

REQUEST FOR AN EMI R&D PROJECT 2027–2029

MEMBER STATE: SPAIN

Principal Investigator¹: CARLOS CALVO-SANCHO, MARCOS MARTÍNEZ-ROIG

Affiliation: Centro de Investigaciones sobre Desertificación, Consejo Superior de Investigaciones Científicas (CIDE, CSIC-UV-GVA)

Address: Ctra. CV-315 km 10,5 46113 Moncada (Valencia)

Other researchers: Cesar Azorin-Molina, Angela Manrique, Nuria P. Plaza, Kevin Monsalvez-Pozo, Shalenys Bedoya-Valestt, Miguel Andres-Martín, Jose Gomez-Reyes, Francisco Granell-Haro, Hairong Li, Mercedes Poggi, Antonio Paulo Ferreira, Cheng Shen, Andres Barrio-Martin, Amar Halifa-Marín

Project Title: Severe Convective Storms in the EuroMediterranean region

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.	SPESSANC	
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 <input checked="" type="checkbox"/>	NO <input type="checkbox"/>

Computer resources required for project year:	2027	2028	2029
High Performance Computing Facility [SBU]	60 000 000	60 000 000	60 000 000
Graphics Processing Unit Cluster-A [GBU]	1 000 000	1 000 000	1 000 000
Graphics Processing Unit Cluster-B [GBU]	1 000 000	1 000 000	1 000 000
Accumulated data storage (total archive volume) ² [GB]	100 000	100 000	100 000

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.

³The 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:

CARLOS CALVO-SANCHO, MARCOS MARTÍNEZ-ROIG

Project Title:

Severe Convective Storms in the EuroMediterranean region

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).

Severe convective storms (SCS) represent one of the most damaging types of extreme weather events, frequently producing downbursts, extreme rainfall, and large hail, which pose major risks to society and infrastructure. Downbursts are strong downdrafts within thunderstorms that rapidly sink towards the ground. Downbursts can generate destructive straight-line winds which are commonly confused with tornadoes as can reach their force and produce similar damages (Fujita, 1978; Wakimoto, 1985). These events are particularly hazardous to aviation and urban areas, where sudden wind shifts can cause severe accidents and structural damage. Extreme rainfall associated with convective systems has intensified in recent decades, with observational evidence showing increases in both the frequency and magnitude of short-duration, high-intensity precipitation events across many regions (IPCC, 2021). Such rainfall not only overwhelms urban drainage systems but also leads to flash floods that result in high mortality rates and extensive economic losses (Prein et al., 2017). Similarly, large hailstorms (≥ 5 cm in diameter) produce significant damage to crops, vehicles, and infrastructure, with recent studies highlighting their increasing frequency in a warming climate (Brooks, 2013; Púčik et al., 2019; Piasecki et al., 2023; Martín et al., 2024). Although numerous downbursts organized in clusters dominated, all these kinds of severe surface impacts related to SCS can occur simultaneously; when they do, the phenomenon is called a derecho.

From a climatological perspective, the occurrence of downbursts, extreme rainfall, and large hail is closely linked to mesoscale convective systems and changing atmospheric thermodynamics. Warming temperatures contribute to higher atmospheric moisture content, which intensifies convective instability and increases the frequency and intensity of SCS involving in extreme rainfall and hail (Trenberth, 2011; Ashley et al., 2023; Battaglioli et al., 2023). Moreover, studies suggest that shifts in vertical wind shear and convective available potential energy (CAPE) under climate change scenarios are likely to enhance the environments conducive to severe hail formation and damaging downdrafts (Brooks et al., 2019; Taszarek et al. 2021). This has significant implications for disaster risk management, as urbanization and population growth in vulnerable areas amplify the societal impacts of these hazards. Thus, advancing high-resolution climate modeling, improving early warning systems, and strengthening infrastructure resilience are critical strategies to mitigate the risks posed by these extreme convective weather phenomena (Prein & Holland, 2018). Moreover, climate change impacts in the sea surface temperature (SST), increasing the likelihood and intensity of marine heatwaves, particularly in the Mediterranean Sea. This can lead in an enhancing of SCS which contributes to significant high-impact weather events, such as the Derecho in the Western Mediterranean in August 2022 (González-Alemán et al. 2023), the giant hail events in Spain and Italy (Martín et al., 2024; De Martín et al. 2025) or the extreme rainfall in Valencia in October 2024 that caused a deadly flash flood (Calvo-Sancho et al., 2026).

