# SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

| | |
|---|---|
| **Reporting year** | 2023 |
| **Project Title:** | High resolution (1.8 km) atmospheric reanalysis, obtained by dynamically downscaling ERA5 reanalysis with the Italian convection permitting meteorological model MOLOCH |
| **Computer Project Account:** | SPITSTOC |
| **Principal Investigator(s):** | Stocchi Paolo |
| **Affiliation:** | ISAC-CNR |
| **Name of ECMWF scientist(s) collaborating to the project** (if applicable) | No one |
| **Start date of the project:** | October 2023 |
| **Expected end date:** | 31/12/2025 |

## Computer resources allocated/used for the current year and the previous one

(if applicable)
Please answer for all project resources

| | | Previous year (jsut 4 month because my project is a late one) | | Current year | |
|---|---|---|---|---|---|
| | | Allocated | Used | Allocated | Used |
| **High Performance Computing Facility** | (units) | 5000000 SBU | 2000000 | 15000000 | 6,485,053 |
| **Data storage capacity** | (Gbytes) | 20000 | 4000 | 50000 | 12000 |

## Summary of project objectives (10 lines max)

The aim of our project is to bring the potential of ERA5 to the local scale with the aim of synergistically exploiting CP-RCM features and ERA5 reliability and creating a very-high resolution climate dataset for past/present climate, using a model setup specific for the areas of interest. This is the rationale adopted in this work project: to create a new additional gridded dataset over Italy, labelled as MORIA (Moloch RIAnalysis), derived from the dynamical downscaling of ERA5 reanalysis from their native resolution (25 km) to a resolution of 1.8 km for the period 1990–2020, more suitable for studies of regional phenomena and for application to vulnerability, impact, and adaptation assessments.

## Summary of problems encountered (10 lines max)

We can confirm that we have not encountered or experienced any technical issues that have hindered our work plan

## Summary of plans for the continuation of the project (10 lines max)

As of today, we have completed the production of the first 12 years of the 30-year reanalysis planned for the project, having utilized approximately 6,485,461 SBU. Currently, we have about 8,514,539 SBU available for this year and an additional 5,000,000 SBU for future use. So far, we have not encountered any technical issues and have maintained consistent simulation. Following the initial validation of the first 15 years produced, which will undoubtedly require substantial machine time for post-processing, we anticipate completing the planned production by the end of the project.

## List of publications/reports from the project with complete references

## Summary of results

In this initial phase of the project, the activities carried out can be summarized as follows:

- First of all, we conducted tests to evaluate the best downscaling strategy and configuration to adopt based on ERA5 data. Essentially, the two downscaling configurations that were evaluated and compared are as follows:

1. A two-step process: starting with ERA5 data, downscaling to 8 km resolution over a MedCordex-type domain using the hydrostatic model BOLAM (https://www.isac.cnr.it/dinamica/projects/forecasts/bolam_short_description_2012.htm), which then provides the initial and boundary conditions for a second downscaling over the Italian domain at 1.8 km resolution (Fig. 1), which we will refer to as **T1** for simplicity from now on
2. Direct downscaling of ERA5 data over the Italian domain at a 1.8 km resolution, using the non-hydrostatic model Moloch (Fig. 2), which we will refer to as **T2** for simplicity from now on

The comparison and evaluation of the two methods are based on a simulation of a reference year, which is 2015

.

