WindFormer:

Pretraining a Spatio-Temporal Transformer for Wind Forecasting

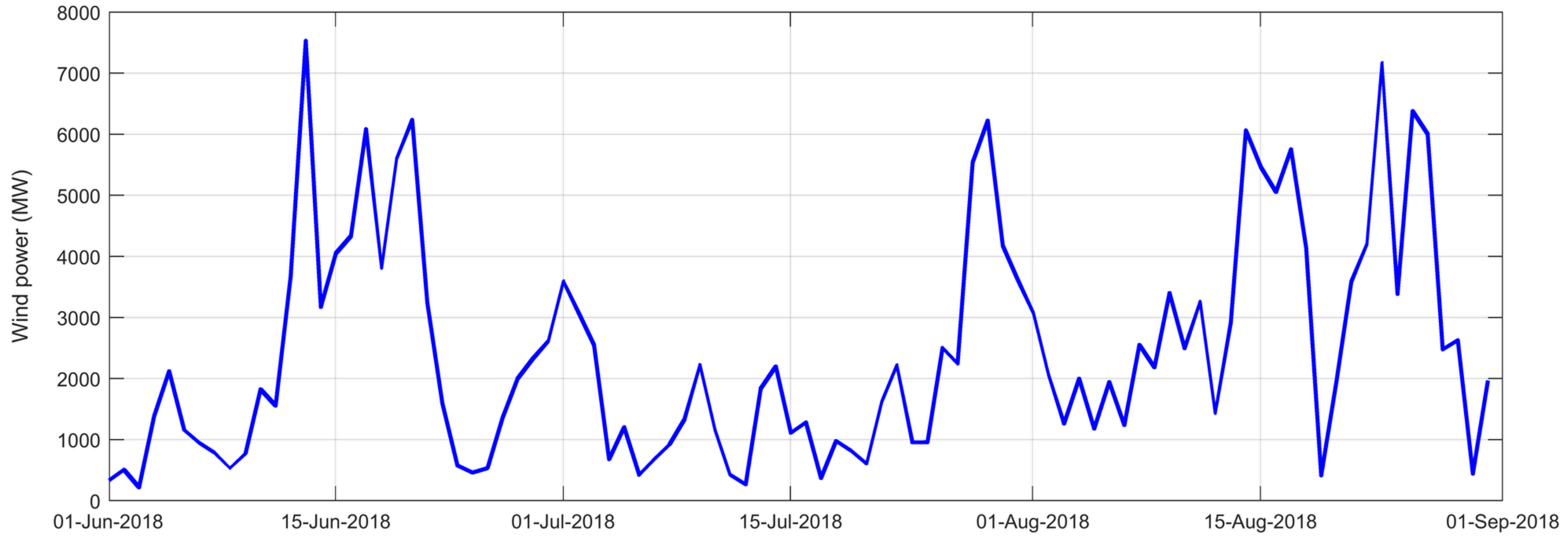
Author Samuel Reyes Sanz

Supervisors Dr. Carlos Camacho Gómez
Dr. Javier Huertas Tato

INDEX

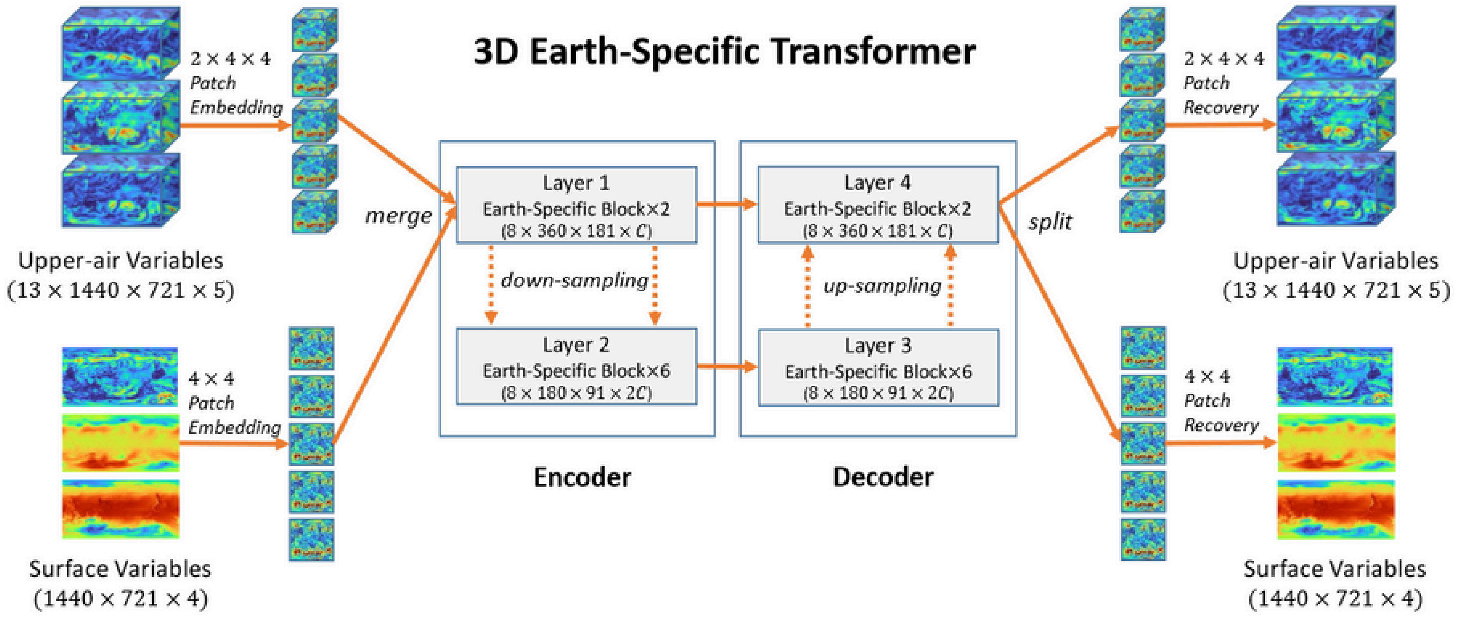
1 INTRODUCTION

MOTIVATION



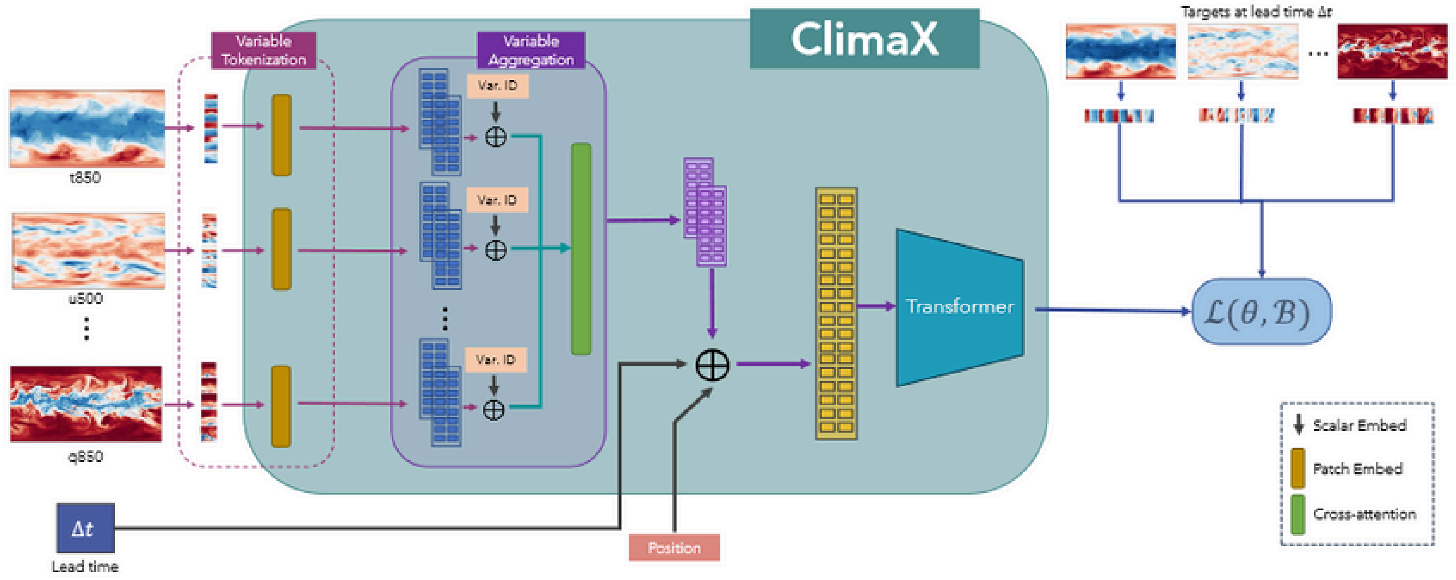
1 INTRODUCTION

STATE OF THE ART



PanguWeather

Kaifeng Bi et al. Pangu□Weather: A 3D High-Resolution Model for Fast and Accurate Global Weather Forecast, November 2022. arXiv:2211.02556 [physics].



