REQUEST FOR AN EMI R&D PROJECT 2027–2029

Country/Organisation: Türkiye ………………………………………………………………………………………..

Principal Investigator[1]: Prof.Dr. Yurdanur Ünal ……………………………………………………………….

Affiliation: Istanbul Technical University (ITU) ………………………………………………………

Address: İTÜ Ayazağa Campus, Rectorate Building, 34467 Maslak-ISTANBUL …

Other researchers:
M.Sc. Ayşegül Ceren Moral (Boğaziçi University, BOUN)
Gökberk Ozan Tiryaki (Istanbul Technical University, ITU)
Gökçenaz Önel (Istanbul Technical University, ITU)
Hilal Barut (Istanbul Technical University, ITU)

Project Title: Convection-Permitting Regional Climate Modelling over Türkiye Using the ICON Model: Sensitivity to Domain Configuration, Resolution, and Physical Parameterizations

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	
Would you accept support for 1 year only, if necessary?	YES X	NO ☐

Computer resources required for project year:		**2027**	**2028**	**2029**
High Performance Computing Facility	[SBU]	3,000,000	3,000,000	2,000,000
Graphics Processing Unit Cluster-A	[GBU]			
Graphics Processing Unit Cluster-B	[GBU]			
Accumulated data storage (total archive volume)[2]	[GB]	30,000	30,000	30,000

[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	[#]			
Total memory	[GB]			
Storage	[GB]			
Number of vGPUs[3]	[#]			

Continue overleaf.

Principal Investigator: Prof.Dr. Yurdanur Ünal ………………………………………………………

Project Title: Convection-Permitting Regional Climate Modelling over Türkiye Using the ICON Model: Sensitivity to Domain Configuration, Resolution, and Physical Parameterizations……………………………

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).

Scientific Plan

As the regional impacts of global warming become increasingly pronounced, Türkiye is undergoing significant transformations in both thermal and hydrological regimes. Its complex topography and location at the intersection of multiple climate systems mean that precipitation and temperature patterns cannot be adequately captured by coarse-resolution models, which systematically fail to resolve terrain-driven gradients and mesoscale convective processes. To bridge this gap, the present project applies the ICOsahedral Nonhydrostatic (ICON) regional climate model at convection-permitting resolutions across Türkiye, examining model performance and the influence of physical parameterization choices on simulated hydroclimatic patterns. Because the computational cost of ICON scales strongly with horizontal resolution, domain size, and simulation length, conducting multi-decadal integrations for each parameterization choice is typically prohibitive. To balance computational efficiency with climatic relevance, three simulation years are selected to represent distinct climatic conditions. First, a representative baseline year is identified by examining 64 years of meteorological data (1960–2024) from 221 stations to find the year most closely aligned with long-term regional averages. Based on this analysis, 2015 was selected, with a mean annual temperature of 13.9 °C (long-term mean: 13.3 °C) and mean annual precipitation of 614.58 mm (long-term mean: 615.46 mm), confirming its suitability as a climatologically representative baseline. Building on this foundation, two anomalous years are further incorporated to probe model robustness beyond mean conditions: 1992, identified as a cold and wet extreme, and 2024, identified as a warm and dry extreme. Together, these three years provide a comprehensive framework for assessing ICON's capacity to reproduce both mean climatological conditions and the full range of interannual variability characteristic of Türkiye.

To provide realistic large-scale atmospheric forcing, initial and boundary conditions are prescribed from ERA5 reanalysis data, which offers high-resolution, temporally consistent fields and has been widely demonstrated to capture synoptic variability over Türkiye. The experimental framework comprises 15 simulations totalling 35 simulation years, structured around four sensitivity dimensions:

1. **Domain size sensitivity** — D_SML, D_MID, and D_LRG are run with identical physical parameterizations to isolate the effect of lateral boundary placement on the interior climate solution. D_MID serves as the control run with default namelist settings.
2. **Dynamical downscaling via ICON-to-ICON nesting (I2I)** — I2I3KM and D_MID3KM employ a two-step nesting strategy in which a 10 km parent domain drives a 3 km convection-permitting inner domain, quantifying the added value of explicit convection representation.
3. **Convection-permitting resolution** — The 3 km configurations are compared against their 10 km counterparts to assess improvements in the representation of precipitation intensity, diurnal cycle, and orographic enhancement over Türkiye's complex terrain.
4. **Physical parameterization sensitivity (D_MED)** — Ten one-at-a-time experiments systematically activate or deactivate individual physical schemes, each run across all three simulation years (2015, 1992, 2024) to assess sensitivity under both average and extreme climate conditions.

The D_MED parameterization experiments target four scheme categories, summarised in Table 1. Each experiment modifies a single namelist parameter relative to the D_MID control, enabling unambiguous attribution of differences in simulated hydroclimate to individual physical processes. The three-year simulation design ensures that parameter sensitivity is quantified not only under mean climatological conditions (2015) but also under a cold and wet extreme (1992) and a warm and dry extreme (2024), providing a robust assessment of scheme transferability across contrasting climate regimes.

Table 1. D_MED parameterization experiments. Each run modifies a single physical scheme relative to the D_MID control and is repeated across all three simulation years.

