Centro Euro-Mediterraneo sui Cambiamenti Climatici

VHR-REA_IT: Transitioning from COSMO-CLM to ICON-CLM at convection-permitting scales

Technical Notes
October 2024
Issue TN0300

REMHI Regional Models and geo-Hydrological Impacts

By ***Leo Loprieno, Mario Raffa, Angelo Campanale, Paola Mercogliano***
Fondazione CMCC Centro Euro-Mediterraneo sui Cambiamenti Climatici
leo.loprieno@cmcc.it
mario.raffa@cmcc.it
angelo.campanale@cmcc.it
paola.mercogliano@cmcc.it

SUMMARY CMCC works on shifting the focus of its climate modeling activities from the COSMO-CLM (CCLM) to the ICON-CLM (ICLM) climate model. Due to an emerging interest in very-high resolution downscaling, we have extended the dynamically downscaled ERA5 reanalysis at a horizontal grid spacing of ~2.2 km, available on the Data Delivery System (DDS) of the CMCC for the period 1981-2022, to the year 2023, whose results are presented in this report. ICLM reduces the warm bias with respect to E-OBS gridded observations that is especially present in warmer months. Furthermore, ICLM improves the representation of daily accumulated precipitation in spring over territories with a complex orography. In this regard, the evaluation of the two model outputs reveal a generally better agreement of ICLM with gridded observations, indicating an added value of this model with respect to CCLM for the year 2023. To confirm this trend further, however, analyses extended to further years are needed.

Keywords Convection-permitting models, dynamical downscaling, reanalysis-driven

=== Acknowledgments ====
The authors thank the Italian SuperComputing Resource Allocation – ISCRA for the necessary computational resources to run simulations on their supercomputer clusters. The authors acknowledge the collaboration with Prof. Giovanni Chillemi of the Dipartimento Innovazione sistemi Biologici Agroalimentari Forestali of the Università degli Studi della Tuscia in Viterbo, Italy. Prof. Chillemi assisted in requesting code hours at the Consorzio Interuniversitario del Nord-Est per il Calcolo Automatico (CINECA), Italy, which had been necessary in the preliminary stages of the project. The authors are further grateful to the CLM-Community, COSMO consortium and to all those who assisted in and contributed to the constant improvement of the models. Lastly, the authors thank Copernicus for the availability of the data used for the evaluation.

1. BACKGROUND

Very-high resolution climate simulations at kilometer-scale horizontal grid spacings are essential for the analysis of extreme weather events in the future and become increasingly affordable (Kendon et al., 2021). In such simulations, convection is often resolved explicitly, which can have a strong impact on land-atmosphere feedbacks such as the one between soil moisture content and precipitation (Taylor et al., 2013) or the influence on the atmospheric flow of urban environments (Trusilova et al., 2013). The understanding of weather and climate extremes can only be enhanced through an effective evaluation of past events, ideally with in-situ observational data. However, such observations are distributed rather heterogeneously in space and time; in this regard, reanalysis data sets are a means to ensure a certain degree of spatio-temporal continuity of the most relevant climate variables (Raffa et al., 2021). The most recent and thus most plausible reanalysis to investigate the past climate until now is the ERA5 reanalysis (Hersbach et al., 2020), released by the European Centre for Medium-Range Weather Forecasts (ECMWF). Its major improvements with respect to previous versions are the enhanced spatial and temporal resolutions, respectively lying at around 31 km and 1h. ERA5 data are publicly available at the Climate Data Store (CDS) of the Copernicus Climate Change Service (C3S) (Copernicus, 2024).

Among the various opportunities enabled by the availability of ERA5 data, one is the downscaling of ERA5 to support local impact and risk analysis. This approach was implemented in the High Performance Computing to Support Smart Land Services (HIGHLANDER) project (Bottazzi et al., 2024) to provide an hourly, convection-permitting scale reanalysis over Italy (Raffa et al., 2021). This new additional gridded data set over Italy labeled as Very High Resolution REAnalysis for ITaly (VHR-REA_IT) was created (Raffa et al., 2021), obtained by dynamically downscaling ERA5 reanalysis to a horizontal grid spacing of 2.2 km for the period 1981-2020.

Adinolfi et al. (2023) have performed an evaluation of the downscaled reanalysis by Raffa et al. (2021) and have shown that the VHR-REA_IT data set is capable of reproducing spatial patterns of observations of mean temperature and precipitation, especially over regions with a complex topography. However, a warm bias over flat topography is still rather usual in Convection Permitting Regional Climate Models (CPRCMs), as confirmed in Sangelantoni et al., (2024). Furthermore, good performance is reached by the implemented urban parameterizations to identify urban heat islands as a local minimum temperature in summer and in the representation of extreme events through a set of climate indicators (Adinolfi et al., 2023).

Due to an emerging interest in very-high resolution downscaling, the CMCC has continued to update these data even after the end of the HIGHLANDER project. As yet, these have been extended to the year 2023 and are expected to be constantly updated. Additionally, in the context of the transition from COSMO-CLM (CCLM, Steppeler et al. 2003; Rockel et al., 2008) to ICON-CLM (ICLM, Zängl et al. 2015; Pham et al., 2021), the CMCC decided to update these data using ICLM. Prior to this, a comparative evaluation of the two models, both ERA5 driven, for the year 2023 was conducted and the main results are presented in this work.

1.1. SUPERCOMPUTER CLUSTERS AND THE DATA DELIVERY SYSTEM (DDS)

The simulations were performed by the Centro euro-Mediterraneo sui Cambiamenti Climatici (CMCC) exploiting its own Supercomputing Center (SCC) clusters, as well as the cluster of the Consorzio Interuniversitario del Nord-Est per il Calcolo Automatico (CINECA). More precisely, within the context of the HIGHLANDER project, CCLM was run on CINECA's *Galileo100 (G100)* cluster, whereas for the extension to the year 2023, COSMO-CLM was run on CMCC SCC's *ZEUS* cluster. *G100* is made up of 528 Intel biprocessor computing nodes with 24 cores each; *ZEUS* is based on 348 Lenovo

biprocessor nodes all interconnected by means of an Infiniband EDR network, exhibiting a computing power of 1.202 TFlops.

ICLM was run on CMCC's *JUNO* cluster, which was installed at CMCC's headquartes in Lecce in 2022. It has a computing power of around 1.134 TFlops and is based on 3rd Generation Intel Xeon Scalable processors as well as on the NVIDIA A100 GPU type.

The CCLM simulations are already available at CMCC's DDS (https://dds.cmcc.it/#/dataset/era5-downscaled-over-italy/hourly), while the ICLM data will be released in the near future. For more detailed information about the DDS, the reader is referred to the following link: https://www.cmcc.it/it/articolo/dati-climatici-aggiornati-e-di-alta-qualita-per-tutti.

This report is structured as follows: Sect. 2 provides a brief description of the models on which the simulations were run and the reference data set involved, Sect. 3 continues with a presentation of the model's performances in simulating key atmospheric variables over the VHR-REA_IT domain, while Sect. 4 emphasizes the most relevant aspects legitimating a transition from the CCLM to the ICLM model.

2. DATA AND METHODS

This work presents a comparison of the performances of CCLM and ICLM in simulating reanalysis-driven very-high resolution climate data over the Italian territory (VHR-REA_IT, Raffa et al. (2021)).

2.1. MODELS AND SIMULATIONS

The simulations have a grid spacing of roughly 0.02° (2.2 km) and were forced by ERA5 reanalysis data. In the high-resolution models, the land surface is represented by the TERRA-ML scheme (Heise et al., 2006), with the TERRA_URB module tailored for

urban environments switched on (Wouters et al., 2015, 2016). This scheme provides a robust canopy-dependent urban physics description with a low computational cost (Trusilova et al., 2016). Below are reported the main features of the high-resolution model setups in more detail. For a summary of the main differences in parameterizations between the two models, the reader is referred to the following table

Model	ICON-CLM	COSMO-CLM
Version	v2.6.7	v5.0.9
Boundary forcing	ERA5 (≃31 km)	ERA5 (≃31 km)
Lateral Boundary Condition (LBC)	update frequency 3h	update frequency 3h
Soil initializations	Temperature and moisture obtained by interpolation from ERA5	Temperature and moisture obtained by interpolation from ERA5
Horizontal resolution	0.0225° (≃2.5 km)	0.02° (≃2.2 km)
Time step	20 s	20 s
N° vertical levels	65 (top model level elevation = 22 km)	60 (top model level elevation = 22 km)
Output frequency	1 h	1 h
Coordinate system	– horizontal: icosahedral grids – vertical: terrain-following Gal-Chen height coordinate (Simmons and Burridge, 1981) and exponential height coordinate (SLEVE) (Schär et al., 2002; Leuenberger et al., 2010)	– horizontal: rotated geographical (lat–lon) – vertical: terrain-following Gal-Chen height coordinate (Gal-Chen and Somerville, 1975) and exponential height coordinate (SLEVE) (Schär et al., 2002)

Radiation scheme	ecRad (Hogan and Bozzo, 2018)	(Ritter and Geleyn, 1992)
Non-orographic gravity wave drag	Wave dissipation at critical level (Orr et al., 2010)	-
Sub-grid scale orographic drag	SSO scheme (Lott and Miller, 1997; Schulz, 2008)	SSO scheme (Lott and Miller, 1997; Schulz, 2008)
Microphysics	One-moment scheme (Doms et al., 2011; Seifert, 2008)	One-moment scheme (Doms et al., 2011, Baldauf and Schulz, 2004)
Convection scheme	Mass flux shallow and deep convection based on Tiedtke-Bechtold (Tiedtke, 1989; Bechtold et al., 2008)	Shallow convection: reduced Tiedtke scheme for shallow convection only (Tiedtke, 1989)
Turbulent transfer	Prognostic TKE based on Raschendorfer (Raschendorfer, 2001)	Prognostic TKE closure (Doms et al. 2011)
Land surface scheme	Tiled TERRA (Schrodin and Heise, 2001; Schulz et al., 2015; Schulz and Vogel, 2020) with TERRA_URB parameterization	TERRA (Schrodin and Heise, 2001; Doms et al., 2011; Schulz et al., 2015) with TERRA_URB parameterization
Land use dataset	GlobCover2009 (Arino et al., 2012)	GlobCover2009 (Arino et al., 2012)

2.1.1. COSMO-CLM

The COnsortium for Small-scale MOdeling Consortium (COSMO) model (Steppeler et al., 2003) is specifically designed for simulations on meso-β to meso-γ scales, i.e. on scales of 2-200 km, and is a non-hydrostatic limited-area model. Vertical advection is simulated by the means of an implicit Crank-Nicholson scheme and horizontal advection using a fifth-order upwind scheme (Baldauf et al., 2011). Additionally, the one-dimensional scheme by (Bott, 1989) is implemented for the multidimensional advection of scalar fields. The model uses rotated geographical and terrain-following height coordinates (Doms and Baldauf, 2011) and the thermodynamical equations are computed on a three-dimensional Arakawa-C grid (Arakawa, 1977). Furthermore, the model implements a split-explicit third-order Runge-Kutta time-stepping scheme (Wicker and Skamarock, 2002). The parameterization of the radiative transfer is based on the δ-two-stream-approach by (Ritter and Geleyn, 1992) and a single-moment bulk cloud-microphysics scheme with five different species (i.e. cloud water, cloud ice, rain, snow, and graupel) is implemented (Reinhardt and Seifert, 2006). To represent turbulence, the model uses a scheme using a turbulent kinetic energy closure for the surface and the planetary boundary layer (Raschendorfer, 2001). Shallow convection is parameterized using the scheme by (Tiedtke, 1989), while deep convection is resolved explicitly. Aerosols are imposed using the AeroCom climatology (Kinne et al., 2006).

