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	2024
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 (2024)		Current year (2025)	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	15 millions SBU	~15 milions SBU	12 milions SBU	~8 milions SBU
Data storage capacity	(Gbytes)	50000	500000	20000	~10000

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 MORE (Moloch REanalysis), derived from the dynamical downscaling of ERA5 reanalysis from their native resolution (25 km) to a resolution of 1.8 km for the period **1990–2024**, 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 successfully completed the production of 30 years of reanalysis data, as originally planned in the submitted project, utilizing nearly the entire allocation of SBUs required for this phase. We have also completed the post-processing for the same period, which required a significant amount of computational time.

Initially, the project aimed to cover the 1990–2020 period; however, we subsequently requested an extension of the reanalysis to 2024 to include recent years marked by severe drought events in Italy—an aspect considered crucial for the ongoing research activities.

At present, we have approximately 3000000 SBUs available until the project's completion. Thus far, simulations have run consistently without any major technical issues. We anticipate completing the production of the remaining 4 years (2021–2024) within the allocated timeframe and resources, meeting the project objectives by its conclusion.

List of publications/reports from the project with complete references

Preliminary results from our project have been presented at the EGU General Assembly.
The abstract is available at the following link: https://doi.org/10.5194/egusphere-egu25-9367

Summary of results

The **MORE** dataset represents the highest spatial resolution (1.8 km) currently available for atmospheric reanalysis over Italy. Developed independently, it adds diversity to the landscape of existing reanalysis products and provides significant value for meteorological and climatological applications. Its high resolution makes it particularly suitable for use in ensemble approaches, which aim to better capture atmospheric processes and increase the robustness of climate and weather-related analyses (Giordani et al., in preparation). This, in turn, enhances the reliability of reanalysis-based applications, including impact modeling and the study of extreme events.

The **MORE** dataset will be made publicly available to the scientific community, offering high-resolution atmospheric fields from 1990 to the present, with continuous updates. This accessibility aims to support a broad range of studies, from regional climate diagnostics to operational decision-making.

To assess its performance, MORE is evaluated using three independent gridded observational datasets, applying spatially distributed diagnostics and statistical summary metrics.

The validation of the **MORE** reanalysis for multiple variables, including precipitation and temperature, is performed using three independent gridded observational datasets: **ARCIS** and **GRIPHO** for precipitation and **UniMi/ISAC-CNR** for temperature.

GRIPHO is an hourly precipitation dataset covering Italy at a spatial resolution of approximately **3 km** (Fantini, 2019). It is derived from quality-controlled rain-gauge measurements and spans the period **2001–2016**. GRIPHO has been widely used in previous studies (Ban et al., 2021; Caillaud et al., 2021, Capecchi et al., 2023) to validate numerical precipitation estimates.

ARCIS (Climatological Archive for Central-Northern Italy; Pavan et al., 2019) is a high-resolution (5×5km) daily gridded precipitation dataset that integrates approximately 1,000 quality-controlled and homogenized rain gauge records from various Italian institutions, including Hydrological Services, Agro-Meteorological Services, and Regional/Local Meteorological Agencies. It provides 24-hour accumulated precipitation data and spans the period 1961–2024, offering comprehensive coverage of central and northern Italy.

UniMi/ISAC-CNR provides daily maximum and minimum temperature data, offering a high-quality, homogenized gridded observational record. The interpolation of observations onto the high-resolution grid follows the anomaly-based approach proposed by Mitchell and Jones (2005). The method reconstructs monthly fields by combining climatological normals (long-term averages for a specified reference period) with corresponding anomalies (deviations from these averages). The specific implementation for the t2m dataset adheres to the methodology described by Brunetti et al. (2014), as applied in Cavalleri et al., (2024). Climatological fields are constructed using data from approximately 1,500 stations. These values are interpolated across the grid via locally weighted linear regression, incorporating elevation and spatial relationships—such as horizontal and vertical distances, slope gradient, aspect, and proximity to the sea—between grid cells and surrounding stations. Anomalies are calculated from a subset of these stations, limited to those with continuous long-term records during the 1991–2020 period (approximately 1,000 stations). Anomaly fields are then interpolated using a weighting scheme that combines radial, vertical, and angular components to correct for the heterogeneous spatial distribution of stations around each grid point.

The validation focuses on the ability of MORE to represent the climatological characteristics of key atmospheric variables over Italy across multiple years. Special attention is given to the dataset's capacity to reproduce the **intensity, frequency, and spatial structure of extreme events**, such as heavy precipitation and heatwaves.

Additionally, MORE's precipitation fields are benchmarked against **LAMMA-HINDCAST**, the previous convection-permitting regional reanalysis produced with the MOLOCH model, while both temperature and precipitation fields are compared with **ERA5-Land** and **CERRA**. This multi-dataset comparison allows us to assess the added value of convection-permitting, high-resolution reanalysis relative to coarser datasets that rely on parameterized convection.

Finally, MORE is also evaluated against **MERIDA-HRES**, a high-resolution reanalysis developed by Ricerca sul Sistema Energetico (RSE) S.p.A. for infrastructure resilience planning in Italy. This comparison serves to contextualize MORE's performance among existing high-resolution reanalyses available for the region.

Table 1 shows a summury of the dataset used in this work

In this progress report, we present just preliminary results from the ongoing validation phase, focusing specifically on precipitation data

For precipitation, the validation period extends from January 1, 2001, to December 31, 2016, using GRIPHO data, and from January 1, 1991, to December 31, 2020, using ARCIS data. The validation includes both datasets because GRIPHO provides full coverage of the Italian domain with both daily and hourly resolution, while ARCIS spans a longer temporal range (1990–today), though it is limited to central-northern Italy

Validation Methods for precipitation

For the precipitation variable our analysis focuses on a set of statistical indices detailed in Table 2, including seasonal mean daily precipitation, seasonal wet-day/hour intensity and frequency, and extreme precipitation events, defined as the 99th and 99.9th percentiles of all daily and hourly precipitation occurrences, respectively. Additionally, model performance is assessed through the probability distribution of precipitation intensities, relative bias, spatial variability, and correlation.