Since just a few works have studied downbursts and hail (particularly in the Euro Mediterranean region), numerous scientific questions remain open. In a warming climate, due to the fossil-fuel caused climate change, there is a growing risk of passing the tipping point due to the amplifying feedbacks, with the hypothesis that extra heat and moisture in the atmosphere enhance the likelihood of a greater frequency, severity and widespread of downbursts and large hail events. Media reports of downbursts and hail have increased over the years; however, according to the IPCC report, there is not a scientific consensus of the historical variability and changes of this hazardous weather event globally and regionally. As discussed in the recent study by Pilguy et al. (2022), convective environment conditions leading to downbursts (macrobursts, microbursts and derechos) due to extra heat and moisture in the atmosphere enhance the likelihood of straight-line winds. Prein (2023) also concluded that powerful storms winds have intensified in central USA over the past 40 years. For the Mediterranean region, Llasat (2021) found a predominant increase of convective precipitation resulting from a strengthening of thermodynamic conditions. This constraint in the state-of-the-art of wind and hail hazards is due to the lack of historical wind and hail observations, and the limitations of current modelling capabilities to resolve localised and short-lived extreme winds and large hail storms.

The main purpose of the current proposal consists to deepen understanding of downbursts and large hail in historical changes and, specially, in the mesoscale dynamics using km-scale simulations. Moreover, to quantify the influence of anthropogenic climate change, we will study it via storylines using Pseudo Global Warming Approach (PGWA) simulations and statistical attribution methodology (e.g., GEV, analogues method, etc.). The PGWA extract the anthropogenic signals from several CMIP6 climate models (to generate an ensemble approach) and perturbing the initial conditions (such as, IFS or ERA5). These anthropogenic signals are obtained for all the prognostic variables needed to initialize a convection-permitting model from the monthly mean CMIP6 data. The WRF-ARW and HARMONIE-AROME models will be used in this project at very high-resolution (< 1 km) to evaluate the mesoscale dynamics and the impact of human-induced climate change in the storm dynamics in downbursts and large hail events.

Despite the socioeconomic and environmental impacts of non-tornadic thunderstorm winds, the state-of-the-art on downburst and large hail event detection, attribution and prediction is still very limited, especially in the Euro Mediterranean region, further research on exploring this severe weather phenomenon is urgently needed. For all these reasons, this proposal will aim to answer the following question: Are non-tornadic winds and large hail events becoming more frequent, severe and widespread due to climate change?

GENERAL AND SPECIFIC OBJECTIVES

The SCS that produce downbursts and large hail may pose a threat to the society in the Euro Mediterranean region and quantify the influence of human-induced climate change in these risks will be assessed in this project:

1. To assess changes in historical downbursts and large hail across the Euro Mediterranean region (specially the Iberian Peninsula) over the last decades. We will use all available observations and simulations to create the first time a comprehensive catalogue of historical downbursts. Regarding large hail events, we will use the Calvo-Sancho et al. (*in preparation*) large hail (≥ 5 cm) dataset in Spain, the national severe weather reports datasets (such as SINOBAS in Spain) and European Severe Weather Database (ESWD) by ESSL to create a comprehensive catalogue of historical large hail events in the last decades. From these both datasets we will quantify the long-term changes of large hail and straight-line winds in the Euro Mediterranean region.
2. To attribute the connection of human-induced climate change with the frequency, severity and widespread of downbursts and large hail events. We will apply an attribution multi-approach to both evaluate the influence of the anthropogenic climate warming on the past observed

changes of these meteorological impacts using statistical attribution methods. On the other hand, we will study the storm dynamics in severe powerful downbursts and large hail events using the PGWA to quantify how the warming climate is enhancing, for example, the microphysics.

3. To improve the limited accuracy to predict large hail events and localised extreme winds using convection-permitting downscaling and ML in operational NWP models. Finally, we will explore the potential of training machine-learning on downscaled NWP to improve the limited skills of NWP in forecasting downbursts and large hail in real time. Specifically, we will deploy and optimize the ECMWF's open-source ANEMOI data-driven framework, testing advanced generative architectures such as Continuous Ranked Probability Score (CRPS) and Latent Diffusion Models (LDMs) to handle the highly stochastic nature of severe convective environments. These AI models will be trained using distributed multi-node GPU setups to evaluate probabilistic ensemble performance. We will experiment with structural configurations, including stretched grid geometries, to dynamically enhance spatial resolution over high-impact convective zones in the Euro-Mediterranean region (e.g., the Iberian Peninsula) without incurring prohibitive global computing costs.

JUSTIFICATION OF THE COMPUTER RESOURCES REQUESTS

We will use mainly two convection-permitting NWP models to achieve the aims of this project: the WRF and HARMONIE-AROME models. Moreover, GPU nodes will be used to launch AI/ML experiments.