2.1.2. ICON-CLM

The development of ICON has been pursued by the collaboration of the German Meteorological Service (DWD), the Max Planck Institute for Meteorology (MPI-M), the German Climate Computing Centre (DKRZ), and the Karlsruhe Institute of Technology (KIT) for establishing a significant advancement in Numerical Weather Prediction (NWP) and climate modeling. It has been demonstrated that ICON is able to provide better conservation properties with respect to previous models, a better scalability on parallel

highperformance computers and the possibility of performing a static mesh refinement (Zängl et al., 2015, 2022; Crueger et al., 2018; Giorgetta et al., 2018). The ICON model offers both a global and a local implementation, the ICON Local Area Model (ICLM), used in this work. In the ICLM model, the governing equations are spatially discretized over an icosahedral-triangular Arakawa-C grid, in the horizontal, and a terrain following height-based SLEVE coordinate formulation (Leuenberger et al., 2010) with Lorenz-type staggering, in the vertical. Each grid of the model is defined with the acronym RnBk, denoting a grid created from an icosahedron whose edges have been initially divided into n parts, followed by k successive edge bisections. The entire set of governing equations is based on the fully compressible non-hydrostatic system for the description of a two-component system (dry air and all the three water phases). The temporal discretization relies on an explicit two-time-level predictor corrector scheme, whereas the terms concerning vertical sound-wave propagation are integrated using an implicit approach. Zängl et al. (2015) found that a particular advantage of the nonhydrostatic fully compressible dynamical core over an icosahedral grid is not only the exact mass conservation and a better tracer mass consistency but also its higher numerical stability over steep slopes, like the steeper mountains of the Alpine domains. Recently, ICLM has been used for regional simulation over Europe with several configurations, both in climate mode (Pham et al., 2021; Chen et al., 2024) and weather mode (Sakradzija and Klocke, 2018).

2.2 E-OBS GRIDDED OBSERVATIONS

The ability of the newly run simulations to reproduce observed fields was assessed by comparing them to the E-OBS data set (Cornes et al., 2018).

E-OBS come as gridded observations of a variety of atmospheric variables covering the period from 01/01/1950 to the present at horizontal resolutions of 0.25° and 0.1° at daily time steps over Europe. This allows for analyses of model performances at

relatively small spatio-temporal scales also in regions that might suffer from a rather scarce coverage of in-situ observations. The simulations have been evaluated by the means of the following variables at a horizontal resolution of 0.1°: mean, minimum, and maximum air temperature at 2m (*TG*, *TN*, and *TX*, resp.), daily accumulated precipitation (*RR*), and the mean wind speed at 10m (*FG*). All variables have been downloaded from the C3S CDS (Copernicus Climate Change Service, 2024). *TG*, *TN*, *TX*, and *RR* have been taken from the latest E-OBS version at the time of analysis, i.e. v29.0e. v29.0e covers data from 1950-01-01 to 2023-12-31. *FG* was not yet available in the latest E-OBS version and thus had to be taken from the previous version v28.0e. As a consequence, the results for *FG* do not cover the whole year 2023, as v28.0e only covers the period from 1950-01-01 to 2023-06-30. With respect to its predecessor versions, v28.0e includes new series for the Campania and Trentino regions, apart from updates for the Emilia-Romagna region (European Climate Assessment Dataset (ECAD), 2024). With respect to v28.0e, v29.0e includes further updates for the Emilia-Romagna region (European Climate Assessment Dataset (ECAD), 2024).

3. RESULTS

This section will report a brief evaluation of the performance of how well CCLM and ICLM represent the air temperature at 2m (Sect. 3.1), the daily accumulated precipitation (Sect. 3.2), and the mean wind speed (Sect. 3.3.). Each subsection starts with a time series displaying the corresponding variable averaged over all land grid points in Italy (see Fig. 1), continues with spatial maps visualizing the monthly difference between ICLM and CCLM, followed by a visualization of the monthly bias with respect to E-OBS v28.0e (for *FG*) and v29.0e (for *TG*, *TN*, *TX*, and *RR*) gridded observations, and concludes with a comparison of the value distributions within E-OBS and the two models. For the spatial maps, monthly biases have been preferred over seasonal ones.

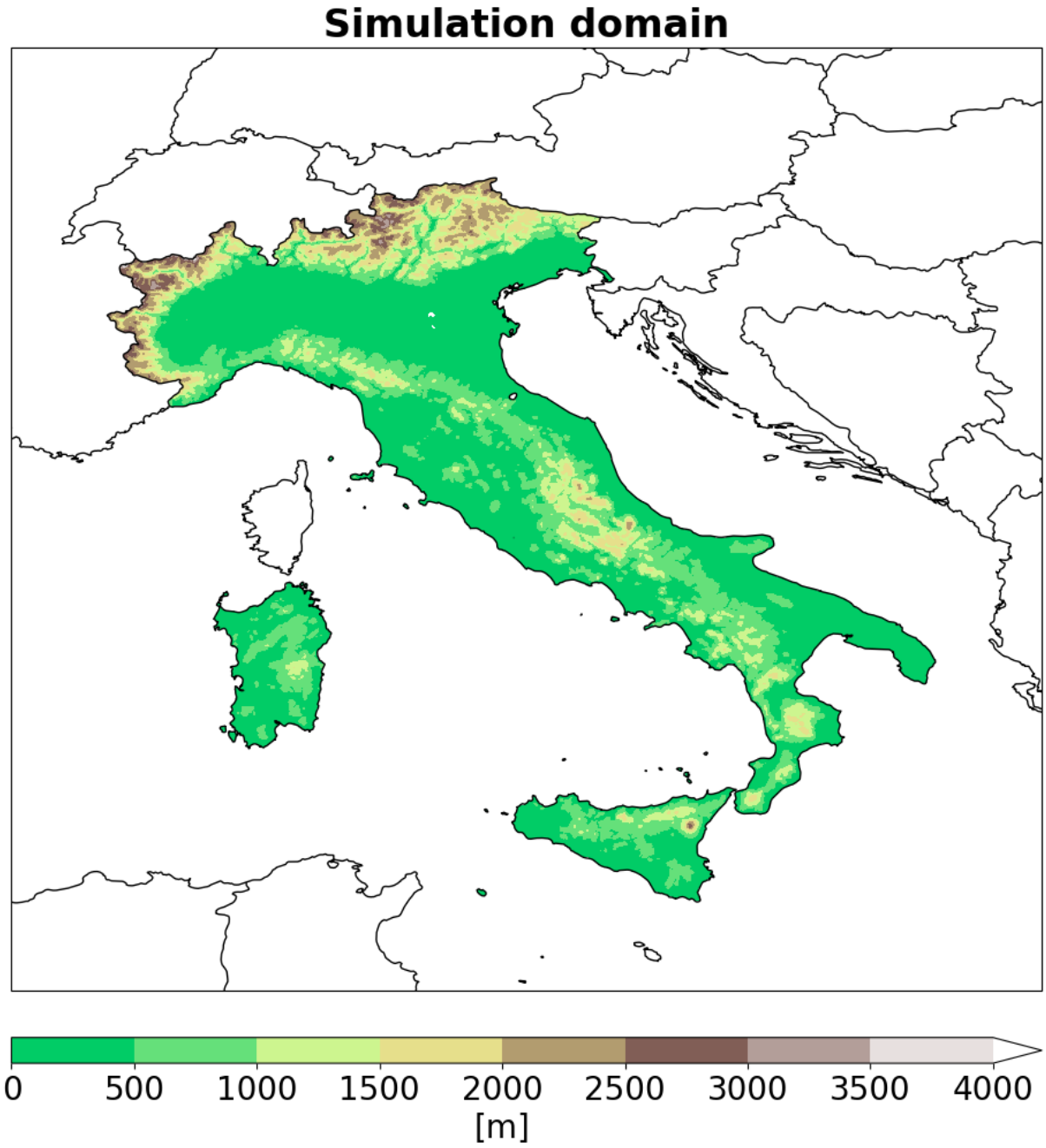

Figure 1: Orography as simulated by both models. The colored grid cells correspond to the points used throughout the analysis. Red dots denote stations used as input for the E-OBS 2m-air temperature data set.

3.1. 2M-AIR TEMPERATURE

The following analysis considers the monthly bias of the daily mean (Sect. 3.1.1), the daily minimum (Sect. 3.1.2), and the daily maximum temperatures (Sect. 3.1.3).

3.1.1. DAILY MEAN TEMPERATURE

Figure 2 shows the time series of the daily mean air temperatures at 2m averaged over all Italian land grid cells as represented by E-OBS gridded observations (black line), by CCLM (blue line), and ICLM (red line). Overall, the models are capable of representing this variable well, especially in the winter season, though some small negative bias persists. In doing so, the bias is reduced in the ICLM model. A more distinct positive bias can be seen in the warmer months, where it can reach up to +4 K.

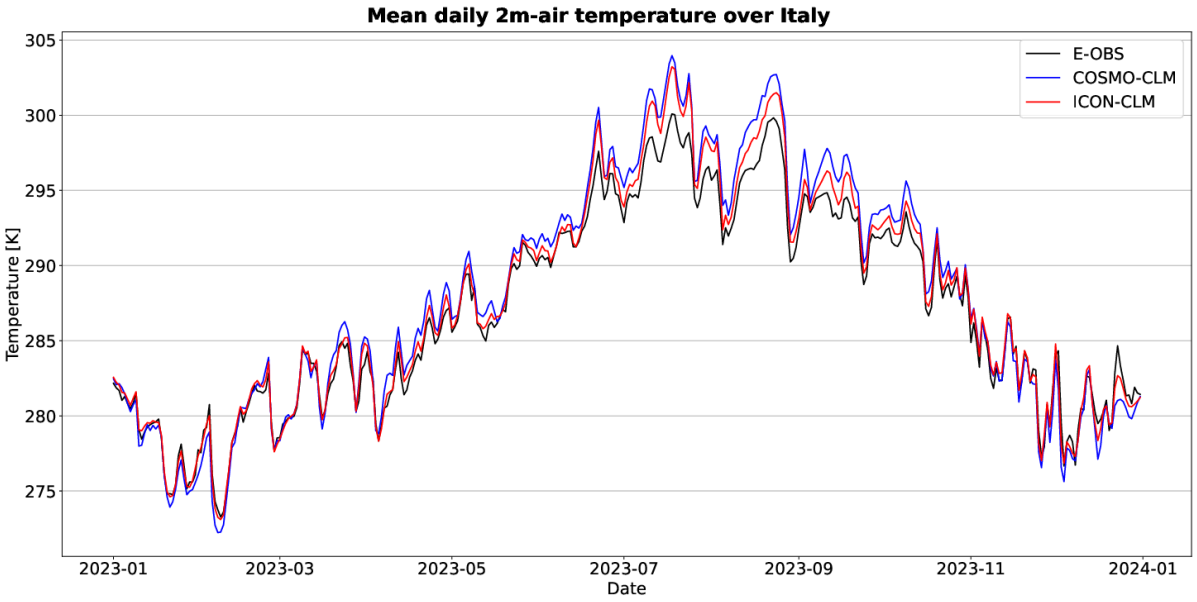

Figure 2: 2m-air temperature series of E-OBS v29.0e (black), CCLM (blue), and ICLM (red) averaged over Italy (see Fig. 1).