ClimaX

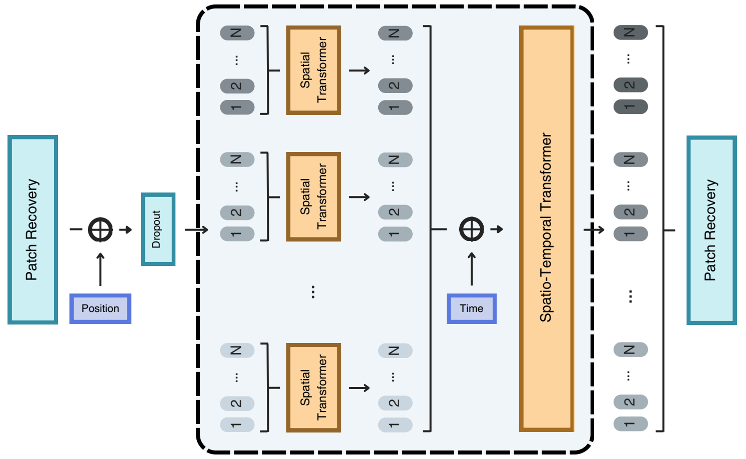
Tung Nguyen et al. ClimaX: a foundation model for weather and climate. In Proceedings of the 40th International Conference on Machine Learning (pp. 25904-25938). July, 2023.

1 INTRODUCTION

PRE-TRAINED MODELS

ERA5

Training →



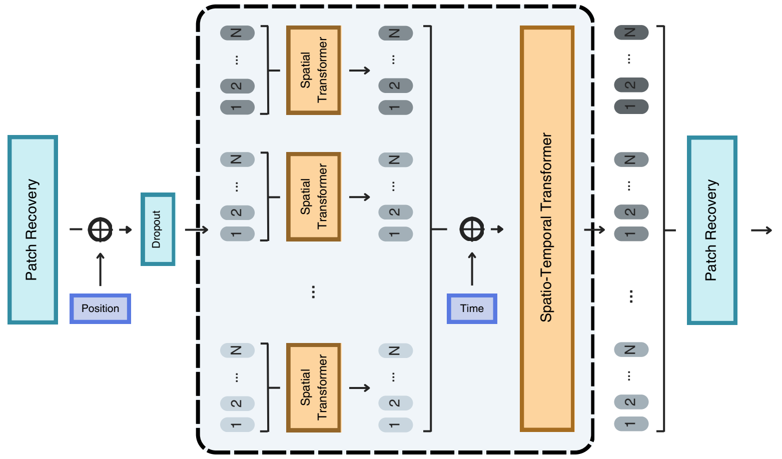
→

Model weights

↓

Observational Data

Fine-tunning →



→

Predictions

1 INTRODUCTION

OBJECTIVES

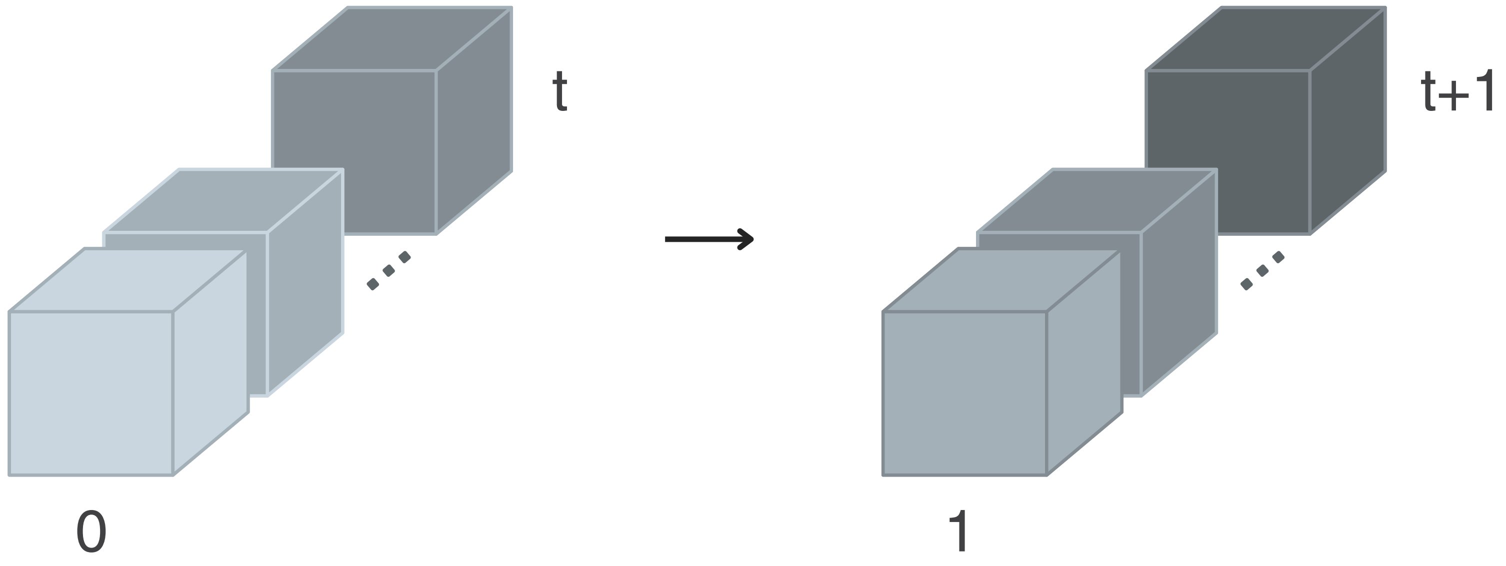
OBJ 1. Create a regional prediction model with a more efficient architecture than the state-of-the-art encoder-decoder.

OBJ 2. Implement a model that operates at high resolutions (0.25º) and various pressure levels on an hourly basis, capable of handling numerous meteorological variables and providing highly accurate short-term predictions.

OBJ 3. Create a versatile model capable of adapting to different regions and variable sets, and serving as a pre-trained model for observational data applications.

2 METHODOLOGY

TASK DEFINITION



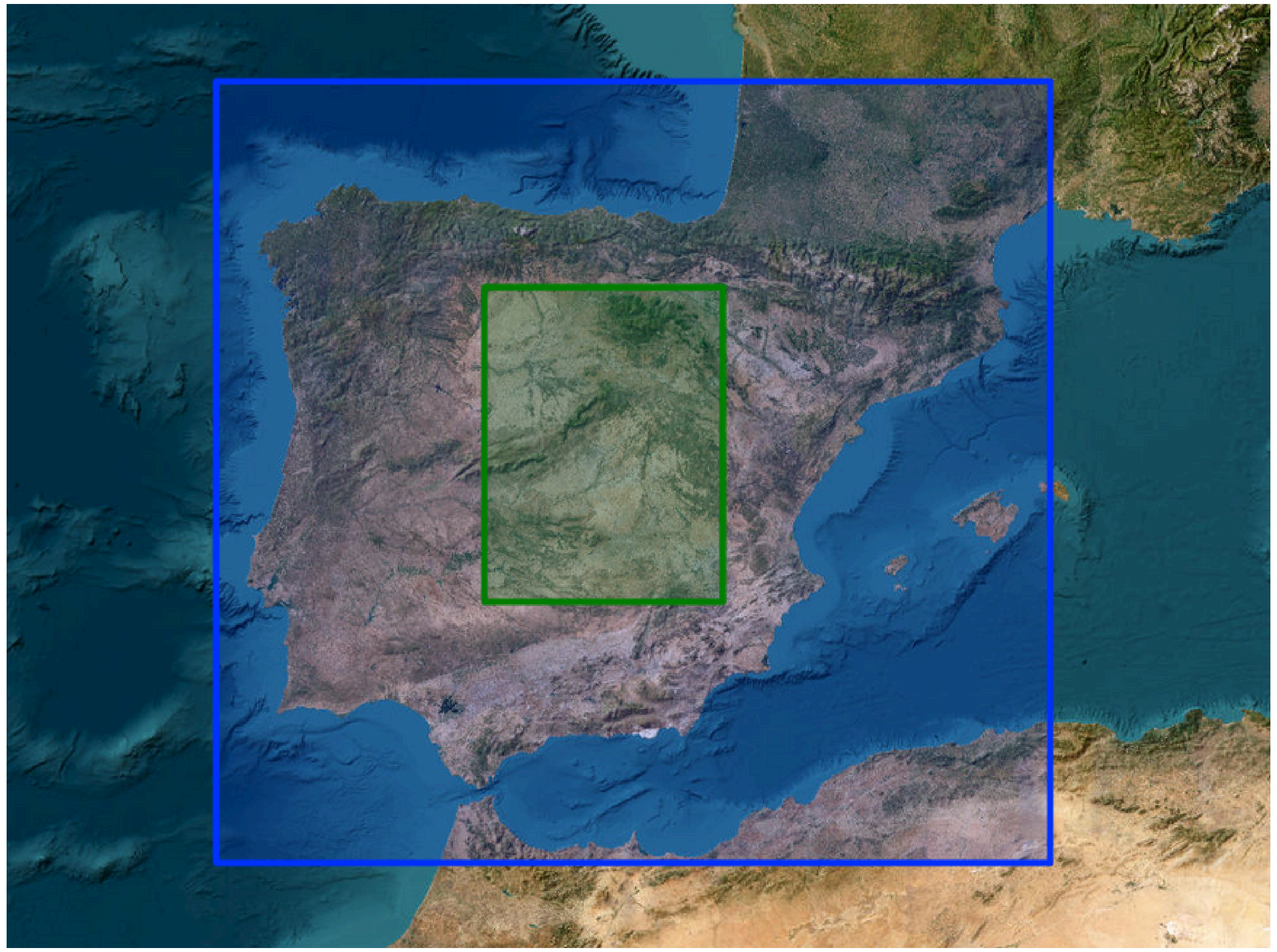
Predicts the future sequence of matrices by shifting them forward by one time step.