Indices are computed for each season: summer (June–July–August), winter (December–January–February), spring (March–April–May), and autumn (September–October–November). To preserve the highest level of detail, both model and observational datasets are maintained on their original grids. Given that different remapping procedures can introduce uncertainties and degrade data quality, we minimize data manipulation whenever possible. However, for metrics requiring direct grid-by-grid comparisons, such as relative bias, spatial correlation, and spatial variability, both observational and model data are remapped onto the observational grids used for comparison to ensure consistency in the analysis.

Name	Kind of data	Temporal resolution	Duration	Spatial resolution
EOBS	Gridded observations	Day	1991-2020	~ 10 km
ARCIS	Gridded observations	Day	1991-2020	~ 5 km
GRIPHO	Gridded observations	Hourly	2001-2016	~ 3 km
ERA5-LAND	Reanalysis	Hourly	1991-2020	~ 9 km
CERRA	Reanalysis	Sub-daily	1991-2020	~ 5.5 km
MOLOCH-ISAC	Downscaling	Hourly	1991-2020	~ 1.8 km
LAMMA	Downscaling	Hourly	1991-2020	~ 2.5 km
MERIDA-HRES	Reanalysis	Hourly	1991-2020	~ 4 km

Table 1: Precipitation datasets used in this study, including their time periods, data sources, and spatial resolutions

Index	**Definition**	**Unit**
Mean	Mean daily precipitation	mm/day
Frequency	Wet day/hour frequency (fraction of wet days/hours per season)	(fraction)
Intensity	Wet day/hour intensity	mm/day - mm/hour
Heavy Precipitation (p99, p99.9)	99th (99.9th) percentile of all daily/hourly precipitation events (wet and dry)	mm/day - mm/hour
Probability Density Function (PDF)	Normalized frequency of precipitation events within a certain bin	-
Relative Bias	Relative difference (model - observation) / observation of spatially averaged values	-
Spatial Variability	Ratio (model / observation) of spatial standard deviations of seasonal values across all grid points	-
Spatial Correlation	The spatial correlation of seasonal values between model and observations across all grid points	-
Mean Absolute Error (MAE)	Average absolute difference between model and observation	-
Mean Absolute Percentage Error (MAPE)	Average percentage error (model - observed /observed)*100	%
Field Correlation (PCORR)	Field correlation coefficient between model and observation	-
Root Mean Square Error (RMSE)	Square root of the mean squared differences between model and observation	-
Relative Error	(model - median (observed)) / median(observed)*100	%
Probability of Detection (POD)	Ratio of correctly detected precipitation events to total observed events (hits / (hits + misses))	-
False Alarm Ratio (FAR)	Ratio of false alarms to total predicted precipitation events (false alarms / (false alarms + hits))	-

Table 2. *Statistical metrics used for the validation of the precipitation variable in the MORE dataset*

Results from the Evaluation

Figures 1, 2, 3 and 4 provide a comprehensive overview of the spatial distribution of the daily and hourly precipitation indices (as defined in Table 1) derived from observational datasets and various reanalyses, highlighting the capability of the MORE dataset to realistically reproduce precipitation climatology. Specifically, seasonal means, precipitation intensity, wet-event frequency, and extreme precipitation, defined by the 99th percentile for daily and 99.9th percentile for hourly precipitation, are analyzed for summer (JJA) and autumn (SON).

As horizontal resolution increases, a more detailed depiction of orographic influences emerges, particularly in complex terrain. Both summer and autumn display enhanced precipitation over regions of elevated topography. However, the contrast between mountainous areas and adjacent lowlands tends to diminish during SON, consistent with the enhanced cyclonic activity in the Gulf of Genoa and Alpine lee-side dynamics.

Figure 1 and 2 demonstrates that MORE captures the spatial patterns of mean daily precipitation, intensity, frequency, and extremes with high fidelity when compared to observations, especially for summer and autumn. In contrast to other reanalyses, such as ERA5-Land and CERRA, which suffer from some biases including overestimation of wet-day frequency in summer and autumn and underestimation of precipitation intensity and extremes in summer, MORE exhibits a more accurate representation. This improved performance is likely due to the explicit representation of convection in MORE and its finer spatial resolution, which allow for a more realistic simulation of localized convective precipitation events, especially during the warm season, resulting in a closer match with observed spatial and temporal precipitation patterns.

Moreover, MORE shows comparable skill to high-resolution regional models such as LAMMA and MERIDA-HRES, particularly in capturing heavy precipitation and its intensity, while outperforming ERA5-Land in both seasons and CERRA during summer.

Figure 3 and 4, focusing on hourly precipitation, further emphasizes the distinction among datasets. ERA5-Land significantly overestimates wet-hour frequency, especially over orographically complex regions, while underestimating intensity, yielding an apparently reasonable daily mean precipitation (Fig.1) that, however, arises from compensating errors. In contrast, MORE (as well as LAMMA and MERIDA-HRES) exhibits superior skill: it mitigates the overestimation of wet-hour frequency and more accurately reproduces hourly precipitation intensity and extremes.

It is also worth noting that, compared to LAMMA and MERIDA-HRES, MORE tends to simulate extreme hourly precipitation with less overestimation, particularly in summer, while still remaining consistent with observed values

