PAUNet: Precipitation Attention-based U-Net for Rain Prediction from Satellite Radiance data

NeurIPS 2023 Competition – Weather4Cast challenge

P. Jyoteeshkumar Reddy
CSIRO

Harish Baki
TU Delft

Sandeep Chinta
MIT

Richard Matear
CSIRO

John Taylor
CSIRO

Background

Input

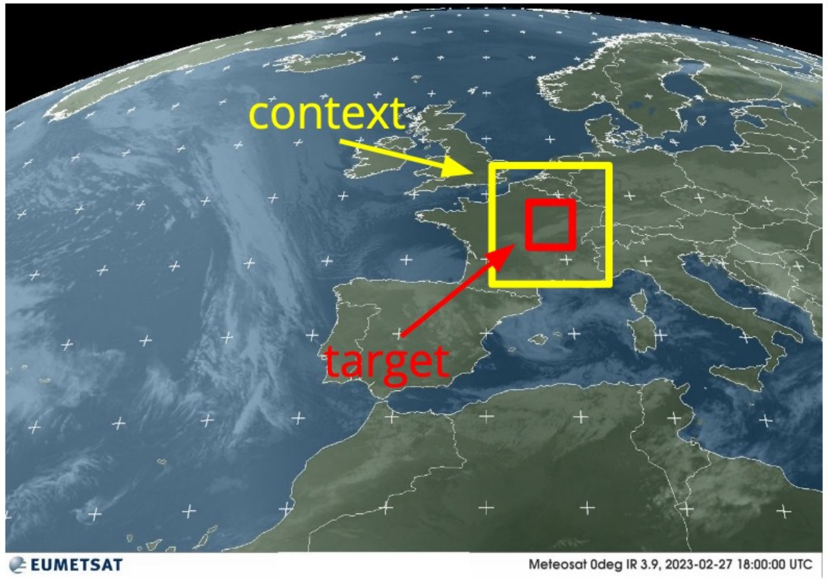
Predict

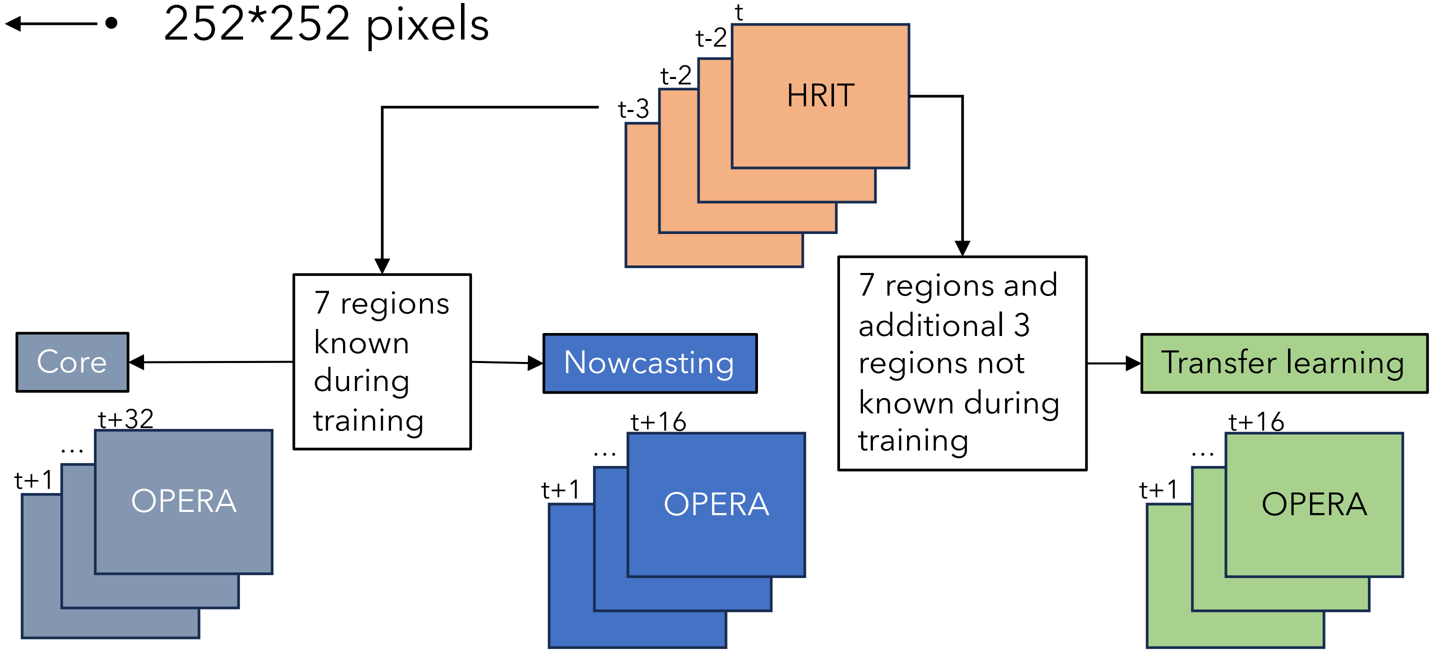
Target

