Evaluation of GraphCast and PanguWeather over India and Downscaling using WRF for Simulating the 2023 North Indian Floods

ECMWF

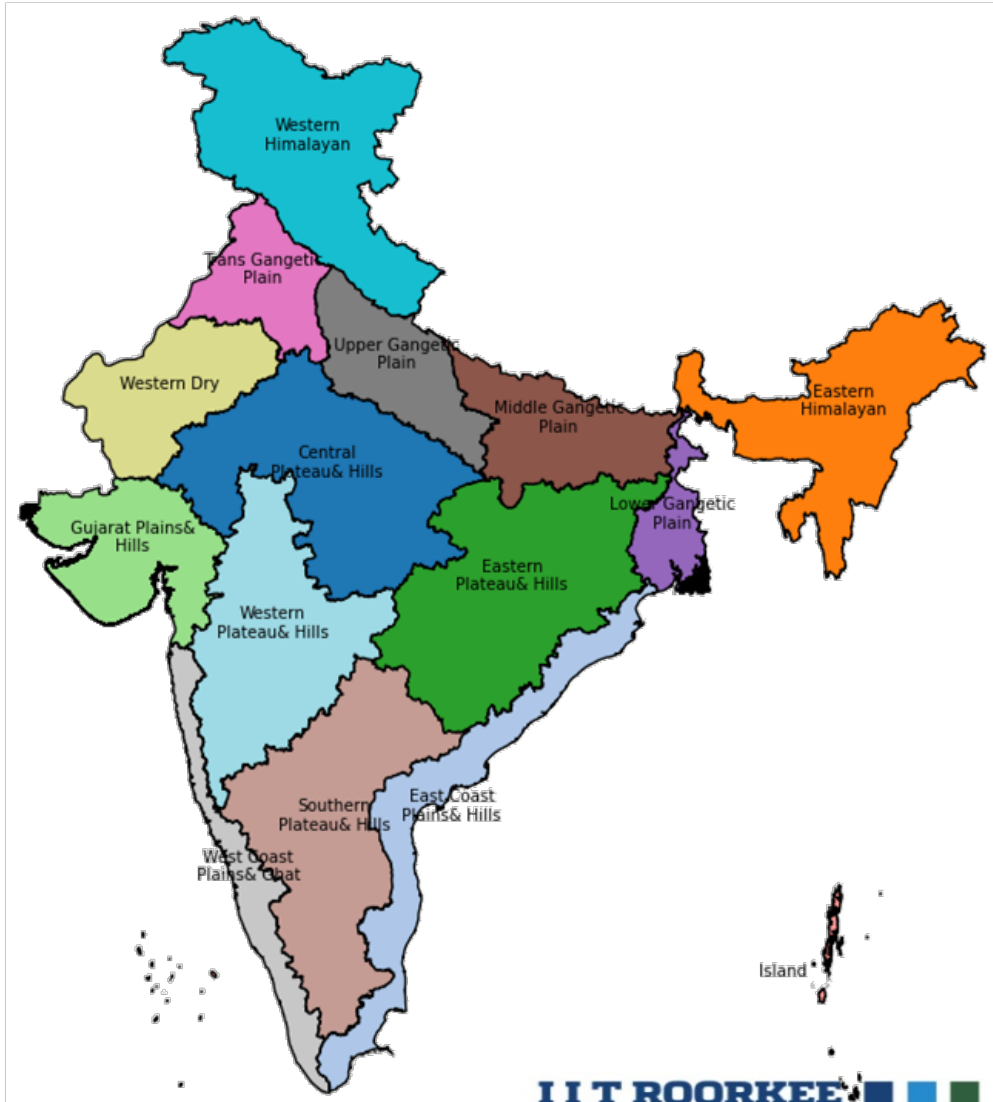
14th April 2026

Presented by:
Balu Shaharica,
Research Scholar, CoEDMM,
Indian Institute of Technology, Roorkee

Introduction: Indian Weather

- India features immense **weather diversity**, ranging from Alpine cold in the Himalayas to arid desert conditions in Rajasthan and tropical rainforests in the south.
- The climate is dominated by a tropical monsoon, bringing heavy summer rain, with regional temperatures ranging from **-45 °C in winter (Leh) to over 55 °C in summer (Rajasthan).**
- India is divided into **15 Agroclimatic Zones**
- **The Indian Meteorological Department (IMD),** the governing body responsible for **weather forecasting**, climate services, and issuing warnings for cyclones, floods, and heatwaves

15 Agroclimatic zones of India



Study Highlights

1. Evaluation of GraphCast and Pangu-Weather for the different Indian Agroclimatic zones

2. Seasonal analysis for GraphCast.
   (JF – Winter ; MAM – Summer ; JJAS - Monsoon ; OND – Post Monsoon)

3. Downscaling of GraphCast output using WRF
   a. Set up WRF to input GraphCast data
   b. Downscale to 5 and 1 km scale to study the 2023 North Indian Floods.

Evaluation of Models: **GraphCast**

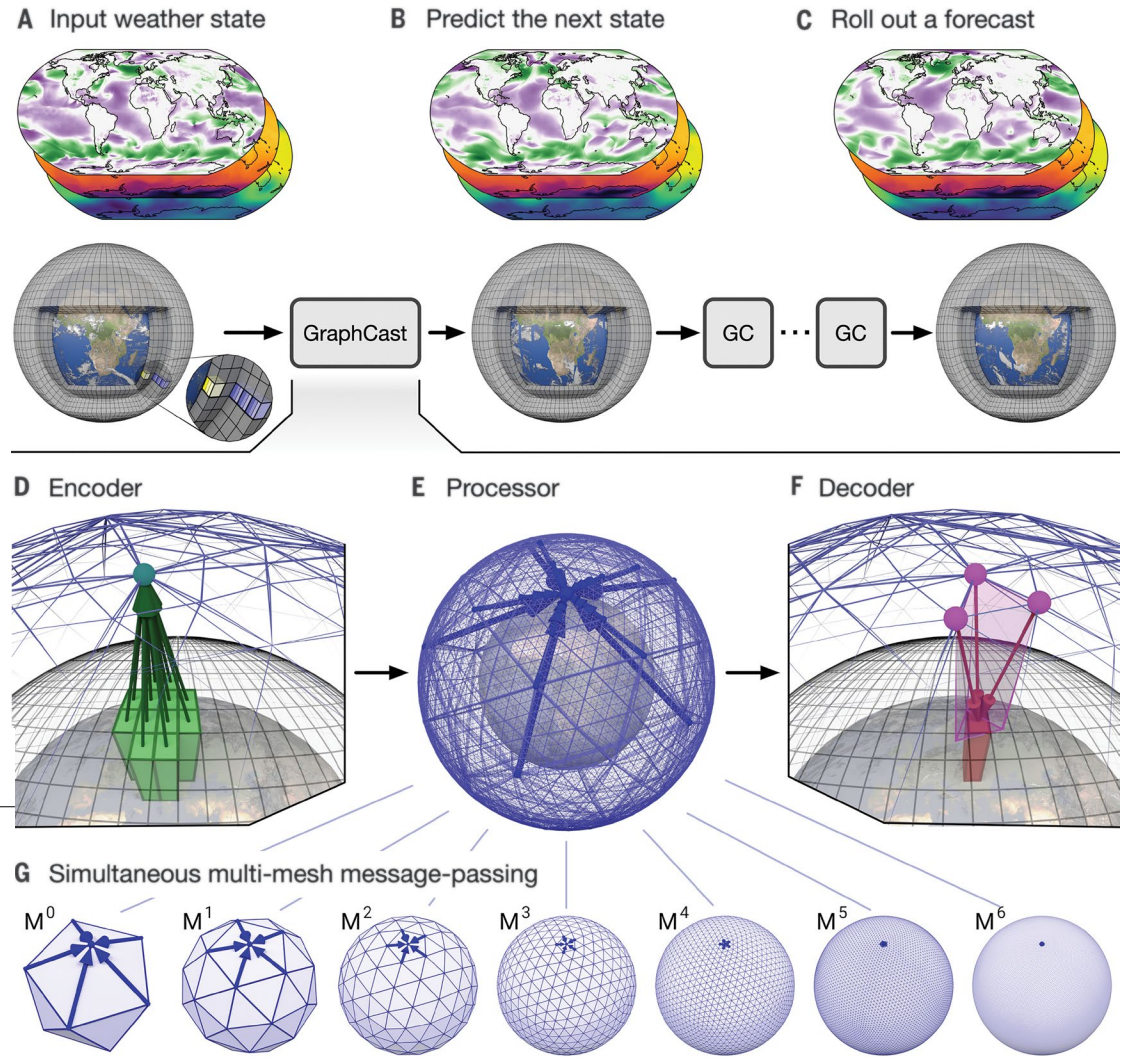
Architecture : GNN with Icosahedral Mesh
Training Dataset : 1979 – 2017 ERA5
Models :

1. ***GraphCast***, the high-resolution model used in the GraphCast paper (0.25 degree resolution, **37 pressure levels**), trained on ERA5 data from 1979 to 2017
2. ***GraphCast_operational***, a high-resolution model (0.25 degree resolution, **13 pressure levels**) pre-trained on ERA5 data from 1979 to 2017 and fine-tuned on HRES data from 2016 to 2021.