Allows outputs to be recursively used as inputs for extended forecast horizons.

Ensures input and output dimensions are identical.

2 METHODOLOGY

DATASET



Reanalysis data (ERA5)

Area: [45N, 10W, 35N, 4E]

Spatial resolution: 0.25º (~ 28 km)

Temporal resolution: Hourly

Pressure Levels (hPa): 950, 925, 900, 850, 800, 700, 600, 500, 250

Time Coverage 1979 to 2023

Surface Variables: u100, v100, u10, v10, d2m, t2m, z, msl, i10fg

Upper-air Variables: u, v, q, t, d, z, w, vo

2 METHODOLOGY

PREPROCESSING

1979-2023

...

Extract anomalies

1. Compute the daily mean for each level and meteorological variable over the 1981-2010 period (climatologies).

2. Subtract the climatology from each data point to obtain the anomalies.

→

Standarize data

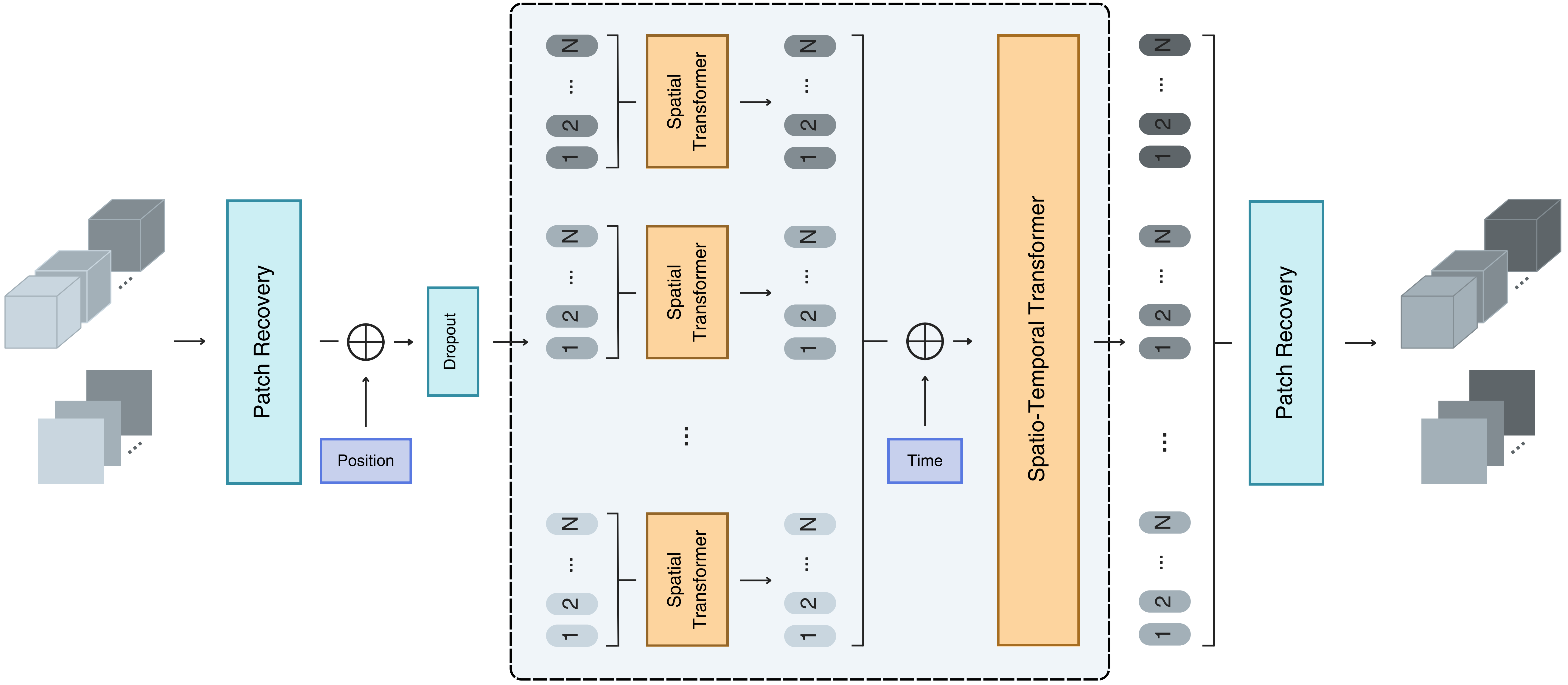
1. Compute the total mean and standard deviation over the training set for each level and meteorological variable.