Satellite radiance ⟶ Ground radar precipitation

- HRIT: EUMETSAT series of satellites
- 11 channels
- 15 minutes temporal resolution
- 12 km spatial resolution
- 252*252 pixels ⟶ Aligned at the center

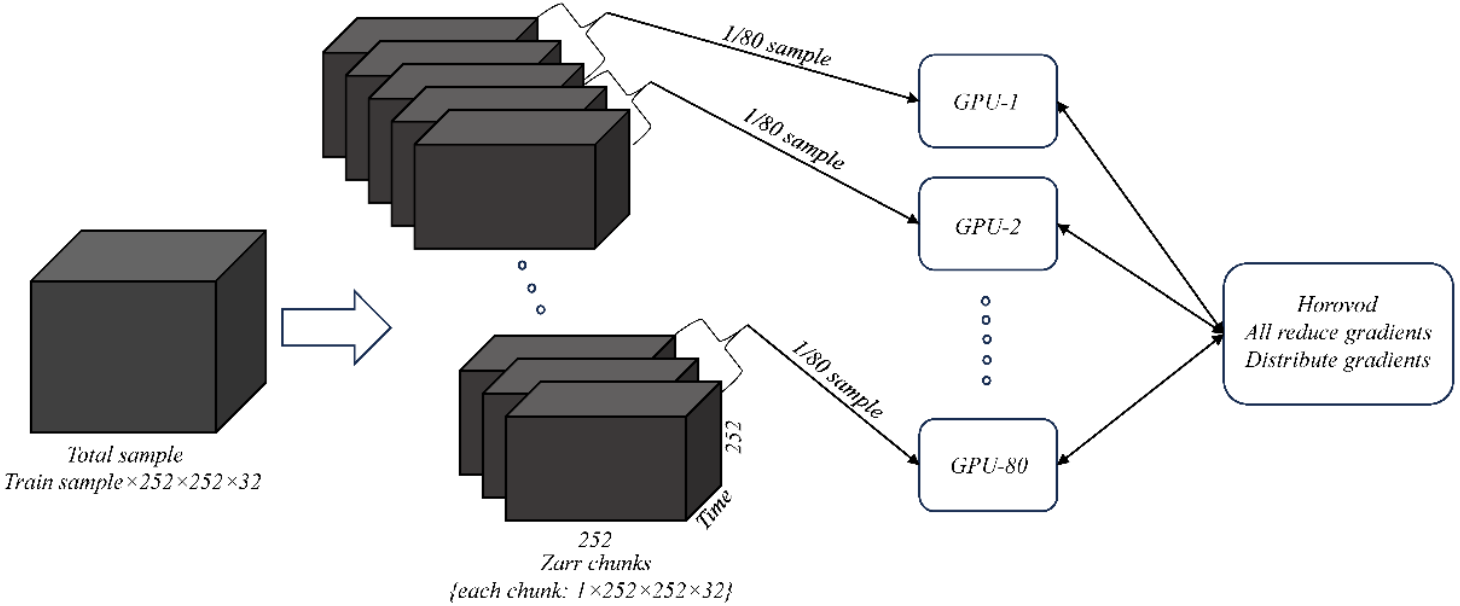
- OPERA: series of radars
- 1 composite precipitation
- 15 minutes temporal resolution
- 2 km spatial resolution
- 252*252 pixels





