REQUEST FOR AN EMI R&D PROJECT 2027–2029

Country/Organisation: Croatia

Principal Investigator[1]: Ivan Vujec (hr1m)

Affiliation: Croatian Meteorological and Hydrological Service

Address: Ravnice 48, 10000 Zagreb, Croatia

Other researchers: Iris Odak (hr1f)

Project Title: Post-processing of regional NWP forecasts

To make changes to an existing project please submit an amended version of the original form

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ………….	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2027-2029	
Would you accept support for 1 year only, if necessary?	YES ☒	NO ☐

Computer resources required for project year:		2027	2028	2029
High Performance Computing Facility	[SBU]	2M	2M	2M
Graphics Processing Unit Cluster-A	[GBU]			
Graphics Processing Unit Cluster-B	[GBU]	6M	6M	6M
Accumulated data storage (total archive volume)[2]	[GB]	5000	7000	8000

[1] The Principal Investigator will act as contact person for this EMI R&D Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.
[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.
[3]The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

EWC resources required for project year:		2027	2028	2029
Number of vCPUs	[#]	1	1	1
Total memory	[GB]	64	64	64
Storage	[GB]	5000	6000	7000
Number of vGPUs[3]	[#]	2	2	2

Continue overleaf.

Principal Investigator: Ivan Vujec (hr1m)…………………………………………

Project Title: Rain post-processing of regional NWP forecasts

Extended abstract

All EMI R&D Project requests should provide an abstract/project description including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The completed form should be submitted/uploaded at https://www.ecmwf.int/en/research/special-projects/special-project-application/special-project-request-submission.

Following submission by the relevant Member State the EMI R&D Project requests will be published on the ECMWF website and evaluated by ECMWF and its Scientific Advisory Committee. The requests are evaluated based on their scientific and technical quality, and the justification of the resources requested. Previous EMI R&D Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

Requests exceeding 5,000,000 SBU should be more detailed (3-5 pages).

Introduction

NWP models have been a cornerstone of weather prediction for several decades. Although the models are constantly improving and the forecasts are now more accurate than ever, the errors the models exhibit can still be relatively large. This is the case both for more predictable variables (e.g. 2-meter temperature or radiation) and for more spatially and temporally variable meteorological fields such as precipitation or wind speed. In addition to improvements in the models themselves, the forecast performance can be enhanced through the post-processing of NWP outputs. The main focus of this research will be the post-processing of gridded regional NWP fields, with the goal of producing more accurate and informative precipitation forecasts for short- and medium-range timescales.

Post-processing of NWP outputs is a way to further improve forecast performance. A relatively straightforward approach is to improve forecasts at locations where observations are available. For example, the analog method is a point-based method in which the current NWP predictors are compared with NWP predictors from the training dataset, and the current forecast is generated from observations from the training period. Improvements in forecast performance (e.g. Vujec and Odak Plenković 2023; Odak Plenković et al. 2020) have been demonstrated for various variables (e.g. 2-m temperature, 10-m wind speed). A more challenging approach to forecast improvement is to post-process the entire gridded forecast fields. For example, Xiang et al. (2022) used the Spatio-Temporal Transformer U-Net to perform bias correction and temporal downscaling for 2-m temperature and the 10-m wind component, while Liu et al. (2024) used a Swin-Transformer based deep learning model to improve precipitation forecasts. However, the input NWP models used in such studies are often global (e.g. Price and Rasp 2022), while many studies focus on ensemble forecasts (e.g. Hess and Boers 2022; Pic et al. 2025).
This research will take a different approach. To make the problem tractable within the available computational and data resources, deterministic regional NWP forecasts on a reduced domain will be post-processed using vision-transformer-based models and available observational and analysis datasets. Although the primary focus of this research is NWP post-processing, the proposed framework is also compatible with observation-based approaches. Such an approach is expected to be particularly suitable for meteorological variables with relatively smooth spatial structure (e.g. surface solar radiation), before being applied to more intermittent fields such as precipitation.

Datasets

The current operational Croatian NWP model is an ALADIN model configuration, which is based on the non-hydrostatic dynamical core with 87 vertical levels and a grid spacing of 2 km, with the lowest level at 10 m. Currently, the sensitivity tests are being performed, and a 1.3 km grid-spacing configuration is being evaluated and planned to become operational in 2027.

