A Deep Learning Framework to Efficiently Estimate Precipitation at the Convection Permitting Scale

Valentina Blasone[1], **Erika Coppola**[2], **Guido Sanguinetti**[3], **Viplove Arora**[3],
Serafina Di Gioia[2], **Luca Bortolussi**[1]

[1] Department of Mathematics, Informatics and Geosciences, University of Trieste
[2] Earth System Physics Section, ICTP, Trieste
[3] Theoretical and Scientific Data Science, SISSA, Trieste
{valentina.blasone@phd.units.it}

Abstract

Precipitation-related extreme events are rapidly growing due to climate change, emphasizing the need for accurate hazard projections. To effectively model the convective phenomena driving severe precipitation, high-resolution estimates are crucial. Existing methods struggle with either insufficient expressiveness in capturing complex convective dynamics, due to the low resolution, or excessive computational demands. In response, we propose an innovative deep learning framework that efficiently harnesses available data to yield precise results. This model, based on graph neural networks, utilises two grids with different resolution and two sets of edges to represent spatial relationships. Employing as input ERA5 reanalysis atmospheric variables on an approximately 25 km grid, the framework produces hourly precipitation estimates on a finer 3 km grid. Findings are promising in accurately capturing yearly precipitation distribution and estimating cumulative precipitation during extreme events. Notably, the model demonstrates effectiveness in spatial regions not included in the training, motivating further exploration of its transferability potential.

1 Introduction

Every year across the world, natural catastrophes cause casualties and significant damage to properties and assets and the trend in weather related extremes is growing, due to climate change (IPCC, 2022). Precipitation-related events (flood, drought, landslides) have a tremendous social and economical impact and are all projected to increase (Van Aalst, 2006; Banholzer et al., 2014). Disaster risk forecasting highly depends on the ability to correctly quantify the hazard related to the natural phenomenon, which is not straightforward, particularly for precipitation. Mainly two types of precipitation exist: weak and severe. The former is caused by strati-form clouds, uniform and stable, with no complicated airflow motion. The latter is related to convective systems, complex and characterised by non-linear airflow motion. Severe precipitation is crucial when considering extreme events, yet difficult to model. Traditional approaches to estimate precipitation on high-resolution grids include Quantitative Precipitation Estimation (QPE) (Cuo et al., 2011), downscaling of low-resolution precipitation projections (CORDEX; EUCP; Laflamme et al., 2016) and Convection Permitting Models (CMPs) (Coppola et al., 2020; Kendon et al., 2021). QPE and downscaling are cost-effective, but struggle in modelling severe precipitation, while CPMs can correctly model convective systems, but high resolution comes at an enormous computational cost. In this setting, models based on Machine Learning (ML) and particularly Deep Learning (DL) may play an important role bringing both improved accuracy in severe events modelling and significant computational speed-up. DL models have been used to reproduce QPEs results (Wang et al., 2022), to improve downscaling (Reddy et al., 2023; Kumar et al., 2023), and few studies addressed the problem of building emulators for precipitation (Wang et al., 2021; Doury et al., 2023,Addison et al., 2022). This study, represents a first attempt in building a DL model that resembles convection permitting dynamical models, with the aim of estimating precipitation distribution at high spatial and temporal resolution. Furthermore, the proposed framework does not use precipitation as predictor and is trained using reanalysis data, which are less biased than model data, conventionally used in training emulators. Nonetheless, climate model data can be used as input for the model during the prediction phase to derive precipitation projections. This will be investigated in a future stage, being particularly meaningful to quantify the impact of climate change on precipitation, especially severe precipitation.

The framework is based on Graph Neural Networks (GNN) (Battaglia et al., 2018; Sanchez-Lengeling et al., 2021; Schlichtkrull et al., 2018) which have recently emerged as a powerful tool in the field of ML, particularly for tasks involving relational data and graph structures. GNNs can capture complex relationships and dependencies within data, making them well-suited for different applications, and have been recently used in climate-related studies (Lam et al., 2023). One of the notable advantages of GNNs is their capacity for transferability across different domains, i.e. GNNs can be trained on one graph and effectively applied to other graphs with similar structural characteristics. This results in a flexible framework, providing the opportunity to investigate predictive capabilities across geographical areas distinct from those encountered during the training.