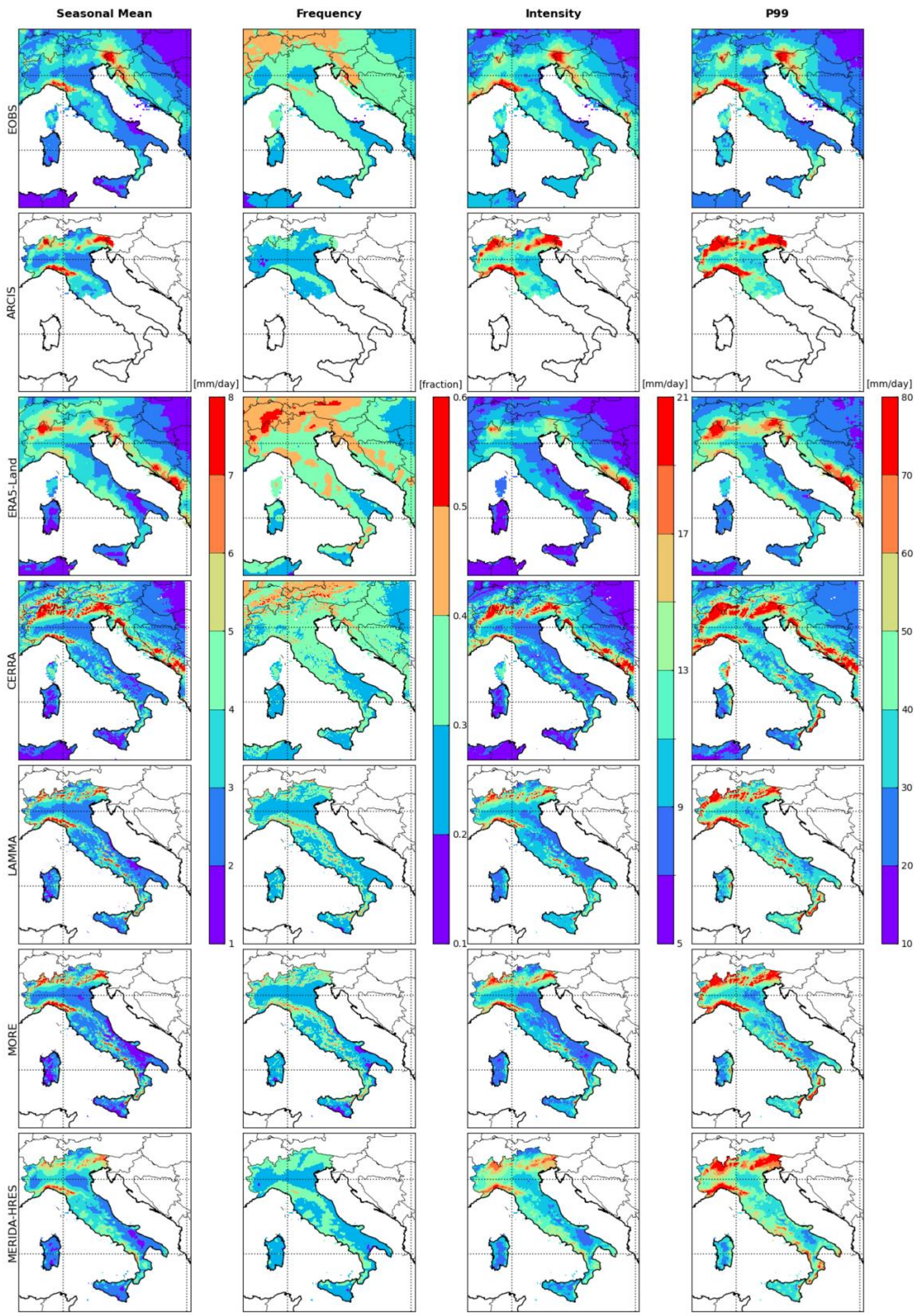

Figure 1: The autumn (SON) seasonal means for various daily precipitation indices: mean precipitation, precipitation frequency, precipitation intensity, and heavy precipitation (defined as the 99th percentile) (Table 2). Results are derived from all observations (OBS) and reanalysis datasets listed in Table 1