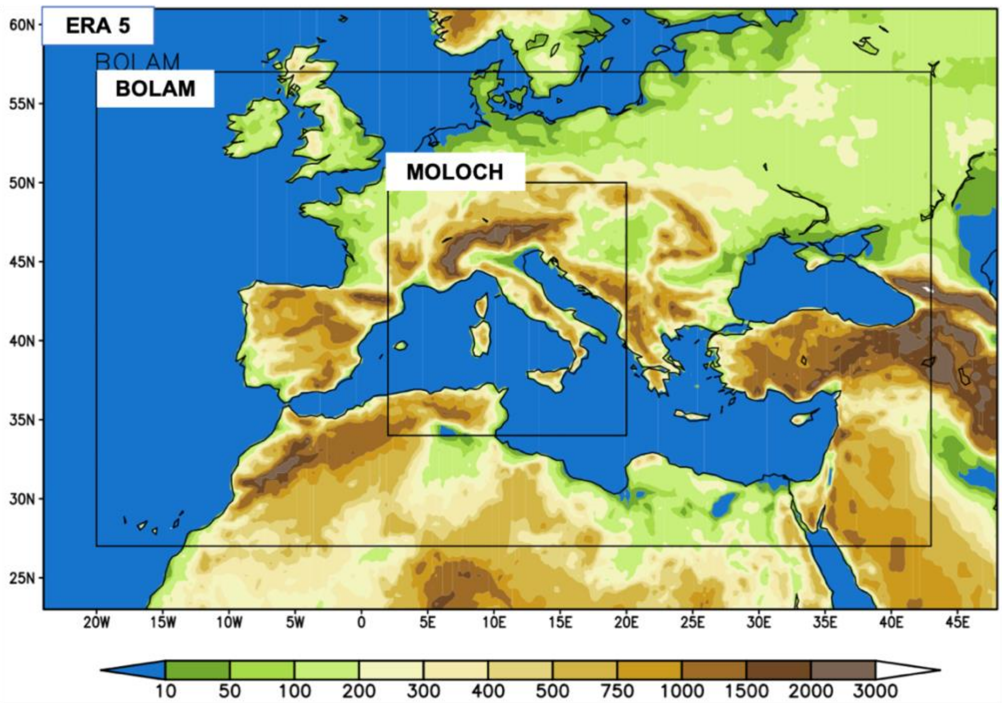


Fig. 1: A two-step Downscaling process : Extent of the ERA5, BOLAM (8 Km) and MOLOCH (1.8 Km) domains of integration with topography (m). The ERA5 domain, intermediate black box, approximately corresponds to the Med-CORDEX area; BOLAM is nested into ERA5 domain over Med_Euro area and the MOLOCH domain, inner black box, covers Italy and surrounding.

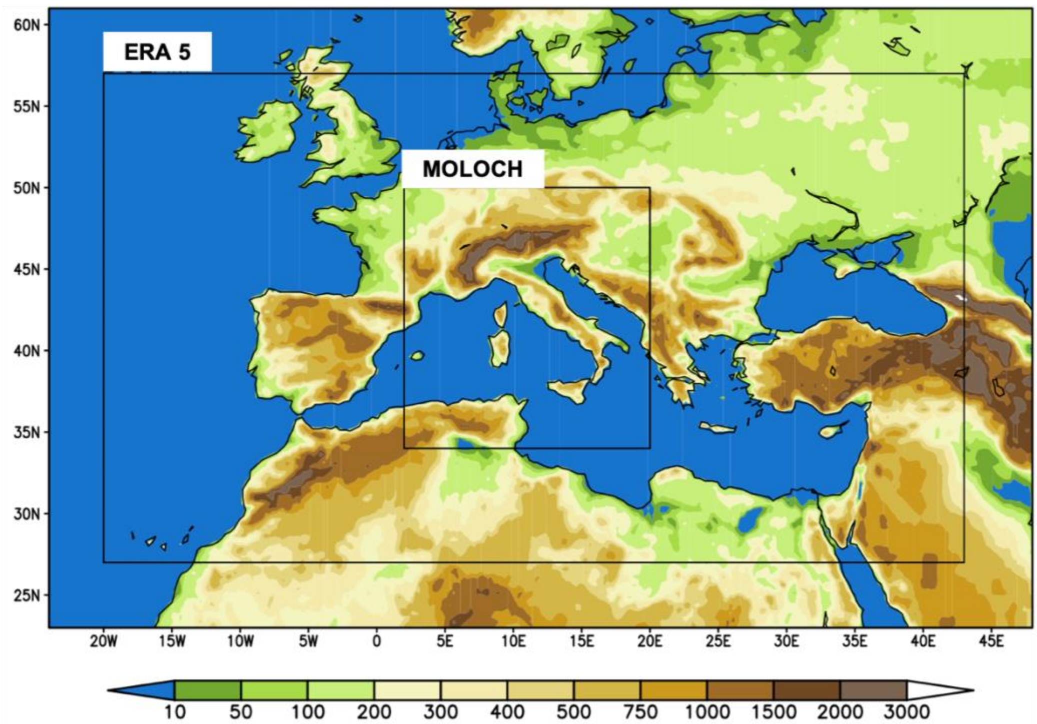


Fig. 2: A 1-step Downscaling process : Extent of the ERA5, and MOLOCH (1.8 Km) domains of integration with topography (m). The ERA5 domain, black box, approximately corresponds to the Med-CORDEX area; MOLOCH domain, inner black box, covers Italy and surrounding.

