nguyen et al., (2023), ClimaX: A foundation model for weather and climate, https://arxiv.org/abs/2301.10343.
9. Nguyen et al., (2023), Scaling transformer neural networks for skillful and reliable medium-range weather forecasting, https://arxiv.org/abs/2312.03876.
10. Xiao et al., (2023), FengWu-4DVar: Coupling the Data-driven Weather Forecasting Model with 4D Variational Assimilation, https://arxiv.org/abs/2312.12455.
11. Han et al., (2024), FengWu-GHR: Learning the Kilometer-scale Medium-range Global Weather Forecasting, https://arxiv.org/abs/2402.00059.
12. Li et al., (2024),CaFA: Global Weather ForeCasting with Factorized Attention on Sphere, https://arxiv.org/abs/2405.07395.
13. Li et al., (2024), Generative emulation of weather forecast ensembles with diffusion models, https://www.science.org/doi/10.1126/sciadv.adk4489.
14. Price et al., (2024), GenCast: Diffusion-based ensemble forecasting for medium-range weather, https://arxiv.org/abs/2312.15796.
15. Bodnar et al., (2024), Aurora: A Foundation Model of the Atmosphere. https://arxiv.org/abs/2405.13063.