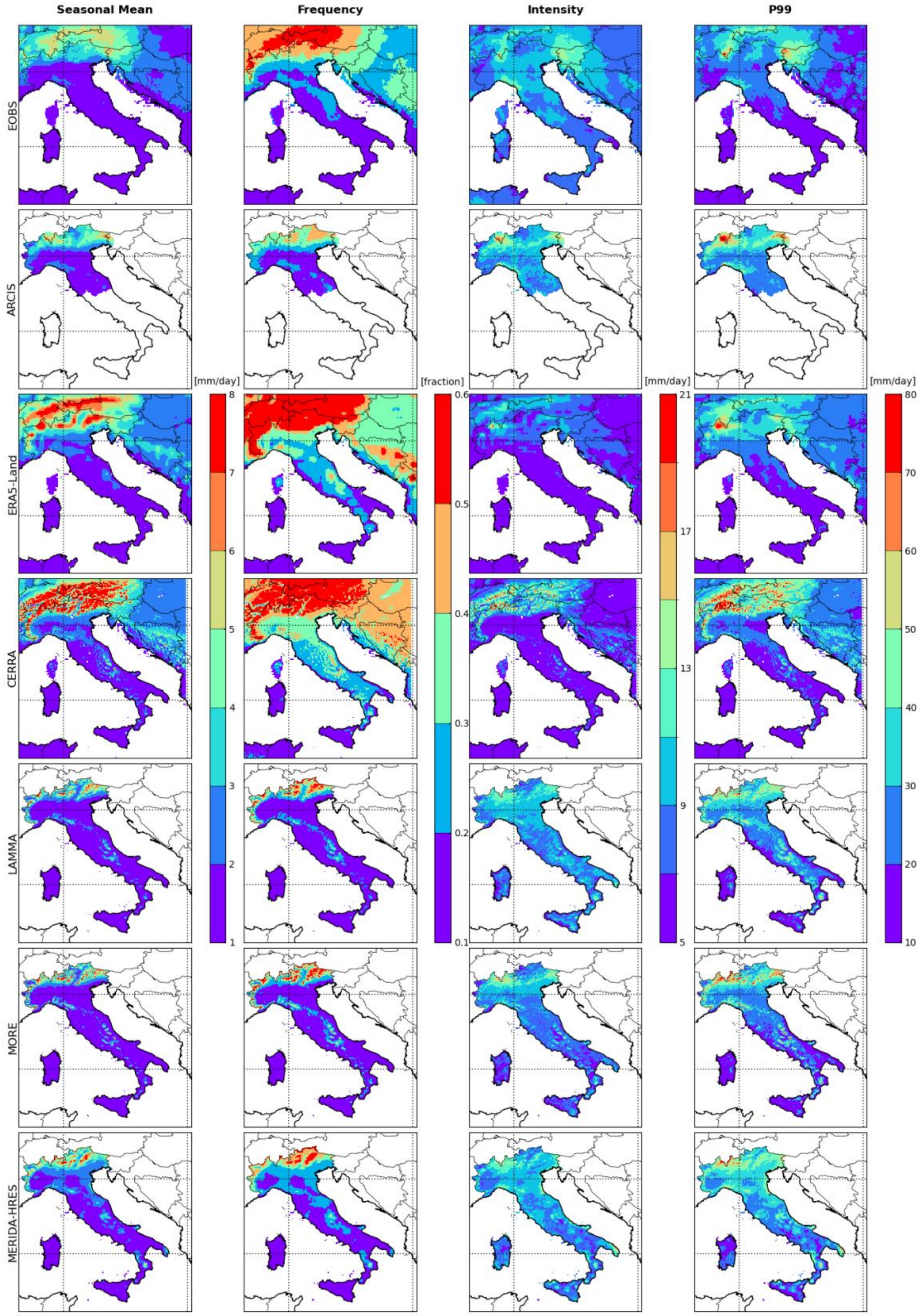

Figure 2: As Fig. 1, but for summer (JJA) season

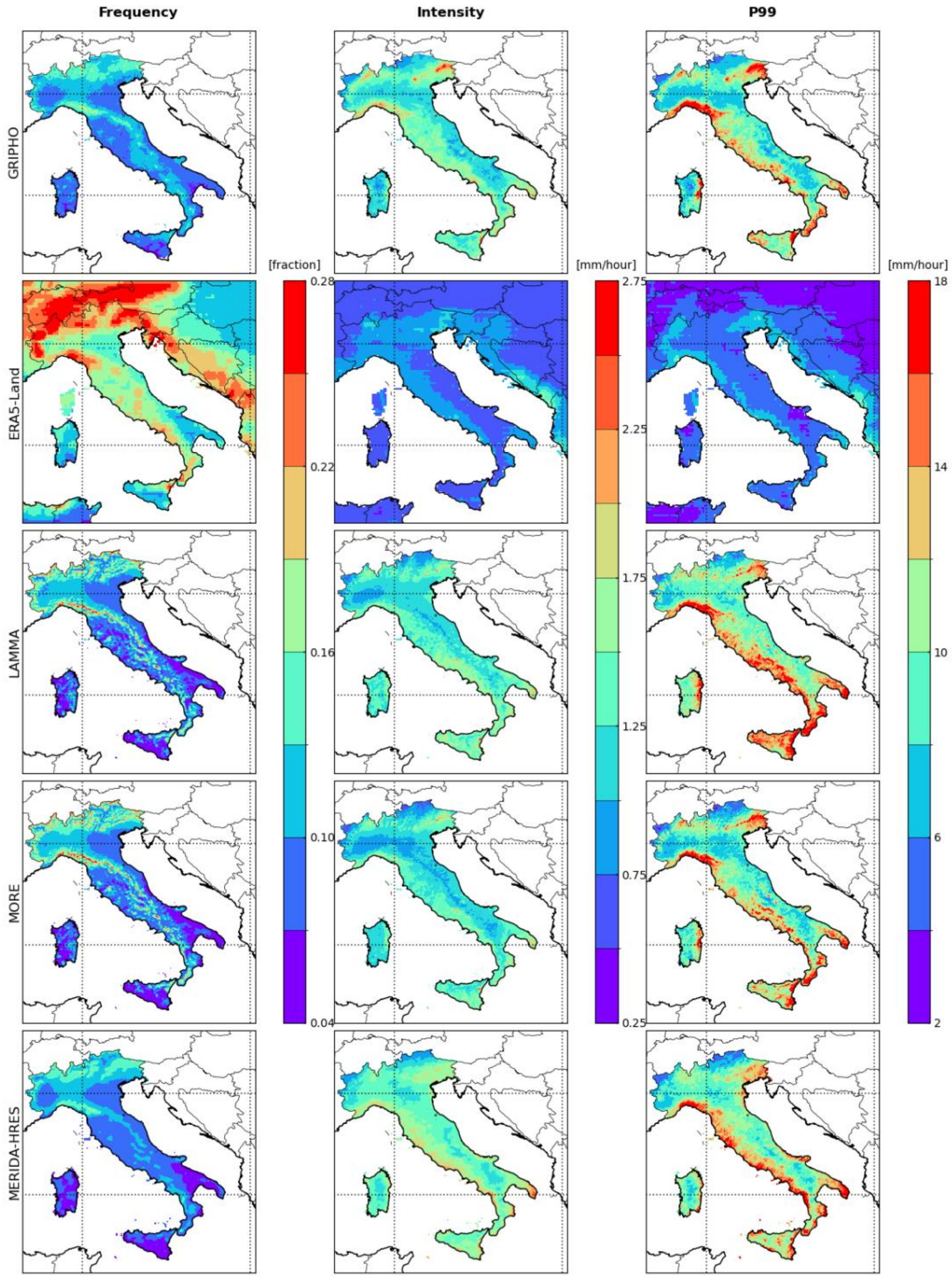

Figure 3: As Fig. 1, but for hourly precipitation. The observation are composed from available gridded hourly precipitation over Italy (Fantini 2019) for the period 2001-2016. Heavy hourly precipitation is defined as the 99.9th percentile of all events.

