Climate Dynamics (2023) 61:1411–1437
https://doi.org/10.1007/s00382-022-06633-2

A convection-permitting and limited-area model hindcast driven by ERA5 data: precipitation performances in Italy

Valerio Capecchi[1] · Francesco Pasi[1,2] · Bernardo Gozzini[1] · Carlo Brandini[1,2,3]

Received: 19 August 2022 / Accepted: 8 December 2022 / Published online: 20 December 2022

Abstract

This study describes the implementation and performances of a weather hindcast obtained by dynamically downscaling the ERA5 data across the period 1979–2019. The limited-area models used to perform the hindcast are BOLAM (with a grid spacing of 7 km over the Mediterranean domain) and MOLOCH (with a grid spacing of 2.5 km over Italy). BOLAM is used to provide initial and boundary conditions to the inner grid of the MOLOCH model, which is set in a convection-permitting configuration. The performances of such limited-area, high-resolution and long-term hindcast are evaluated comparing modelled precipitation data against two high-resolution gridded observational datasets. Any potential added-value of the BOLAM/MOLOCH hindcast is assessed with respect to ERA5-Land data, which are used as benchmark. Results demonstrate that the MOLOCH hindcast, compared to observations, provides a lower bias than ERA5-Land as regards both the mean annual rainfall (−1.3% vs 8.7%) and the 90th percentile of summer daily precipitation, although a wet bias is found in southern Italy (bias $\simeq$ 17.1%). Improvements are also gained in the simulation of the 90th percentile of hourly precipitations both in winter and, to a minor extent, in summer. The diurnal cycle of summer precipitations is found to be better reconstructed in the Alps than in the hilly areas of southern Italy. This study also presents the rainfall peaks obtained in the simulation of two well-known severe precipitation events that caused floods and damages in north-western Italy in 1994 and 2011. Finally, it is discussed how the demonstrated reliability of the BOLAM and MOLOCH models associated to the documented low computational cost, promote their use as a valuable tool for downscaling not only reanalyses but also climate projections.

Keywords ERA5 · MOLOCH · Hindcast · Precipitation · Italy

1 Introduction

In November 2017, the Copernicus Climate Change Service promoted the International Conference on Reanalysis (Buizza et al. 2018) to bring together reanalysis producers and observation providers. The conference gave the opportunity to the user community to assess the status of available products and discuss future developments. Later, in autumn 2018, the release of the first stream of ERA5 data (Hersbach et al. 2020) pushed forward the use of reanalyses in the public and private sectors not only for research purposes but also for climate services and policy making (Buizza et al. 2018). ERA5 is the cutting-edge dataset among global reanalyses with an unprecedented resolution, both spatial and temporal. The dataset benefits from a decade of advances in model physics and data assimilation methods with respect to its predecessor ERA-Interim (Dee et al. 2011). However, how it has been argued in recent studies (Bandhauer et al. 2022; Jiang et al. 2021; Lavers et al. 2022; Singh et al.

✉ Valerio Capecchi
capecchi@lamma.toscana.it

Francesco Pasi
pasi@lamma.toscana.it

Bernardo Gozzini
gozzini@lamma.toscana.it

Carlo Brandini
brandini@lamma.toscana.it

1 LaMMA, Laboratorio di Meteorologia e Modellistica Ambientale per lo sviluppo sostenibile, Via madonna del piano 10, Sesto Fiorentino (Firenze) 50019, Italy

2 Consiglio Nazionale delle Ricerche, Istituto di Bioeconomia, Via madonna del piano 10, Sesto Fiorentino (Firenze) 50019, Italy

3 Present Address: Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine, Via madonna del piano 10, Sesto Fiorentino, (Firenze) 50019, Italy