2 METHODOLOGY

2.1 DATA

Five atmospheric variables are used as predictors, at five levels of pressure (Table 1) each represented on a low-resolution grid of $0.25°$ longitude-latitude ($\sim$25 km for Europe). Hourly values are considered for predictors. As training input data, these variables are taken from the ERA5 reanalysis dataset (Hersbach et al., 2020) from the European Centre for Medium-Range Weather Forecasts (ECMWF). The use of reanalysis data to train the DL model presents several advantages over relying on climate model data. Reanalysis data assimilate a wide range of observational data, providing a more accurate representation of historical climate conditions. This approach helps mitigate biases and uncertainties that may be inherent in climate model simulations. Conversely, the use of climate models data is beneficial when the objective is to derive future projections and may be used during the prediction phase. Additionally, a remapping of the global multi-resolution terrain elevation data (Danielson & Gesch, 2011) to a grid of 3 km is used as a predictor. Input atmospheric variables and topographic altitude are both normalized to zero-mean unit-variance. Atmospheric variables data at five time instants are considered for an individual hourly prediction: the prediction time and the $6, 12, 18$ and 24 hours before.

Table 1: Atmospheric variables used as predictors.

Variable	**Symbol**	**Unit**	**Pressure Levels [hPa]**	**Space res.**	**Time res.**
Specific humidity	q	[kg kg-1]	1000; 850; 700; 500; 200	$25° \times 25°$	1hr
Temperature	t	[K]	1000; 850; 700; 500; 200	$25° \times 25°$	1hr
Eastward wind	u	[m/s]	1000; 850; 700; 500; 200	$25° \times 25°$	1hr
Northward wind	v	[m/s]	1000; 850; 700; 500; 200	$25° \times 25°$	1hr
Geopotential	z	[m^2/s^2]	1000; 850; 700; 500; 200	$25° \times 25°$	1hr

The GRidded Italian Precipitation Hourly Observations (GRIPHO), a high-resolution hourly precipitation dataset for Italy (Fantini, 2019), was selected as the target for this framework. Originally developed to support hydrological models and validate Regional Climate Models (RCMs) simulations, GRIPHO utilises raw station data from CETEMPS (Verdecchia, 2019), the sole high-resolution station-based dataset covering the entire country from 2001 to 2016. Following cleaning and re-gridding processes, the data is represented over a 3 km Lambert Conformal Conic grid, which is not orthogogal, neither regular in longitute-latitude coordinates.

2.2 Graph Conceptualization

Each point within the low-resolution and high-resolution grids corresponds to a specific geographical location, suggesting to model both grids as a unified heterogeneous graph featuring two node- and two edge-types:

- *Low* nodes: first set of nodes, generated from the points on the low-resolution grid with spatial resolution of approximately 25 km.
- *High* nodes: second set of nodes, created from the points on the high-resolution grid with spatial resolution of 3 km.
- *Low-to-High* edges: unidirectional edges, which connect *Low* nodes to *High* nodes, with each *High* node linked to a fixed number $k = 9$ of *Low* nodes, chosen through k-Nearest Neighbours (k-NN); these edges model the downscaling of atmospheric variable information from the *Low* nodes to the *High* nodes (Figure 1a).
- *High-within-High* edges: bidirectional edges that capture relationships among *High* nodes based on an 8-neighbors approach, ensuring each node is connected to its eight nearest neighbors (Figure 1b).

