Project: **1481**
Project title: **Climate-Resilient Development Pathways in Metropolitan Regions of Europe – CARMINE**
Principal investigator: **Elena Xoplaki**
Report period: **2025-11-01 to 2026-10-31**

CARMINE delivers practical roadmaps for research and innovation priorities on adaptation and mitigation, risk assessment and cost-effectiveness for eight European metropolitan Case Study Areas (CSAs), which implement Digital Twin use cases. Its scientific core is a framework of high-resolution risk assessment models that link climate dynamics, Earth system processes and socioeconomic drivers. To bridge the gap between global climate datasets and risk assessments at local scale, meteorological and climate variables are downscaled to metropolitan resolution (about 1 km) with AI-based methods, which also correct the systematic biases of the regional climate models. The resulting datasets allow the robust detection of extreme events relevant to the economy and to society, in particular heatwaves, wildfires and floods.

The downscaling and bias-correction framework was established over the EURO-CORDEX domain (22° W–45° E, 25° N–72° N). It was trained on 1990–2014, the period set by the availability of EMO1 observations, and it uses large-scale predictors chosen for their physical relation to each target variable: 500 hPa geopotential height, 200 hPa vorticity and surface precipitation for precipitation, 500 hPa geopotential height for temperature, and 500 hPa geopotential height with mean sea level pressure for wind speed. The HYRAS dataset is used to test the results. The preprocessing of historical data and projections, the computation of empirical orthogonal functions, the training of the models and the bias correction required a high level of parallelisation and a large amount of memory on Levante, and the same work has to be applied to the future projections to keep resolution and quality comparable.

The approach learns the spatial relations between low-resolution, large-scale variables and high-resolution surface fields from historical data and applies them to climate model projections for heat risk assessment. This allows the identification and spatial delineation of heatwaves with a focus on urban areas, based on high-resolution temperature products generated by the statistical machine-learning model. For wildfire, the risk is estimated with a next-day fire occurrence prediction model developed by NOA/BEYOND, which uses daily weather data, and the resulting products support the dialogue with local stakeholders on transition risks and on urban planning and resilience strategies.

As a first result, the chain takes the output of MPI-ESM1-2-HR from its native grid of about 100 km down to ~1 km. Figure 1 follows the hottest day of 2003 in the historical run (1 August 2003) over a downscaling window in central Europe (about 9–17° E, 48–54° N). On the native grid the window is covered by a few coarse cells, while at 1 km the temperature field resolves the contrasts related to relief and valleys. The correction applied by the network with respect to the baseline is spatially coherent and within about ±1 °C, with warmer and cooler adjustments in different parts of the window, and the final step adds quantile delta mapping (QDM) to align the statistics of the output with the observational reference. The same chain will be applied to the other variables and to the projections, so that heat risk can be assessed at metropolitan scale.

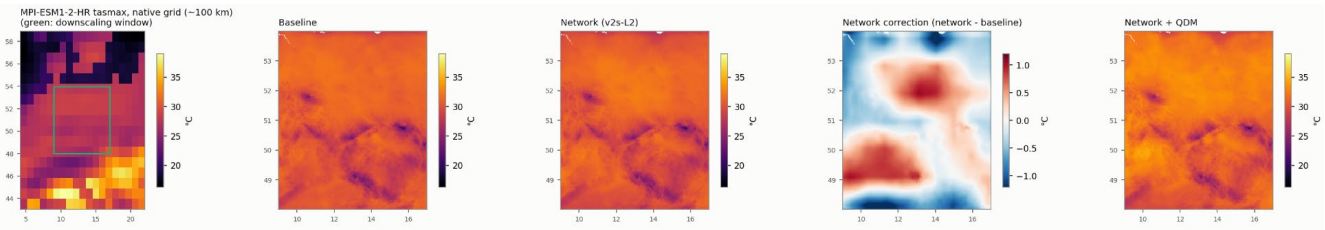


Figure 1. The downscaling chain applied to the hottest MPI-ESM1-2-HR day of 2003 (historical run, 1 August 2003), from the native grid of about 100 km (green box: downscaling window) to 1 km: baseline, network (v2s-L2), network correction (network – baseline) and network + QDM.

The main lesson of the work of the last year concerns extreme events. Models trained on the historical period see few extremes, so the climate signal associated with them is not well identified, and this limits what they can say about the future. We addressed this on two sides. The search of analogues over the entire North Atlantic window was carried out with the autoencoder-based VA-AM method, which runs on GPUs and works only in the latent space, instead of in the full-resolution fields, making the search much faster. Different configurations and dimensions of the latent space are being tested to find the one that best estimates extreme events, which is the main limitation of analogue approaches. On the downscaling side, different loss functions were tested for precipitation. The daily RMSE turned out to be unsuited, because it penalises extremes and degrades the spatial pattern of rainfall (Rampal et al., 2026), whereas a Bernoulli-gamma loss, which models the occurrence and the intensity of precipitation separately, looks promising. Several downscaling and bias-correction models are being trained and compared with these choices.

The next step is the hydrological component to assess floods over the greater Leipzig area and hydrological droughts over the Po Basin and Slovenia. The LISFLOOD workflow was developed and tested on Levante in the MedEWSa Slovakia pilot, including runs with raw and downscaled forcings, and its measured cost is the basis of the resource estimates for CARMINE. For Leipzig, which focuses on wildfire, heatwaves and floods, LISFLOOD will be run over the Saale basin (about 24,000 km²), with seasonal hydrological forecasts driven by downscaled SEAS5.1 (25-member reforecasts 1993–2016 and 51-member forecasts 2017–2025, 7 lead months) and a continuous projection up to 2050 driven by downscaled MPI-ESM1-2-HR. The same chain will be applied to Bologna, over the Po basin at Pontelagoscuro and the Reno basin (about 80,000 km²), where the 2022 Po drought and the May 2023 floods in Emilia-Romagna show the relevance of both droughts and floods. A smaller case study over Slovenia (about 20,000 km²) will be configured and tested. In Athens, the analysis addresses heatwave-related mortality with upscaled Copernicus C3S datasets, and for Birmingham two scenarios (SSP3-7.0 and SSP5-8.5) will be provided to evaluate a phenology model and the effect of temperature and soil moisture on vegetation.

These developments define the computing needs of 2027. The training of the autoencoder and of the downscaling models requires GPU time, while the post-processing, the computation of indicators and the LISFLOOD simulations run on CPU nodes and account for most of the CPU time requested. The hydrological runs are independent across ensemble members, scenarios and time slices, and a 25% contingency is included for re-runs after quality control and with the revised precipitation downscaling, so that the work scales well on Levante.

CARMINE builds on the datasets and knowledge of DAKI (bb1245), NUKLEUS (bb1203), BioValue (bb1480), MedEWSa (bb1478) and its methods and case-study results will inform NUKLEUS and UDAG (bb1364). The results also have the potential to address climate issues in the Mediterranean region and to provide information for hazard mitigation decisions to the local authorities of the case study areas.

References

Rampal, N. et al., 2026: CORDEX-ML-Bench: A Benchmark for Data-Driven Regional Climate Downscaling-Experiment Design and Overview. arXiv preprint arXiv:2606.29172