CONFIGURATION:

Processor : Param Ganga
Input : Current and 6 hours before
Output : 6 hourly forecast
Post processing : Clipped for India and analysed for various agro-climatic zones

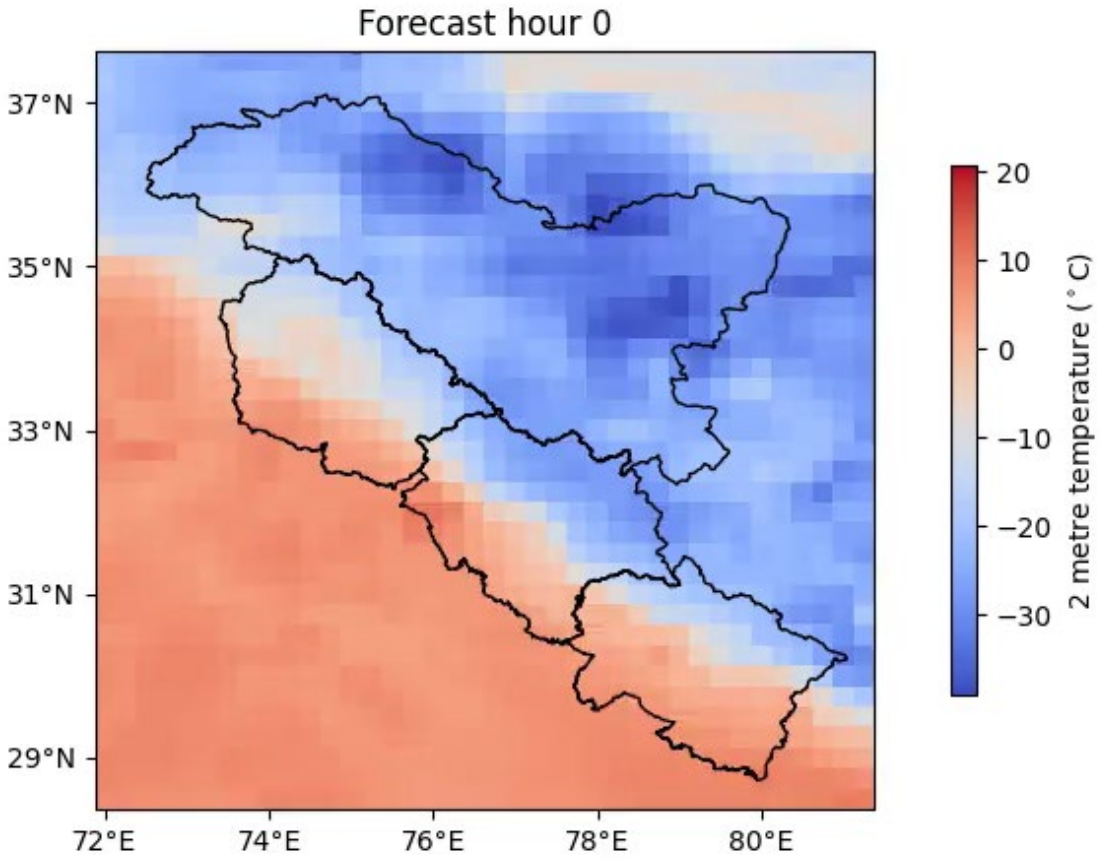
Surface variables (5)	Atmospheric variables (6)	Pressure levels (37)
2-m temperature (2*T*)	**Temperature** (*T*)	1, 2, 3, 5, 7, 10, 20, 30, **50**, 70,
10-m u wind component (10*U*)	**U component of wind** (*U*)	**100**, 125, **150**, 175, **200**, 225,
10-m v wind component (10*V*)	**V component of wind** (*V*)	**250**, **300**, 350, **400**, 450, **500**,
Mean sea level pressure (*MSL*)	**Geopotential** (*Z*)	550, **600**, 650, **700**, 750, 775,
Total precipitation (*TP*)	**Specific humidity** (*Q*)	800, 825, **850**, 875, 900, **925**,
	Vertical wind speed (*W*)	950, 975, and **1000** hPa



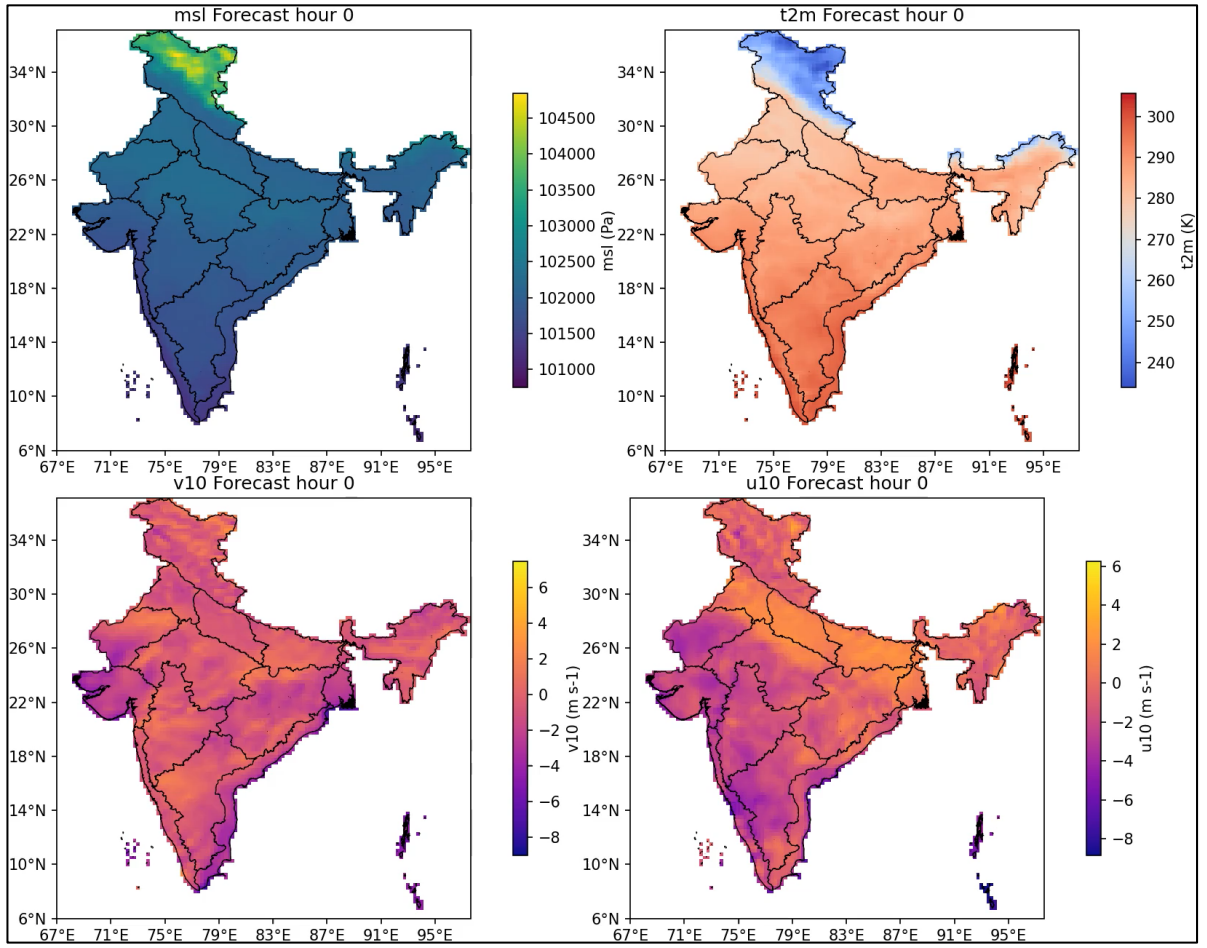
Lam, Remi, et al. "Learning skilful medium-range global weather forecasting." *Science* 382.6677 (2023): 1416-1421.