2. $$Z = \frac{X - \mu}{\sigma}$$

→

2 METHODOLOGY

ARCHITECTURE



Inspired by Video Vision Transformers

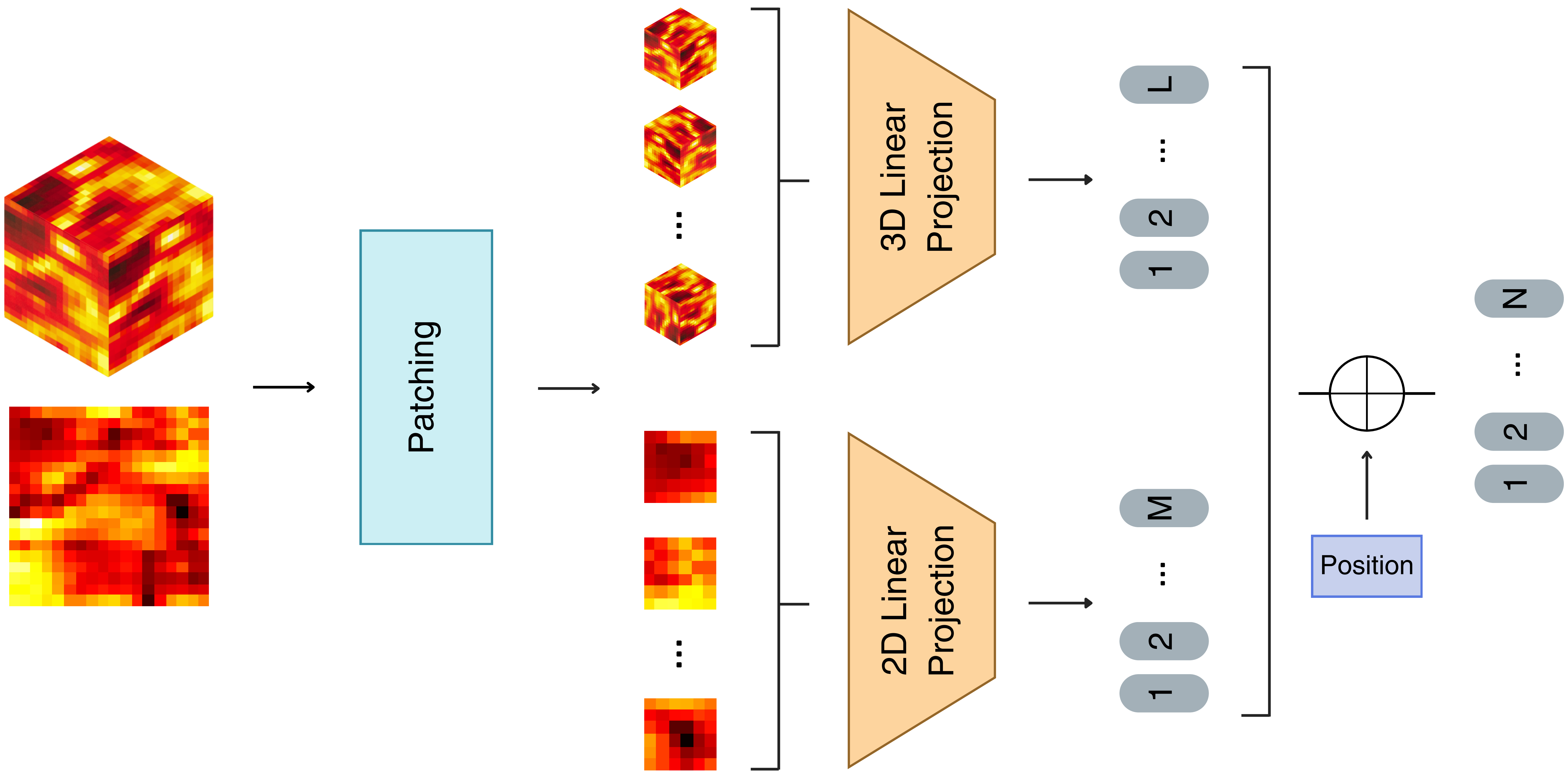
Decoder-only

Spatial and temporal attention

Same outputs as inputs

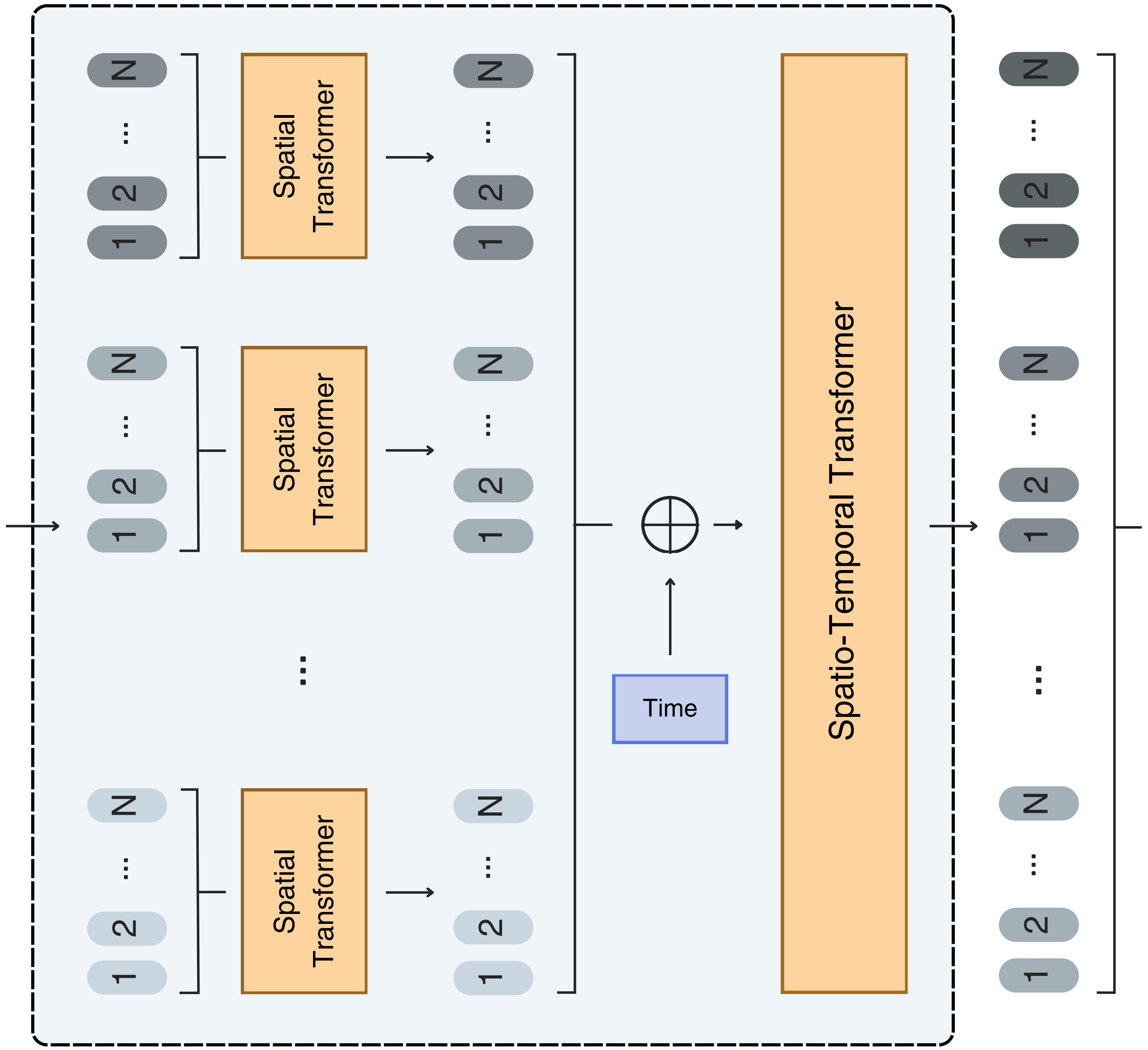
2 METHODOLOGY

PATCH EMBEDDING

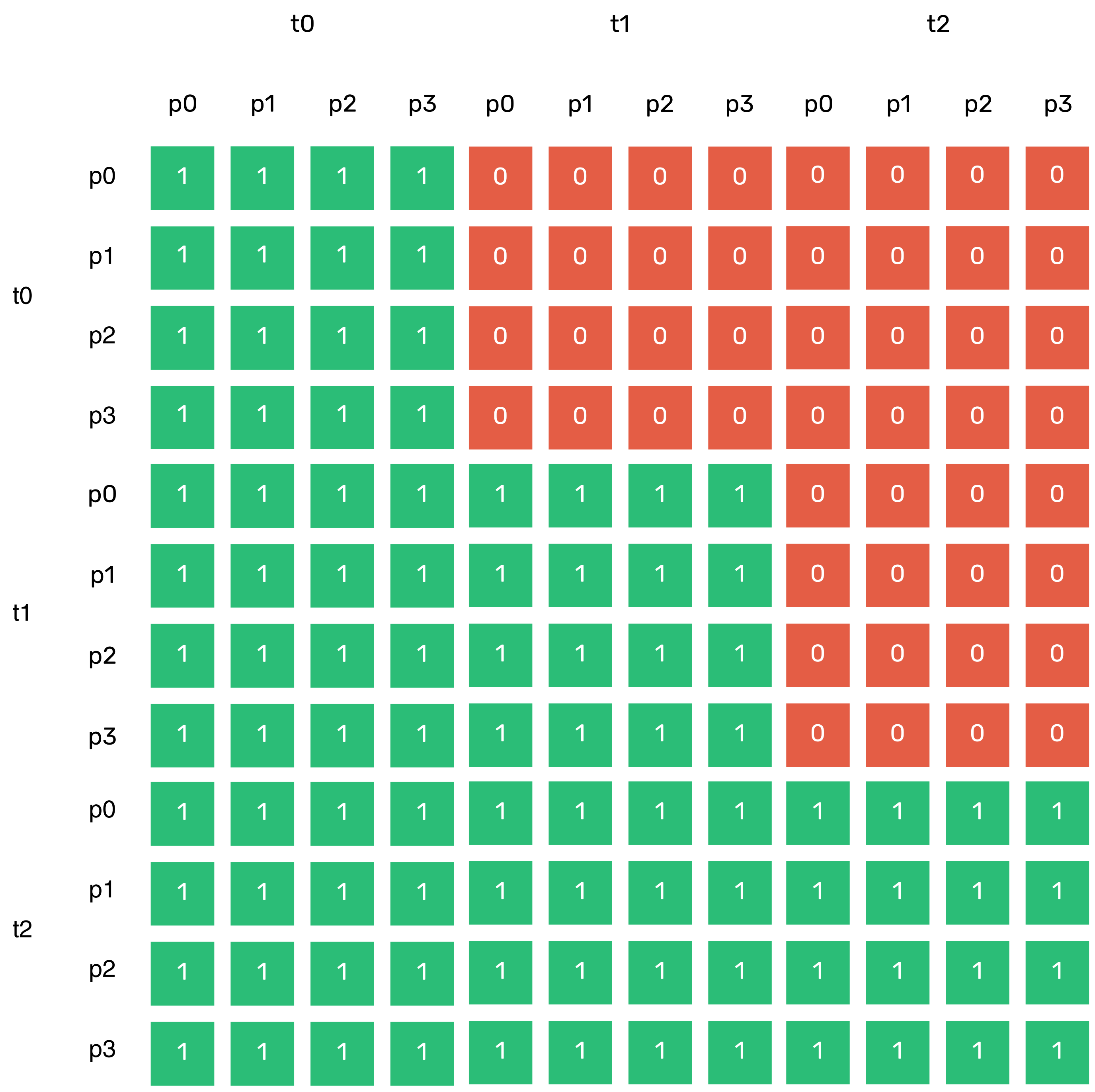


2 METHODOLOGY

TRANSFORMER MODULES



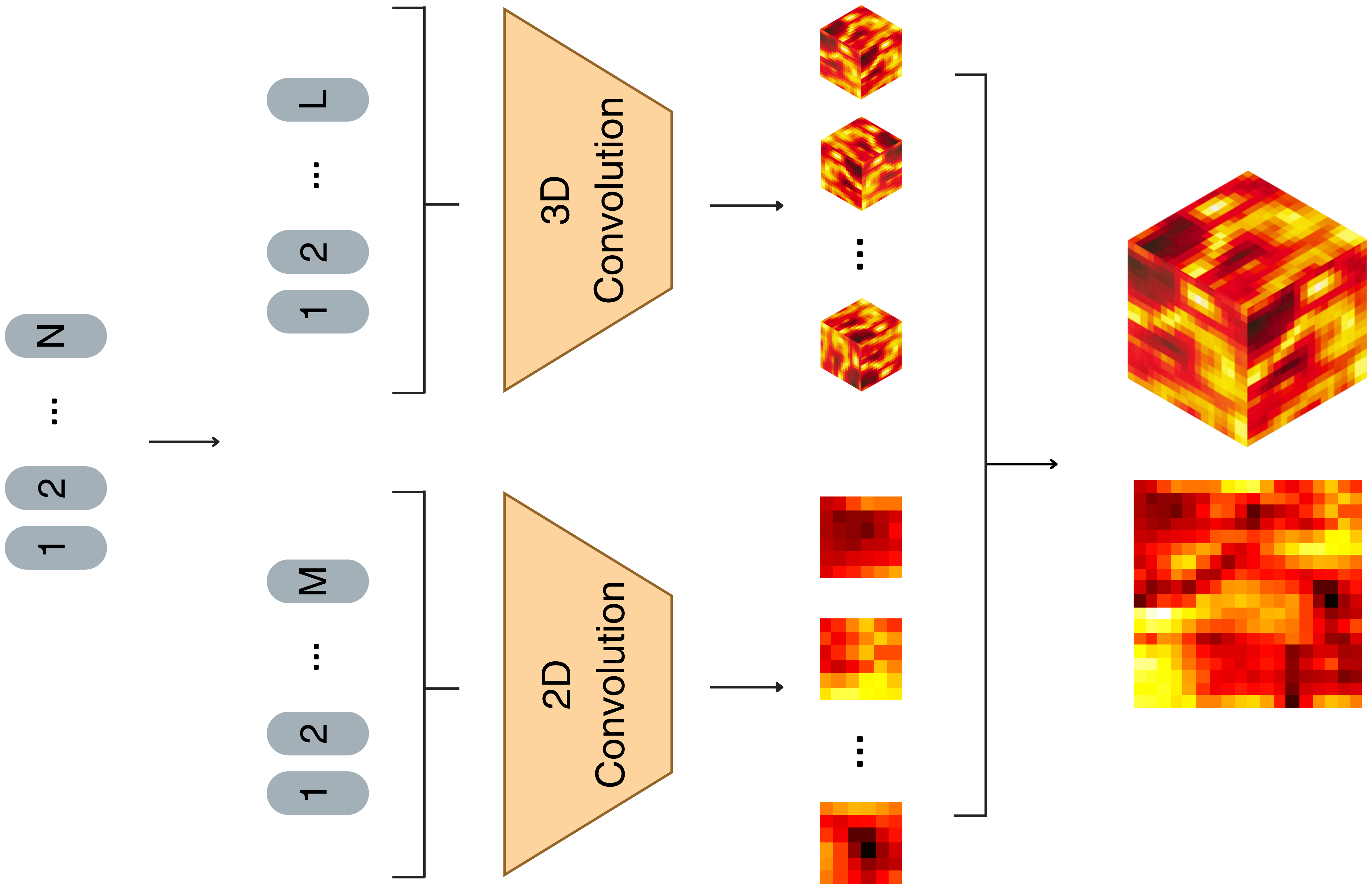
Spatial and Spatio-Temporal Transformers



Spatio-temporal attention mask

2 METHODOLOGY

PATCH RECOVERY



3 EXPERIMENTS AND RESULTS

HYPERPARAMETERS

Training Data Split	1979 to 2019
Validation Data Split	2020 to 2021
Testing Data Split	2022 to 2023

Model Image Size (3D)	9x18x18x8
Patch Size (3D)	3x6x6
Model Image Size (2D)	18x18x9
Patch Size (2D)	6x6
Sequence Length	12

Batch Size	64
Learning Rate	0.0001

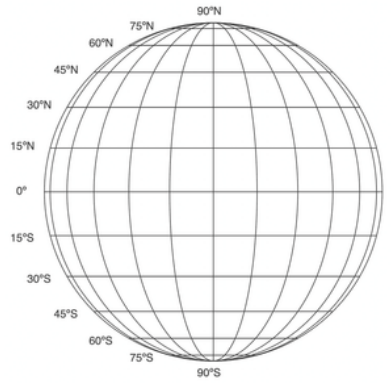
Small	Base	
512	768	Model Dimension