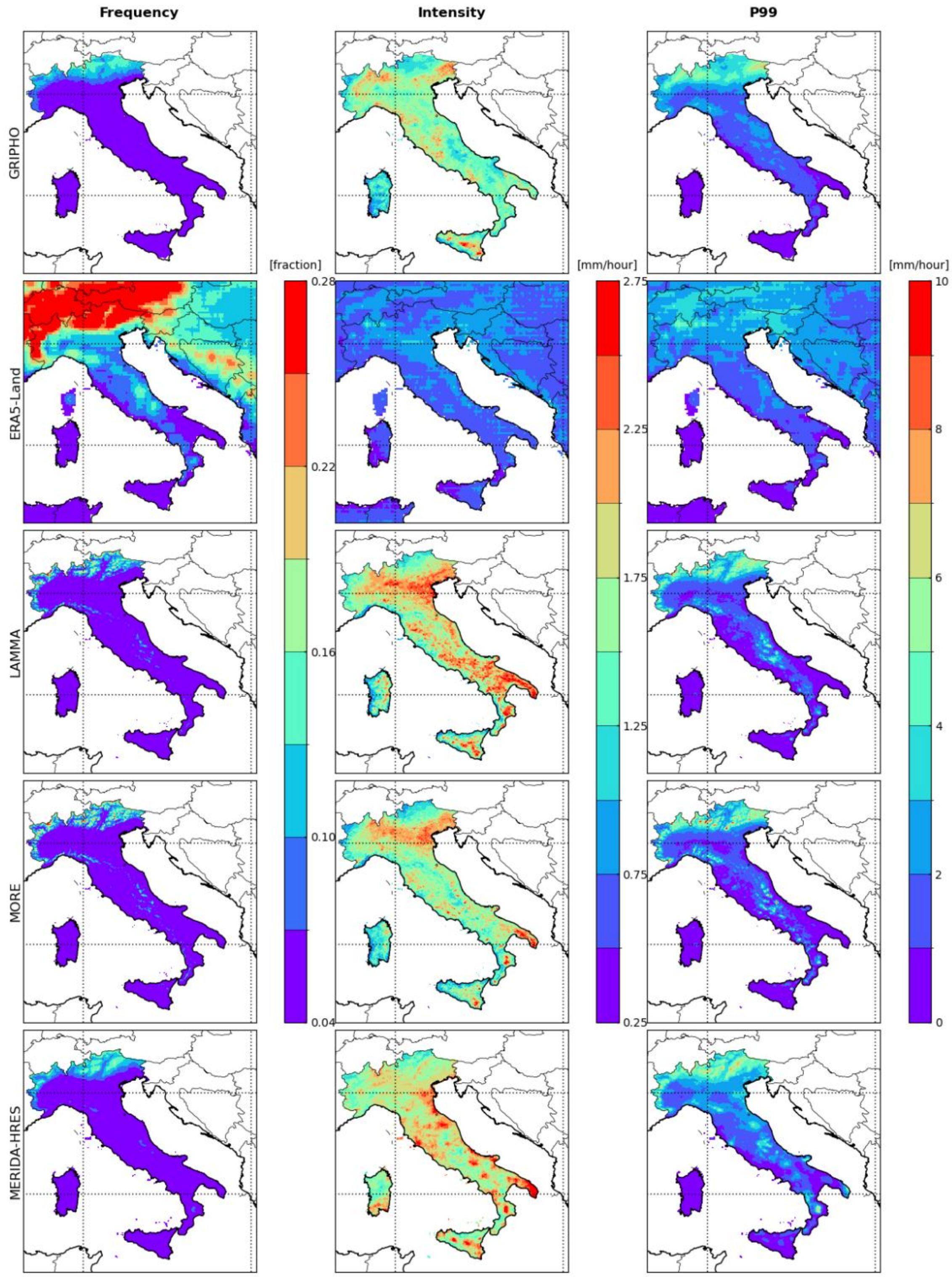

Figure 4: As Fig. 2, but for summer season.

Figures 5, 6, 7 and 8 provide a comprehensive overview of the relative biases associated with all daily and hourly precipitation indices (as defined in Table 2) for the autumn (SON) and summer (JJA) seasons. Figures 5 and 6 depict the relative biases of various daily precipitation indices (see Table 1) with respect to the ARCIS observational dataset, evaluated specifically over the ARCIS domain. Figures 5 and 6 illustrate analogous results for hourly precipitation indices, assessed over the Italian territory using GRIPHO, currently the only available high-resolution gridded hourly precipitation dataset for this region.

For both daily and hourly metrics, the acceptable bias range is to be considered spanning from −5% to +25%, to account for observational uncertainties. The biases shown in Figures 3 and 4 highlight model performance over the ARCIS domain. In the Alpine region, during both SON and JJA seasons, the relative biases in the MORE reanalysis are generally modest and substantially lower compared to those observed in ERA5-LAND and CERRA, particularly during the summer season. The performance of MORE is comparable to that of high-resolution regional reanalysis such as LAMMA and MERIDA-HRES, with biases mostly within the acceptable range.

On average, ERA5-LAND, CERRA, and MERIDA-HRES tend to overestimate mean daily precipitation and wet-day frequency, especially during the JJA season. This overestimation is more pronounced in ERA5-LAND and CERRA, while MORE and LAMMA both based on the MOLOCH model demonstrate improved agreement with observations. Conversely, all reanalysis datasets generally underestimate precipitation intensity and extreme precipitation events, particularly in summer. An exception is MERIDA-HRES, which exhibits a systematic overestimation of heavy precipitation events in both seasons, with a stronger signal during JJA.

In conclusion, MORE exhibits notably improved performance in terms of relative bias. It mitigates the overestimation of wet-hour frequency characteristic of coarser resolution products like ERA5-LAND and CERRA and more accurately reproduces hourly precipitation intensity, especially during the summer season.

We now focus on the relative biases in hourly precipitation, presented in Figures 5 and 6, computed over the Italian territory. As already observed at the daily scale, this behavior becomes even more evident at the hourly scale: the MORE reanalysis exhibits relatively small biases and a clearly improved performance in terms of relative bias, particularly during the summer season.

ERA5-LAND systematically overestimates the frequency of wet hours while underestimating precipitation intensity across almost all seasons. In contrast, MORE and the other higher-resolution datasets show a better ability to reduce wet-hour frequency and to simulate more intense precipitation events (see Figs. 5 and 6).

During summer, ERA5-LAND shows a marked overestimation of wet-hour frequency and a clear underestimation of precipitation intensity. Conversely, the MORE reanalysis slightly underestimates wet-hour frequency but captures both precipitation intensity and extreme events (e.g., heavy precipitation) quite well, showing a clear improvement compared to LAMMA.

Although MERIDA-HRES, similarly to the other high-resolution datasets, generally exhibits small relative biases, it represents an exception during the summer season. In particular, it tends to overestimate wet-hour frequency, especially over southern Italy, and overestimate extreme precipitation events, as indicated by the 99.9th percentile (Fig.6).