Figure 2 further reveals that the deviations between the two models are never as large as between the one single model and the gridded observations. The inter-model difference for each month is shown in Fig. 3. It can be seen that ICLM is distinctly colder than CCLM throughout large parts of the year, except in the months January-February and November-December. The temperature over the Alps, however, are constantly higher in ICLM. On the other hand, a small area in northwestern Italy is simulated slightly colder in ICLM than in CCLM.

Figures 4 and 5 show the monthly bias with respect to E-OBS data of the mean air temperature at 2m as simulated by the CCLM and ICLM models, respectively. Throughout the whole year, in the CCLM model there is a constant cold bias visible in northeastern Italy (Fig. 4). Conversely, a warm bias can be seen in all months in northeastern Italy at the transition to the Po Valley. These biases are also adopted in the ICLM model, however taking the local topographic features more into account (Fig. 5). Both models show a relatively large overestimation of the temperature especially in the months July and August over the Po Valley, which yields a very flat topography, and Sicily, where the positive bias exceeds 4 K. In the winter months, on the other hand, a distinct cold bias is visible over large parts of the country, most pronounced in southern and northwestern Italy. The trends observable for the different months are kept by the ICLM model, albeit to an overall lesser extent (Fig. 5).

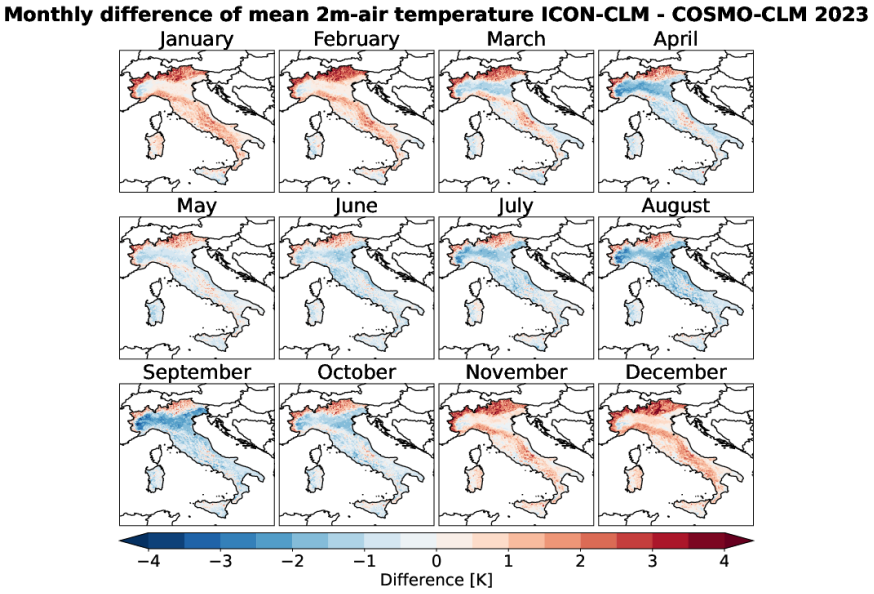

Figure 3: Difference in mean 2m air temperature for each month between ICLM and CCLM.

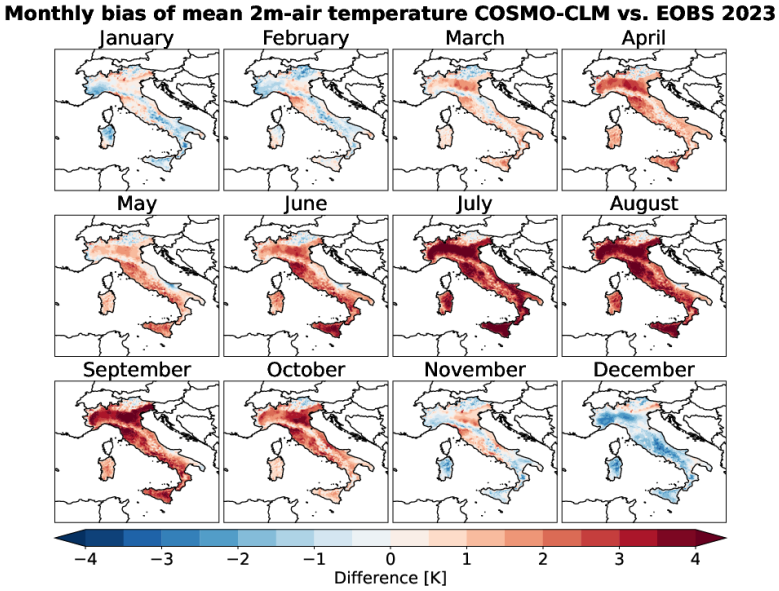

Figure 4: Bias of the air temperature at 2m in CCLM for each month of the year 2023. Reference: E-OBS v29.0e.

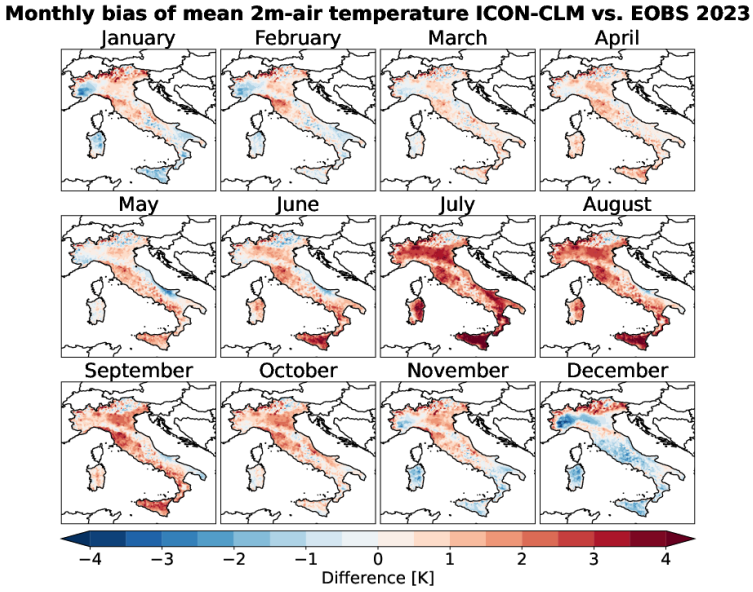

Figure 5: As Fig. 4, but for ICLM.

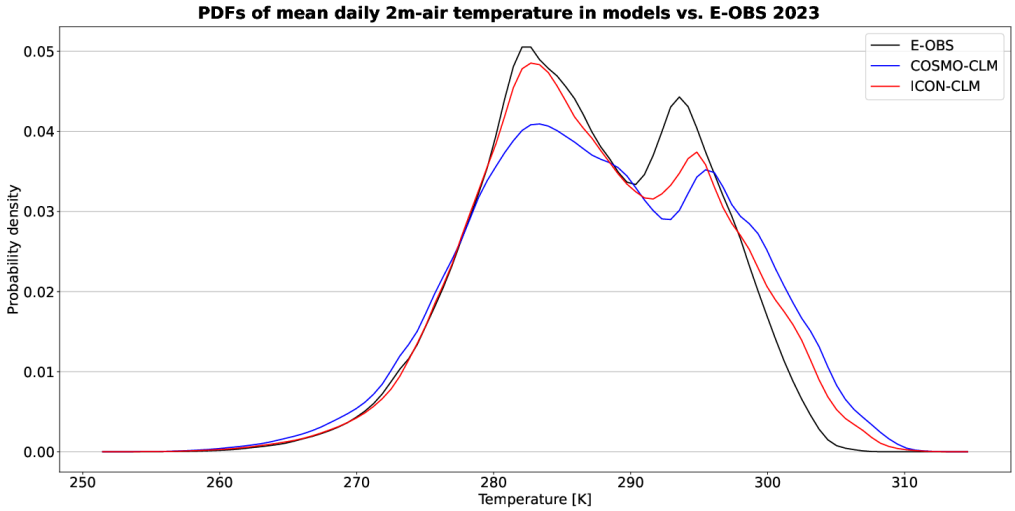

Figure 6: Probability density functions (PDFs) of daily averaged 2m-air temperature of E-OBS v29.0e (black), CCLM (blue), and ICLM (red). Only land grid cells (see Fig. 1) were considered in the calculation.

The better performance of ICLM is consistent when comparing the probability density functions (PDFs) of the observations and the models. Figure 6 shows the PDFs of E-OBS data (black line), the CCLM model (blue line), and the ICLM model (red line). The distribution considers daily values of land grid cells over Italy, corresponding to the colored grid cells in Fig. 1. In all three cases, the majority of the values are focused on two temperatures, likely representing the mean summer and winter temperatures, respectively. The first peak, captured by both observations and models, lies at temperatures of around 281 K. The second peak, on the other hand, is slightly overestimated by the models; while it lies around temperatures of 292 K for E-OBS, in the models it appears around 294 K (ICLM, red) and 295 K (CCLM, blue) (Fig. 6). The reduced density of the peaks in both models compared to E-OBS come at the expense of too many warm values, going as far as 310 K.

3.1.2. DAILY MINIMUM TEMPERATURE

Figure 7 displays the time series of the daily minimum temperatures at 2m, again averaged over Italy considering only land grid cells. The black line represents the E-OBS series, while the blue and red lines stand for the CCLM and ICLM time series, respectively. Though the agreement of the models with E-OBS data largely remains, in some months there is a slightly larger bias from what can be observed in the mean temperature (Fig. 2). The most distinct difference appears in June, where in some cases the bias adds up to 4 K (CCLM, blue) and 2 K (ICLM, red), respectively. The best agreement between observations and models is reached in fall and December.

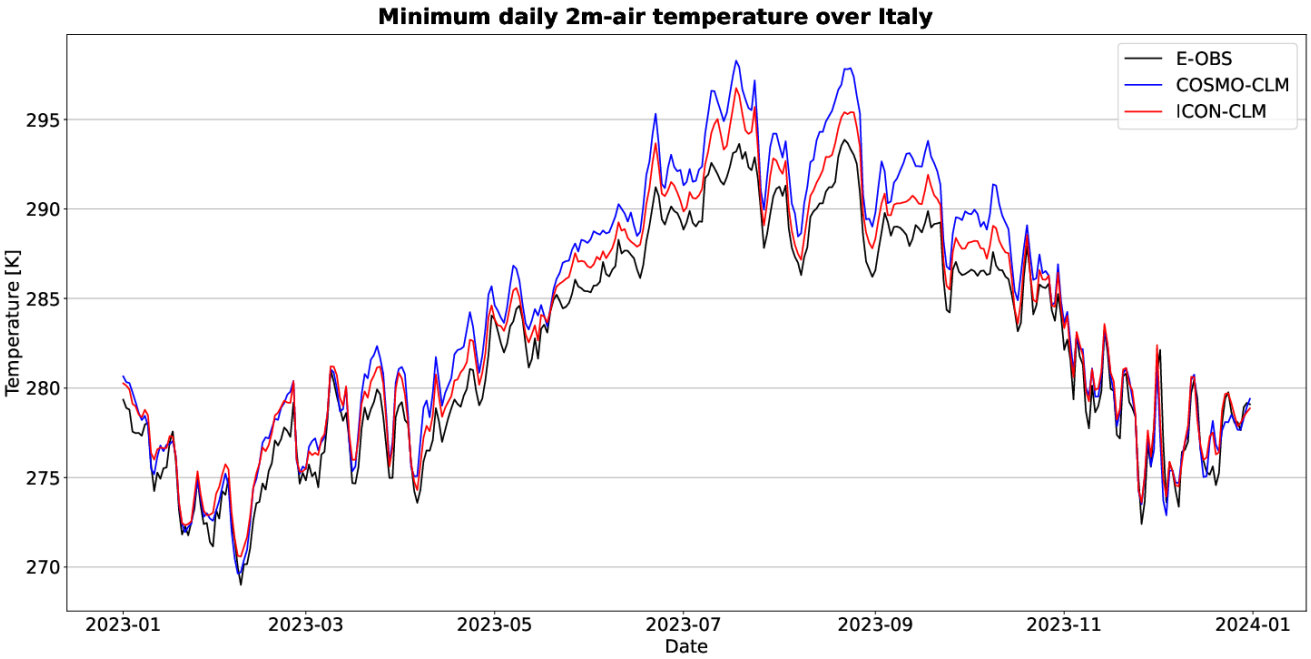

Figure 7: Daily minimum 2m-air temperature series of E-OBS v29.0e (black), CCLM (blue), and ICLM (red) averaged over Italy (see Fig. 1).