6	12	Depth
6	12	Heads
32	64	Dimension per Head
102	307	MLP Size
4	2	Dropout Rate
0.1	0.1	Embedding Dropout
0.1	0.1	Reconstruction
0.1	0.1	Dropout

3 EXPERIMENTS AND RESULTS

EVALUATION METRICS

$$RMSE = \sqrt{\frac{\sum_{i,j}^{N_{\text{lat,lon}}} L(i)\left(\hat{A}_{i,j} - A_{i,j}\right)^2}{N_{\text{lat}} \times N_{\text{lon}}}},$$

$$L(i) = N_{\text{lat}} \times \frac{\cos\phi_i}{\sum_{i=1}^{N_{\text{lat}}} \cos\phi_i}$$



$$ACC = \frac{\sum_{i,j}^{N_{\text{lat,lon}}} L(i)\hat{A}_{i,j}A_{i,j}}{\sqrt{\sum_{i,j}^{N_{\text{lat,lon}}} L(i)\left(\hat{A}_{i,j}\right)^2 \times \sum_{i,j}^{N_{\text{lat,lon}}} L(i)\left(A_{i,j}\right)^2}},$$

Kaifeng Bi et al. Pangu□Weather: A 3D High-Resolution Model for Fast and Accurate Global Weather Forecast, November 2022. arXiv:2211.02556 [physics].

Root Mean Square Error (RMSE)

Measures the average magnitude of the prediction errors.

Lower is better.