Evaluation of Models: **GraphCast Preliminary Run**

16 day hindcast of Jan 2022 using GraphCast, for surface variables, T2M, U10, V10, MSLP.

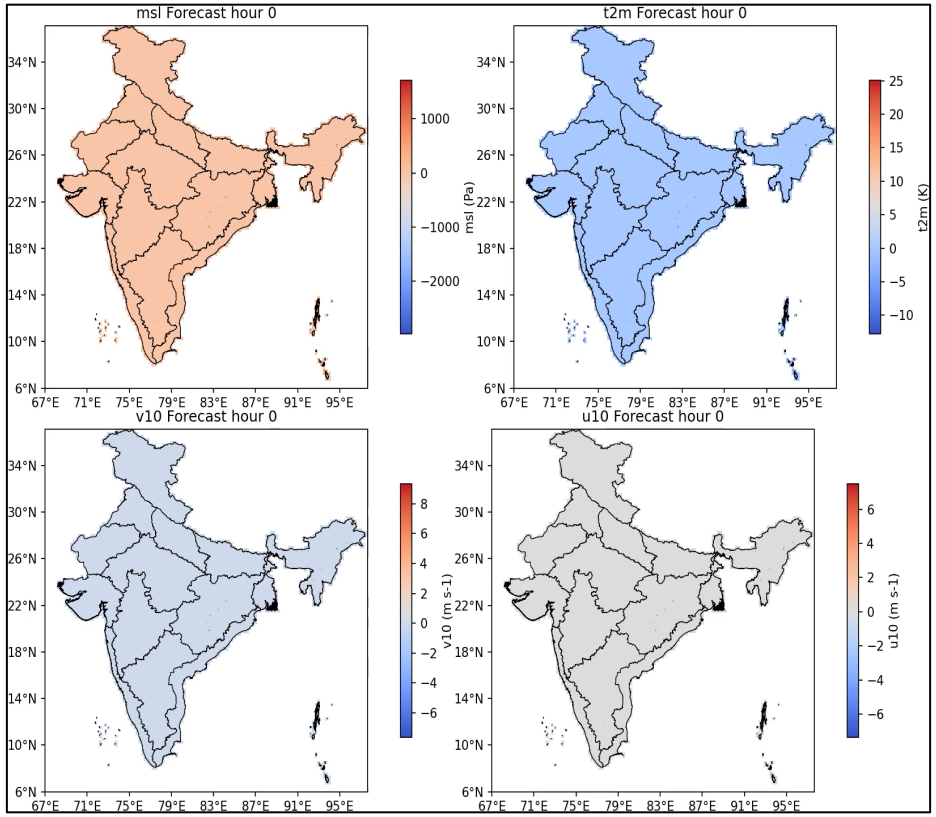


2-m temperature over Western Himalayas

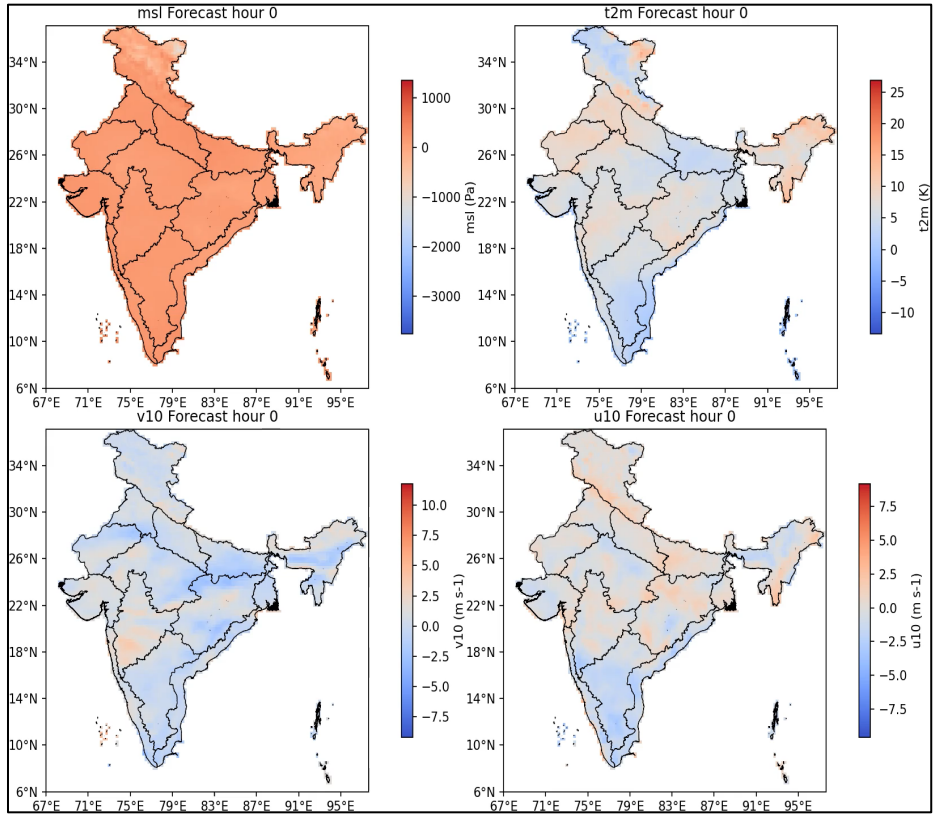


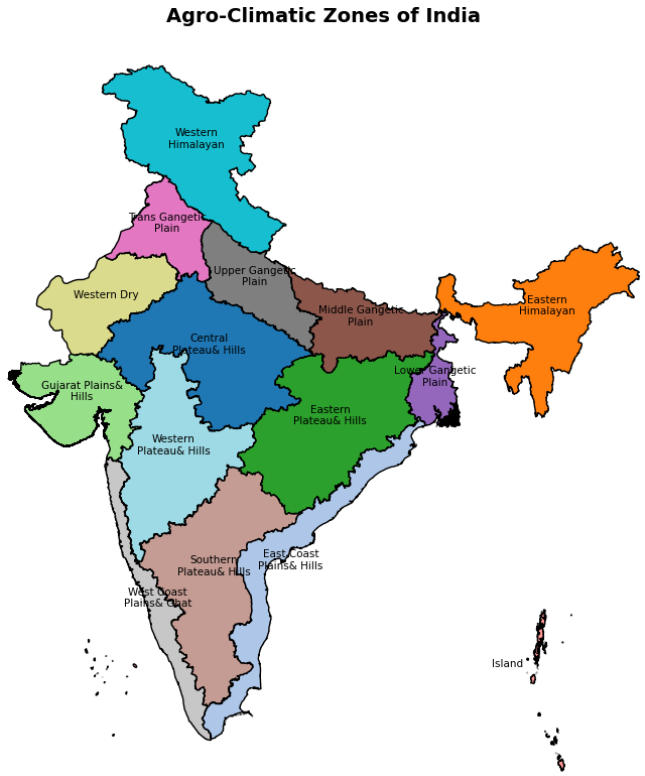
Evaluation of Model : **Bias Plots**

GraphCast