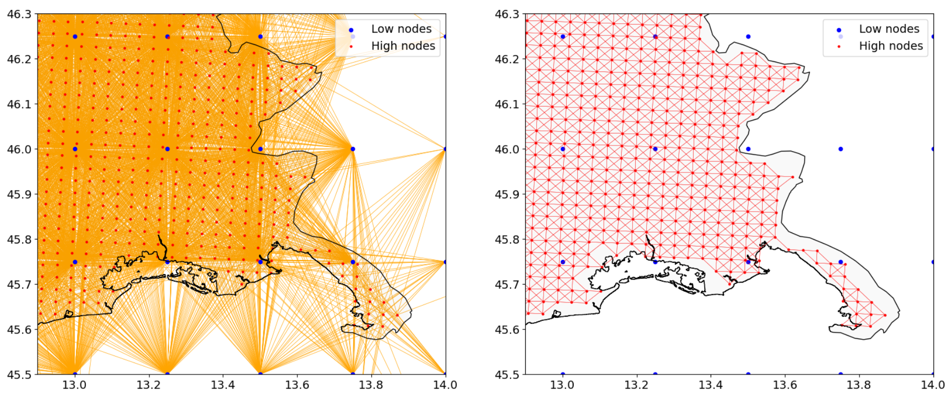

Figure 1: Graph conceptualisation, close-up of (a) Low-to-High edges, (b) High-within-High edges

2.3 DL Model

A notable characteristic of the target dataset is that 90% of its values fall below the meteorological threshold for precipitation, assumed as 0.1 mm, effectively rendering them as zeros. This inherent imbalance significantly impacts the training of DL models. To mitigate this challenge, the Hurdle approach (Cragg, 1971) is adopted, through the construction of two distinct models: a *Classifier* and a *Regressor*. The *Classifier* is trained on the entire dataset and discerns between two classes: 0, i.e. precipitation below the threshold, and 1, i.e. precipitation above the threshold. Conversely, the *Regressor* is exclusively trained on targets where precipitation values exceed the threshold, and provides a quantitative estimation of hourly precipitation. During the evaluation phase, predictions from both models are computed and subsequently multiplied to yield a singular estimate of the precipitation value. A scheme of the Hurdle approach adopted in this framework is depicted in Figure 2a. Models share the same GNN structure, which consists of three primary modules (see Figure 2b). First, a *downscaler* module, which employs a single message passing layer to map atmospheric variables, represented as *Low* node features, to learned attributes on the *High* nodes. Following this, a *processor* module that updates information on the *High* nodes, starting from the learned attributes and the topographic altitude, incorporating spatial relationships through a sequence of five message passing layers. Finally, the *predictor* module, which utilises a feed-forward network to make the ultimate prediction based on the processed information from the preceding modules. The model is entirely implemented in PyTorch, using as GNN layers graph convolutions with attention (Veličković et al., 2017), particularly the GATv2Conv layer from (Brody et al., 2021). The choice of using GNNs is particularly effective for mainly two reasons: overcome the issue of having different coordinate systems of input/output grids, causing them to be not nested, and the potential tranferability to different spatial domains, allowing predictions on graphs with different number of nodes from that of the training graphs.

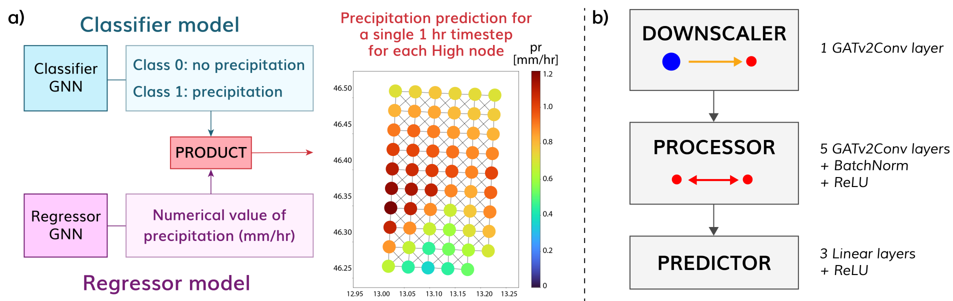

Figure 2: Schematic views of (a) Hurdle approach, (b) GNN model.