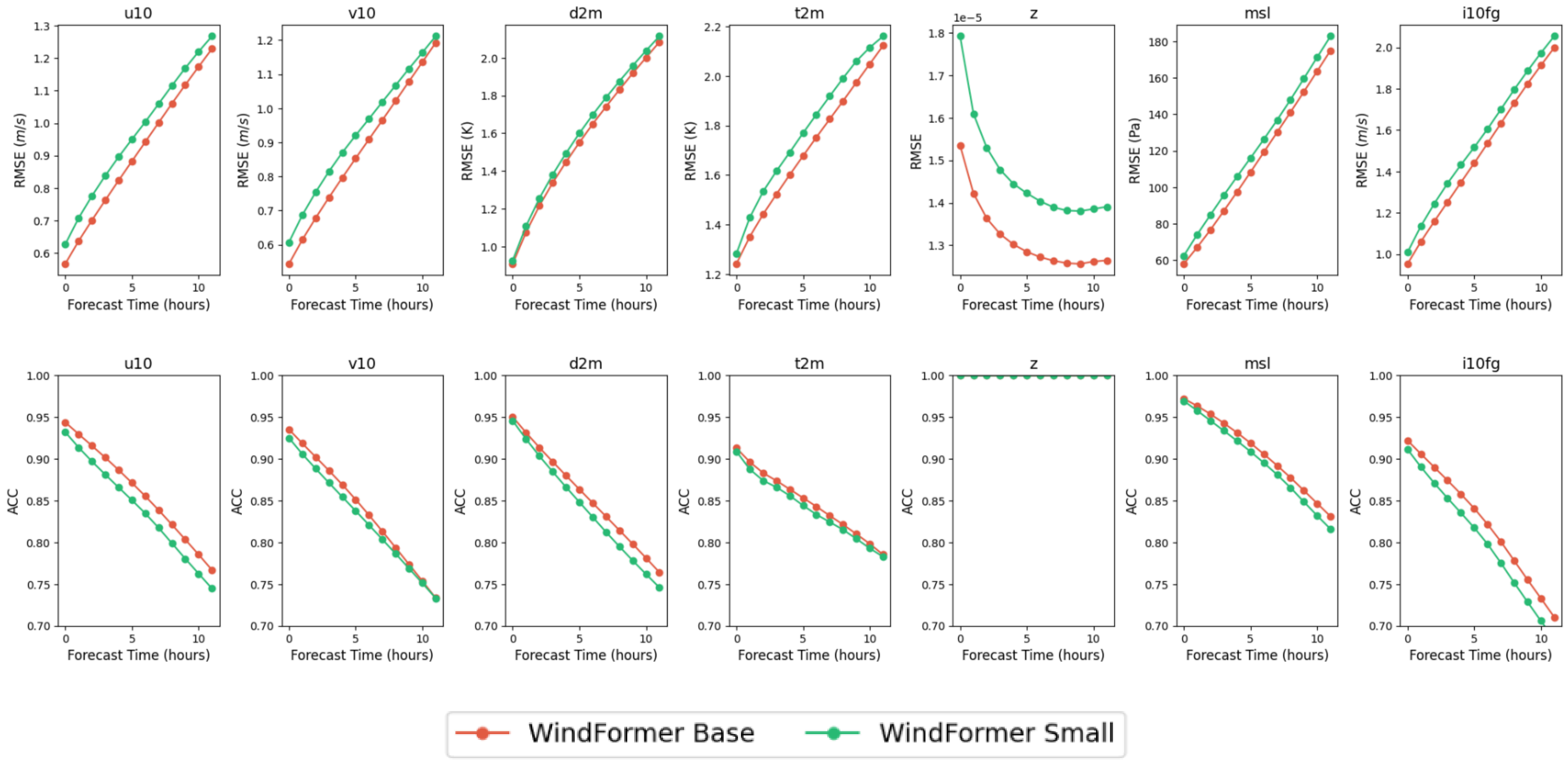
Anomaly Correlation Coefficient (ACC)

Represents the spatial correlation between forecast anomalies and real anomalies.

Higher is better.

3 EXPERIMENTS AND RESULTS

SMALL VS BASE



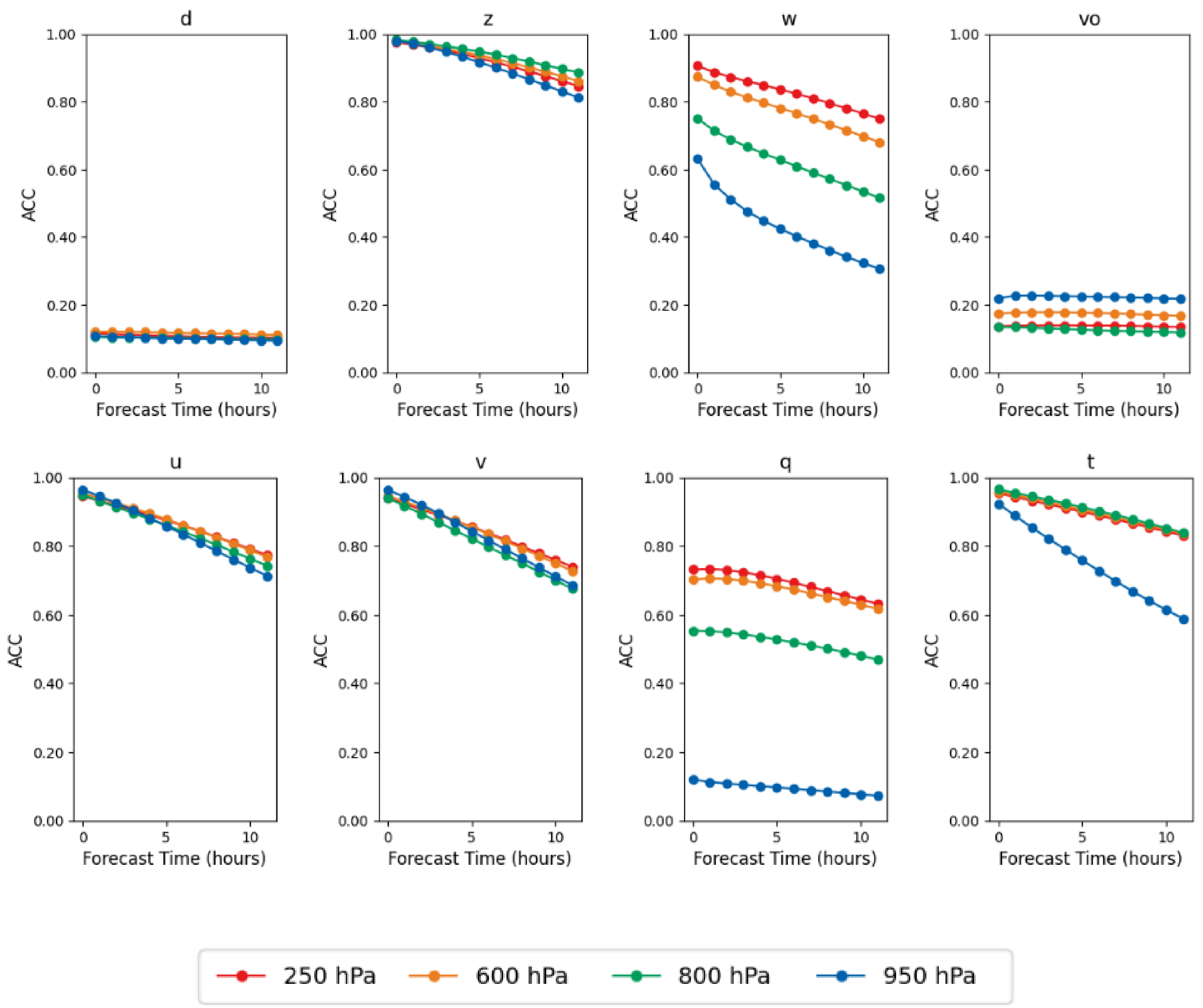
The base model outperforms the small model.

As lead time increase, RMSE rises and ACC decreases.

The variable z (geopotential) shows better performance.