Data handling

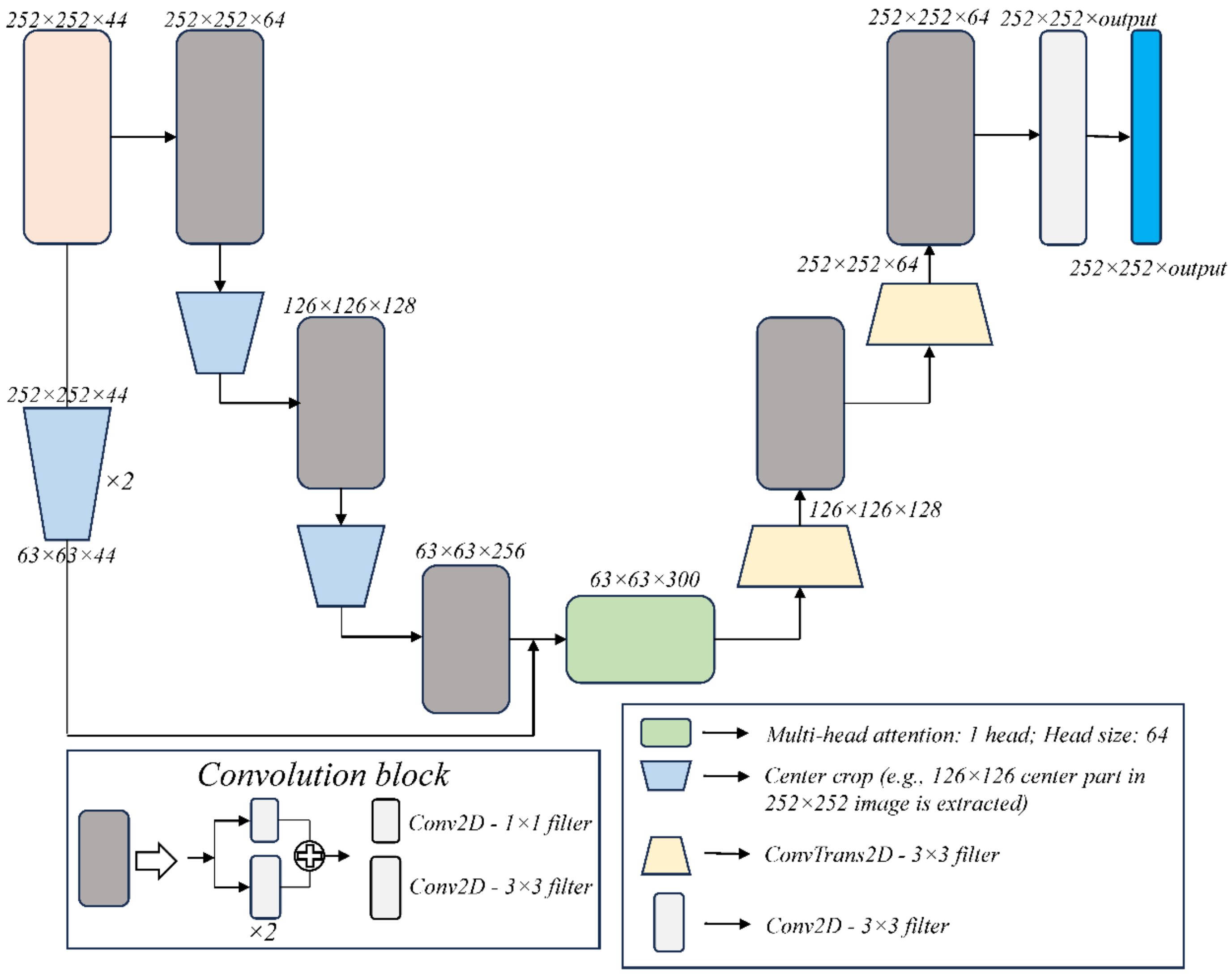
- Inputs: $samples\times 252\times 252\times 11\times 4$ Targets: $samples\times 252\times 252\times 32$
- Original training samples are 228928. Out of which, 69693 were removed due to number of rainy pixels less than 100 [1]
- Final samples: 159235 training and 840 validation
- 80 Nvidia V100 GPUs, 32 GB memory, on the Gadi supercomputer at the National Computational Infrastructure in Canberra



[1] Gruca, A., Serva, F., Lliso, L., Rípodas, P., Calbet, X., Herruzo, P., Pihrt, J., Raevskyi, R., Šimánek, P., Choma, M., Li, Y., Dong, H., Belousov, Y., Polezhaev, S., Pulfer, B., Seo, M., Kim, D., Shin, S., Kim, E., Ahn, S., Choi, Y., Park, J., Son, M., Cho, S., Lee, I., Kim, C., Kim, T., Kang, S., Shin, H., Yoon, D., Eom, S., Shin, K., Yun, S.-Y., Le Saux, B., Kopp, M.K., Hochreiter, S., Kreil, D.P., Ciccone, M., Stolovitzky, G., Albrecht, J., sro, M.: Weather4cast at NeurIPS 2022: Super-Resolution Rain Movie Prediction under Spatio-temporal Shifts. In: Proceedings of Machine Learning Research. pp. 292–312 (2023)