The two types of downscaling approaches based on the 2015 simulation were carried out as follows:

In T1, the 1-year simulation has been divided into daily runs, daily data of the high-resolution atmospheric downscaling were produced as follows: every day at 18:00 UTC, a BOLAM simulation is started using global reanalyses as initial conditions; boundary conditions are provided every 6 h for the following 30 h. The domain of integration is shown in Fig. 1. Hourly outputs from the BOLAM simulation provide the initial and boundary conditions to the MOLOCH simulation, which starts each day at 21:00 UTC and has a forecast length equal to 27 h. The MOLOCH model produces outputs every hour over the domain of integration shown in Fig. 1 (inner rectangle). The daily data of the BOLAM/MOLOCH downscaling are built using the last 24 h of the two model simulations, while the first 6 and 3 h of integration of the BOLAM and MOLOCH model, respectively, are considered as spin-up times and thus discarded; a scheme of the experimental design is shown in Fig. 3

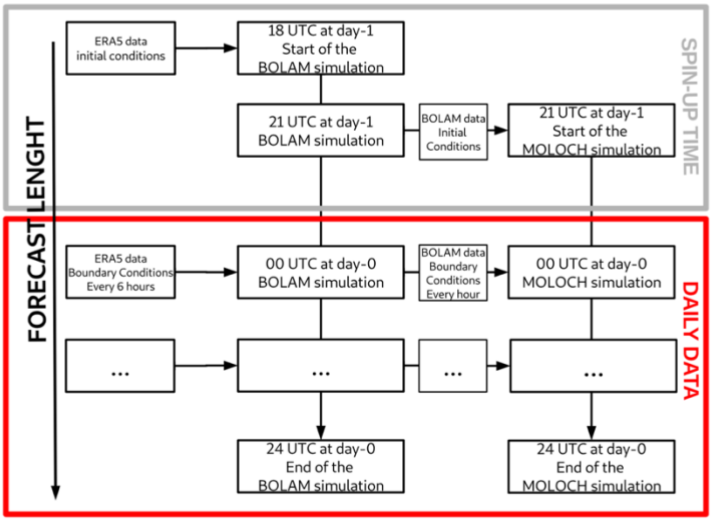


Fig. 3: Modelling setup for the production of one single day of the A two-step Downscaling process

In T2, the 1-year simulation has been divided into daily runs, where each MOLOCH run is initialized at 18:00 UTC by ERA5 atmospheric fields, while soil temperature, water content and snow height, are taken from the previous simulation; boundary conditions are provided every hour. The MOLOCH model produces outputs every hour over the domain of integration shown in Figure 2 (inner rectangle). The MORIA dataset is then built using the last 24 hours of each model simulation, while the first six hours of integration, are considered as spin-up times and thus discarded. A scheme of the experimental design is shown in Figure 4

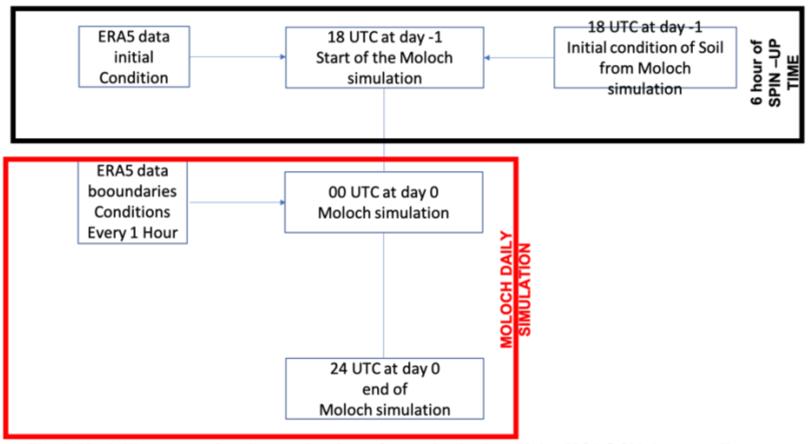


**Fig.4: Modelling setup for the production of one single day of the MOLOCH downscaling.**

Figures 5, 6, and 7 present examples of the results from our comparison between T1 and T2 concerning the precipitation variable, utilizing various statistical indices detailed in Table 1. As demonstrated in Figures 5, 6, and 7, the two approaches exhibit no significant differences. Consequently, we opted to proceed with T2 due to its considerably lower computational costs, thus avoiding intermediate simulations with the hydrostatic model BOLAM