2.4 Training and Testing

The geographical area considered for training is approximately 120.000 km^2 (see Figure 3a) with 720 *Low* nodes and 14036 *High* nodes, representing the north of Italy. A time span of 15 years, from 2001 to 2015 is considered for training. The time span was subdivided as 90% training set and 10% validation set. The year 2016 was instead used for testing. Focal Loss (FL) (Lin et al., 2017) (see Appendix A.1) was used to train the *Classifier* model, while the *Regressor* was trained using Weighted Mean Square Error (WMSE) loss as in Wang et al. (2022) (see Appendix A.2). Both *Classifier* and *Regressor* models were separately trained for 24 hours on 4$\times$ NVIDIA Ampere GPUs on Leonardo, the new pre-exascale Tier-0 EuroHPC supercomputer hosted by CINECA and currently built in the Bologna Technopole, Italy (Turisini et al., 2023). The model was then tested by deriving precipitation estimates for all the Italian territory, approximately 300.000 km^2 (see Figure 3a), with 2646 *Low* nodes and 33153 *High* nodes, in order to check its effectiveness also in spatial regions beyond its training scope. The trained framework demands only few minutes for getting precipitation estimates for an entire year, making it one of its main advantages. This in fact significantly contrasts with the time demands of convection-permitting dynamical models, which necessitate days to yield results on an equivalent high-resolution grid. The framework was verified in its ability of estimating annual precipitation distribution, as well as cumulative precipitation during extreme events.

3 Results

Figure 3 shows the results obtained utilising the proposed framework on the testing area and period, compared to observations taken from the GRIPHO dataset. More specifically, Figure 3b and Figure 3c show the aggregated results for the testing year 2016, respectively in terms of cumulative precipitation and precipitation intensity distribution. Results are overall good, with predictions that sometimes slightly underestimate or overestimate the observations, particularly in the areas beyond the training scope. Figure 3d shows the time series of average hourly precipitation. This plots shows how the DL results are able to match quite well the observation at the desired high temporal resolution, specifically the peaks, with a low Root Mean Square Error (RMSE). Note that the RMSE is given with sign to show where the DL results overestimate (positive RMSE) or underestimate (negative RMSE) the observation at the considered hourly time step. Figure 3e shows instead the cumulative precipitation for a set of six extreme events that took place in the time span and areas considered in the study[1]. Also in this case, the framework is able to capture the event total precipitation, both in terms of area extension and intensity. As expected, a worsening in performance is observed as the considered region moves away from the training area. Nevertheless, the ability of the framework to generalise both in space and in time suggests it learned an approximation of the physical relation between the large scale dynamic and the local scale precipitation, possibly improvable by fine-tuning the model over a wider area.

[1]The event of 13-14 Sept 2015 is part of the training set, while all others are part of the testing set.