PAUNet model architecture

- Center cropping –contextual feature extraction
- Residual network based convolution block
- Multiheaded attention
- Deconvolution based downscaling

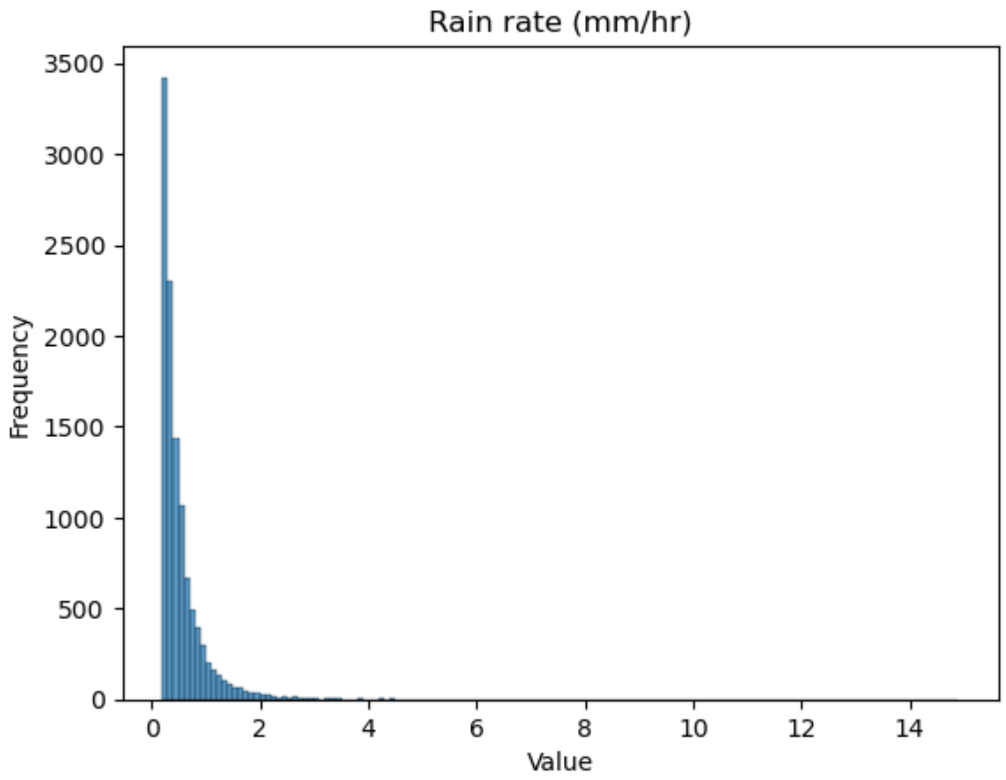


A suitable loss function is all you need

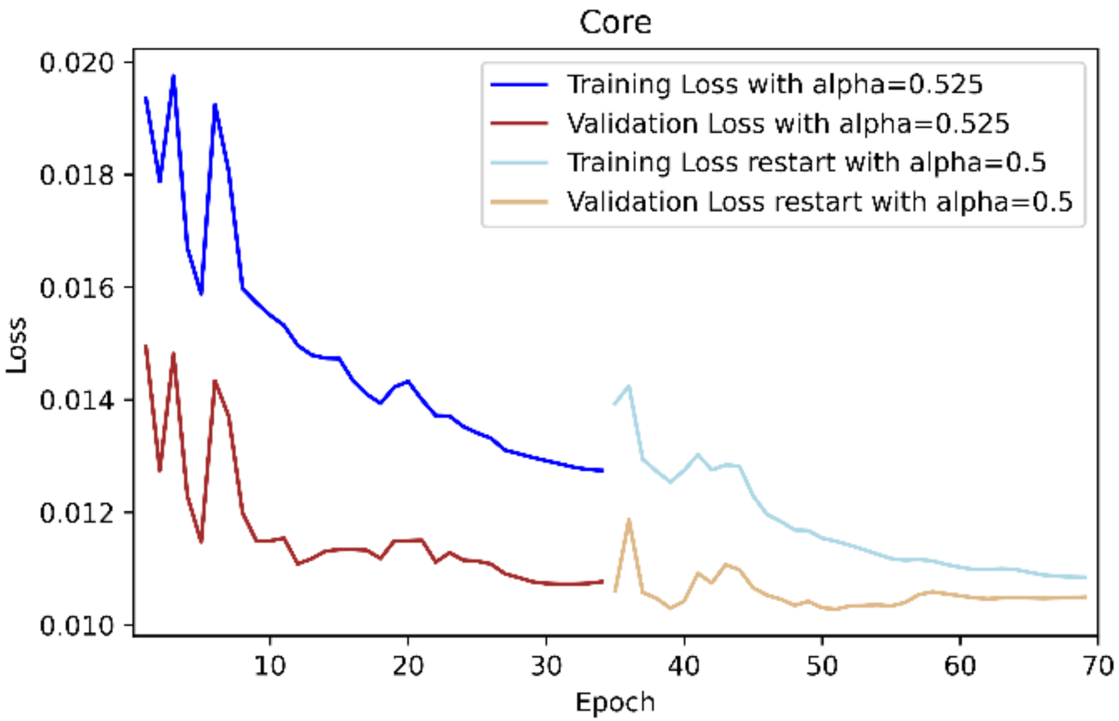
- Experimented with
- Mean Squared Error loss (L2)
- Mean Absolute Error loss (L1)
- L1 + L2 with thresholds 0.2, 1, 5, 10, and 15
- Binary Cross Entropy loss with thresholds 0.2, 1, 5, 10, and 15