Figure 8 largely confirms the trends visible in Fig. 3, whereupon the months April-October appear as much warmer in CCLM than in ICLM. In August in northwestern Italy, ICLM is up to 4K colder than CCLM. The bias between the models and E-OBS, especially present in the warmer months, can also be identified looking at the monthly values. Figures 9 and 10 display the monthly mean bias of the minimum air temperature at 2m with respect to E-OBS gridded observations of the CCLM and ICLM model, respectively. While there is no region that experiences a constant cold bias throughout all months, there is a constant overestimation of the minimum temperature in northeastern Italy in both the CCLM and ICLM model (Figs. 9 and 10, resp.). Especially in the CCLM model (Fig. 9), the bias is strongest throughout the months June-October, more precisely above the Po Valley and in central Italy, except for the areas in the Apennines. This anomaly appears spatially reduced in the ICLM model (Fig. 10). Furthermore, as opposed to the mean temperature, the minimum temperature does not show a clearly negative bias

affecting the whole country. Yet, a feature that is observable throughout all months only in the ICLM model (Fig. 10), is a cold bias over northwestern Italy, which is not always present in the mean temperature (Fig. 5). The strongest cold bias in the CCLM model, on the other hand, is visible in December over the central Apennines (Fig. 9).

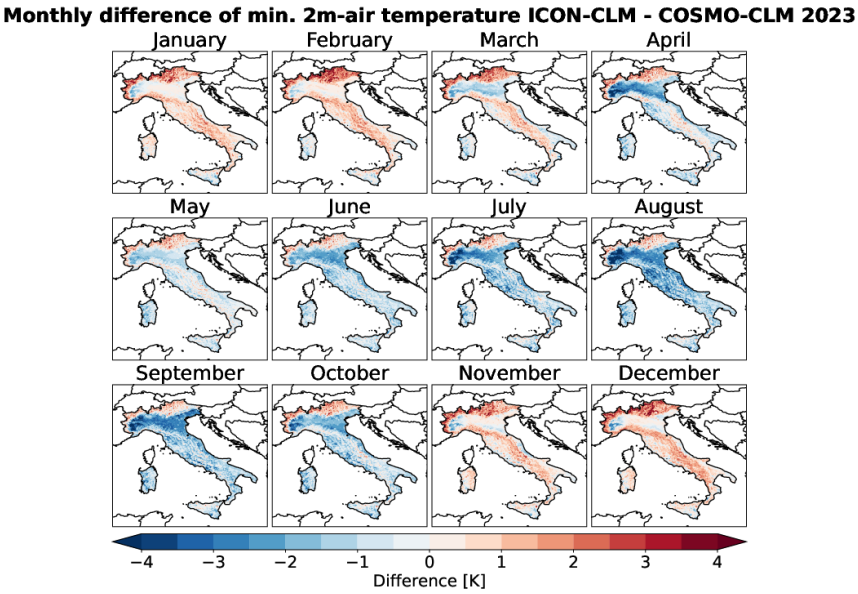

Figure 8: Difference in minimum 2m air temperature for each month between ICLM and CCLM.

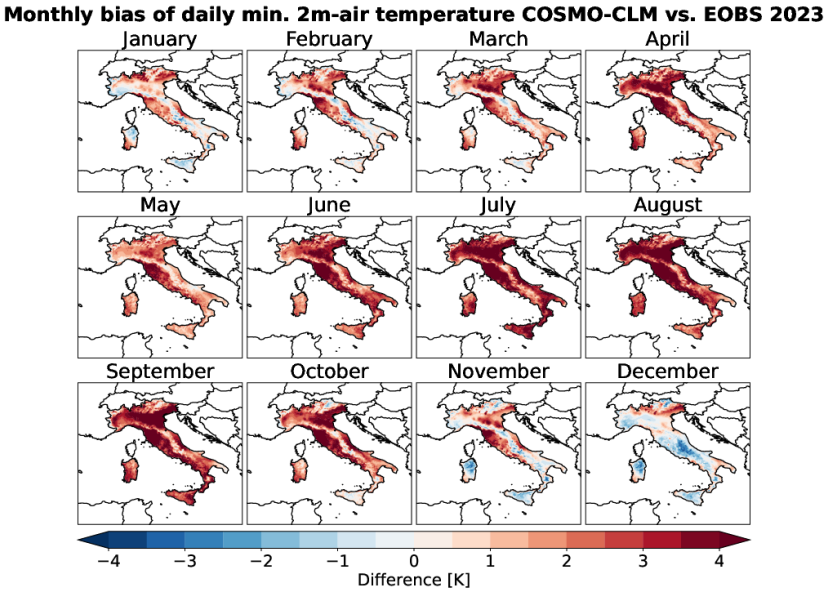

Figure 9: Monthly bias of the daily minimum air temperature at 2m in CCLM for each month of the year 2023. Reference: E-OBS v29.0e.

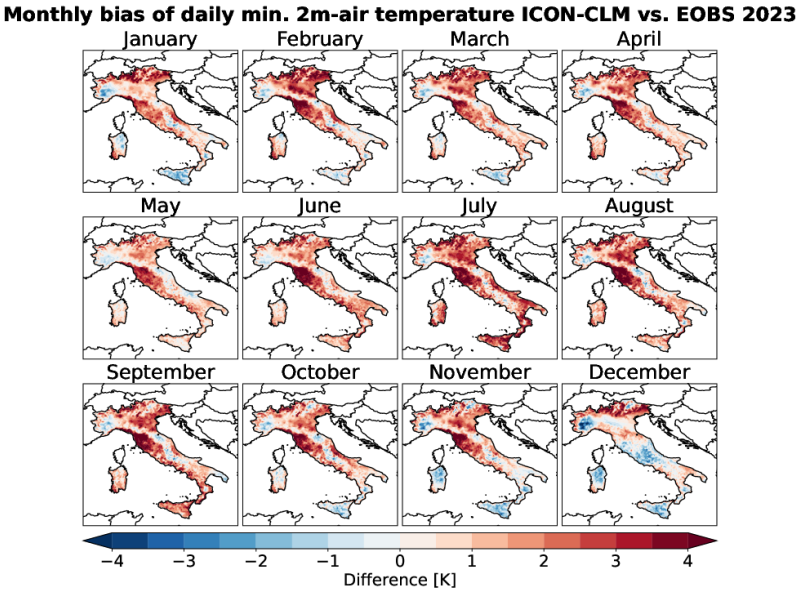

Figure 10: As Fig. 9, but for ICLM.

Figure 11 displays the PDFs of the daily minimum temperatures present in E-OBS (black), the CCLM (blue), and the ICLM model (red), and shows a slight warm bias throughout the whole set of temperature values. The overall shapes of the curves, however, are close to the one for the daily mean temperature (Fig. 6). The distribution of the values in the gridded observations (Fig. 11, black) is close to normal with two similar peaks situated at temperatures of around 279 and 289 K. Both models vaguely imitate the shape, whereupon the distribution in ICLM (Fig. 11, red) is slightly closer to the reference distribution than the one in CCLM (Fig. 11, blue). In ICLM, the majority of values lie around 280 K, while in CCLM they lie in the range 277-282 K and around 292 K. Furthermore, as opposed to E-OBS, the data of both models yield values clearly exceeding 300 K.

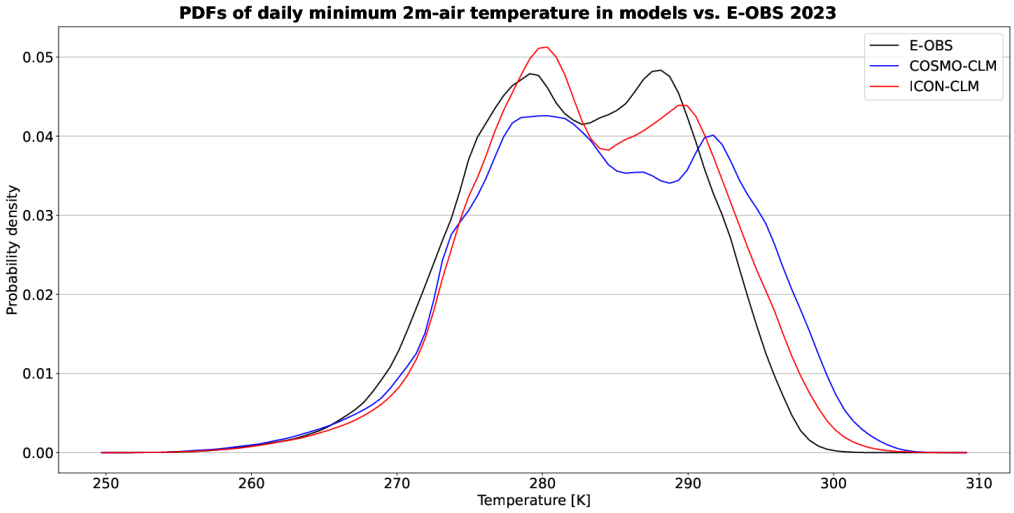

Figure 11: As Fig. 6, but for the daily minimum temperature.

3.1.3. DAILY MAXIMUM TEMPERATURE

Figure 12 visualizes the time series of the daily maximum temperatures at 2m, averaged over all land grid cells over Italy. The agreement between model and

observational data is slightly higher in this variable than in the daily minimum temperature (Fig. 7), being more alike to the series of the mean temperature (Fig. 2). The only period where the simulated values exhibit a clear bias with respect to E-OBS is again the summer season, whereupon biases of 4-5 K are reached. Furthermore, the trends of the two models are somewhat closer in this case than they are in the mean and maximum temperature. The best agreement between the models and gridded observations occurs again during the fall period.