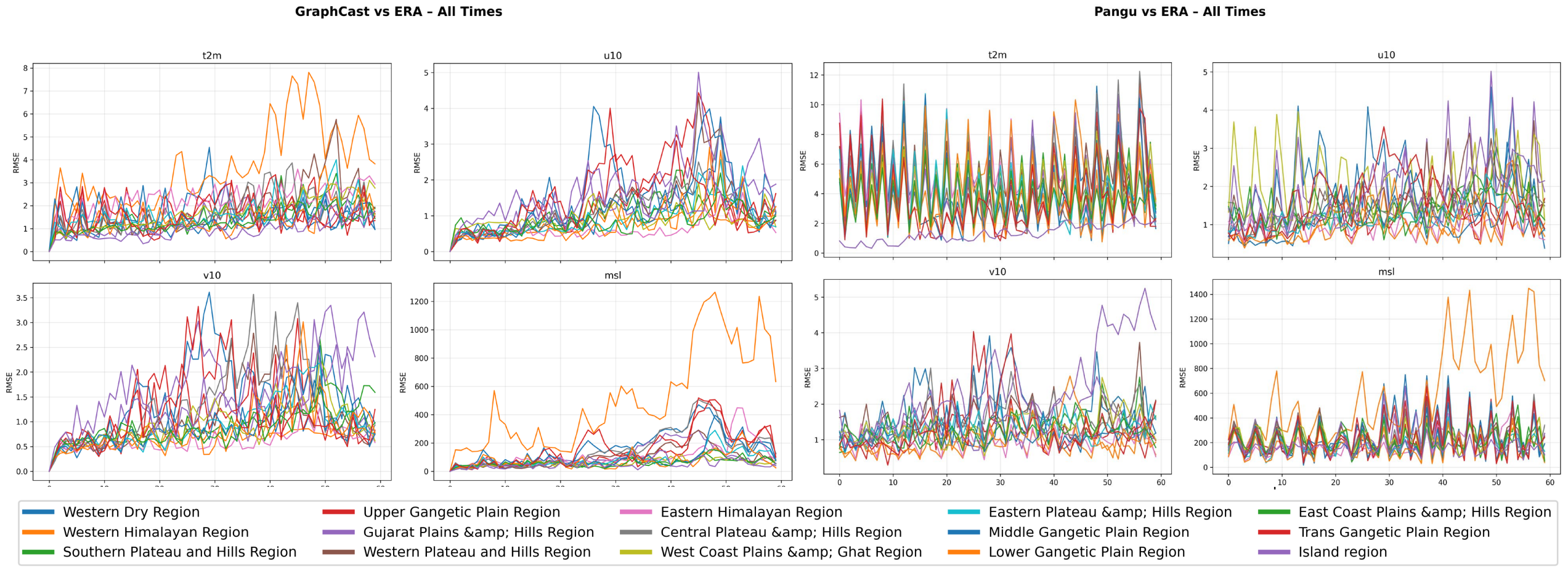
Pangu-Weather





Difference plots of GraphCast and Pangu-Weather with ERA5 data, for different agroclimatic zones of India

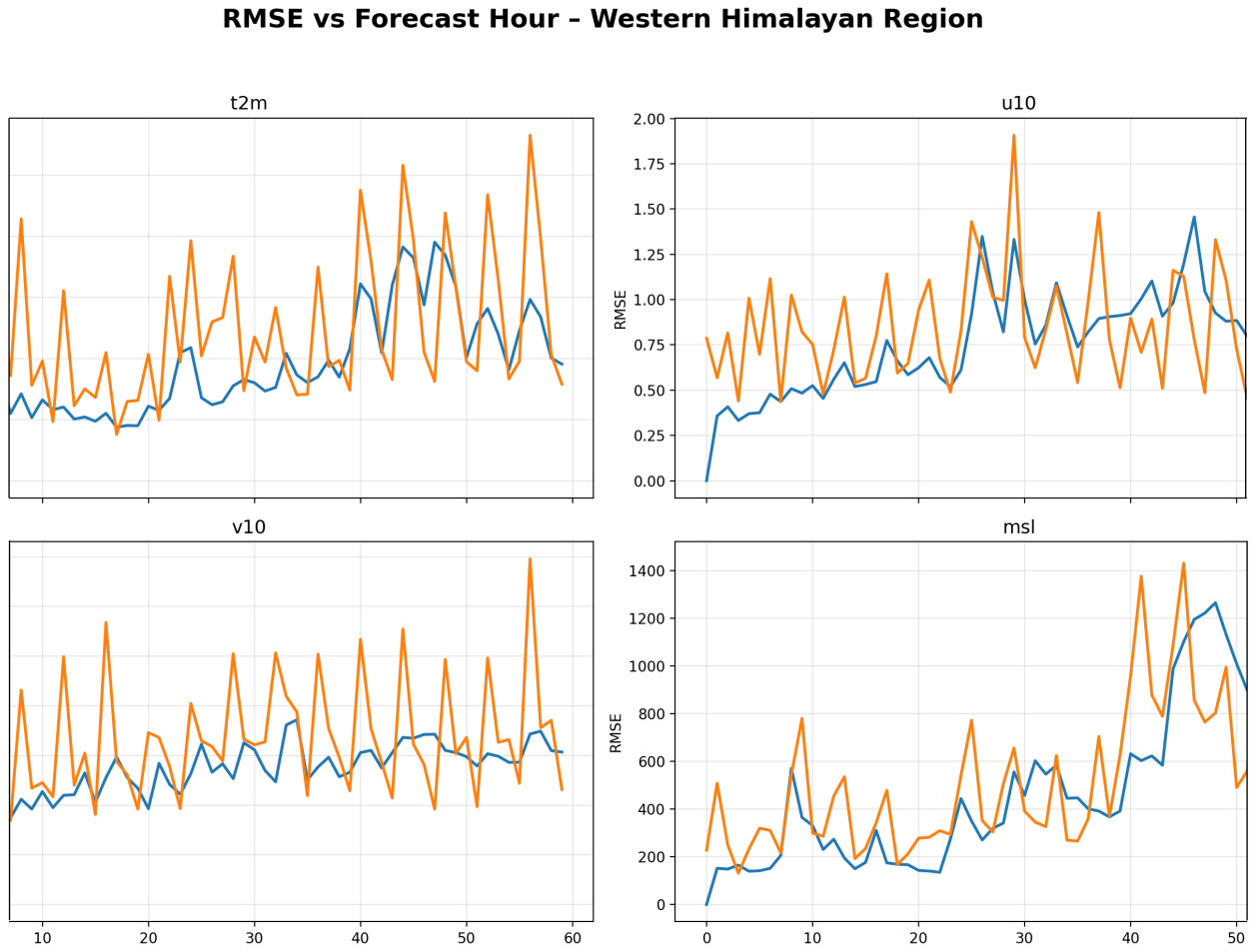
Evaluation of Models : **Agroclimatic zones**



RMSE of GraphCast and Pangu-Weather with ERA5 for each agroclimatic zone plotted against forecast steps