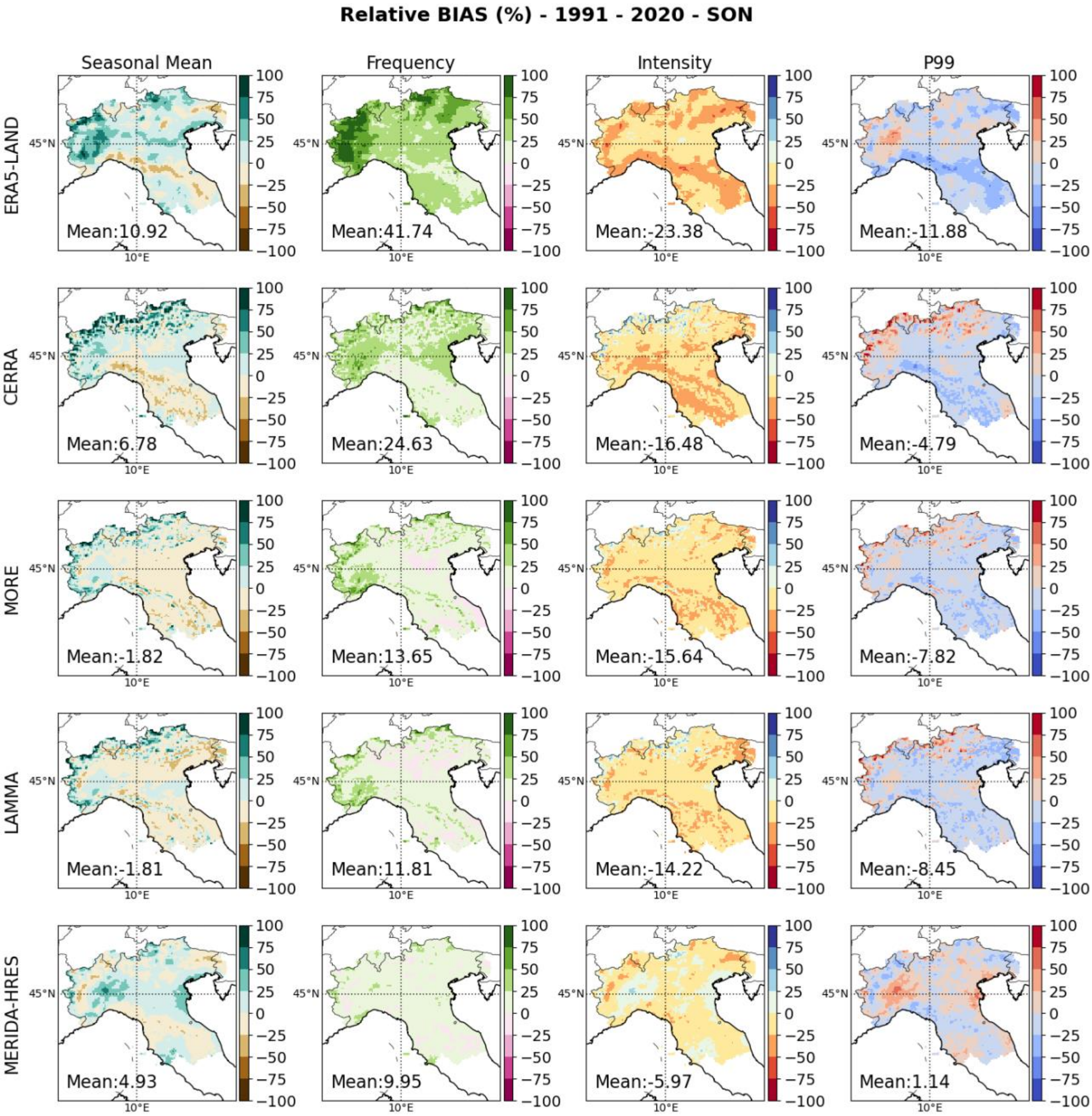

Figure 5: The relative bias of daily precipitation indices in the autumn season (SON). Bias is calculated for each index (Table 2) with respect to ARCIS observations, over the ARCIS observation area, considering all reanalysis datasets listed in Table 1.

