REQUEST FOR A SPECIAL PROJECT 2024–2026

MEMBER STATE:	Netherlands
Principal Investigator[1]:	Dr. M.J. Schmeits
Affiliation:	KNMI
Address:	P.O. Box 201, 3730 AE De Bilt
Other researchers:	Dr. K. Whan (KNMI) Dr. Y. Shapovalova (deep-learning expert, Radboud University Nijmegen)
Project Title:	High-resolution data-driven weather forecasts for Western Europe

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.)	2024	
Would you accept support for 1 year only, if necessary?	YES ■	NO

Computer resources required for project year:		**2024**	**2025**	**2026**
High Performance Computing Facility	[SBU]			
Accumulated data storage (total archive volume)[2]	[GB]			

EWC resources required for project year:		**2024**	**2025**	**2026**
Number of vCPUs	[#]	35	35	35
Total memory	[GB]	350	350	350
Storage (accumulated)	[GB]	10000	20000	25000
Number of vGPUs[3]	[#]	2	2	2

Continue overleaf.

[1] The Principal Investigator will act as contact person for this Special 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.

Principal Investigator: Dr. M.J. Schmeits

Project Title: High-resolution data-driven weather forecasts for Western Europe

Extended abstract

All Special 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 Special 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 Special 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

Exceptionally fast progress in the development of data-driven weather forecasting models over recent years has caught many weather institutes by surprise. Tech companies (e.g., Google, Huawei and NVIDIA) have been leading the way in this field, demonstrating that deep learning models (typically graph neural networks or transformers) can produce deterministic forecasts that are comparable in skill to the IFS HIRES for a number of variables (Keisler, 2022, Pathak et al., 2022, Bi et al., 2022, Lam et al., 2022). Recently, Chen et al. (2023) have demonstrated skill comparable to the IFS ensemble mean. The work so far on data-driven forecasts with deep learning has focused solely on global models from ERA5 reanalysis.

Scientific plan

In this project, we plan to train a deep learning model over Western Europe, with a focus on the Netherlands, using the 10-year DOWA reanalysis (https://www.dutchoffshorewindatlas.nl/about-the-atlas/dowa-data/data-info). The DOWA reanalysis (Dutch Offshore Wind Atlas, Wijnant et al., 2019) is a limited area reanalysis product that is made using the high-resolution (2.5 km) non-hydrostatic NWP model Harmonie-Arome. To our knowledge, this will be the first high-resolution limited-area data-driven forecasting project. We will focus on predicting surface variables that are most important to national meteorological institutes (like the Royal Netherlands Meteorological Institute (KNMI)) for general forecasts and warnings, such as wind speed/gusts, precipitation, air pressure, and temperature. The project contains several steps. We plan to first develop data-driven (preferably ensemble/probabilistic) forecasts out to +48 hours using deep learning methods. Given the focus on warnings, we will extend previous work by exploring how the skill for extremes can improve with different loss functions (i.e., balanced or weighted loss). We further plan to compare these forecasts with the operational Harmonie-Arome (ensemble) forecasts that are used at KNMI. Finally, we plan to compare the data-driven (ensemble) forecasts with post-processed (probabilistic) forecasts from the operational Harmonie-Arome model.

We have existing collaborations with deep learning experts at universities in the Netherlands, and we are seeking collaboration with other meteorological institutes for this project. We also plan to acquire additional funding for this project.

Justification of computer resources

Earlier studies based on ERA5 data (Keisler, 2022, Pathak et al., 2022, Bi et al., 2022, Lam et al., 2022, Chen et al., 2023) have used between 1 and 192 GPUs, so we think (at least) 2 vGPUs would be needed for this project. The other numbers in the Table above are set to the maximum as well, because this project needs as much computing power as possible. Additionally, a large storage capacity is needed due to the large size of the data sets.

Technical characteristics of the code

Our code will use standard libraries for deep learning (such as PyTorch or Tensorflow) and will be built using (elements of) architecture(s) of previous studies (e.g. Keisler, 2022, Pathak et al., 2022, Bi et al., 2022, Lam et al., 2022, Chen et al., 2023).

References

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X. and Tian, Q., 2022. Pangu-Weather: A 3D High-Resolution Model for Fast and Accurate Global Weather Forecast. _arXiv preprint arXiv:2211.02556_.

Chen, L., Zhong, X., Zhang, F., Cheng, Y., Xu, Y., Qi, Y. and Li, H., 2023. FuXi: A cascade machine learning forecasting system for 15-day global weather forecast. _arXiv preprint arXiv:2306.12873_.

Keisler, R., 2022. Forecasting global weather with graph neural networks. _arXiv preprint arXiv:2202.07575_.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., Ravuri, S., Ewalds, T., Alet, F., Eaton-Rosen, Z. and Hu, W., 2022. GraphCast: Learning skillful medium-range global weather forecasting. _arXiv preprint arXiv:2212.12794_.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K. and Hassanzadeh, P., 2022. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. _arXiv preprint arXiv:2202.11214_.

Wijnant, I., van Ulft, B., van Stratum, B., Barkmeijer, J., Onvlee, J., de Valk, C., Knoop, S., Kok, S., Marseille, G., Baltink, H.K. and Stepek, A., 2019. The Dutch Offshore Wind Atlas (DOWA): Description of the dataset. _KNMI Tech. Rep._