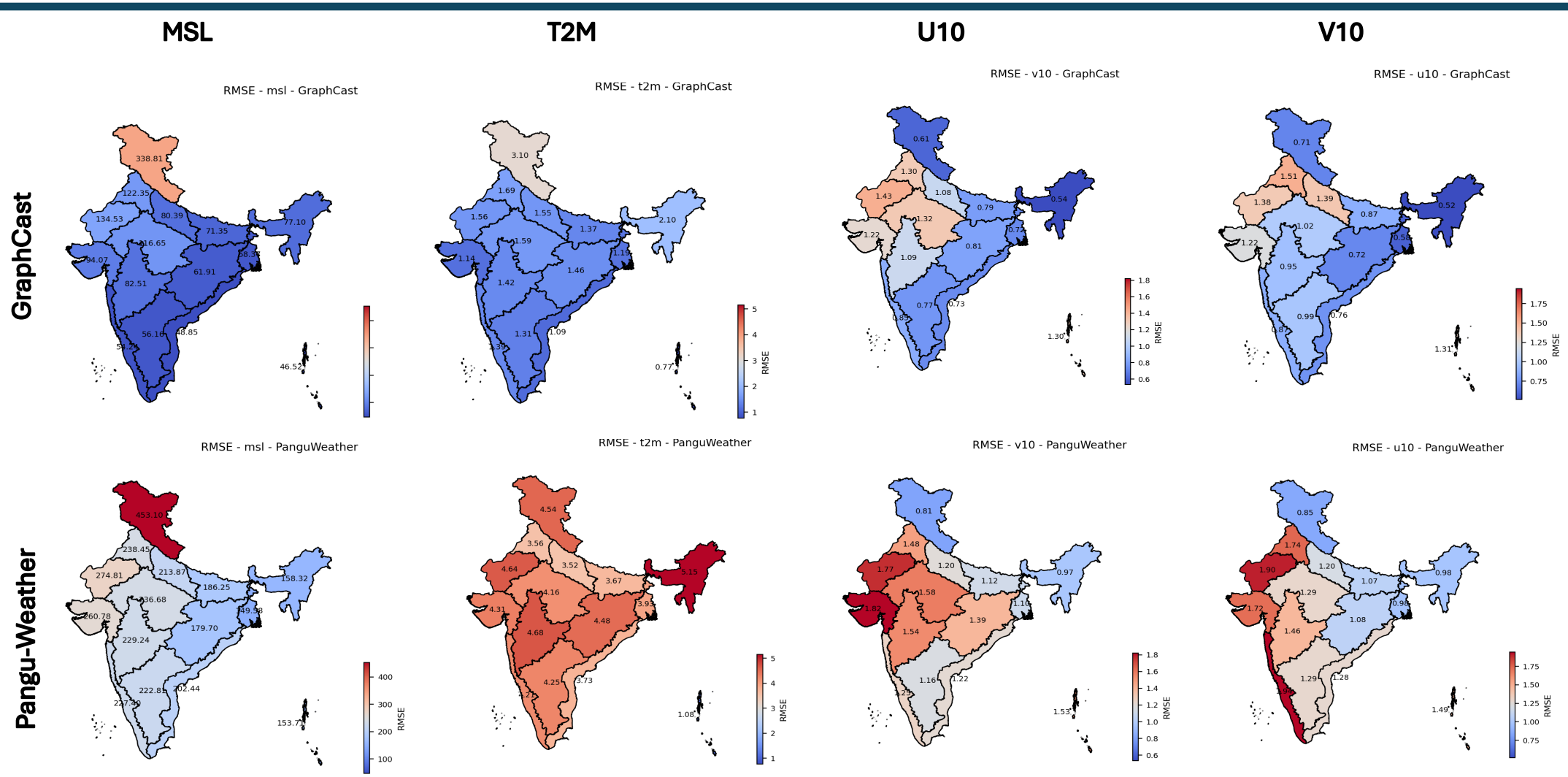
Evaluation of Models : **Lead-time Analysis**

- GraphCast in general is performing better than Pangu-Weather over the Western Himalayas
- The error is fluctuating for Pangu-Weather – The peaks are at 0000 hrs.
- MSL is diverging more compared to other variables.
- Further evaluation with other models to be included



Comparing performance of the T2M, MSL, U10, V10 in GraphCast and Pangu-Weather over the Northwestern Himalayas
(—GraphCast, —Pangu-Weather)

Evaluation of Models : **Area-averaged RMSE**

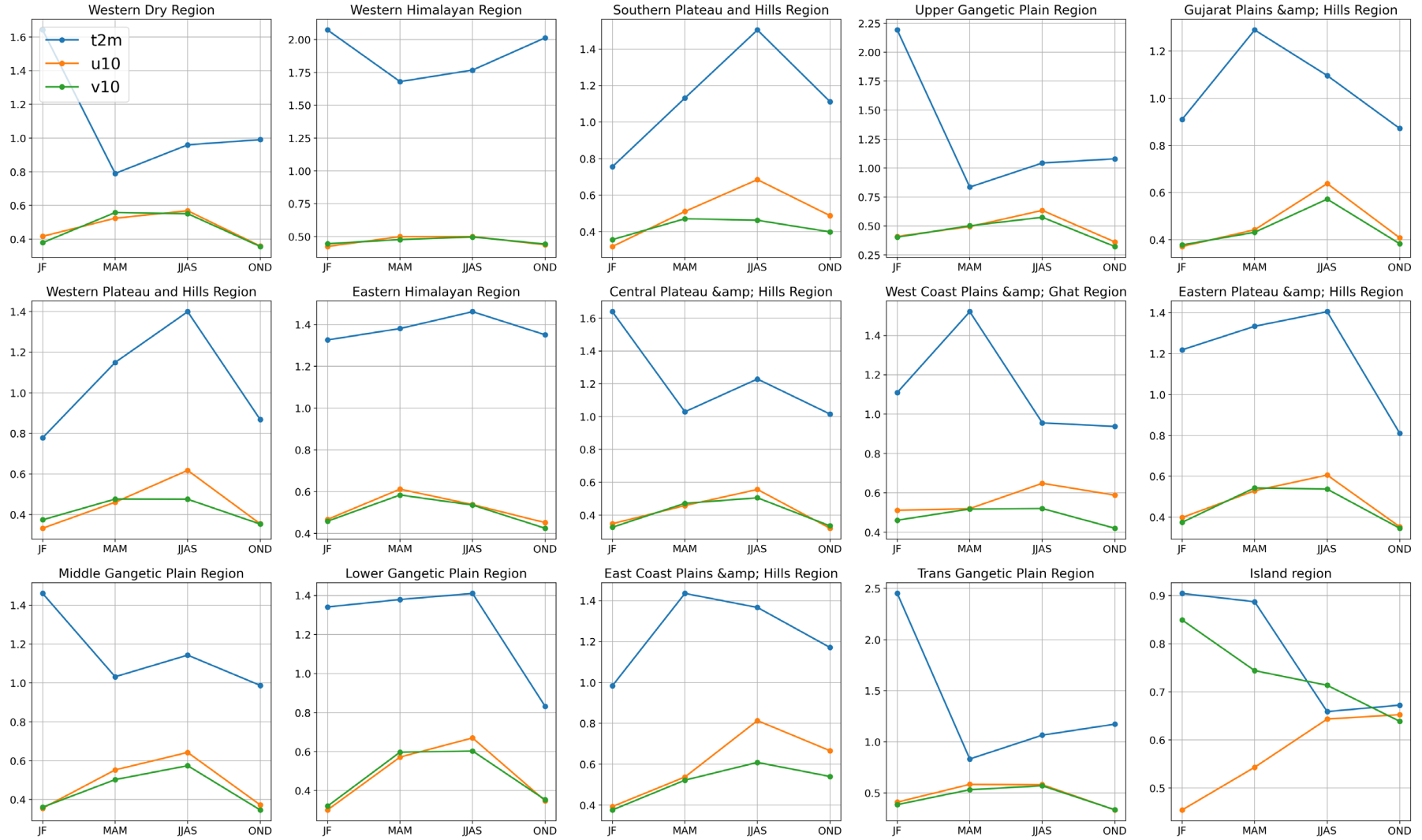


Time-averaged RMSE for each zone

MSL, T2M by both GraphCast and Pangu-Weather over the Western Himalayan region has high er RMSE