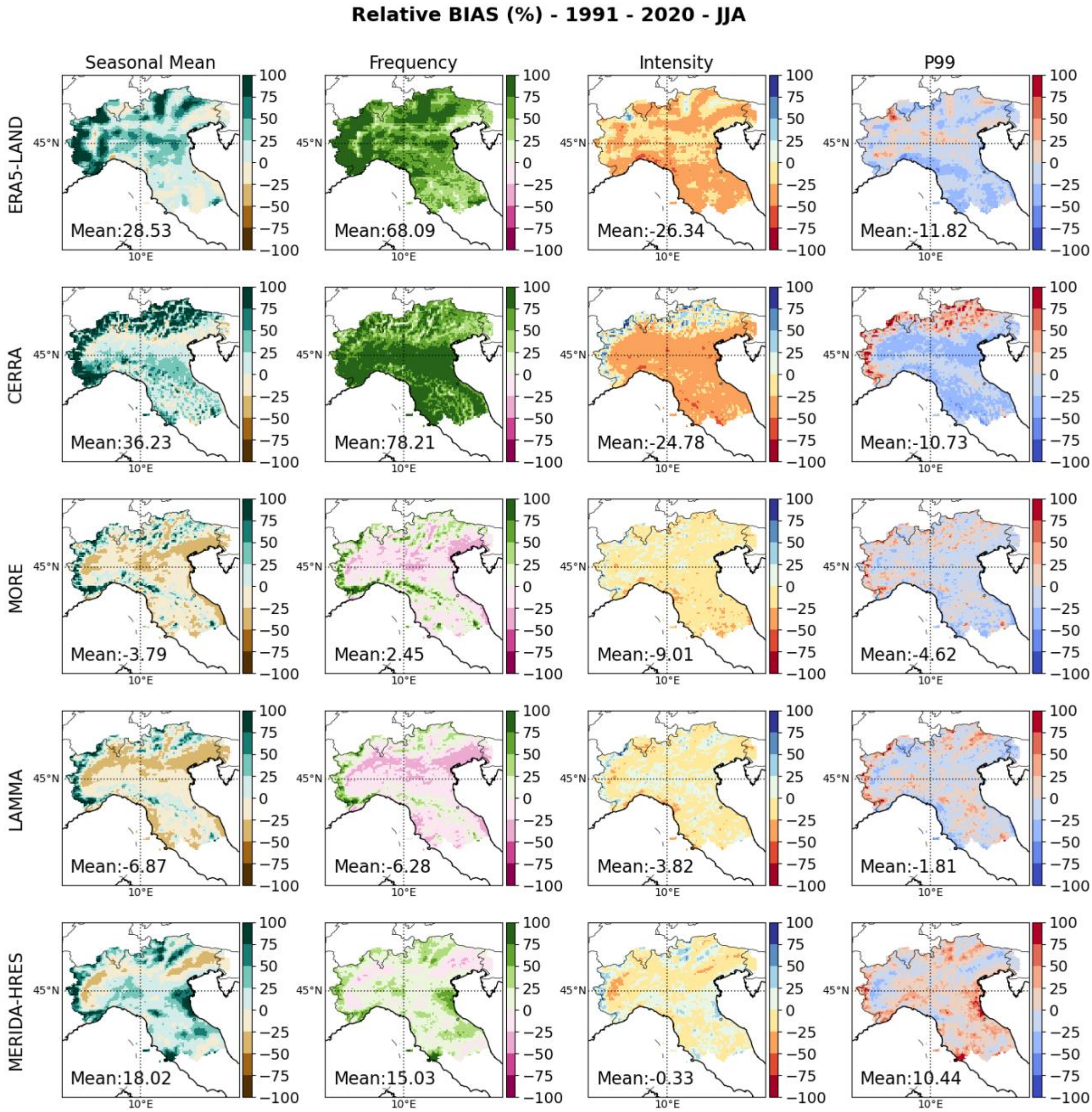

Figure 6: As Figure 5 but for summer season (JJA)

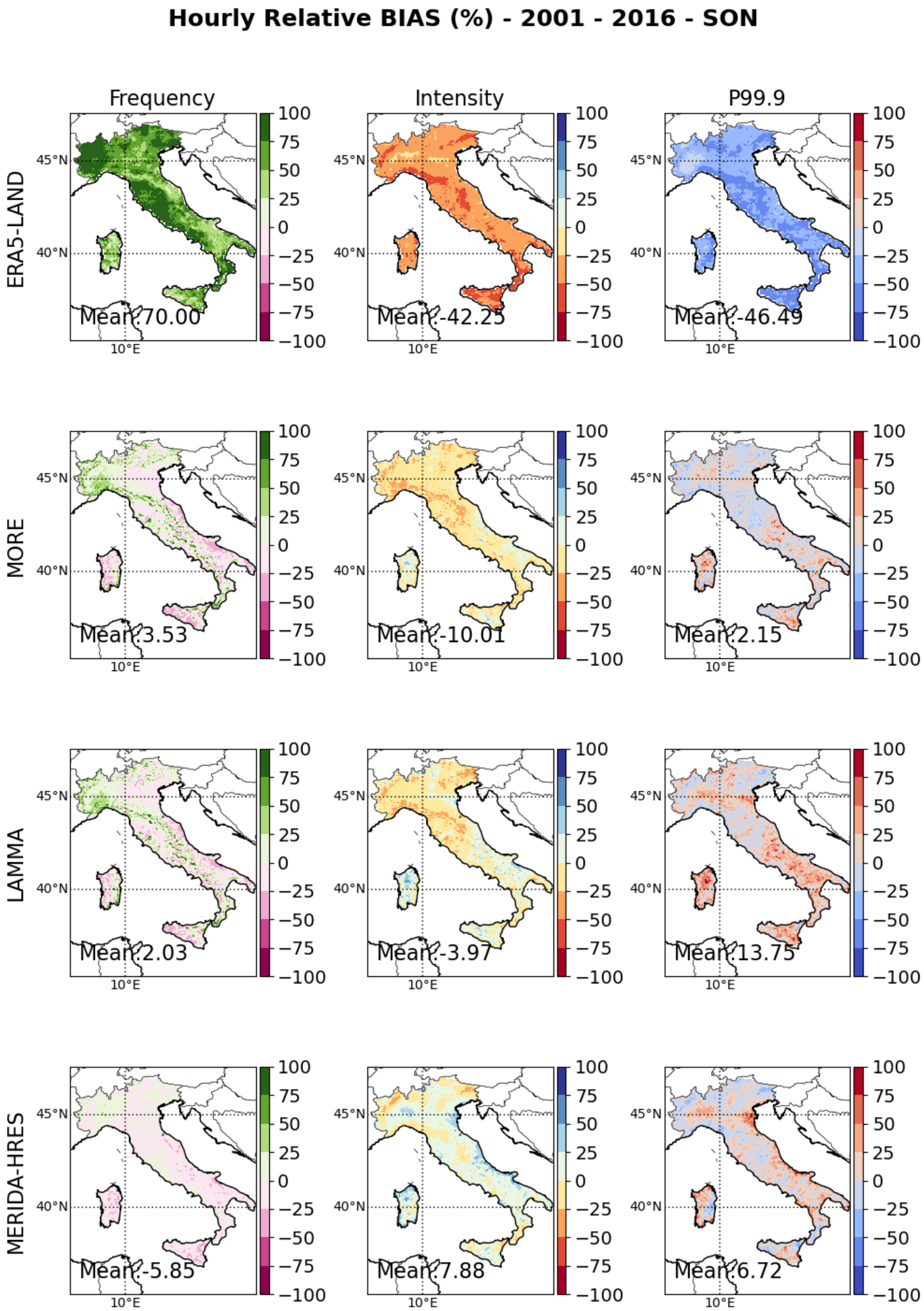

Figure 7: the relative bias of hourly precipitation indices in the autumn season (SON). Bias is calculated for each index (Table 1) with respect to GRIPHO observations, over the GRIPHO observation area, considering all the hourly reanalysis datasets listed in Table 1.