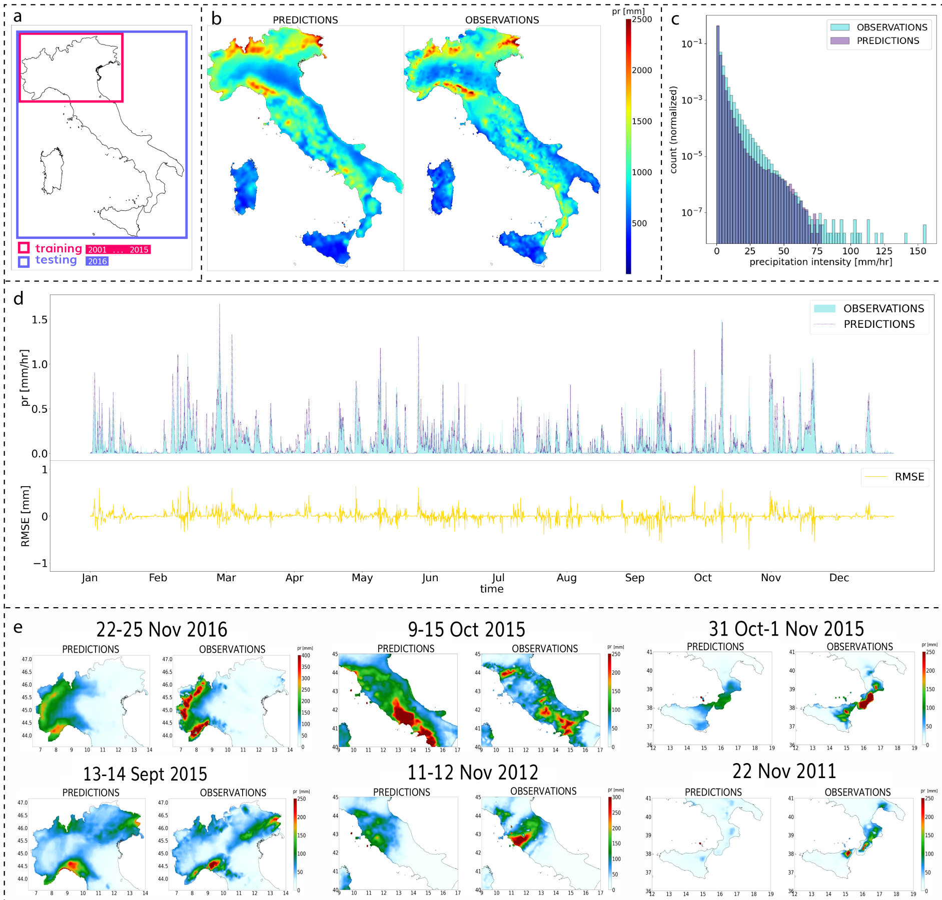

Figure 3: Results utilizing the proposed framework and comparison with observations; (a) training/testing areas, (b) cumulative precipitation, (c) histogram of precipitation intensity, (d) time series of average hourly precipitation, (e) extreme events.

4 Conclusion

The developed framework has demonstrated promising outcomes, accurately capturing yearly precipitation distribution and closely representing total precipitation during examined extreme events. Findings indicate good generalization capabilities in spatial and temporal domains, motivating further investigation into its transferability potential. Moreover, results can still be improved by refining the architecture, thanks to the flexibility provided by the framework. The subsequent phase involves incorporating diverse regions beyond the Italian territory in both training and testing. Future project stages aim to extend the framework's application to predict High Precipitation weather Events (HPEs) by utilising convection-permitting RCMs ensemble simulations within the central Mediterranean domain. This introduces a notable challenge, as the prediction phase involves the adoption of model-generated predictors for a framework trained on reanalysis data. Additionally, adaptability across various ensemble members of RCMs will be assessed, and outcomes will be compared with those of conventional dynamical downscaling methods. This thorough evaluation aims to investigate the effectiveness and reliability of the framework in simulating extreme events, a crucial aspect in addressing the challenges posed by climate change.

REFERENCES

Henry Addison, Elizabeth Kendon, Suman Ravuri, Laurence Aitchison, and Peter Watson. Machine learning emulation of a local-scale uk climate model. In *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*, 2022. URL `https://www.climatechange.ai/papers/neurips2022/21`.

Sandra Banholzer, James Kossin, and Simon Donner. *The Impact of Climate Change on Natural Disasters*, pp. 21–49. Springer Netherlands, 2014. ISBN 978-94-017-8598-3. doi: 10.1007/978-94-017-8598-3_2.

Peter W Battaglia, Jessica B Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018. doi: 10.48550/arXiv.1806.01261.

Shaked Brody, Uri Alon, and Eran Yahav. How attentive are graph attention networks? *arXiv preprint arXiv:2105.14491*, 2021.

Erika Coppola, Stefan Sobolowski, E Pichelli, F Raffaele, B Ahrens, I Anders, N Ban, Sophie Bastin, Michael Belda, D Belusic, et al. A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over europe and the mediterranean. *Climate Dynamics*, 55(1):3–34, 2020. doi: 10.1007/s00382-018-4521-8.

CORDEX. `https://cordex.org/`.

John G. Cragg. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5):829–844, 1971. ISSN 00129682, 14680262. URL `http://www.jstor.org/stable/1909582`.

Lan Cuo, Thomas C Pagano, and QJ Wang. A review of quantitative precipitation forecasts and their use in short-to medium-range streamflow forecasting. *Journal of hydrometeorology*, 12(5): 713–728, 2011.

J. J. Danielson and D. B. Gesch. Global multi-resolution terrain elevation data 2010 (gmted2010), 2011.

A. Doury, S. Somot, S. Gadat, Ribes A., and Corre . Regional climate model emulator based on deep learning: concept and first evaluation of a novel hybrid downscaling approach. *Climate Dynamics*, 60(10):1751–1779, 2023. doi: https://doi.org/10.1007/s00382-022-06343-9.

EUCP. `https://www.eucp-project.eu/`.