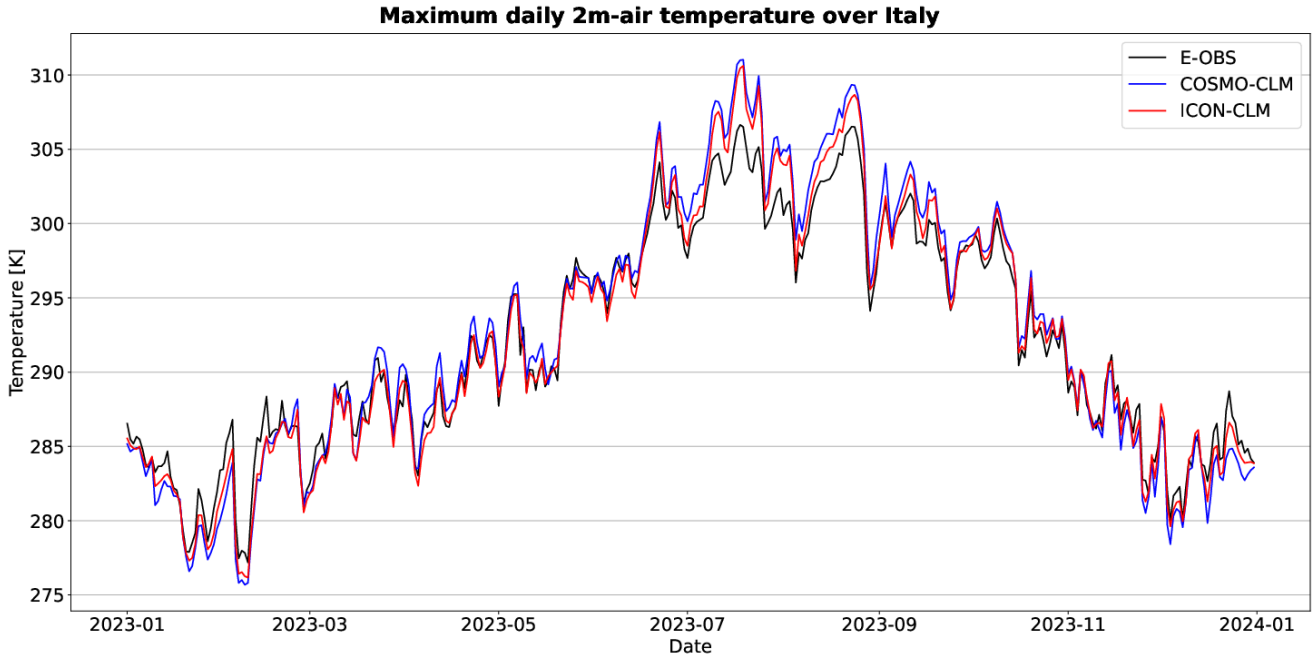

Figure 12: Daily maximum 2m-air temperature series of E-OBS v29.0e (black), CCLM (blue), and ICLM (red) averaged over Italy (see Fig. 1).

The reduced anomaly with respect to E-OBS is further visible in the spatial distribution of monthly temperature values. The difference in maximum temperature between ICLM and CCLM (Fig. 13) is of similar magnitude as the one visible in the difference in mean temperature (Fig. 3). Figures 14 and 15 show the bias of the daily maximum temperature against E-OBS for each month as simulated by the CCLM and ICLM models, respectively. In CCLM, the Alpine regions close to the Austrian border

yield a slight cold bias in all months (Fig. 14). In both models, the warming is most pronounced in July and August in the Po Valley and Sicily, where it is larger than 4 K (Figs. 14 and 15). In CCLM, the spring months have a relatively low overall bias (Fig. 14). Furthermore, both CCLM and ICLM show a distinct cold bias in January and December over northwestern Italy. In contrast to CCLM, however, ICLM has a relatively distinct cold bias that extends across large parts of the country in the months January-May (Fig. 15), whereas in CCLM this cold bias is limited to the months January-February and November-December (Fig. 14).

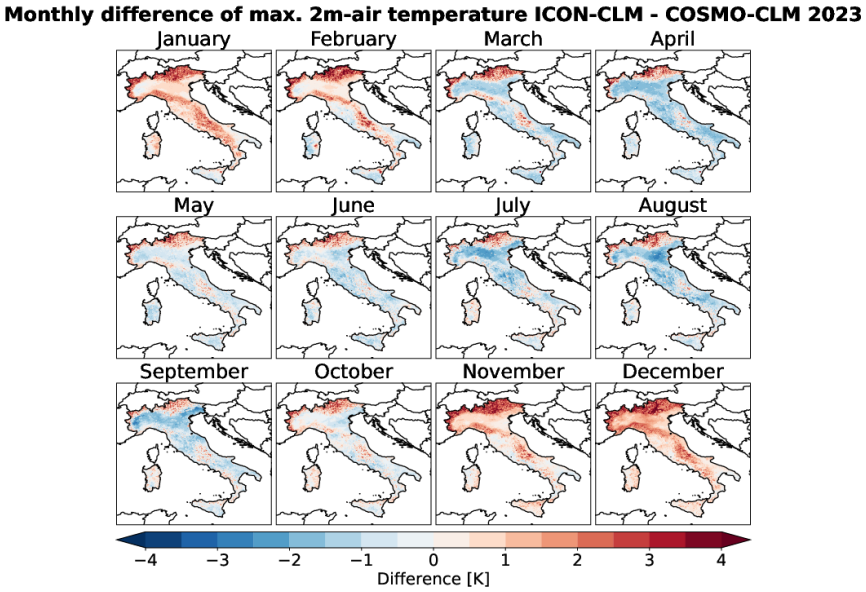

Figure 13: Difference in maximum 2m air temperature for each month between ICLM and CCLM.

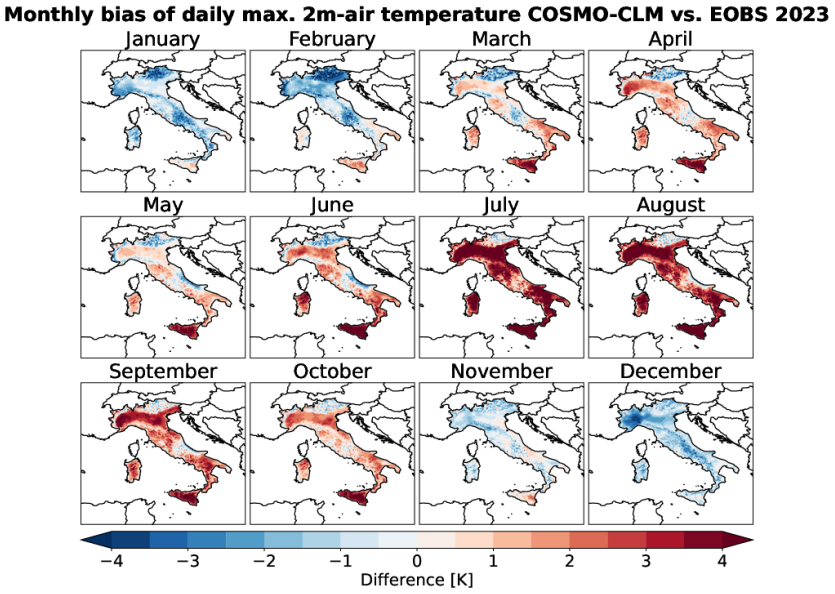

Figure 14: Monthly bias of the daily maximum air temperature at 2m in CCLM for each month of the year 2023. Reference: E-OBS v29.0e.

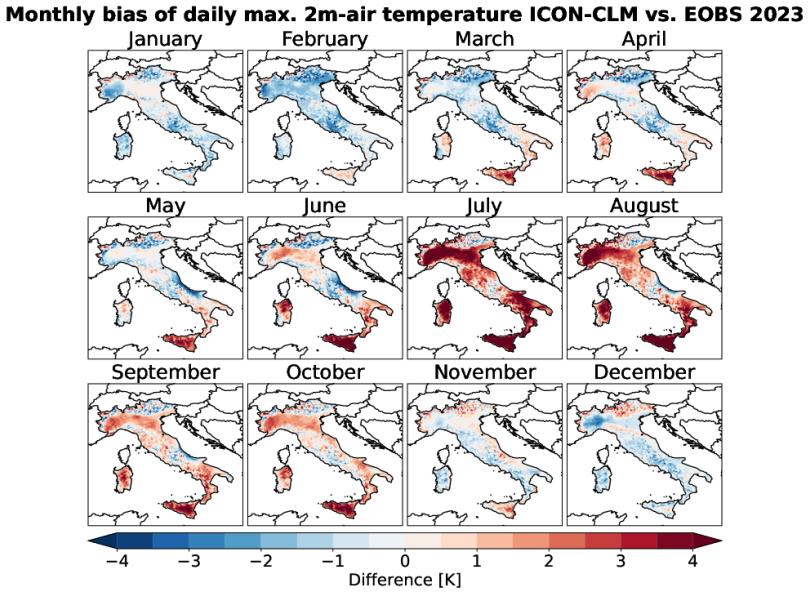

Figure 15: As Fig. 14, but for ICLM.

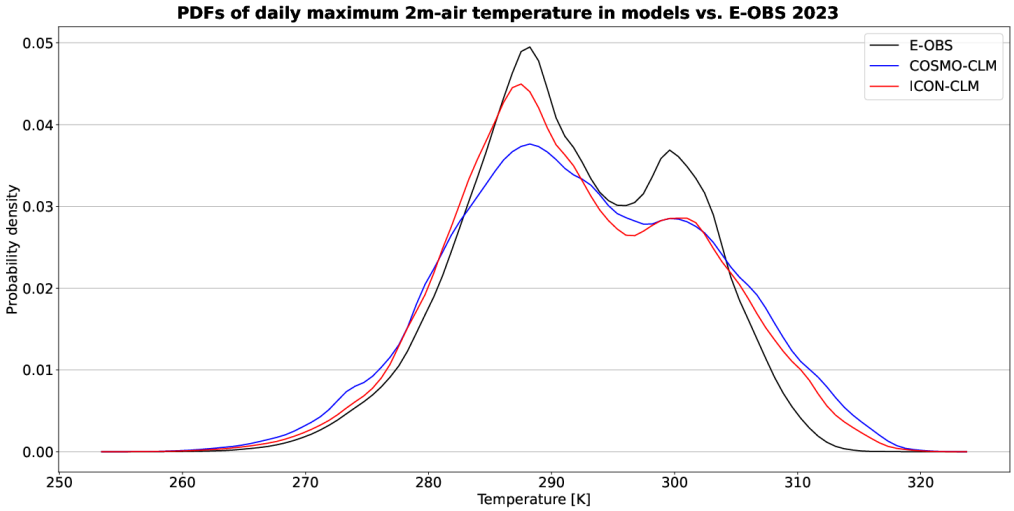

Figure 16: As Fig. 6, but for the daily maximum temperature.

Figure 16 shows how the daily values of this variable are distributed in the three different data sets. For daily maximum temperatures colder than 280 K, all three distributions lie close to one another. Similar to what can be seen in the daily mean and minimum temperatures, the E-OBS distribution (Fig. 16, black line) is characterized by two clear density peaks of different magnitudes. The higher one marks temperature values at around 289 K, whereas the other local maximum lies at around 300 K. The PDF of CCLM (Fig. 16, blue line) is defined by just one clear maximum, located at a value very close to the maximum of the E-OBS PDF. The shape of the ICLM function (Fig. 16, red line), on the other hand, includes features of both the E-OBS and the CCLM PDFs; while for values below 295 K E-OBS observations and ICLM model show similar densities, the distribution of ICLM becomes very close to the one of CCLM progressing towards higher values. This figure shows, moreover, how both models overestimate high temperatures with respect to E-OBS, reaching the 320 K mark.

3.2. DAILY ACCUMULATED PRECIPITATION

Figure 17 depicts a time series of the daily accumulated precipitation, taking into account all grid cells located over Italy. The black line denotes the time series for E-OBS, while the blue and red lines represent the series for the CCLM and ICLM model, respectively. Unlike in the temperature bias, the periods with a clear precipitation bias are not strictly limited to months or seasons. Yet, both models and gridded observations record a period of constant precipitation activity in the months May-June, likely partly influenced by the Emilia-Romagna heavy precipitation event that took place in May. The precipitation intensities within the time series go as high as 21 mm/d in the ICLM model and E-OBS, albeit related to different events. The ICLM model simulates one event at the end of August, while E-OBS observations register an other event of similar magnitude at the beginning of November.

