# REQUEST FOR AN EMI R&D PROJECT 2027–2029

| | |
|---|---|
| **MEMBER STATE:** | Türkiye |
| **Principal Investigator[1]:** | Ahmet Mümin Üyüklü |
| **Affiliation:** | Climate Science and Meteorological Engineering Department, Istanbul Technical University |
| **Address:** | İTÜ Ayazağa Campus Reşitpaşa, 34485 Sarıyer/İstanbul, Türkiye |
| **Other researchers:** | Ahmet Mümin Üyüklü, Umur Dinç, Prof. Dr. Hüseyin Toros |
| **Project Title:** | Machine Learning Downscaling of Precipitation Forecasts for the Marmara Region Using the Anemoi Framework |

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 2028 | |
| 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] | 60,000,000 | 8,000,000 | - |
| Graphics Processing Unit Cluster-A | [GBU] | - | - | - |
| Graphics Processing Unit Cluster-B | [GBU] | 4,320,000 | 2,880,000 | - |
| Accumulated data storage (total archive volume)[2] | [GB] | 12,000 | 15,000 | - |

| **EWC resources required for project year:** | | **2027** | **2028** | **2029** |
|---|---|---|---|---|
| Number of vCPUs | [#] | - | - | - |
| Total memory | [GB] | - | - | - |
| Storage | [GB] | - | - | - |
| Number of vGPUs[3] | [#] | - | - | - |

*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:** Ahmet Mümin Üyüklü

**Project Title:** Machine Learning Downscaling of Precipitation Forecasts for the Marmara Region Using the Anemoi Framework

# Extended abstract

The Marmara Region, home to around 25 million people, is one of the most densely populated and economically active areas in Türkiye. Accurate precipitation forecasts at high spatial resolution are critical for water resource management, urban planning, and disaster management across the region. Precipitation forecasting over the region is difficult. The Sea of Marmara, the surrounding terrain, and the convergence of Mediterranean and continental air masses produce precipitation patterns at scales below what global models can represent. AIFS operates on an approximately 0.25° grid over this domain, meaning that the orographic channelling, coastal convergence, and convective triggering that drive high-impact rainfall events largely occur at sub-grid scales.

Convection-permitting models running at 3 km resolve these features but are expensive to run at medium range. A machine learning downscaling model trained on historical paired data can produce the same spatial detail at a fraction of the cost. Once trained, it is expected to produce high-resolution fields at substantially lower computational cost than convection-permitting WRF simulations. This project builds such a model for the Marmara Region using ECMWF's Anemoi framework, with WRF-ARW hindcasts driven by ERA5 as the high-resolution training reference and station observations from Turkish State Meteorological Service (TSMS) for post-processing and verification.

This work follows from Nipen et al. (2024), who showed that an Anemoi-based model with a stretched-grid architecture can match or outperform an operational NWP system for several forecast parameters over the Nordic region. Our approach is different: rather than modifying the global AIFS architecture, we train a dedicated encoder-decoder downscaling model on its output, using WRF as a physically consistent high-resolution training reference. No comparable system currently exists for the Marmara Region or Türkiye, and the use of Anemoi as the common framework connecting all steps is new. The methodology can be extended to the full Türkiye domain in a follow-on project.

**Approach**

A 10-year WRF-ARW hindcast (2015–2024) at 3 km over the Marmara Region will be produced using ERA5 as boundary conditions, updated every 6 hours. The simulation will use two nested domains: a 9 km outer domain covering Türkiye and the surrounding seas, and a 3 km inner domain centred on Marmara. Cumulus parametrization will be disabled at 3 km. Before the full run, a 3-month sensitivity experiment (October–December 2019) will test up to six physics combinations, varying microphysics and boundary layer scheme, to identify the best-performing configuration for the region. The WRF simulations will be run in monthly continuous segments, each preceded by a one-day spin-up period whose output will be excluded from the training dataset. Only the surface variables needed for model training will be written to disk: 6-hourly accumulated precipitation, 2 m temperature, 2 m dewpoint temperature, and 10 m wind speed and direction. Pressure-level fields will not be stored as these are provided by AIFS as model input. The period 2015–2023 will be used for training, 2024 for validation, and 2025 for independent testing.