A. Fantini. Climate change impact on flood hazard over italy, 2019.

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

IPCC. *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Summary for Policymakers. Cambridge University Press, Cambridge, UK and New York, USA, 2022. ISBN 9781009325844.

EJ Kendon, AF Prein, CA Senior, and A Stirling. Challenges and outlook for convection-permitting climate modelling. *Philosophical Transactions of the Royal Society A*, 379(2195):20190547, 2021. doi: 10.1098/rsta.2019.0547.

Bipin Kumar, Kaustubh Atey, Bhupendra Bahadur Singh, Rajib Chattopadhyay, Nachiketa Acharya, Manmeet Singh, Ravi S Nanjundiah, and Suryachandra A Rao. On the modern deep learning approaches for precipitation downscaling. *Earth Science Informatics*, 16(2):1459–1472, 2023.

Eric M. Laflamme, Ernst Linder, and Yibin Pan. Statistical downscaling of regional climate model output to achieve projections of precipitation extremes. *Weather and Climate Extremes*, 12:15–23, 2016. ISSN 2212-0947. doi: https://doi.org/10.1016/j.wace.2015.12.001. URL `https://www.sciencedirect.com/science/article/pii/S221209471530058X`.

Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, Alexander Merose, Stephan Hoyer, George Holland, Oriol Vinyals, Jacklynn Stott, Alexander Pritzel, Shakir Mohamed, and Peter Battaglia. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, 2023. doi: 10.1126/science.adi2336.

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.

P. Jyoteeshkumar Reddy, Richard Matear, John Taylor, Marcus Thatcher, and Michael Grose. A precipitation downscaling method using a super-resolution deconvolution neural network with step orography. *Environmental Data Science*, 2:e17, 2023. doi: 10.1017/eds.2023.18.

Benjamin Sanchez-Lengeling, Emily Reif, Adam Pearce, and Alexander B. Wiltschko. A gentle introduction to graph neural networks. `https://distill.pub/2021/gnn-intro/`, 2021.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings 15*, pp. 593–607. Springer, 2018.

Matteo Turisini, Giorgio Amati, and Mirko Cestari. Leonardo: A pan-european pre-exascale supercomputer for hpc and ai applications, 2023.

Maarten K Van Aalst. The impacts of climate change on the risk of natural disasters. *Disasters*, 30 (1):5–18, 2006. doi: 10.1111/j.1467-9523.2006.00303.x.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

V. Verdecchia. Cetemps hydrological model, 2019.

Cong Wang, Ping Wang, Pingping Wang, Bing Xue, and Di Wang. A spatiotemporal attention model for severe precipitation estimation. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022. doi: 10.1109/LGRS.2021.3084293.

J. Wang, Z. Liu, I. Foster, W. Chang, R. Kettimuthu, and V. R. Kotamarthi. Fast and accurate learned multiresolution dynamical downscaling for precipitation. *Geoscientific Model Development*, 14 (10):6355–6372, 2021. doi: 10.5194/gmd-14-6355-2021.

A Appendix

A.1 Focal loss

The *Classifier* model was trained using FL as in (Lin et al., 2017). The formulation of the loss is in Equation 1. Parameters where set as $\alpha_t = 0.75$ and $\gamma = 2$.

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \tag{1}$$

Considering $y \in \{\pm 1\}$ the ground-truth class and $p \in [0.1]$ the model's estimated probability for the class with label $y = 1$, p_t is defined as:

$$p_t = \begin{cases} p & \text{if} \quad y = 1 \\ 1 - p & \text{otherwise} \end{cases} \tag{2}$$

A.2 Weighted mean square error loss

The *Regressor* model was trained using WMSE loss as in Wang et al. (2022). The formulation of the loss is in Equation 3, while the adopted weights and thresholds are in Table 2.

$$L_{WMSE}(y, \hat{y}) = \sum_{h,w} weight_{h,w} \times (y_{h,w} - \hat{y}_{h,w})^2 \tag{3}$$

Table 2: Weights and thresholds of the WMSE loss.

Weight	1	2	5	10	20	50
Threshold	0mm	1mm	5mm	10mm	20mm	50mm