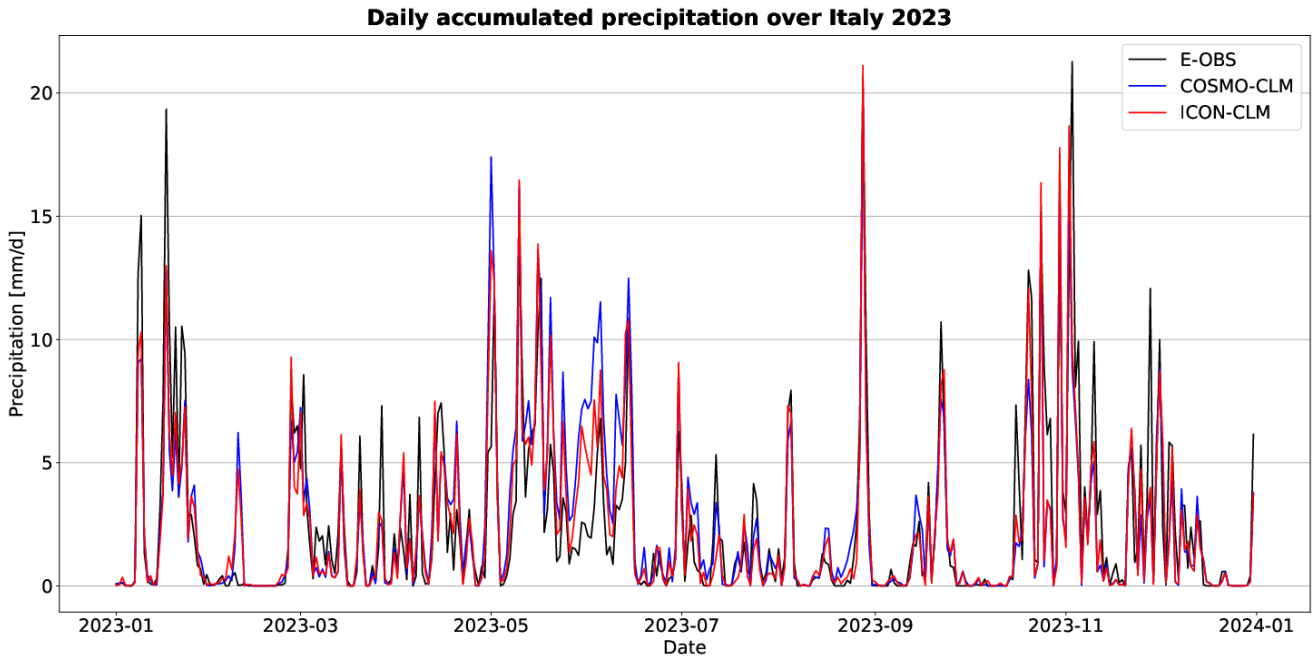

Figure 17: Series of daily accumulated precipitation E-OBS v29.0e (black), CCLM (blue), and ICLM (red) averaged over Italy (see Fig. 1).

In terms of spatial distribution, one can see how ICLM is much drier over the Alpine domain in the months May-August (Fig. 18). In May, however, the Emilia-Romagna region stands out as an isolated wet area in northern Italy, reflecting the two heavy precipitation events that occurred in the region at the beginning and middle of May. Figures 19 and 20 depict the monthly bias of the daily accumulated precipitation with respect to E-OBS in the CCLM and the ICLM model, respectively. Both models show a wet bias in the months May and June over large parts of the country, whereupon this signal is strongest over the Alpine domain and in the CCLM model (Fig. 19). This bias is still present in the ICLM model, but substantially improved (see e.g. June in Fig. 20). Furthermore, both models experience a distinct dry bias in January, October, and November, over northwestern, central, and southern Italy. In the CCLM simulation, one can observe a slight yet constant wet bias throughout the year in northeastern Italy, approximately in the same region that experiences a constant positive mean temperature bias (see Fig. 4). More precisely, the driest bias is observable over southern Italy in both models.

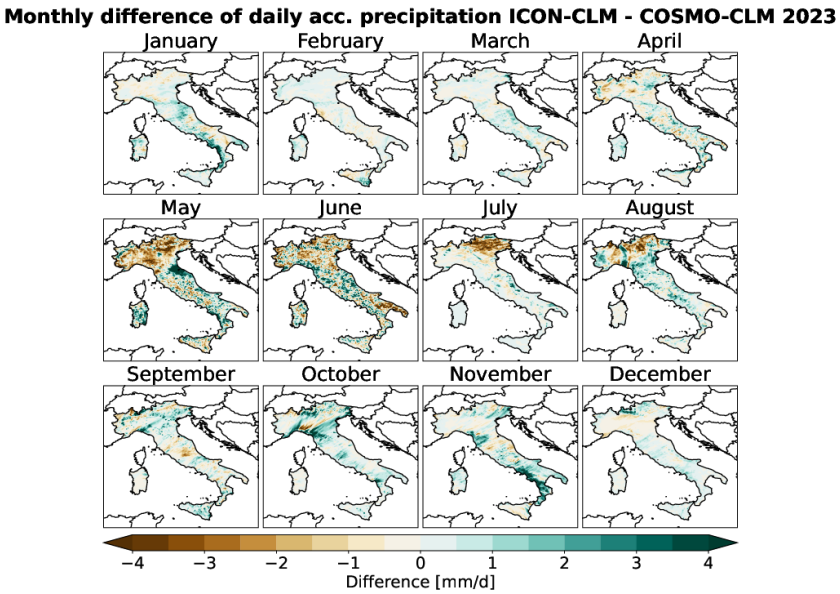

Figure 18: Difference in daily accumulated precipitation for each month between ICLM and CCLM.

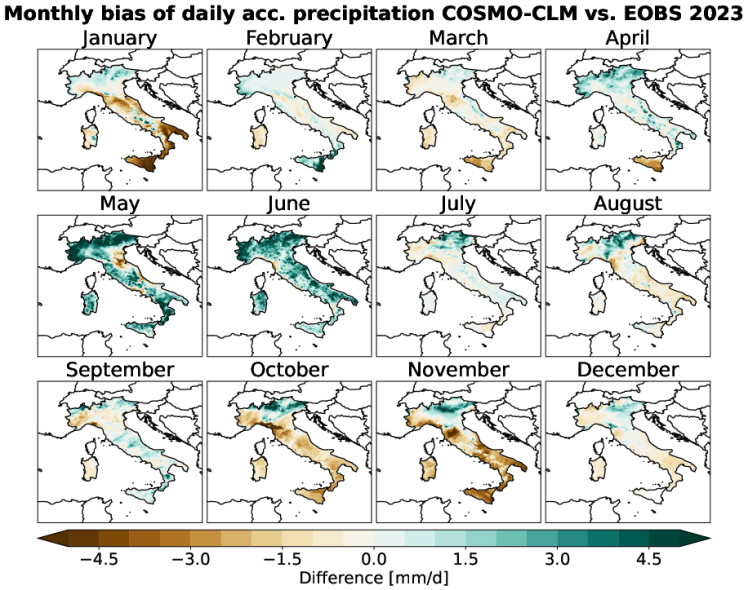

Figure 19: Bias of daily accumulated precipitation in CCLM for each month of the year 2023. Reference: E-OBS v29.0e.

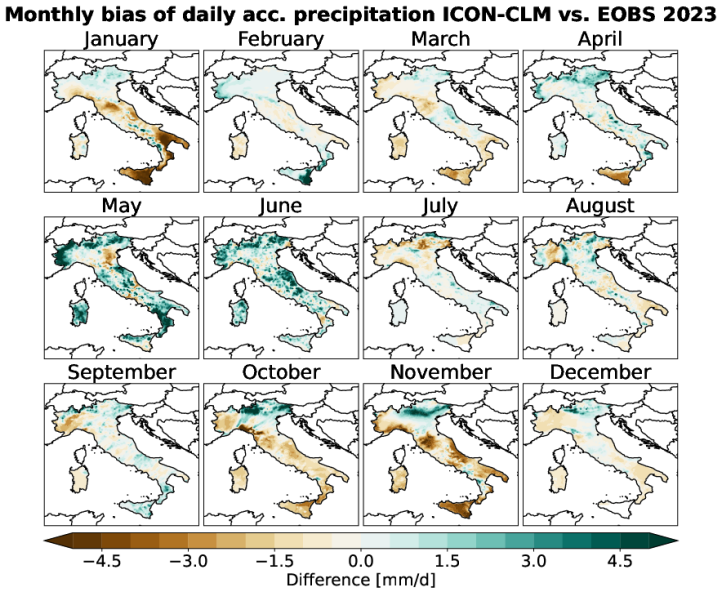

Figure 20: As Fig. 19, but for ICLM.

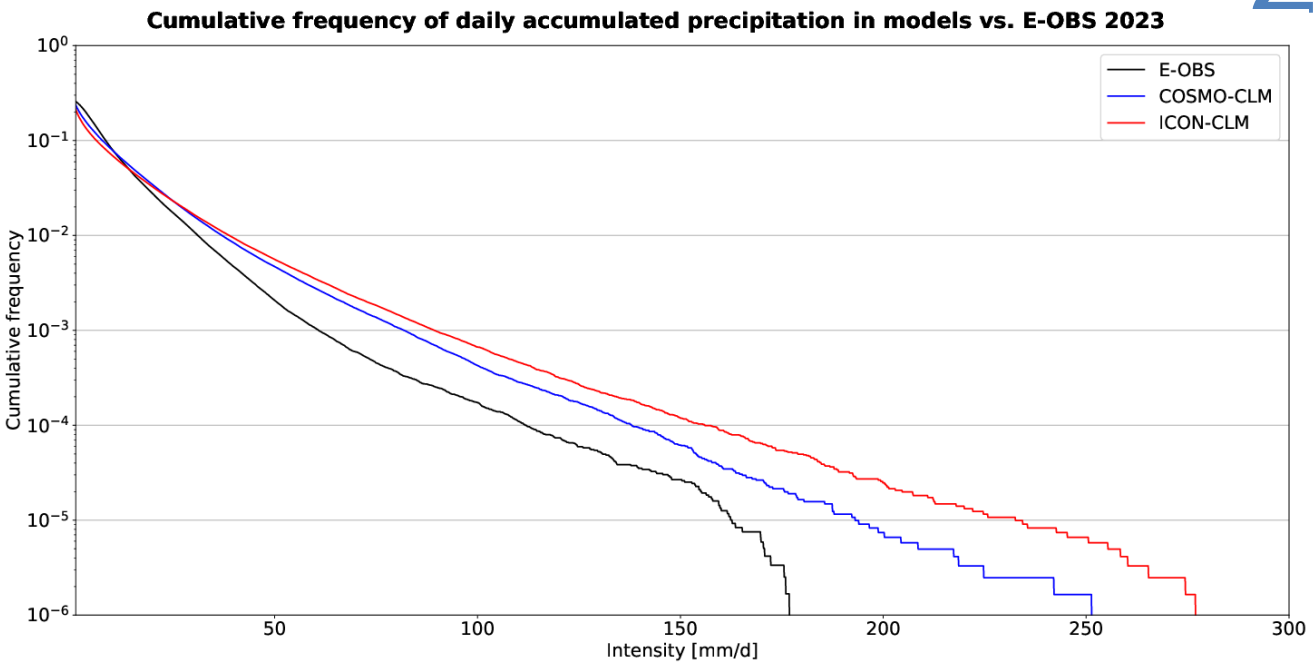

Figure 21: Cumulative frequency of daily accumulated precipitation as present in E-OBS v29.0e (black), CCLM (blue), and ICLM (red). The y-axis intersects the x-axis at a value of 1 mm/d. Only Italian land grid points were considered in the calculation (see Fig. 1).