3 EXPERIMENTS AND RESULTS

UPPER LEVELS



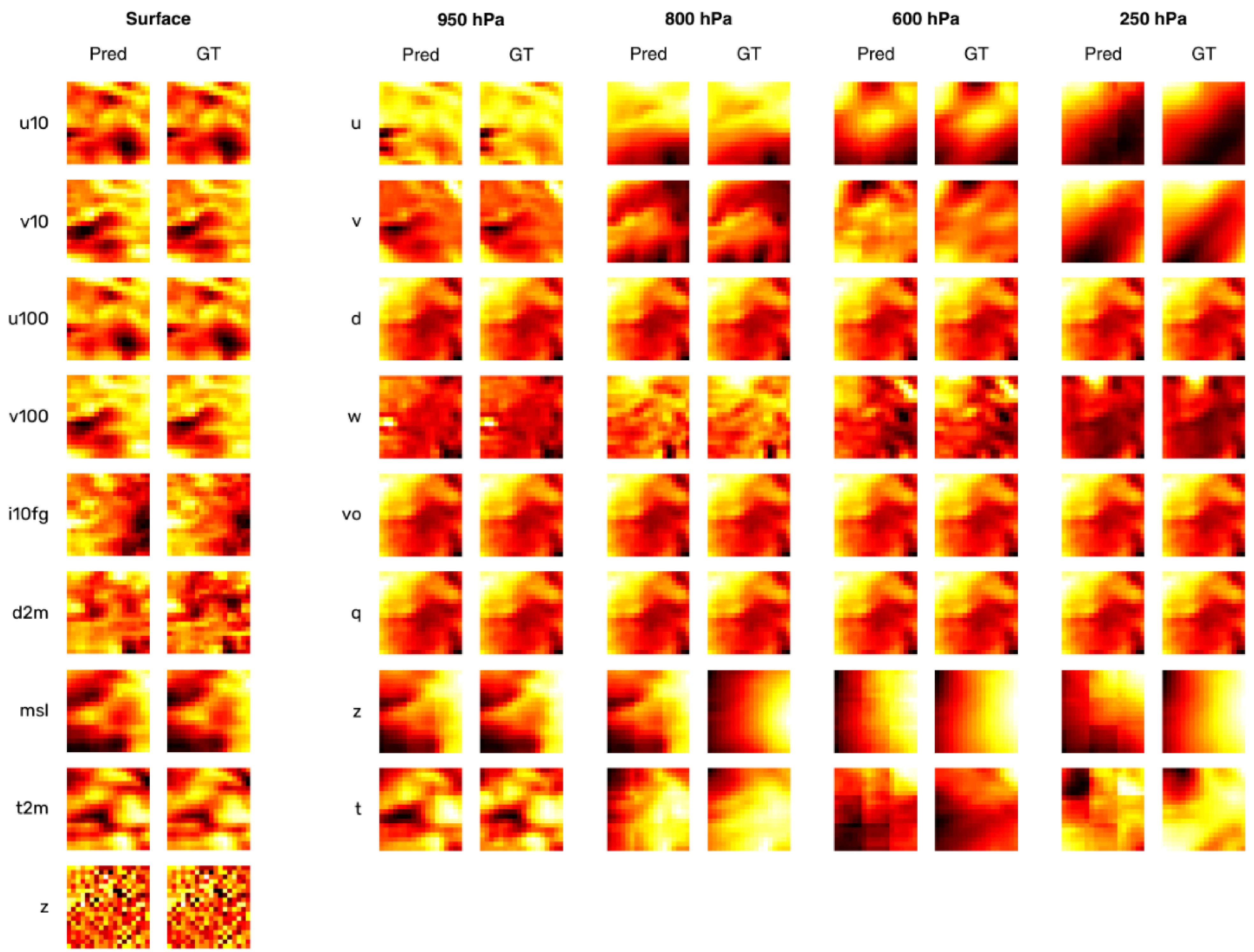
Depending on the variable, better or worse results are obtained when comparing different pressure levels.

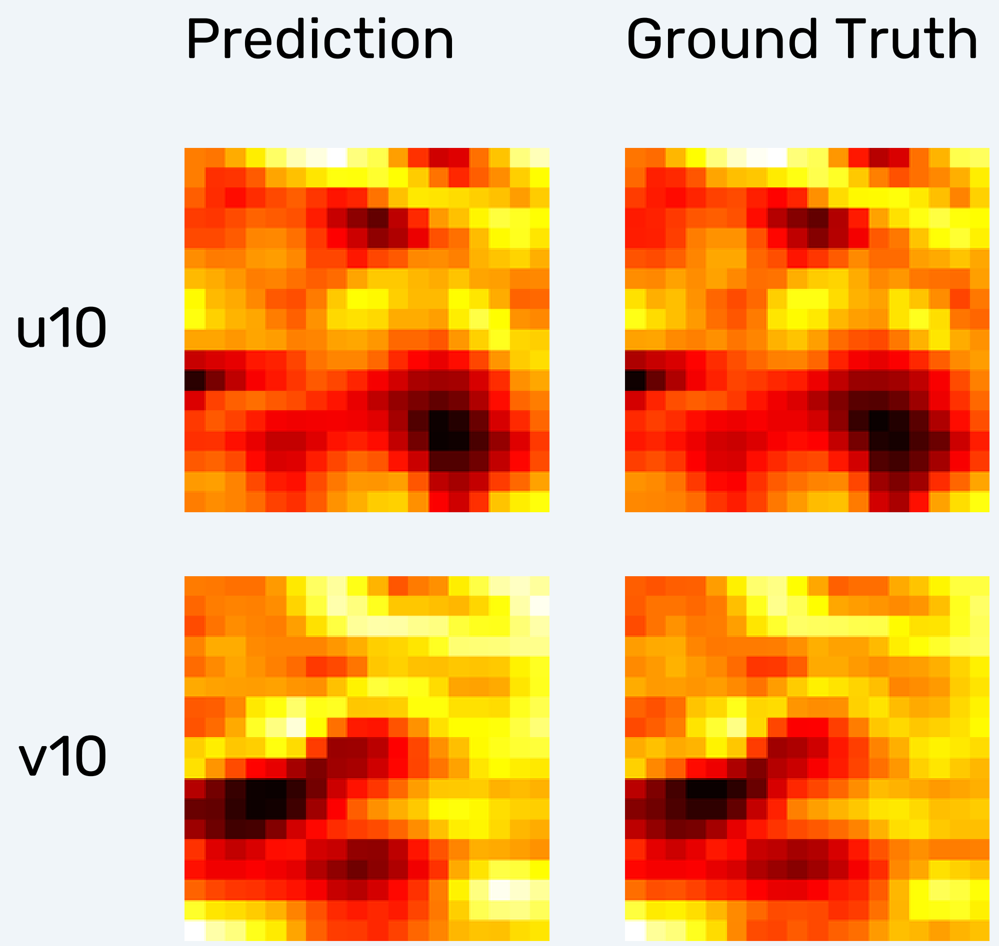
Divergence (d) and vorticity (vo) exhibit poorer performance.

RMSE is not a good metric for comparing a variable across different levels.