AIFS forecasts will be generated using anemoi-inference with ERA5 initial conditions once daily at 00 UTC over the 2015–2025 period, with forecast lead times from +6 to +240 hours at six-hour intervals. AIFS forecasts will be generated globally, after which the fields covering Türkiye and the surrounding seas will be extracted as regional inputs to the downscaling model. The model will operate as a single-step spatial downscaling system: for each timestep, AIFS fields at approximately 0.25° resolution will serve as input and the corresponding WRF 3 km fields will serve as the training target. Forecast lead time will be included as an input feature so that the model can distinguish between different forecast horizons. Since AIFS already encodes global atmospheric dynamics in its output fields, the input domain will be restricted to a regional extent covering Türkiye and surrounding seas, providing sufficient spatial context to capture the large-scale circulation patterns that influence Marmara precipitation. Selected AIFS pressure-level variables, including geopotential, wind components, specific humidity, and temperature, together with selected surface variables such as 2 m temperature, 10 m wind, and mean sea-level pressure, will be used as input. The paired AIFS–WRF dataset for 2015–2024 will be prepared using anemoi-datasets. AIFS forecasts for 2025 will be used only for independent evaluation against TSMS station observations. An encoder-decoder graph will be constructed with anemoi-graphs: input data nodes will correspond to the AIFS grid points over the regional domain and output data nodes will correspond to the WRF 3 km grid points over the Marmara domain. Encoder edges will connect the AIFS data nodes to a hidden mesh; decoder edges will connect the hidden mesh to the WRF 3 km output nodes using KNN connections. Static terrain fields (elevation, slope, and land-sea mask from a 90 m DEM) will be added as node attributes on the output grid. The model will be built with anemoi-models and trained with anemoi-training for a single-step spatial downscaling task. The training objective will cover the selected surface target variables at 3 km, with higher loss weights assigned to 6-hourly accumulated precipitation. GPU memory use and training throughput will be assessed during the initial benchmark experiments. Anemoi training will initially be tested on a single GPU, while selected full-scale experiments may use two GPUs through distributed training or model sharding if justified by the measured performance and memory requirements.

The trained model will be applied to the historical period, and its output will be compared with TSMS station observations using temporally separated subsets for model fitting and evaluation to ensure independent validation. A lightweight post-processing model will learn the difference between model output and observed values as a function of forecast amount, season, and station location. This will correct systematic errors that remain after the WRF-based downscaling, including location- and season-dependent biases. Keeping this step separate from the main model will make it straightforward to update when new observations become available, and will allow the contribution of each component to be measured independently during evaluation.

**Evaluation**

The full pipeline will be evaluated on the independent 2025 test period. Since the WRF hindcast covers 2015–2024, the 2025 evaluation relies entirely on TSMS station observations as ground truth, none of which were used in any training step. The primary evaluation will focus on precipitation: RMSE, bias, Equitable Threat Score at several thresholds, and quantile-quantile plots to examine heavy events. Results will be reported for three configurations: AIFS only, AIFS with downscaling, and AIFS with downscaling and post-processing, so that the contribution of each step is clear. Temperature and wind speed will be reported as supplementary verification. Several high-impact case studies will also be included.

**Work Plan**

The first year covers data production and model training. ERA5 data for the Marmara domain will be extracted from the MARS archive and the WRF physics sensitivity experiment will be run first. Once the best physics configuration is identified, the full 10-year hindcast will be produced. In parallel, AIFS will be run over 2015–2024 to build the paired training and validation dataset, while AIFS forecasts for 2025 will be generated for independent evaluation against TSMS station observations. The Anemoi downscaling model will be trained in the second half of the first project year.

The second year begins with an assessment of the Year 1 model results. Based on the verification scores, targeted adjustments will be made, for example, changing the loss weighting for heavy precipitation or testing additional input variables. The station post-processing model will then be trained and the complete pipeline will be evaluated on the 2025 test period. The manuscript will be written in parallel and submitted before the project ends.