- Combination of BCE and Focal loss with thresholds 0.2, 1, 5, 10, and 15
- Combinations of all the above, with thresholds 0.2, 1, 5, 10, and 15
- **Exponential Focal Precipitation Loss (e-FPL)**
- $\boldsymbol{eFPL} = [\mathbf{min}(\boldsymbol{\beta e^{\alpha x}}, \mathbf{128})]\times[\boldsymbol{MAE}(\boldsymbol{x}', \boldsymbol{y}')]\times\left[\left(\mathbf{1.01} - \boldsymbol{e}^{-\boldsymbol{MAE}}\right)^{\boldsymbol{\gamma}}\right]$

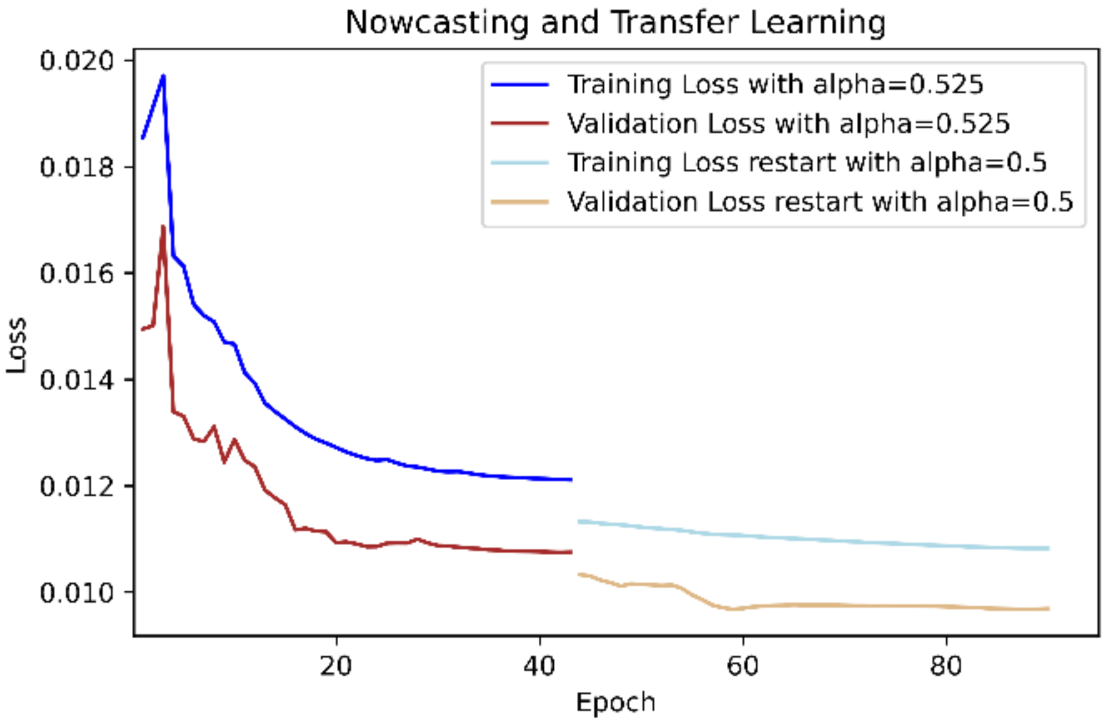
$$\alpha, \beta, and\ \gamma\ are\ hyperparameters$$