2021; Zhang et al. 2021), the relatively coarse resolution of ERA5, grid spacing is 31 km, and the assumptions on the parameterisation of convection, discourage its direct use in regional or local applications; especially those involving rainfall as input. To bridge the gap between the resolution of global reanalyses and regional scale impact studies, several notable initiatives have investigated the use of high-resolution and limited-area models fed by global reanalyses to produce more detailed climate databases. To stick to the European context, we mention few studies and datasets. The project Uncertainties in Ensembles of Regional Re-Analyses (UERRA) delivers a regional reanalysis dataset for the CORDEX EUR-11 (European Coordinated Regional Climate Downscaling Experiment, Giorgi et al. (2009)) domain. It is based on the ALADIN model (Bubnová et al. 1995), applied using the hydrostatic assumption, as regards the atmospheric part and on the SURFEX model (Masson et al. 2013) as its surface counterpart. The horizontal resolution is 11 km for the atmosphere and 5.5 km for near-surface variables. A 3D-Var method (Gustafsson et al. 2001) is used to ingest conventional data into a short-term forecast to produce hourly analysis, whereas ERA-Interim data are used to provide large scale constrains to the limited-area integrations. Data are available for the period 1961–2019 and no further update is foreseen since ERA-Interim was discontinued in August 2019. The recent release (August 2022) of the Copernicus European Regional ReAnalysis (CERRA) aims at filling the gap left by UERRA. CERRA uses the HARMONIE-ALADIN data assimilation system fed by ERA5 lateral boundary conditions to produce a pan-European reanalysis dataset; its horizontal resolution is 5.5 km and it has 106 vertical levels. Data are available from the early 1980s to the delayed present. COSMO-REA6 (Bollmeyer et al. 2015) is a high-resolution reanalysis dataset produced by the German Meteorological Service using version 4.25 (released in September 2012) of the mesoscale model COSMO, which is nested in the global reanalysis fields of ERA-Interim for the period 1995–2015. The subgrid convection processes are parameterised using the Tiedtke scheme (Tiedtke 1989); the spatial resolution is approximately 6 km and the simulations cover the CORDEX EUR-11 domain. A data assimilation system based on a continuous nudging (Schraff and Hess 2003) is used to ingest radiosondes, aircraft and near-surface station data. A 1-year test period is used to evaluate the added value of assimilating observed data with respect to a dynamical downscaling experiment (i.e., an identical model setup without assimilation). As regards the total yearly accumulated precipitation, the dynamical downscaling experiment has very similar spatial patterns to the assimilated experiment, with an overestimation in Scandinavia Peninsula, Russia, Turkey and part of the Alps and UK. As regards the mean diurnal cycle of precipitation in summer, both experiments exhibit a shift by approximately 3 h with respect to rain-gauge data, possibly because of a delayed initiation of convection. As regards extreme precipitation events (rainfall rates higher than 20 mm in 3 h and 50 mm in 24 h), the dynamical downscaling experiment is proven to slightly outperforms the assimilated experiment with a frequency distribution similar to that of observed rainfall data. A few years later, the update of the project (Wahl et al. 2017) used a model which runs without parameterization of deep moist convection and which is able to improve the spatial representation of local precipitation over Germany at different time scales, from monthly to annual time scale. The Irish Meteorological Service has carried out a 38-year very high-resolution regional climate reanalysis for Ireland (Whelan et al. 2018) using the non-hydrostatic ALADIN-HIRLAM numerical weather prediction system (Seity et al. 2011) and driven by ERA-Interim for boundary conditions. This reanalysis (hereinafter MÉRA) covers Ireland, the United Kingdom, and northern France and spans the period from 1981 to 2019. The resolution used (2.5 km horizontal grid) allows deep convection to be partially resolved. To ingest observations, the authors utilised an optimal interpolation method (Giard and Bazile 2000) for surface variables and a three-dimensional variational data assimilation technique (Fischer et al. 2005) for upperair data. The authors claim that such high-resolution and convection-permitting reanalysis dataset is more suitable to study the trends of near-surface parameters and the frequency distribution of weather extremes. Doddy Clarke et al. (2021) showed that MÉRA doesn't consistently outperform ERA5 data when compared to ground measurements for renewable energy purposes. The authors also proposed the development of a new MÉRA model with ERA5 as the driving global reanalysis, which should lead to an improved regional reanalysis for Ireland and nearby countries. MERIDA (Bonanno et al. 2019) is a meteorological reanalysis dataset for the Italian domain; it is aimed at providing information to energy market stakeholders. The WRF model (Skamarock et al. 2008) supplied by ERA5 data as initial and boundary conditions is used to run 3-days long numerical forecasts. The period covered is 1990–2020 and the grid spacing is 7 km smoothly relaxing larger-scale features at the boundaries of the domain using a spectral nudging. As regards 2-m temperature and precipitation, a higher-resolution dataset (MERIDA-OI with a grid spacing of approximately 4 km) is achieved by postprocessing and merging numerical data with observations by means of an optimal interpolation technique (Uboldi et al. 2008). To test the impact of the cumulus parameterisation, the authors realised a 1-year long simulation with the explicit treatment of convection processes and a grid spacing of approximately 4 km. Surprisingly, they noted slightly better performances of MERIDA than the alternative configuration; in fact, they found that the experiment without the parameterisation of convection leads to a higher number of false alarms. They do not support this conclusion with any particular explanations. More recently, Raffa et al. (2021a) compared two different strategies to downscale the ERA5 data with the COSMO-CLM model (Rockel et al. 2008). In the first configuration a 2.2 km resolution simulation is directly one-way nested in ERA5, whereas in the second experiment the traditional two-step nesting strategy is adopted, that is the simulation at 2.2 km resolution is one-way nested in a 12 km grid spacing which in turn is one-way nested in ERA5. The sensitivity tests were carried out for the period 2007–2011. By comparing rainfall outputs with the E-OBS (Cornes et al. 2018) gridded dataset as ground truth, the authors argued that the direct nesting should be preferred. In a second paper (Raffa et al. 2021b), the authors presented the dataset, which covers the Italian domain, for the period 1989–2020. More recently, Reder et al. (2022), demonstrated that this dataset provides more reliable data than global reanalyses in characterising extreme precipitations at the urban scale. Several studies focused on the reliability of climate data to represent precipitation over the Alpine area. When evaluating results over a 10-year long validation period or less (Ban et al. 2015; Lind et al. 2016), the consensus is that deploying convection-permitting models improves the representation of subdaily summer rainfall and the frequency of both daily and hourly precipitation events. Other studies (Adinolfi et al. 2020) provide evidence that convection-permitting models fit betterthe statistical distribution of Alpine heavy precipitations than the driving convection-parameterised models; this holds in particular for hourly extreme precipitations.