#	Scheme category	Configuration	Description
1	Shallow convection	shallow_conv ON	Shallow convection parameterization activated
2	Shallow convection	shallow_conv OFF	Shallow convection parameterization deactivated
3	Radiation	ECRAD	ECMWF radiation scheme (ecRad)
4	Radiation	RRTM	Rapid Radiative Transfer Model (RRTM)
5	Cloud microphysics	3-cat ice	Three-category ice: cloud ice, snow, graupel
6	Cloud microphysics	Two-moment (Seifert)	Two-moment bulk microphysics scheme (Seifert & Beheng)
7	Cloud cover	Köhler	diagnostic cloud cover (Koehler)
8	Cloud cover	turbdiff	clouds as in turbulence (turbdiff)
9	Turbulence	COSMO diffusion and transfer	COSMO-derived diffusion and surface transfer
10	Turbulence	Classical Smagorinsky	Classical Smagorinsky diffusion scheme

Simulation Design

The full simulation matrix is summarised in Table 2. The five configuration runs (D_SML, D_MID, D_LRG, I2I3KM, D_MID3KM) each cover 2015 only. The ten D_MED parameterization runs each cover 2015, 1992, and 2024, yielding 30 simulation years for the sensitivity analysis. The total across all experiments is 35 simulation years.

Table 2. Full simulation matrix. "N Runs" denotes the number of individual model integrations per configuration group.

Run ID	Domain	Resolution	Nesting	Sim. Years	N Runs	Primary Purpose
D_SML	Small domain	10 km	ERA5 → ICON	2015	1	Domain size sensitivity (small)
D_MID	Medium domain	10 km	ERA5 → ICON	2015	1	Control run (default settings)
D_LRG	Large domain	10 km	ERA5 → ICON	2015	1	Domain size sensitivity (large)
I2I3KM	Large → nested	10/3 km	ERA5 → ICON → ICON	2015	1	ICON-to-ICON nesting; CPM impact
D_MID3KM	Medium → nested	10/3 km	ERA5 → ICON → ICON	2015	1	CPM resolution over complex terrain
D_MED	Medium domain	10 km	ERA5 → ICON	2015, 1992, 2024	10 × 3	Physical parameterization sensitivity across climate regimes

The experimental design follows a two-stage approach: in the first stage, default namelist settings are applied to isolate the impact of spatial configuration; in the second stage, the optimal domain configuration is applied across all parameterization experiments and extended to the anomalous years. To assess out-of-sample performance, the selected configurations were additionally tested under these anomalous years, probing model transferability beyond the climatologically average reference year.

Simulation outputs from all configurations are assessed against observations from 221 meteorological stations obtained from the General Directorate of Meteorology (MGM). Model performance is quantified at monthly and seasonal scales using root mean square error (RMSE), mean bias, Pearson correlation coefficient, and the index of agreement, providing a comprehensive picture of systematic errors and the consistency of simulated spatial patterns and temporal evolution across seasons.

Justification of Computer Resources Requested

The computational requirements are determined by the number of active configurations, simulation years, domain sizes, and horizontal resolutions. The project comprises 35 simulation years in total: 5 single-year configuration runs and 30 simulation years for the 10 D_MED parameterization experiments (each run over 3 years). A summary is provided in Table 3; precise SBU estimates will be inserted following benchmark test runs on HPC and SBU calculator.

Table 3. Resource summary by configuration group. SBU values to be completed after benchmark runs.

Configuration group	Domains / nesting	Simulation years	SBU estimate
Domain size sensitivity (D_SML, D_MID, D_LRG)	3 × 10 km	3 runs × 1 year	1.000.000
I2I nesting (I2I3KM, D_MID3KM)	2 × (10+3 km)	2 runs × 1 year	1.000.000
Parameterization sensitivity (D_MED)	10 km, 1 domain	10 runs × 3 years	6.000.000
Total (35 simulation years)			**8.000.000 SBU**

Grid cell counts for the 10 km and 3 km domains vary considerably with domain extent; exact values will be provided by the PI following domain definition. Storage estimates follow a similar structure to comparable ECMWF Special Projects, scaled to the number of output variables and temporal sampling frequency adopted in this study.

Storage requirements are estimated based on output from the D_MID control run, which produces approximately 10 TB of raw model output per simulation year when all diagnostic variables are retained. To reduce storage demands, the output strategy will be restricted to surface and near-surface variables relevant to hydroclimatic evaluation, excluding three-dimensional pressure-level fields and intermediate diagnostic outputs. This selective approach is expected to reduce per-run storage to approximately 3–4 TB. Additionally, spin-up periods will be deleted upon completion of each simulation. Where institutional storage capacity permits, completed simulation output will be transferred to local servers, further reducing the accumulated footprint on ECMWF archiving infrastructure. Accordingly, the storage estimates presented in Table 4 reflect this reduced per-run volume rather than the full raw output size.

Technical Characteristics of the Code

The atmospheric model ICON-CLM (https://code.mpimet.mpg.de/projects/iconpublic), including the land surface scheme Terra, is run within the SPICE framework (https://hereon-coast.atlassian.net/wiki/spaces/SPICE/overview), with ERA5 reanalysis fields providing initial and lateral boundary conditions.

ICON-to-ICON nesting (I2I configurations) uses the standard online nesting capability of ICON-CLM, with the 3 km inner domains run at convection-permitting resolution without a deep convection parameterization scheme.

The D_MED parameterization experiments follow a one-at-a-time namelist modification approach, with all other settings held constant relative to the D_MID control (see Table 4).

Model output is verified against 221 MGM stations using RMSE, mean bias, Pearson r, and index of agreement at monthly and seasonal timescales.

Table 4. Technical components of the modelling system.

Component	Version / Source
Atmospheric model	ICON-CLM (latest stable release, CLM community)
Land surface scheme	Terra (Schulz & Vogel, 2020)

Run framework	SPICE v2.1 (https://hereon-coast.atlassian.net/)
Boundary conditions	ERA5 reanalysis (ECMWF)
Radiation schemes	ecRad; RRTM
Cloud microphysics schemes	3-category ice; Two-moment Seifert & Beheng; Köhler
Clous cover	Köhler; turbdiff
Turbulence schemes	COSMO diffusion and transfer; Classical Smagorinsky
Verification dataset	MGM 221-station network (daily and hourly T, P)