Table 1

| Index | Definition | Unit |
| --- | --- | --- |
| Seasonal Mean | Mean daily precipitation for season | mm/day |
| Frequency | Wet-day/hour frequency** (fraction of number of wet-days/hour per season) | Fraction |
| Intensity | Wet day/hour intensity ** | mm/day-mm/hour |
| Heavy precipitation | Heavy precipitation defined as the 99th or 99.9th percentile of all daily or hourly precipitation events (mm/day or mm/h) | mm/day – mm/hour |
| PDF | Probability density function (PDF) defined as the normalized frequency of occurrence of precipitation events within a certain bin. | |

***Defining a wet-day as a day with at least 1 mm of rain and a wet-hour as an hour with at least 0.1 mm of rain*

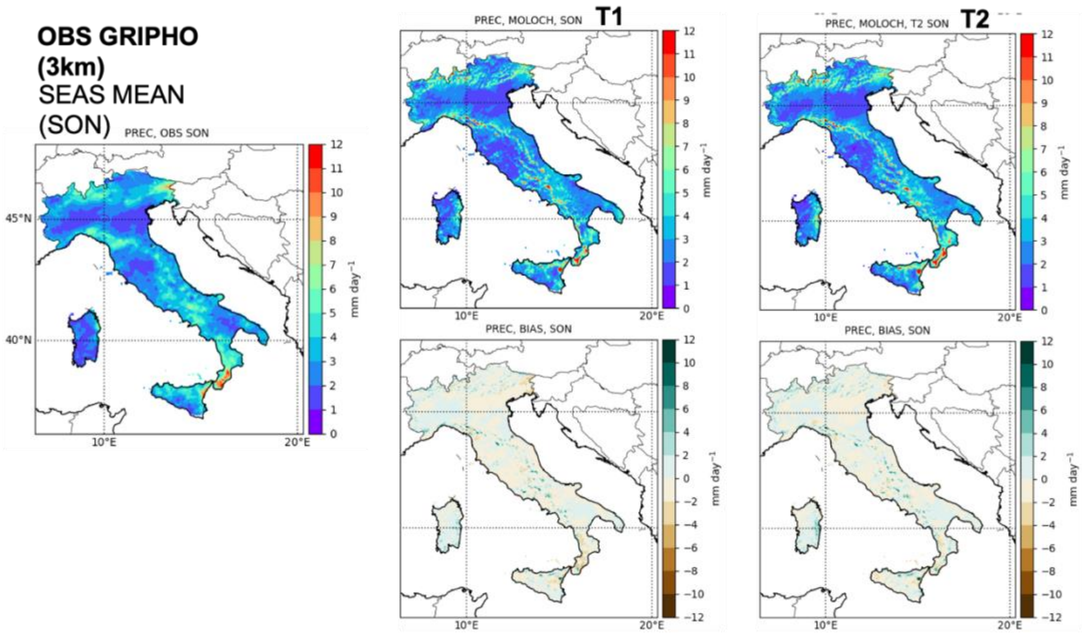


**Fig.5: Autumn Seasonal Mean precipitation for T1, T2, and Bias calculated with respect to the Gripho Observation dataset**