3 EXPERIMENTS AND RESULTS

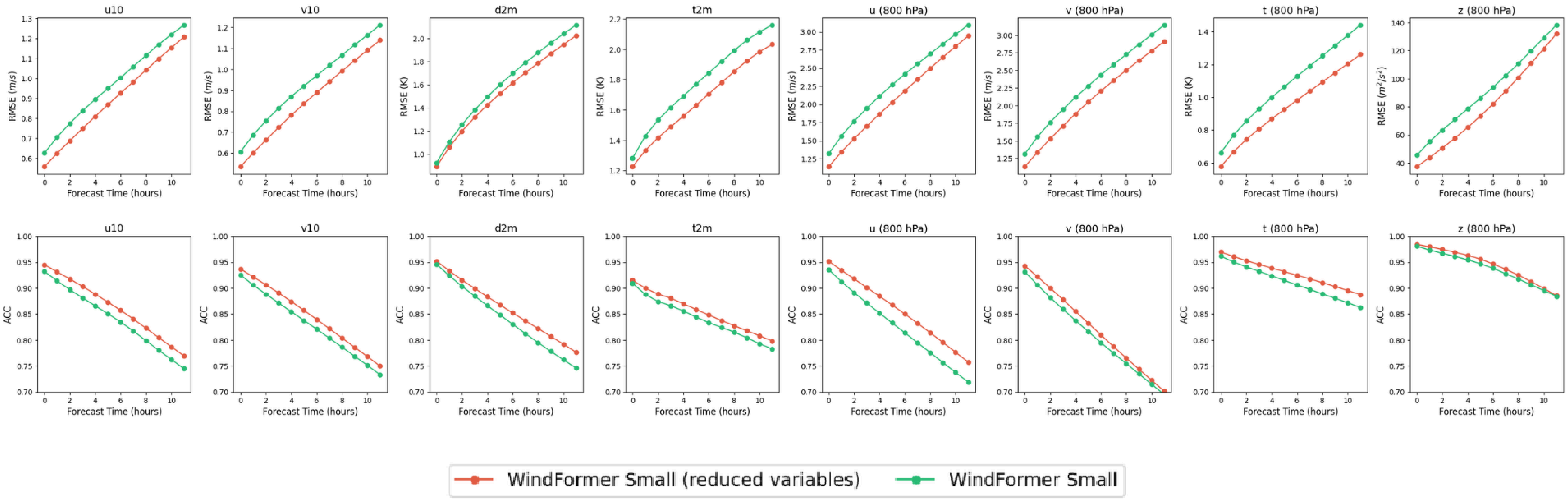
VISUALIZATION





3 EXPERIMENTS AND RESULTS

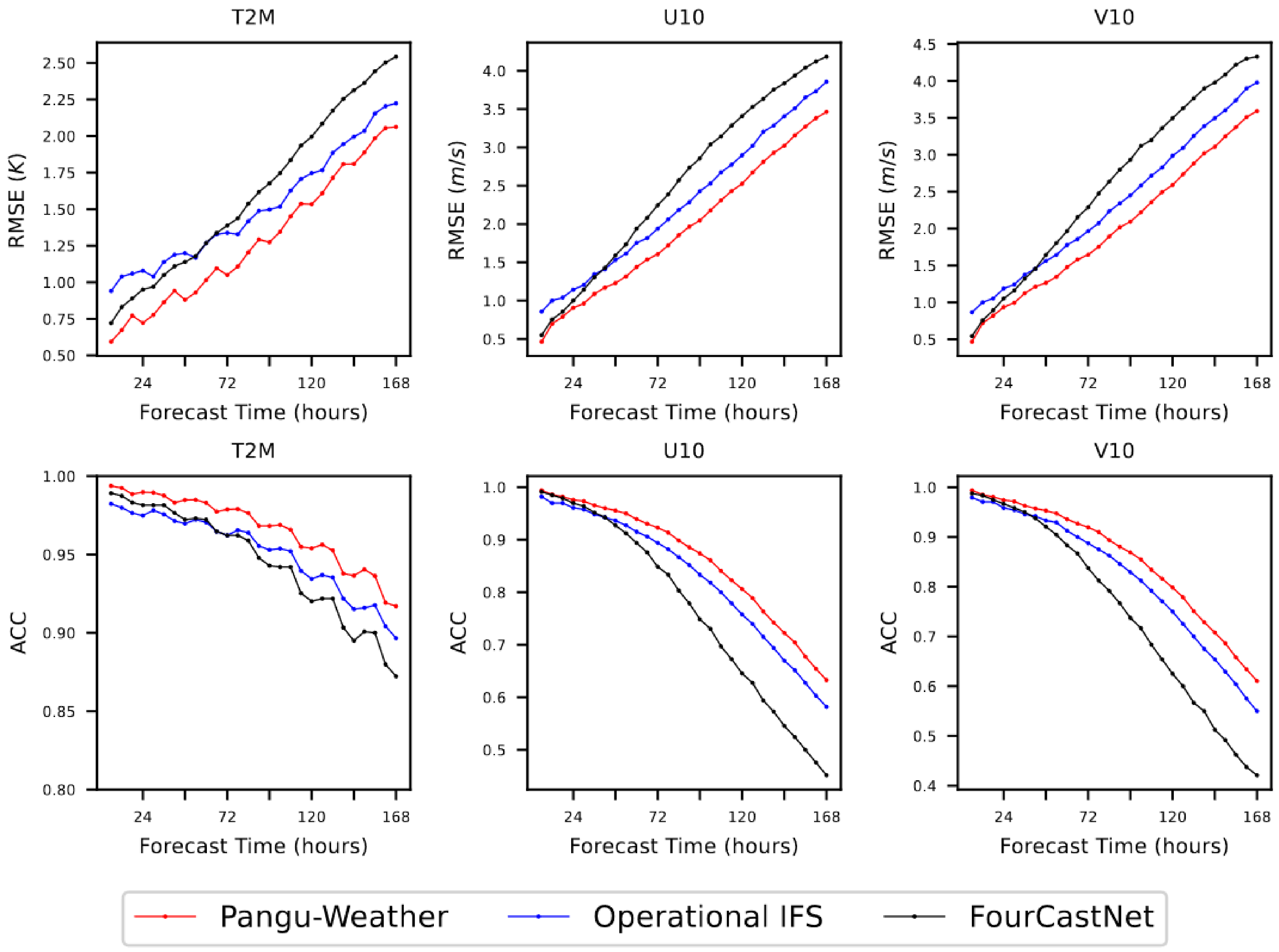
REDUCING VARIABLES SET

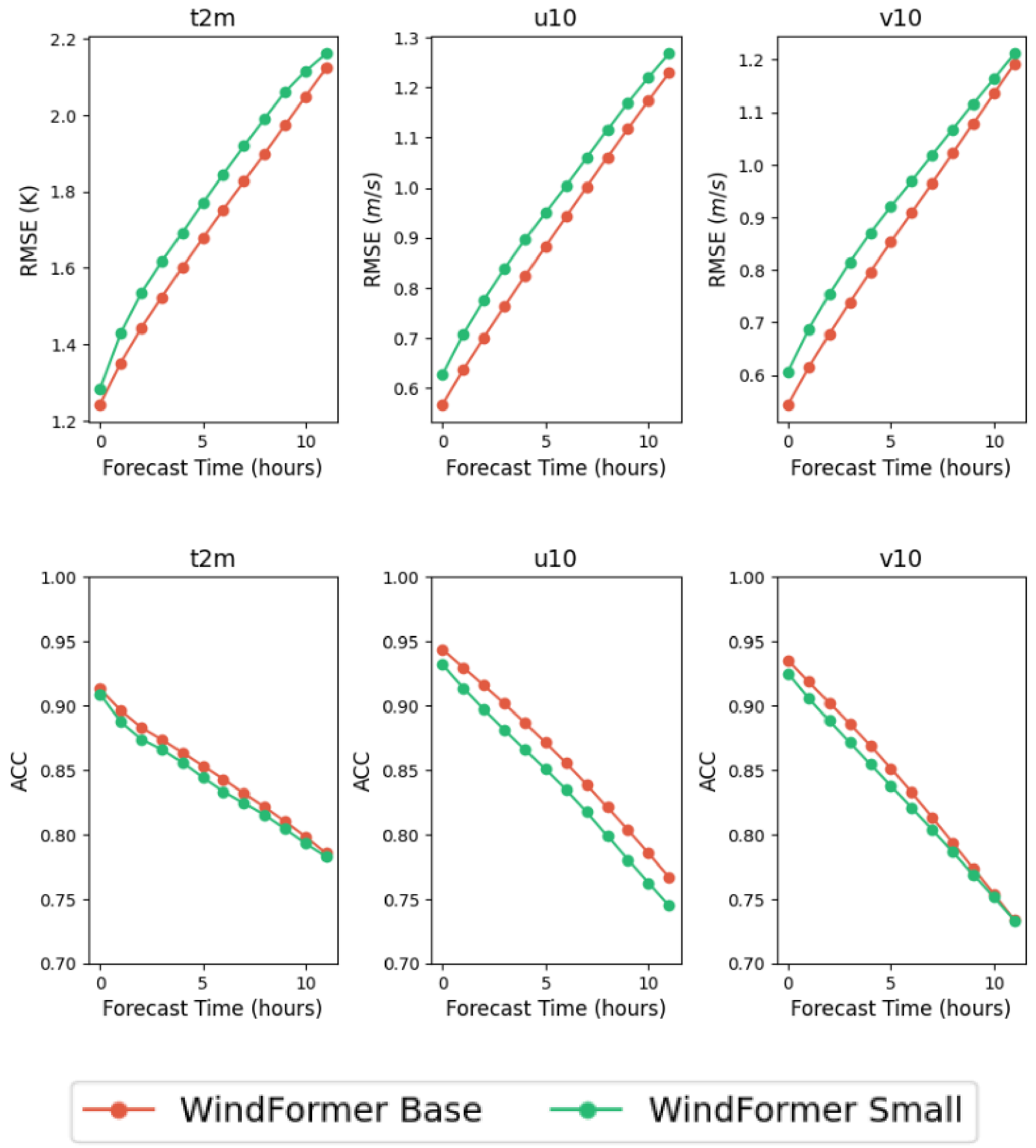


The z (surface), q and vo variables are removed.

3 EXPERIMENTS AND RESULTS

STATE-OF-THE-ART COMPARISON





Kaifeng Bi et al. Pangu-Weather: A 3D High-Resolution Model for Fast and Accurate Global Weather Forecast, November 2022. arXiv:2211.02556 [physics].

4 CONCLUSIONS AND FUTURE WORK

FUTURE WORK

- More **experiments** are required.
- Better **comparisons** against the state-of-the-art are needed.
- Test with **observational data**.
- Reduce the **accumulation of errors**.

4 CONCLUSIONS AND FUTURE WORK

CONCLUSIONS

Development of WindFormer, a transformer-based model with a **decoder-only** architecture, **more efficient** than current encoder-decoder models.

Implementation of a **high-resolution model** (0.25°) at various pressure levels, capable of **precise short-term predictions** for many meteorological variables.

Demonstrating the model's **versatility** to adapt to different regions and variable sets.

Highlighting the model's potential as a **pre-trained model** on reanalysis data, fine-tunable with observational data.

Validating the model's **performance** through experiments, achieving similar RMSE and ACC values compared to state-of-the-art models with less efficient architectures.

Thanks!

@samuel-reyes-sanz

@SamuReyes

samuel.reyes.sanz@alumnos.upm.es

POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

AI+DA