An overview of the precipitation values present in the three data sets is given in Fig. 21. This figure shows the cumulative frequency of a precipitation event as a function of its intensity. This kind of visualization was chosen over a conventional PDF, since it is able to better highlight low frequency-high intensity events. The black line marks the curve for E-OBS gridded observations, the blue line the one for CCLM, and the red line the one for ICLM. The y-axis corresponds to the intensity threshold of 1 mm/d defining non-drizzle daily precipitation events used in many works (e.g. Ban et al., 2021). While the models and the gridded observations mostly agree with one another for intensities of around 20 mm/d and less, albeit slightly underestimating the frequency of precipitation, both models drastically overestimate extreme precipitation events. The

maximum intensity recorded by E-OBS lies around a value of 175 mm/d. In CCLM and ICLM, the most intense precipitation adds up to roughly 250 mm/d and 275 mm/d, respectively.

3.3. WIND SPEED

Figure 22 displays the difference between ICLM and CCLM in simulating the mean wind speed each month, and reveals that ICLM has higher wind fields throughout the year and across the whole country. Figures 23 and 24 visualize the monthly bias of the months January-June of the daily mean wind speed with respect to E-OBS v28.0e as simulated by the CCLM and the ICLM models, respectively. Since the wind speed data has not yet been updated in the E-OBS dataset, the latest available data goes only until June 2023. While wind speed is made available as a pre-calculated variable within E-OBS data, it had to be inferred separately from the u- and v-wind components of the model output. Both models exhibit a bias of up to -4 m/s in northwestern Italy, which gradually smooths out in the following months, albeit to a lesser extent in the ICLM model (Fig. 24). Although the negative bias is more pronounced in ICLM, the overall bias is reduced compared to the CCLM model when considering the whole country; the CCLM simulations yield a constant positive bias of around 3 m/s throughout the majority of the Italian territory.

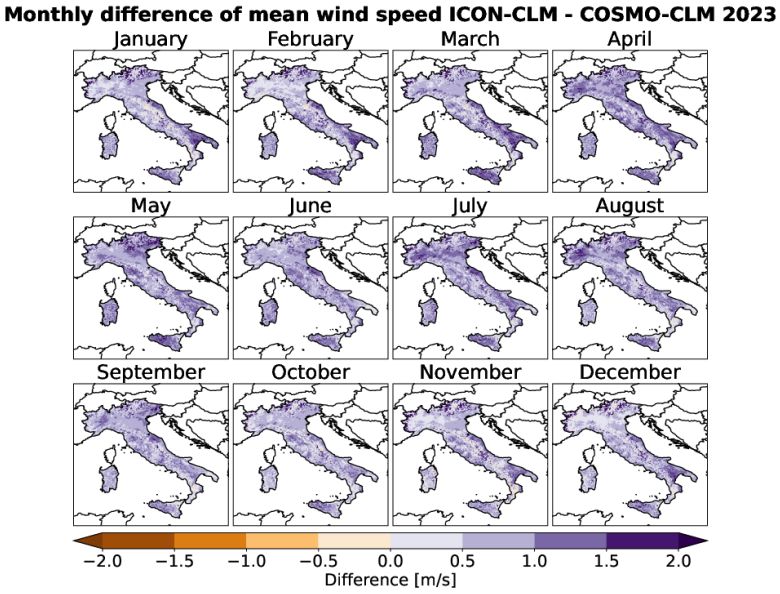

Figure 22: Difference in mean wind speed for each month between ICLM and CCLM.

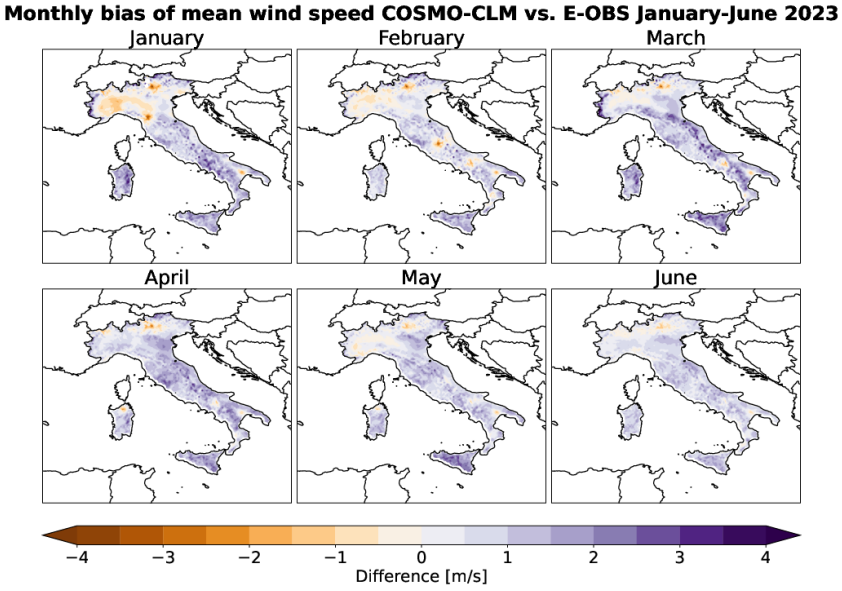

Figure 23: Monthly bias of mean wind speed of CCLM with respect to E-OBS. Due to the incompleteness of wind data in the E-OBS dataset, only the months January-June 2023 are shown.

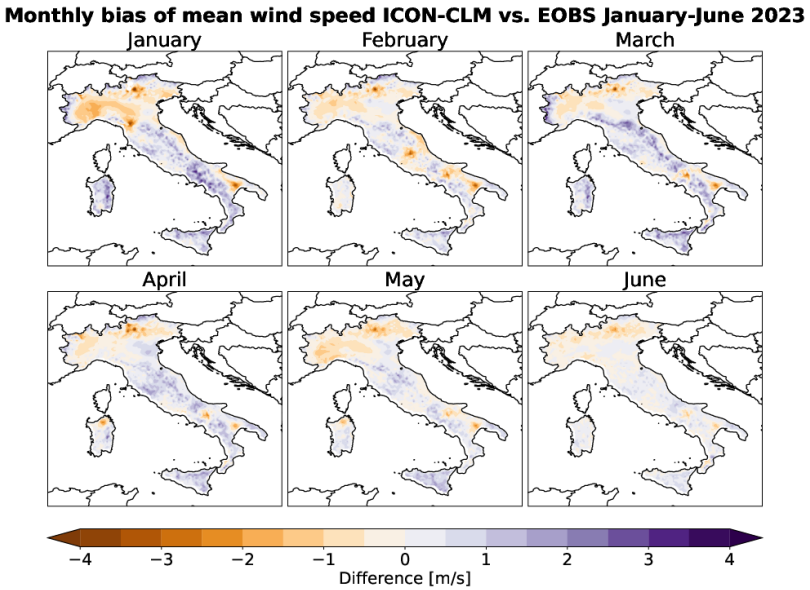

Figure 24: As Fig. 23, but for ICLM.

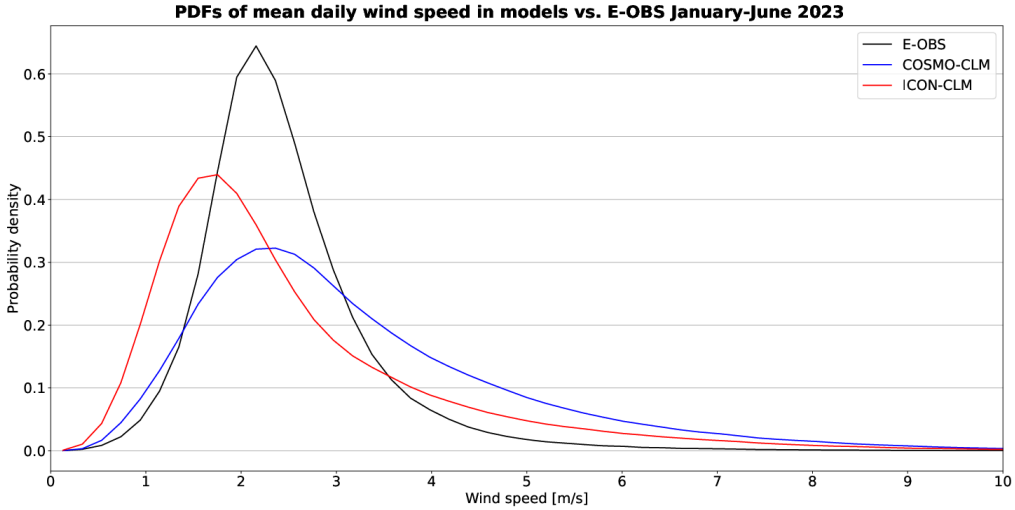

Figure 25: Probability density functions (PDFs) of mean daily wind speed as represented in E-OBS v29.0e (black), CCLM (blue), and ICLM (red). Due to the incompleteness of wind data in the E-OBS dataset, the calculations cover only the months January-June 2023. The wind speed axis was adapted to focus on the bulk data. Only land grid cells were included in the calculation (see Fig. 1).

Looking at the PDFs of the wind speed for the three data sets (Fig. 25), one can see that in all cases the majority of the grid points is concentrated around values of 2-3 m/s. Furthermore, the slightly weaker bias of ICLM (Fig. 25, red line) is apparent, approaching the maximum density of the E-OBS wind speeds (Fig. 25, black line) with respect to the CCLM function (Fig. 25, blue line) and exhibiting slightly more values above the 5 m/s mark. At the same time, the overestimation of wind speeds by CCLM with respect to E-OBS is consistent with what can be observed in the spatial maps (Figs. 23 and 24).

4. SUMMARY AND CONCLUDING REMARKS

The data sets evaluated here contain dynamically downscaled ERA5 reanalysis, originally available at a horizontal resolution of 31 km, to a grid spacing of 2.2 km. The downscaling has been conducted directly with CCLM v5.0.9 and ICLM v2.6.7. The simulations output have hourly resolution. The runs have been performed for the year 2023 and are centered over Italy. The output of the CCLM run is ingested into the data delivery system (DDS, https://dds.cmcc.it/#/dataset/era5-downscaled-over-italy/hourly) of the CMCC, from where it can be easily downloaded and ideally assist in research activities and downstream services.

The evaluation and comparison of the CCLM and ICLM model outputs reveal a generally better agreement of the ICLM simulation with gridded observations, and thus highlight the added value of ICLM with respect to CCLM for the year 2023. To further confirm this added value, the analysis would need to be extended to the years before; the results presented in this work, however, allow for a positive preliminary judgement of the ICLM performance. Concerning daily mean temperature values, ICLM has proven to substantially reduce the warm bias present especially in warm months and over complex topography; yet, a distinct overestimation of temperatures remains over rather flat topography (see Figs. 4 and 5). Such findings have also been found and discussed by

(Adinolfi et al., 2023; Sangelantoni et al., 2024), backing them up with additional observational data sets. Focusing on the daily accumulated precipitation, ICLM delivers a substantial added value in the representation over orographically complex territories (see Figs. 19 and 20). This effect is especially apparent in May and June over the Alps and the southern Apennines.

A major limitation of the analysis conducted here is the absence of subregions for a more detailed investigation, which could potentially give more detailed insight into the behavior of the assessed variables. Since for the time series the mean was calculated spatially over the whole Italian territory (see Figs. 2, 7, 12, and 17), which is marked by a rather complex topographic structure, some potentially present orography-dependent signals might have been canceled out.