Observational data are available from two main sources. INCA (Integrated Nowcasting through Comprehensive Analysis) is a high-resolution meteorological analysis and nowcasting system that has recently been acquired by DHMZ and is currently being tested. The system generates hourly and sub-hourly precipitation field analyses, which can be used as gridded reference precipitation data. Additionally, through the METMONIC project, DHMZ has obtained three new and three upgraded radar stations. In addition to improved products, previously uncovered parts of Croatia are now covered by the radar network, and the transferability of the machine learning models can also be investigated. Satellite-derived observations could also be incorporated in future as an additional source of observational data for machine learning applications.

Scientific plan and computational resources

Since heavy rainfall can have severe impacts causing substantial economic losses and risks to human safety, the improvement of operational precipitation forecasts will be performed using vision-transformer-based models applied to gridded regional ALADIN fields. While quantitative precipitation forecasting is the primary objective of this project, the proposed framework is intentionally designed to be sufficiently general to accommodate other gridded meteorological variables and future observation-based or hybrid modelling approaches.
Taking into account the computational/GPU constraints, the study will be limited to a smaller subdomain of the full operational ALADIN domain. In this way, the transferability of the ML model can be examined on another part of the operational domain. The transferability of the ML models is particularly important in a situation of limited computational resources. In this way, a model trained on one part of the domain could be used on the whole domain, and other meteorological services could potentially apply the same ML model on their domain, reducing the need for excessive model training. Also, the model trained on one radar dataset could be verified on the dataset from the other radar station. This could potentially be an indicator of whether or not the single-station data are sufficient for the training, or the radar composite is necessary for the algorithm to work. Quantitative precipitation forecast is the primary variable of interest, while the probability of precipitation would be a logical next step, and will be investigated if sufficient resources are available. This project will generally focus on smaller datasets and reduced number of variables and degrees of freedom, putting more focus on accessibility and ease of use, than on the marginal improvements achieved by larger models.

Since this is the initial stage of the experiment, the required computational resources are difficult to estimate precisely. Based on discussions with a colleague working on a project involving Vision Transformers and taking into account the computational requirements of that project, 6 million GBU units were requested, together with two vGPUs on EWC. Since the disk-space requirements may vary due to the need to create custom datasets for machine-learning training, 5-8 TB of data storage was requested. If fewer resources are available for this project, the analysis will be scaled accordingly.

References

Hess, P., & Boers, N. (2022). Deep learning for Improving Numerical Weather Prediction of Heavy Rainfall. Journal of Advances in Modeling Earth Systems, 14, e2021MS002765. https://doi.org/10.1029/2021MS002765

Liu, H., Fung, J. C. H., Lau, A. K. H., & Li, Z. (2024). Enhancing Quantitative Precipitation Estimation of NWP Model With Fundamental Meteorological Variables and Transformer Based Deep Learning Model. Earth and Space Science, 11, e2023EA003234. https://doi.org/10.1029/2023EA003234

Odak Plenković, I., Schicker, I., Dabernig, M., Horvath, K., & Keresturi, E. (2020). Analog-based post-processing of the ALADIN-LAEF ensemble predictions in complex terrain. Q J R Meteorol Soc 146:1842–1860. https://doi.org/10.1002/qj.3769

Pic, R., Dombry, C., Naveau, P., & Taillardat, M., (2025). Distributional Regression U-Nets for the Postprocessing of Precipitation Ensemble Forecasts. Artif. Intell. Earth Syst., 4, 240067, https://doi.org/10.1175/AIES-D-24-0067.1.

Price, I., & Rasp, S. (2022). Increasing the accuracy and resolution of precipitation forecasts using deep generative models. In *International conference on artificial intelligence and statistics* (pp. 10555-10571). PMLR. https://doi.org/10.48550/arXiv.2203.12297

Vujec, I., & Odak Plenković, I. (2023). Kalman filter sensitivity tests for the NWP and analog-based forecasts post-processing. Meteorol Atmos Phys 135, 1. https://doi.org/10.1007/s00703-022-00939-w

Xiang, L., Guan, J., Xiang, J., Zhang, L., & Zhang, F. (2022). Spatiotemporal model based on transformer for bias correction and temporal downscaling of forecasts. Front. Environ. Sci. 10:1039764. https://doi.org/10.3389/fenvs.2022.1039764