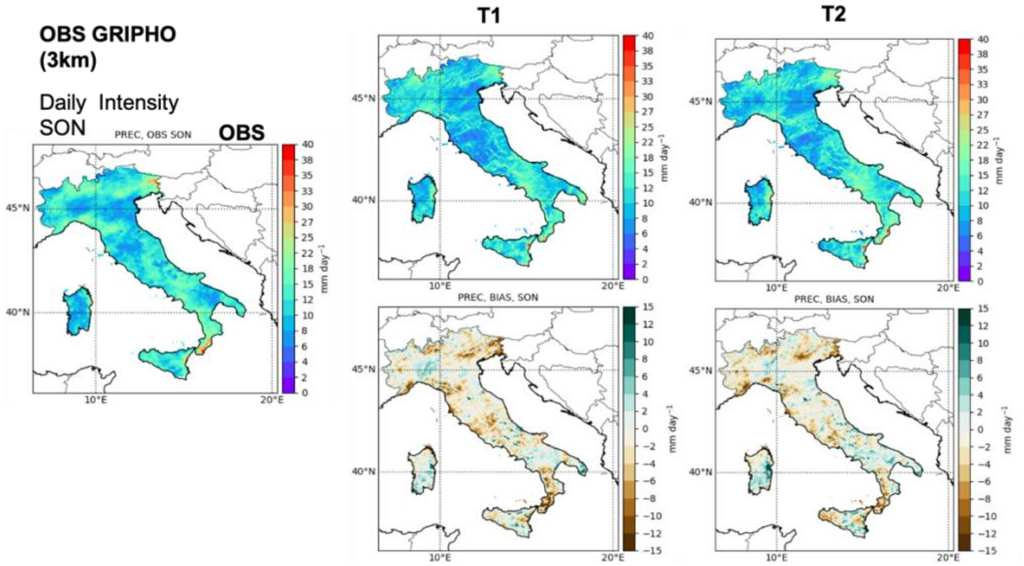


**Fig.6: Daily precipitation intensity for autumn for T1, T2, and Bias calculated with respect to the Gripho Observation**

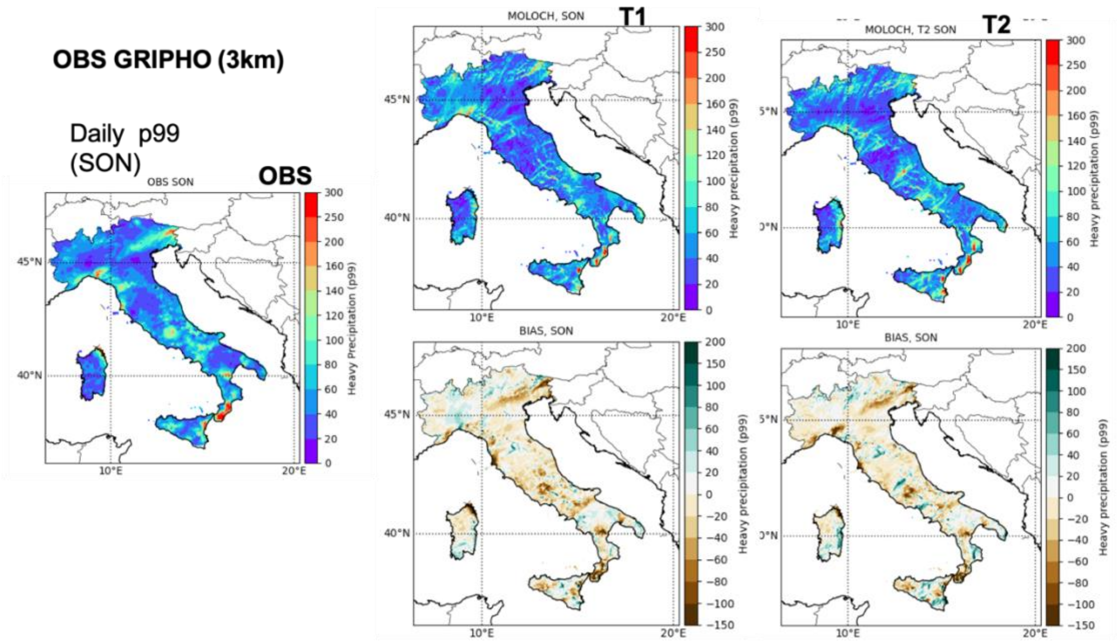


**Fig.7: Daily precipitation p99 percentile for autumn for T1, T2, and Bias calculated with respect to the Gripho dataset**

- Once the optimal downscaling strategy was established and the best configuration identified, the activity continued with the implementation of an automated procedure for running the simulations needed to produce the 30 years of reanalysis data (1990-2020) required by the project. The automated procedure developed during this period essentially consists of a series of bash scripts that allow for the sequential submission of sabtch jobs, both for input production and for running the Moloch model, all while imposing the necessary job dependencies nd the setup showed in Fig.4.
- During this period, post-processing procedures were also implemented to convert the output files from the simulations into GRIB2 and NetCDF formats, as well as for transferring the data from SCRATCH to the ECFS archive once produced

- Finally, we have initiated the production of our reanalyses and have thus far completed the first 12 years out of the 30-year dataset that we aim to finish by the end of the project