An other fact to consider is that over a territory such as a whole country, gridded observations like E-OBS provide one of the most reliable reference data sets for evaluation; however, since gridded data sets have undergone a certain degree of interpolation, whose accuracy is proportional to the availability of station observations, the quality of the information delivered by E-OBS might not always be as homogeneous as the data suggests (Ministero dell'Ambiente e della Sicurezza Energetica, 2023). Since the management of in-situ observations in Italy is mostly organized on a regional basis, the station density varies largely across the different regions (see red dots in Fig. 1). In doing so, the availability in the Calabria region, for example, is much higher than in Sicily, which is located very closely to the former. Thus, it is always advisable to have a minimum knowledge of the observation station availability in the investigated area when working with E-OBS, considering that gridded observations themselves can sometimes also be prone to errors.

BIBLIOGRAPHY

Adinolfi, M., Raffa, M., Reder, A., and Mercogliano, P. (2023). Investigation on potential and limitations of era5 reanalysis downscaled on italy by a convection-permitting model. *Climate Dynamics*, 61(9):4319–4342.

Arakawa, A. (1977). Computational design of the basic dynamical processes of the ucla general circulation model. In Davidson, R. B. et al., editors, *Methods in Computational Physics*, pages 173–265. Academic Press.

Arino, Olivier; Ramos Perez, Jose Julio; Kalogirou, Vasileios; Bontemps, Sophie; Defourny, Pierre; Van Bogaert, Eric (2012): Global Land Cover Map for 2009 (GlobCover 2009) [dataset]. © *European Space Agency (ESA) & Université catholique de Louvain (UCL)*, *PANGAEA*, https://doi.org/10.1594/PANGAEA.787668.

Baldauf, M., & Schulz, J. P. (2004). Prognostic precipitation in the Lokal-Modell (LM) of DWD. *Cosmo Newsletter*, 4, 177-180.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the cosmo model: Description and sensitivities. *Monthly Weather Review*, 139(12):3887–3905.

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S., et al. (2021). The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part i: evaluation of precipitation. *Climate Dynamics*, 57:275–302.

Bechtold, P. (2008, September). Convection parameterization. In *Proc. Seminar on Parameterization of Subgrid Physical Processes* (pp. 63-86).

Bottazzi, M., Rodríguez-Muñoz, L., Chiavarini, B., Caroli, C., Trotta, G., Dellacasa, C., Marras, G. F., Montanari, M., Santini, M., Mancini, M., et al. (2024). High

performance computing to support land, climate, and user-oriented services: The highlander data portal. *Meteorological Applications*, 31(2):e2166.

Bott, A. (1989). A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes. *Monthly weather review*, 117(5):1006–1016.

Chen, S., Poll, S., Hendricks Franssen, H.-J., Heinrichs, H., Vereecken, H., and Goergen, K. (2024). Convection-permitting icon-lam simulations for renewable energy potential estimates over southern africa. *Journal of Geophysical Research: Atmospheres*, 129(6):e2023JD039569.

Copernicus Climate Change Service (2024). In-situ Gridded Observations for Europe. https://cds.climate.copernicus.eu/cdsapp#!/dataset/insitu-gridded-observations-europe?tab=form. Accessed: 2024-09-24.

Copernicus (2024). ERA5 hourly data on pressure levels from 1979 to present. https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-complete?tab=overview. viewed 16/09/2024.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., and Jones, P. D. (2018). An ensemble version of the e-obs temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17):9391–9409.

Crueger, T., Giorgetta, M. A., Brokopf, R., Esch, M., Fiedler, S., Hohenegger, C., Kornblueh, L., Mauritsen, T., Nam, C., Naumann, A. K., et al. (2018). Icon-a, the atmosphere component of the icon earth system model: Ii. model evaluation. *Journal of Advances in Modeling Earth Systems*, 10(7):1638–1662.

Doms, G. and Baldauf, M. (2011). A description of the nonhydrostatic regional cosmo-model part i: dynamics and numerics. *Deutscher Wetterdienst, Offenbach*.

European Climate Assessment & Dataset (ECA&D) (2024). ENSEMBLES Data Download. https://www.ecad.eu/download/ensembles/download.php. Accessed: 2024-09-24.

Giorgetta, M. A., Brokopf, R., Crueger, T., Esch, M., Fiedler, S., Helmert, J., Hohenegger, C., Kornblueh, L., Köhler, M., Manzini, E., et al. (2018). Icon-a, the atmosphere component of the icon earth system model: I. model description. *Journal of Advances in Modeling Earth Systems*, 10(7):1613–1637.

Heise, E., Ritter, B., Schrodin, R., and Wetterdienst, D. (2006). *Operational implementation of the multilayer soil model*. Citeseer.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., Poli, P., Radu, R., Schepers, D., Simmons, A., Soci, C., Laloyaux, P., Owono, A. O., Ribas, R., and Suttie, M. (2023). Era5: A full-observing-system global reanalysis for the atmosphere, land surface, and ocean waves. European Centre for Medium-Range Weather Forecasts (ECMWF). Accessed: 2024-10-12.

Hogan, R. J., & Bozzo, A. (2018). A flexible and efficient radiation scheme for the ECMWF model. *Journal of Advances in Modeling Earth Systems*, 10(8): 1990-2008.

Kendon, E., Prein, A., Senior, C., and Stirling, A. (2021). Challenges and outlook for convection-permitting climate modelling. *Philosophical Transactions of the Royal Society A*, 379(2195):20190547.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T., Boucher, O., Chin, M., et al. (2006). An aerocom initial assessment–optical properties in aerosol component modules of global models. *Atmospheric Chemistry and Physics*, 6(7):1815–1834.

Leuenberger, D., Koller, M., Fuhrer, O., and Schär, C. (2010). A generalization of the sleve vertical coordinate. *Monthly Weather Review*, 138(9):3683–3689.

Lott, F., & Miller, M. J. (1997). A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, 123(537), 101-127.

Ministero dell'Ambiente e della Sicurezza Energetica (2023). Piano Nazionale di Adattamento ai Cambiamenti Climatici. Ministero dell'Ambiente e della Sicurezza Energetica. Accessed: 2024-09-16.

Orr, A., Bechtold, P., Scinocca, J., Ern, M., & Janiskova, M. (2010). Improved middle atmosphere climate and forecasts in the ECMWF model through a nonorographic gravity wave drag parameterization. *Journal of Climate*, 23(22): 5905-5926.

Pham, T. V., Steger, C., Rockel, B., Keuler, K., Kirchner, I., Mertens, M., Rieger, D., Zängl, G., and Früh, B. (2021). Icon in climate limited-area mode (icon release version 2.6. 1): a new regional climate model. *Geoscientific Model Development*, 14(2):985–1005.

Raffa, M., Reder, A., Marras, G. F., Mancini, M., Scipione, G., Santini, M., and Mercogliano, P. (2021). Vhr-rea_it dataset: very high resolution dynamical downscaling of era5 reanalysis over italy by cosmo-clm. *Data*, 6(8):88.

Raschendorfer, M. (2001). The new turbulence parametrization of lm. cosmo newsletter no. 1: 90–98.

Reinhardt, T. and Seifert, A. (2006). A three-category ice scheme for lmk. *Cosmo Newsletter*, 6:115–120.

Ritter, B. and Geleyn, J.-F. (1992). A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Monthly weather review*, 120(2):303–325.

Rockel, B., Will, A., & Hense, A. (2008). The regional climate model COSMO-CLM (CCLM). *Meteorologische zeitschrift*, 17(4): 347.

Sakradzija, M. and Klocke, D. (2018). Physically constrained stochastic shallow convection in realistic kilometer-scale simulations. *Journal of Advances in Modeling Earth Systems*, 10(11):2755–2776.

Sangelantoni, L., Sobolowski, S., Lorenz, T. (2024). Investigating the representation of heatwaves from an ensemble of km-scale regional climate simulations within CORDEX-FPS convection. *Clim Dyn,* 62: 4635–4671. https://doi.org/10.1007/s00382-023-06769-9.

Schär, C., Leuenberger, D., Fuhrer, O., Lüthi, D., & Girard, C. (2002). A new terrain-following vertical coordinate formulation for atmospheric prediction models. *Monthly Weather Review*, 130(10): 2459-2480.

Schrodin, R., & Heise, E. (2001). *The multi-layer version of the DWD soil model TERRA_LM*. DWD.

Schulz, J. P. (2008). Introducing sub-grid scale orographic effects in the COSMO model. *COSMO Newsl*, 9, 29-36.

Schulz, J. P., Vogel, G., Becker, C., Kothe, S., & Ahrens, B. (2015, April). Evaluation of the ground heat flux simulated by a multi-layer land surface scheme using high-quality observations at grass land and bare soil. In *Egu general assembly conference abstracts* (p. 6549).

Schulz, J. P., & Vogel, G. (2020). Improving the processes in the land surface scheme TERRA: Bare soil evaporation and skin temperature. *Atmosphere*, 11(5): 513.

Seifert, A. (2008). On the parameterization of evaporation of raindrops as simulated by a one-dimensional rainshaft model. *Journal of the Atmospheric Sciences*, 65(11), 3608-3619.

Steppeler, J., Doms, G., Schättler, U., Bitzer, H., Gassmann, A., Damrath, U., and Gregoric, G. (2003). Meso-gamma scale forecasts using the nonhydrostatic model lm. *Meteorology and atmospheric Physics*, 82:75–96.

Taylor, C. M., Birch, C. E., Parker, D. J., Dixon, N., Guichard, F., Nikulin, G., and Lister, G. M. (2013). Modeling soil moisture-precipitation feedback in the sahel: Importance of spatial scale versus convective parameterization. *Geophysical Research Letters*, 40(23):6213–6218.

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly weather review*, 117(8):1779–1800.

Trusilova, K., Früh, B., Brienen, S., Walter, A., Masson, V., Pigeon, G., and Becker, P. (2013). Implementation of an urban parameterization scheme into the regional climate model cosmo-clm. *Journal of Applied Meteorology and Climatology*, 52(10):2296–2311.

Trusilova, K., Schubert, S., Wouters, H., Fruh, B., Grossman-Clarke, S., Demuzere, M., & Becker, P. (2016). The urban land use in the COSMO-CLM model: a comparison of three parameterizations for Berlin. *METEOROLOGISCHE ZEITSCHRIFT*, 25(2): 231-244.

Wicker, L. J. and Skamarock, W. C. (2002). Time-splitting methods for elastic models using forward time schemes. *Monthly weather review*, 130(8):2088–2097.

Wouters, H., Demuzere, M., Blahak, U., Fortuniak, K., Maiheu, B., Camps, J., Tielemans, D., and van Lipzig, N. P. (2016). The efficient urban canopy dependency parametrization (sury) v1. 0 for atmospheric modelling: description and application with the cosmo-clm model for a belgian summer. *Geoscientific Model Development*, 9(9):3027–3054.

Wouters, H., Demuzere, M., De Ridder, K., and van Lipzig, N. P. (2015). The impact of impervious water-storage parametrization on urban climate modelling. *Urban Climate*, 11:24–50.

Zängl, G., Reinert, D., and Prill, F. (2022). Grid refinement in icon v2. 6.4. *Geoscientific Model Development*, 15(18):7153–7176.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M. (2015). The icon (icosahedral non-hydrostatic) modelling framework of dwd and mpi-m: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141(687):563–579.