$$\alpha\ is\ a\ super - hyperparameter$$



Learning curves





- Loss curves ensures models are not overfitted
- Pseudo-ensemble: models saved at different training epochs are selected as ensemble members

Results

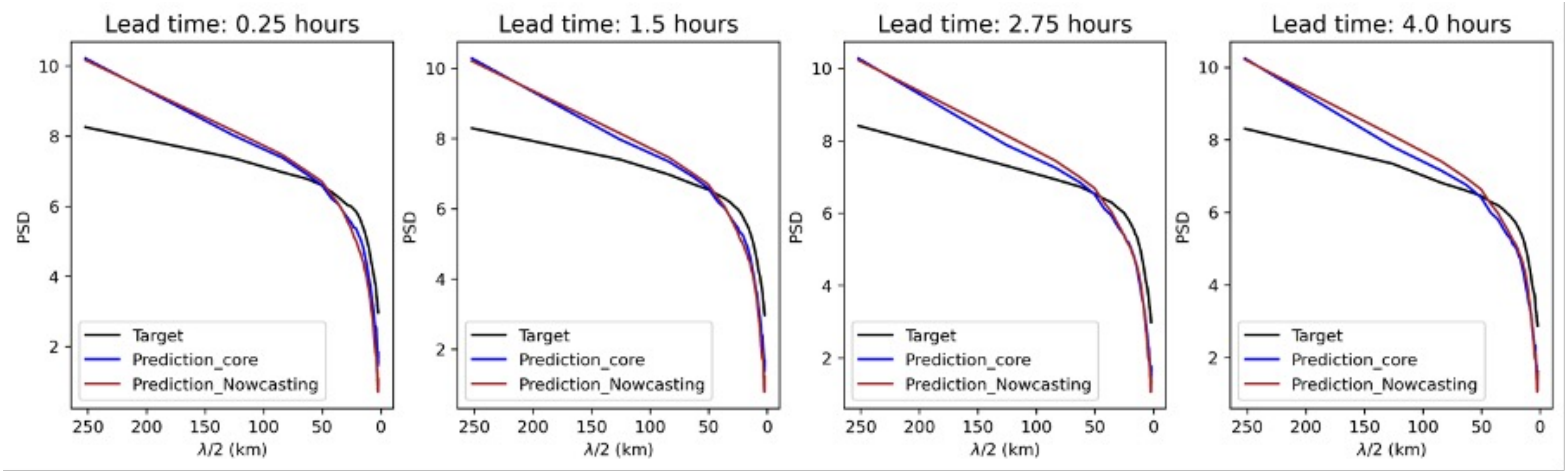
Test scores

CSI score	Core	Nowcasting	Transfer Learning
Leaderboard official baseline	0.0444	0.0482	0.0459
PAUNet	0.0488 (10%)	0.0569 (18%)	0.0575 (25%)

Validation scores

Metric	Core	Nowcasting	Transfer Learning
CSI (0.2)	0.195 (73.01%)	0.218 (74.26%)	0.222 (76.02%)
CSI (1)	0.062 (23.14%)	0.066 (22.45%)	0.061 (20.81%)
CSI (5)	0.006 (2.40%)	0.007 (2.23%)	0.006 (2.05%)
CSI (10)	0.003 (1.01%)	0.002 (0.60%)	0.002 (0.65%)
CSI (15)	0.001 (0.44%)	0.001 (0.45%)	0.001 (0.47%)
CSI (overall)	0.053	0.059	0.058
HSS (overall)	0.064	0.072	0.072