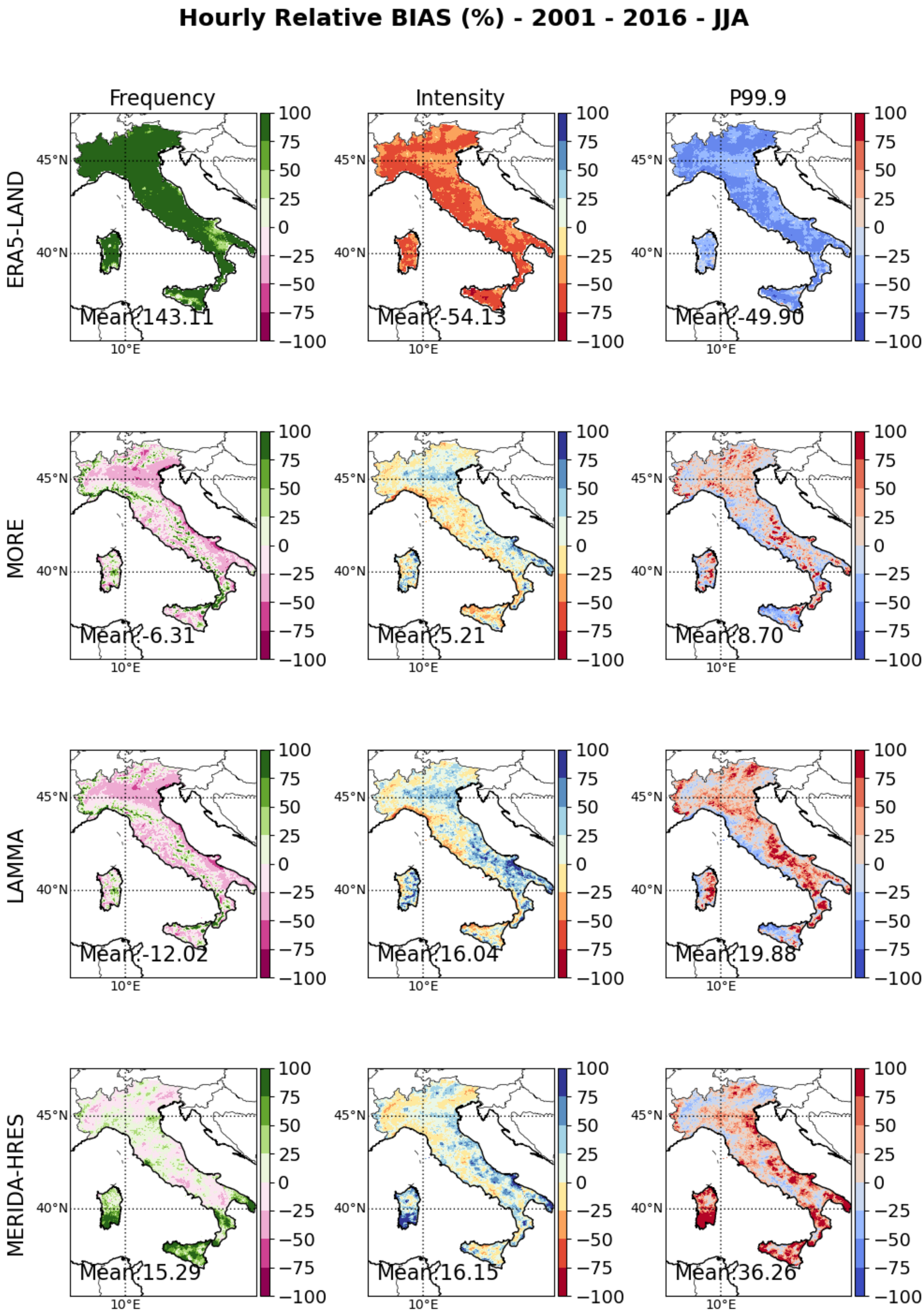

Figure 8: As Figure 7 but for summer season (JJA)

References

Ban, N. et al. The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation. *Clim. Dyn.* **57**, 275–302 (2021)

Caillaud C, Somot S, Alias A, Bernard-Bouissières I, Fumière Q, Laurantin O, Seity Y, Ducrocq V (2021) Modelling Mediterranean heavy precipitation events at climate scale: an object-oriented evaluation of the CNRM-AROME convection-permitting regional climate model. Clim Dyn 56:1717–1752. https://doi.org/10.1007/s00382-020-05558-y

Capecchi, V., Pasi, F., Gozzini, B., & Brandini, C. (2023). A convection-permitting and limited-area model hindcast driven by ERA5 data: Precipitation performances in Italy. Climate Dynamics, 61(3), 1411–1437. https://doi.org/10.1007/s00382-022-06633-2

Cavalleri, F., Lussana, C., Viterbo, F., Brunetti, M., Bonanno, R., Manara, V., Lacavalla, M., Sperati, S., Raffa, M., Capecchi, V., Cesari, D., Giordani, A., Cerenzia, I. M. L., & Maugeri, M. (2024). Multi-scale assessment of high-resolution reanalysis precipitation fields over Italy. Atmospheric Research, 312, 107734. https://doi.org/10.1016/j.atmosres.2024.107734

Cavalleri, F., Viterbo, F., Brunetti, M., Bonanno, R., Manara, V., Lussana, C., Lacavalla, M., and Maugeri, M., 2024. Inter-comparison and validation of high-resolution surface air temperature reanalysis fields over Italy. International Journal of Climatology, 44(8), 2681–2700.

Fantini A (2019) Climate change impact on flood hazard over Italy. Ph.D. thesis, Universita degli Studi di Trieste. http://hdl.handle.net/11368/2940009

Pavan, V., Antolini, G., Barbiero, R. *et al.* High resolution climate precipitation analysis for north-central Italy, 1961–2015. *Clim Dyn* 52, 3435–3453 (2019). https://doi.org/10.1007/s00382-018-4337-6

Stocchi, P.; Pichelli, E.; Torres Alavez, J.A.; Coppola, E.; Giuliani, G.; Giorgi, F. Non-Hydrostatic Regcm4 (Regcm4-NH): Evaluation of Precipitation Statistics at the Convection-Permitting Scale over Different Domains. Atmosphere 2022, 13, 861. https://doi.org/10.3390/atmos13060861