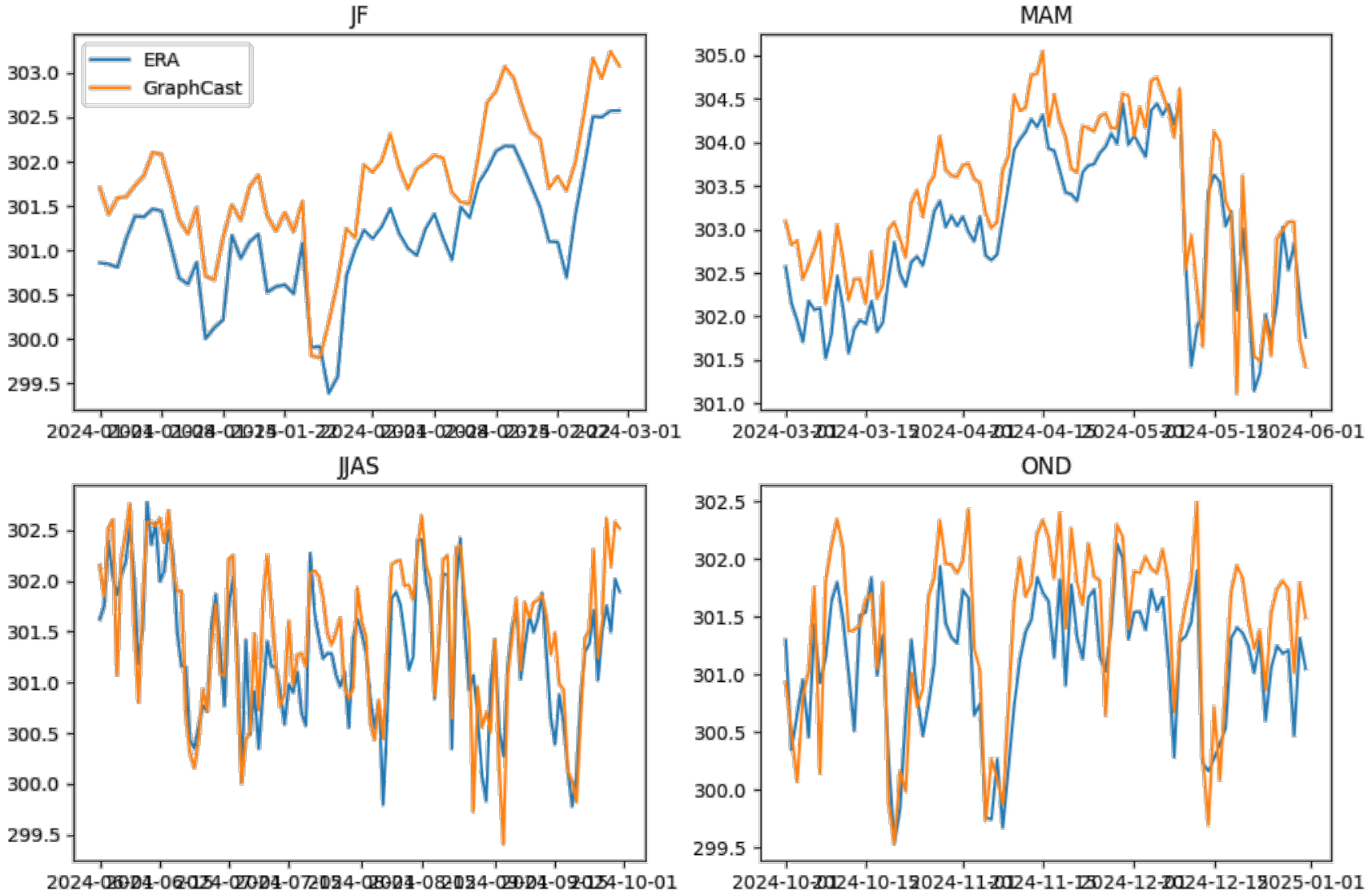
Evaluation of Models : A Seasonal Analysis

- GraphCast was run for 2024 all 366 days
- Analysis is done for **6 hour** lead time for **T2m, U10, V10,** for **JF/MAM/JJAS/OND.**
- **RMSE** is calculated w.r.t to ERA5
- For monsoon season (JJAS), wind speed has higher error. Indian monsoons need better forecasting and hence modelling.
- For the Western Himalayan region, the temperature for winter months has higher error

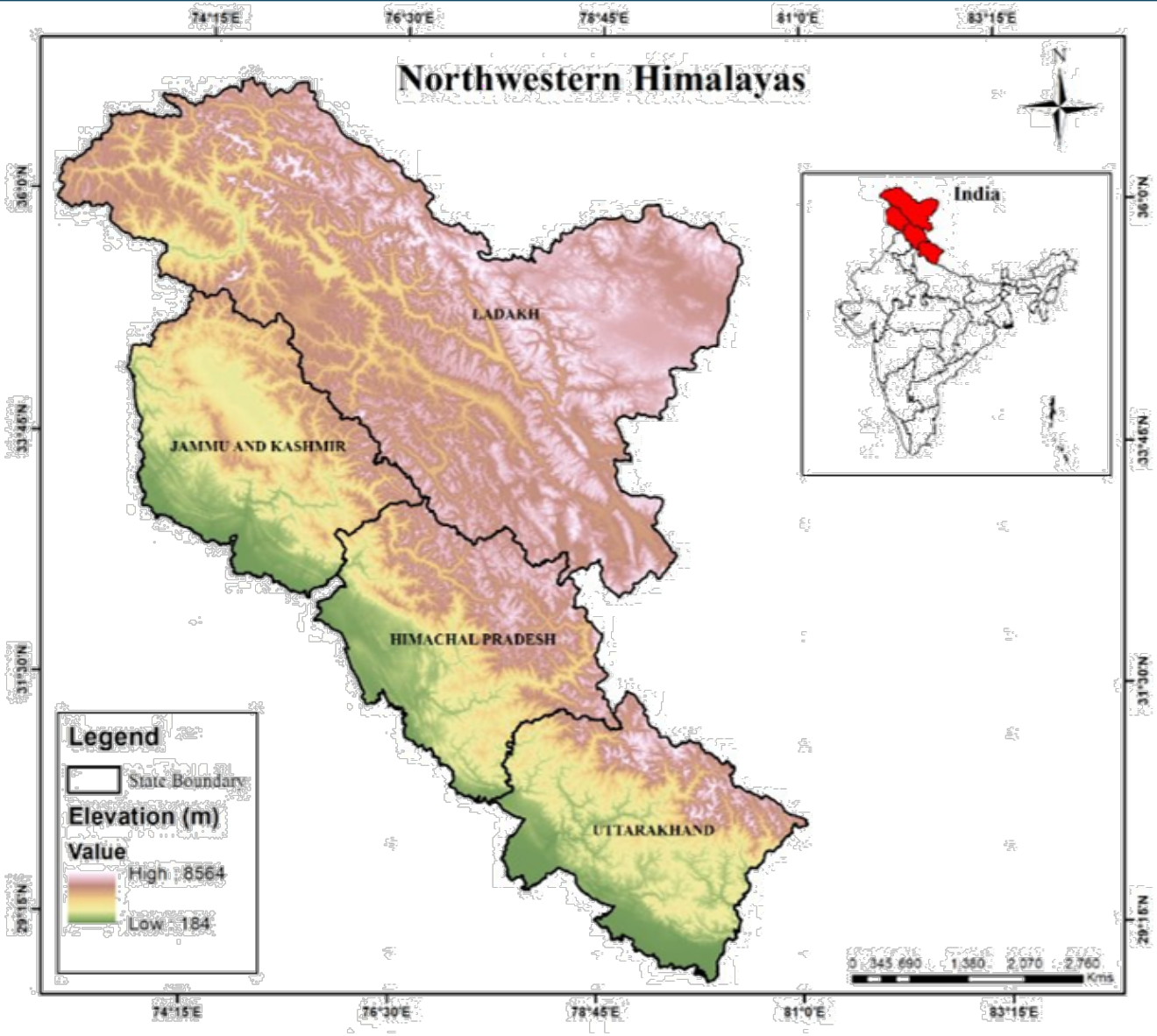


Evaluation of Models : A Seasonal Analysis

Western Himalayas daily 2m temperature trend for the whole year (0600 hours). The trend is matching with an offset, which is greater for JF.



Extreme Event Simulation : **Study Area**



- From the model evaluation, the Western Himalayan region was chosen to further improve the forecasting accuracy, This is mainly due to the complex terrain and increasing occurrences of weather relather natural disasters
- Hosts part of the Himalayas where the weather changes rapidly and is hard to forecast.
- Due to climate change, there is an increase in frequency of extreme precipitation, GLOF, landslides, cloud bursts, and other disasters.

Extreme Event Simulation : North Indian Floods (2023)

Source: The Washington Post

Heavy rainfall during the 2023 monsoon season resulted in severe flooding and landslides across Northern India, primarily affecting residents in Himachal Pradesh, Punjab, Chandigarh, Uttarakhand, Jammu and Kashmir, Haryana, Rajasthan and Delhi.

The event occurred in the monsoon month of **July** in **2023,** where maximum rainfall was recorded on **9th and 10th** in different places in the Western Himalayas. Cascading effects like landslides, flash floods, etc followed.

9,572 people were evacuated and 101,764 affected. In addition, 460 buildings were partially or fully damaged across 51 Districts of nine states.