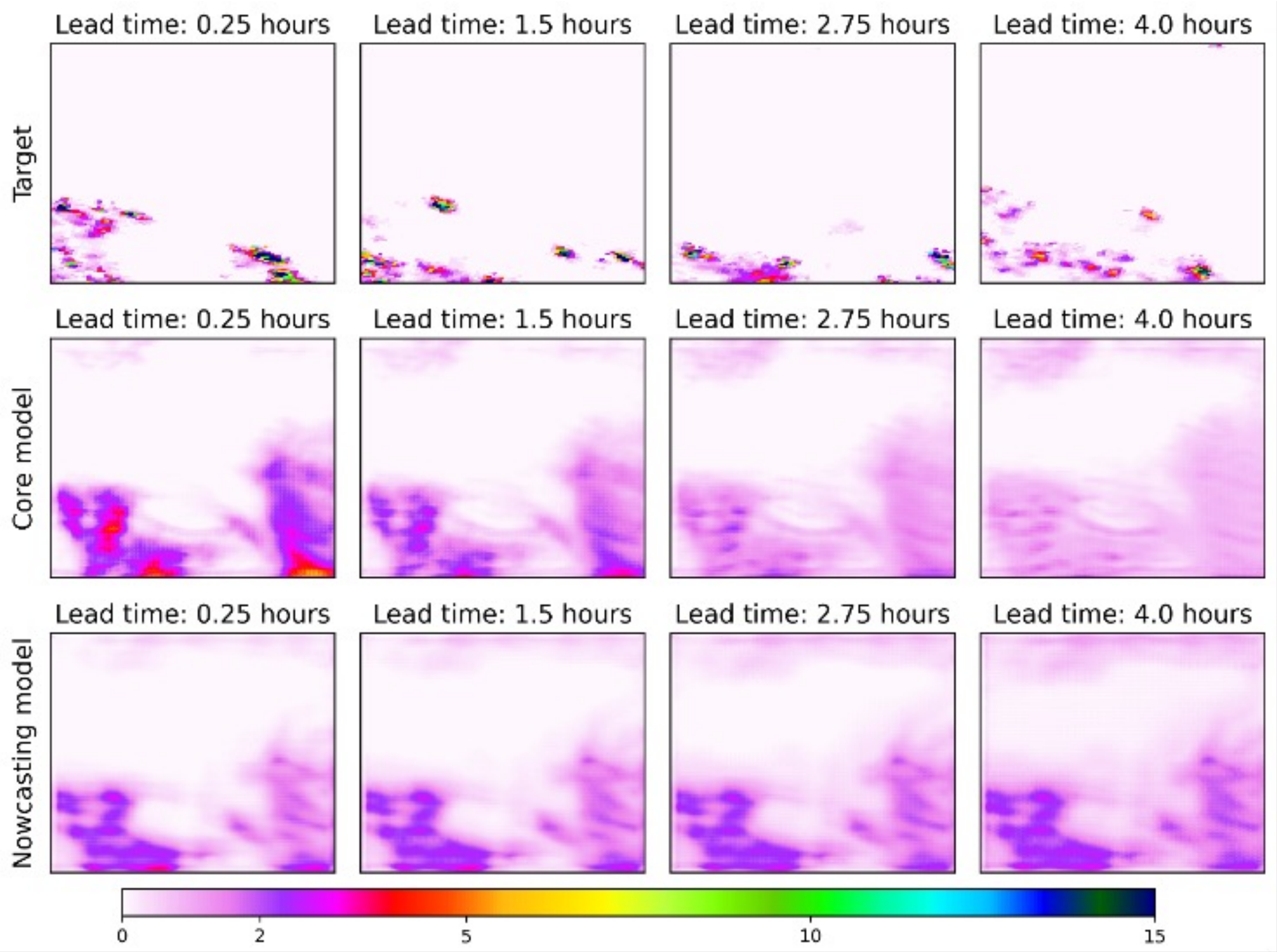
Power Spectral Density – A potential evaluation metric



- This suggests that model produces much rain and substantially underestimates the extremes

Model predictions

- Model at a short lead time able to predict the spatial structure and moderate-high rainfall intensities
- At long lead times, predictions smear out
- Cannot accurately represent the high-resolution local scale extreme rainfall intensities



Conclusions and future scope

- Zarr allows efficient data handling across CPUs on GPU nodes
- Basic U-Net with center cropping and multiheaded attention
- Heart of the model accuracy lies in the choice of loss function
- Exponential focal precipitation loss substantially improved model accuracy

- Power spectral density will provide valuable insights as an evaluation metric
- Fine tuning the super-hyperparameter
- Inclusion of orography
- Exploring various architectures, including GANs and Diffusion models

Thank you for you attention