**Resource Justification**

The main computational workload will be carried out on the ECMWF HPCF. The CPU estimate is based on a ten-year nested WRF-ARW hindcast using a 9 km outer domain and a 3 km Marmara inner domain, together with up to six three-month physics sensitivity experiments. Including preprocessing, spin-up, failed or repeated simulations, and performance uncertainty, 60 million SBU are requested for 2027. A further 8 million SBU are requested for 2028 for data processing, verification, targeted reruns, and final evaluation.

GPU Cluster-B resources are requested for AIFS inference and Anemoi model development and training. The estimate assumes one AIFS forecast per day from the 00 UTC initial condition, with lead times from +6 to +240 hours at six-hour intervals over 2015–2025. The allocation also covers preliminary benchmarks, model-development experiments, several full training runs, final inference, and repeated experiments. As no project-specific GPU benchmarks are currently available, the request is based on a planning estimate of 1,200 total GPU-hours in 2027 and 800 total GPU-hours in 2028, corresponding to 4.32 million and 2.88 million GBU-B, respectively. Initial benchmarks will be used to refine the training configuration and determine the appropriate number of GPUs for individual jobs.

The accumulated archive requirement is estimated at 12 TB by the end of 2027 and 15 TB by the end of 2028. This covers the regional ERA5 subset, AIFS output, WRF surface output, paired training data, model checkpoints, evaluation products, and intermediate datasets. As no benchmark runs have yet been performed on ECMWF systems, these values are planning estimates; initial WRF and Anemoi benchmark runs will be used to verify the assumed throughput before full production.

**Software**

The project uses the Anemoi framework throughout. AIFS is run with anemoi-inference using publicly released ECMWF checkpoints. Training data is prepared in zarr format with anemoi-datasets. The encoder-decoder graph is built with anemoi-graphs. The downscaling model uses anemoi-models for the GNN architecture and anemoi-training for the training loop. WRF-ARW 4.x will be run on the ECMWF HPCF, and ERA5 data will be processed through WPS. Data handling will use xarray, zarr, and cfgrib. All configuration files, training recipes, post-processing scripts, and evaluation code will be made publicly available on GitHub to support reproducibility and reuse.

**Expected Outcomes**

- A precipitation downscaling pipeline for the Marmara Region built with the Anemoi framework. All configuration files, training recipes, post-processing scripts, and evaluation code will be made publicly available on GitHub to support reproducibility and reuse.
- A 10-year, 3 km WRF hindcast archive over Marmara (2015–2024), available for future research.
- A peer-reviewed publication in Weather and Forecasting or Artificial Intelligence for the Earth Systems.
- A documented codebase ready to be scaled to the full Türkiye domain.

**References**

Lang, S., et al. (2024). AIFS – ECMWF's data-driven forecasting system. arXiv:2406.01465.

Nipen, T. N., et al. (2024). Regional data-driven weather modeling with a global stretched-grid. arXiv:2409.02891.

Skamarock, W. C., et al. (2019). A description of the Advanced Research WRF Model version 4. NCAR/TN-556+STR.

Özen, C., Dinç, U., Deniz, A., & Karan, H. (2020). Wind power generation forecast by coupling numerical weather prediction model and gradient boosting machines in Yahyalı wind power plant. Wind Engineering. https://doi.org/10.1177/0309524X20972115

Özen, C., Kaplan, O., Özcan, C., & Dinç, U. (2019). Short Term Wind Speed Forecast By Using Long Short Term Memory. 9th International Symposium on Atmospheric Sciences (ATMOS 2019), İstanbul, Türkiye.

Ünal, Z. F., Dinç, U., Özen, C., & Toros, H. (2019). Air Pollution Forecasting for Ankara with Machine Learning Method. 9th International Symposium on Atmospheric Sciences (ATMOS 2019), İstanbul, Türkiye.

Baño-Medina, J., et al. (2025). A regional high resolution AI weather model for the prediction of atmospheric rivers and extreme precipitation. npj Climate and Atmospheric Science, 8, 385.