Source: Business Today

Extreme Event Simulation : **WRF Setup**

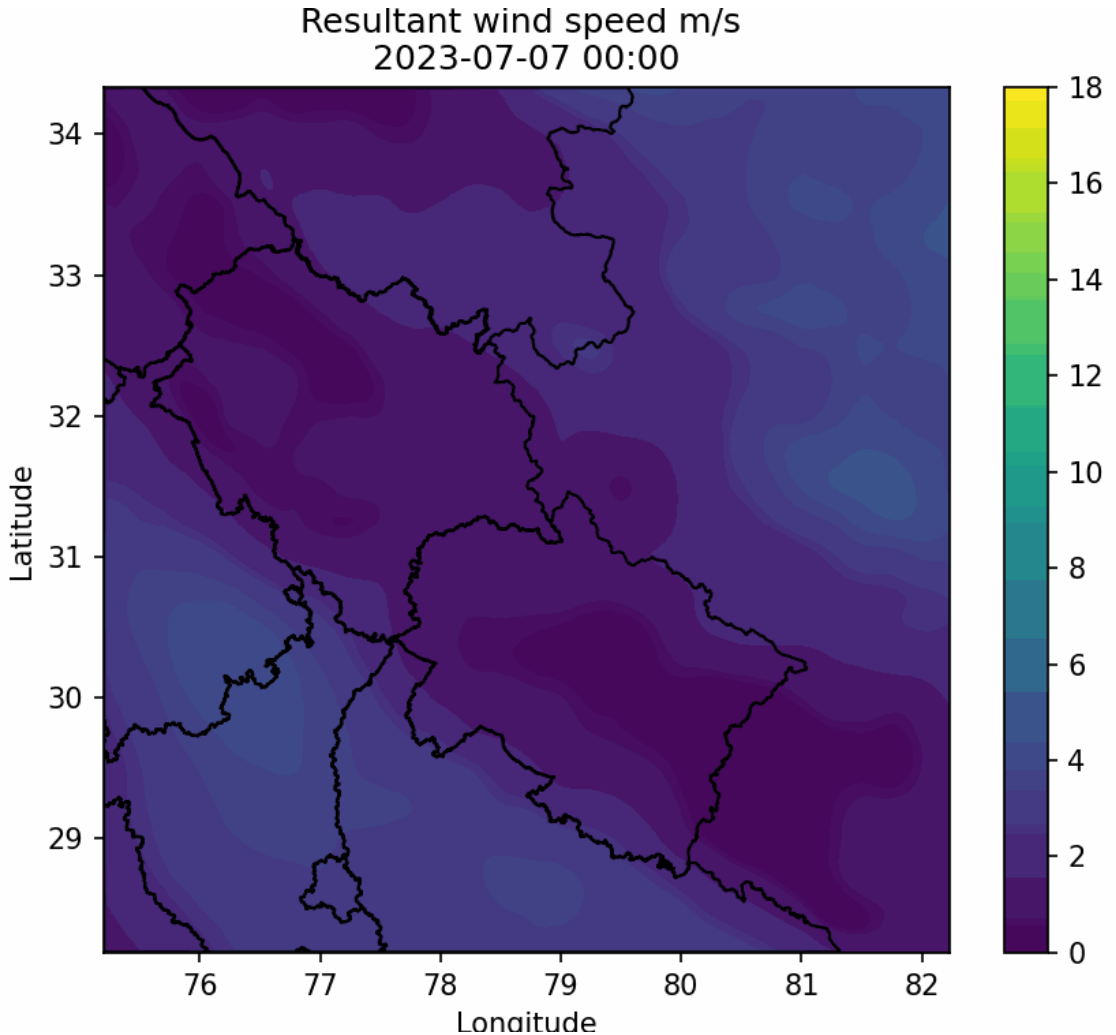
INPUT DATA : GRAPHCAST + ERA5

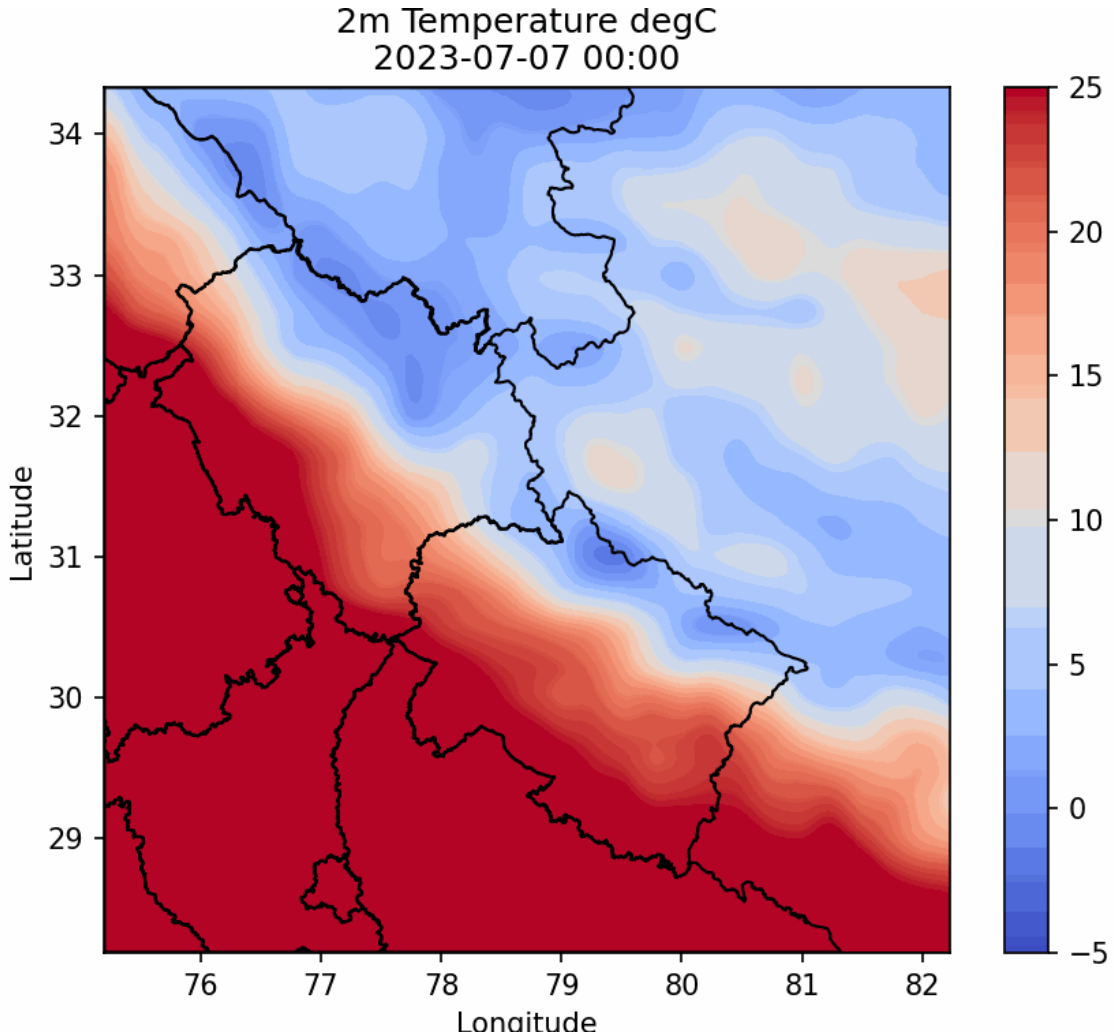
1. **GraphCast**, the high-resolution model used in the GraphCast paper (0.25 degree resolution, **37 pressure levels**), trained on ERA5 data from 1979 to 2017
    - Model set up: Run from 6th July to 12th July, 2023
    - Lead time : 24 hours forecast was extracted
    - Data: 7th July 0000 hrs 12th July 0000 hrs **6 hourly data**
2. **Missing Data :** Soil data is required for WRF to run using the surface physics.
    - 4 level soil temperature
    - 4 level soil moisture