WRF model is a state-of-the-art numerical model highly approved and great support by the modelling community. This model is characterized for weather simulation solving the fully compressible, non-hydrostatic Euler equations by the numerical method finite differences. In this project, we will use the WRF in 4.7.1. version and will be configured with two domains using one-way nesting approach: the outer domain with 3 km of grid resolution and the high resolution one with 1 km, using 1000 grid points in both directions and domains. The vertical resolution will set-up in 65 sigma model levels unequally spaced, with larger number of levels (i.e., higher vertical resolution) in the lower troposphere for a better representation of the convective planetary boundary-layer processes due to in the PBL is occurring the most mesoscale dynamics that affects the SCS. We will use adaptive time steps to be more efficient in computing time and resources. We will conduct a sensitivity analysis of the WRF physics options to determine the most suitable configuration for simulating downbursts and large hail. Since convection will be explicitly resolved by the model at the chosen resolution, the cumulus parameterization scheme will remain deactivated. Initial and boundary conditions will be retrieved from the ERA5 reanalysis and IFS analysis.

HARMONIE-AROME model is conceived and supported by the ACCORD consortium, being the operational NWP model in several European meteorological offices. This model is characterized by a non-hydrostatic spectral dynamical core, semi-Lagrangian, semi-implicit discretization of the equations. In this project we will use the HARMONIE-AROME default physics option assessed by the members of the ACCORD consortium in several meteorological situations. The model has a unique domain which the user can choose the horizontal and vertical resolution. In this project, we will simulate in very high-resolution (< 1 km) directly from the IFS initial and boundary conditions due to the low benefit to do a nesting approach (Díaz-Fernández et al., *under review*). The grid dimensions will be 1000 grid points in both directions with 65 hybrid model levels in the vertical.

For each SCS event, 250000 SBUs approximately have been used using the WRF model, while around 650000 SBUs have been required when running very high-resolution in HARMONIE-AROME.

Regarding the AI/ML component of the project, we will deploy the ECMWF's open-source ANEMOI data-driven framework and evaluate advanced configurations of the Artificial Intelligence Forecasting System (AIFS). Data-driven models present an unprecedented opportunity to drastically

reduce inference time while maintaining convective-permitting skills, but their training and fine-tuning stages are highly resource-intensive.

Our experimental pipeline is structured in progressive phases, scaling from 4 to 16 parallel NVIDIA A100 (40GB) nodes (consuming 20 to 80 GBUs per wall-clock hour, respectively):

1. Global Baseline & Transfer Learning: We will initialize the framework using global ANEMOI/AIFS models pre-trained on ERA5 and operational IFS data at standard resolution (~30 km). This initial phase requires 4-GPU nodes (20 GBUs/h) for rapid prototyping and architecture validation.

2. Stretched-Grid Convection-Permitting Fine-Tuning: To resolve severe convective storms (SCS) at the kilometer scale (1 km) over the Euro-Mediterranean target area without the prohibitive cost of global high-resolution training, we will exploit ANEMOI's native stretched-grid geometry. Fine-tuning the model under this multi-resolution configuration demands scaling up to 16 parallel A100 GPUs (80 GBUs/h) to accommodate the massive spatial tensor dimensions, dense graph-convolutions, and high-frequency vertical levels required to simulate localized downbursts and hail mechanisms.

3. Deterministic vs. Probabilistic Frameworks (CRPS & Diffusion): Initial runs will focus on deterministic regional models. Subsequently, we will transition to probabilistic frameworks to quantify forecast uncertainty. We will train and compare two state-of-the-art approaches: an optimized AIFS variant utilizing Continuous Ranked Probability Score (CRPS) loss minimization, and a Latent Diffusion Model (LDM) setup for generative ensemble member production.

4. Hyperparameter Tuning & Sensitivity Analysis: Intensive parallel training suites will be executed to optimize neural network parameters (e.g., learning rates, attention heads, and message-passing depths in GNNs).

To sustain this institutional-scale pipeline, the compute requirements are estimated as follows:

- Global baselines and 4-GPU prototyping: ~10,000 wall-clock hours distributed across multiple experiments = 200,000 GBUs/year.

- High-resolution stretched-grid training and multi-node (16-GPU) scaling for CRPS and Diffusion architectures: ~10,000 wall-clock hours of intensive training and hyperparameter optimization = 800,000 GBUs/year.

Therefore, a total allocation of 1,000,000 GBUs per project year is strictly requested to guarantee model convergence, avoid Out-Of-Memory (OOM) limitations, and determine the optimal AI architecture for real-time severe weather forecasting.