WRF SET-UP : GRAPHCAST + ERA5

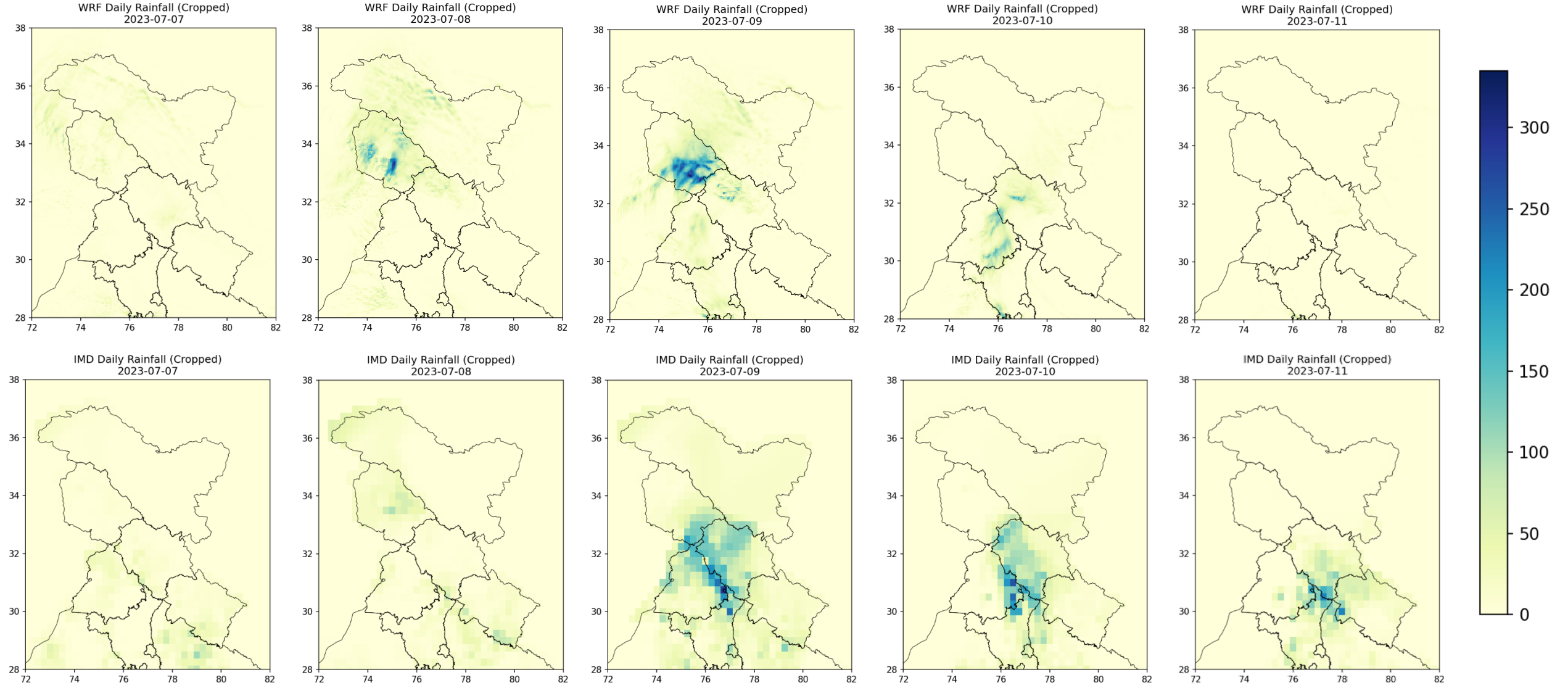
1. **Domain :** 2 nested domains over the Western Himalayas
    1. D01 : 5 x 5 KM resolution
    2. D02 : 1 x 1 KM resolution
2. **Time :** Hourly
3. **Ungrib :** The GraphCast and the missing soil data were “ungribed” independently using a custom Vtable.

Extreme Event Simulation : **WRF output 1km resolution**





Extreme Event Simulation : **WRF vs IMD rainfall comparison**

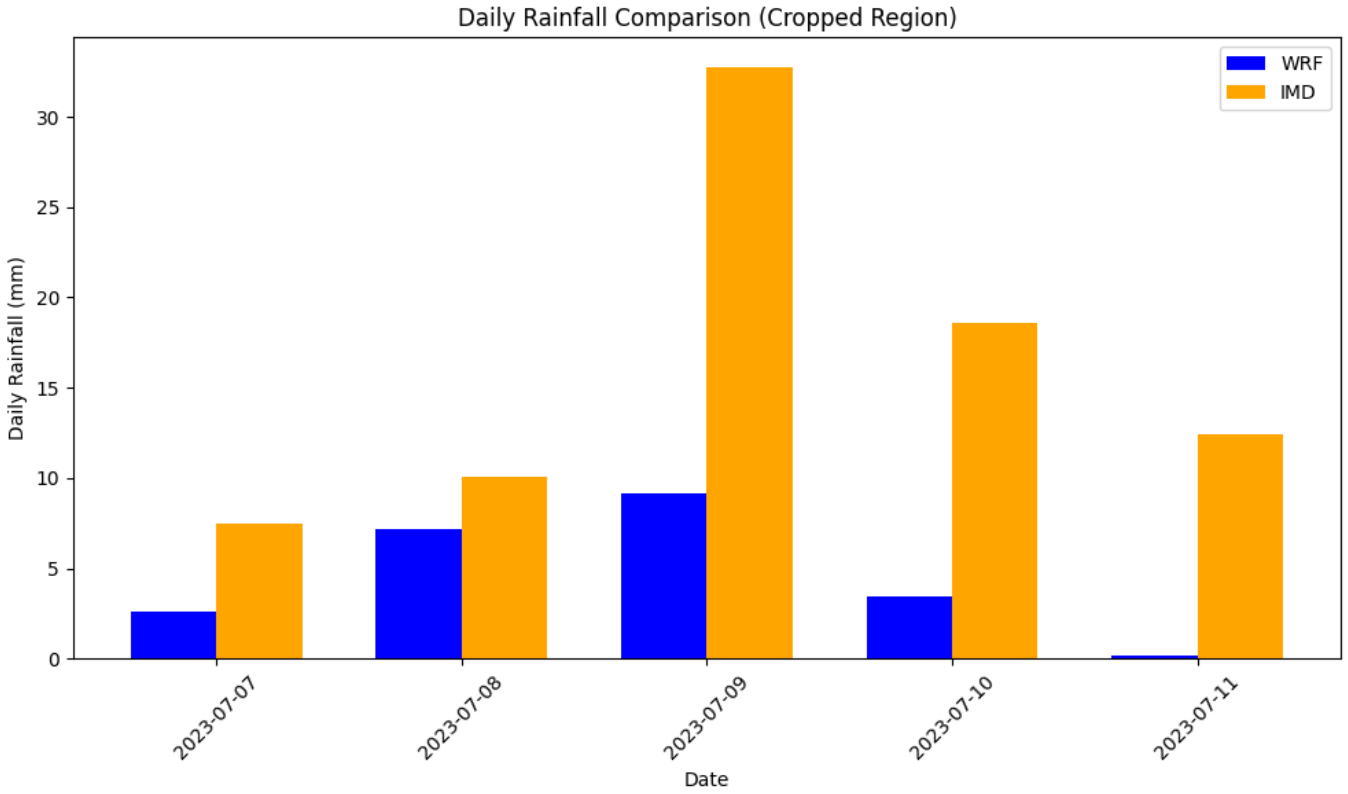


Extreme Event Simulation : **Conclusion**

WRF initialized with GraphCast data has underpredicted the rainfall for the 2023 North Indian Floods. The IMD data is gridded reanalysis data which is the closest to the observation.

For better evaluation, it could be compared to results of other models, like GFS.

To deploy GraphCast to be used for regional forecast using WRF, openly available GFS forecast data can be used to fill in the missing the missing soil data



References

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., ... & Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416-1421.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2022). Pangu-weather: A 3d high-resolution model for fast and accurate global weather forecast. *arXiv preprint arXiv:2211.02556*.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., ... & Huang, X. Y. (2019). A description of the advanced research WRF version 4. *NCAR tech. note ncar/tn-556+ str*, *145*(10.5065).

Raoult, B., Pinault, F., Mertes, G., Dramsch, J. S., Cook, H., & Chantry, M. ai-models [Computer software]. https://github.com/ecmwf-lab/ai-models


Thank
You!