From the brief review of previous works, some points emerge as critical to provide reliable pictures of regional (i.e., mesoscale) past climates: (i) the use of high-quality global reanalyses to feed limited-area models, (ii) the use of convection-permitting models to account for a more accurate description of convection processes and related precipitations, (iii) the implementation of data assimilation techniques to optimally merge observations to short-range weather forecasts. In this work we present the precipitation performances of a new hindcast produced by downscaling ERA5 data for the period 1979–2019 by means of the BOLAM/MOLOCH limited-area numerical weather models. ERA5 is considered the cutting-edge dataset as regards global reanalyses, and the availability of ERA5 up to the 1950s allows for a back-extension of the BOLAM/MOLOCH hindcast to obtain more robust statistically sound conclusions about trends in the recent climate. One innovative aspect of our work relies on the use of the convection-permitting MOLOCH model over the Italian domain (as recently done by Raffa et al. (2021b) for the period 1989–2020); major gains are expected to be achieved in the simulation of precipitation extremes and mesoscale convective systems. In fact, switching off the convective parameterisation has potential benefits in long-term simulations for the representation of hourly extreme precipitations and the mean diurnal cycle of precipitations (Prein et al. 2013; Fosser et al. 2015; Coppola et al. 2020; Ban et al. 2021; Reder et al. 2022). This holds in particular during summer, when convection plays a major role in controlling the occurrence of rainfall. Regarding the implementation of a data assimilation scheme, we claim that, with limited computational resources, such task would have overly slow down the production of the hindcast. This holds besides the technical and scientific challenges related to the implementation the data assimilation at such high-resolution grid spacing (2.5 km for the MOLOCH model). However, although no assimilation is performed, the BOLAM/MOLOCH hindcast is intended to provide reliable data for downstream services and models. In fact, numerical weather hindcasts were frequently used in previous works to force wave models (Mentaschi et al. 2013; Ferrari et al. 2020; Osinski and Radtke 2020; Vannucchi et al. 2021) or for climatological studies (Giorgi and Gutowski Jr 2015; Ruti et al. 2016; Ban et al. 2021). Moreover, Bollmeyer et al. (2015) demonstrated, for a 1-year long test experiment, that the simple dynamical downscaling of global reanalyses (i.e., no observations assimilated at starting time) provides rainfall predictions as accurate as those produced with the more CPU-demanding continuous nudging scheme.

The rainfall data we present are the result of a pragmatic approach which takes full advantage of the improved quality of ERA5 data by using a suite of reliable (Buzzi et al. 2014; Davolio et al. 2020) and fast (Malguzzi and Tartagione 1999; Capecchi 2021) limited-area numerical models. The remainder of the paper is structured as follows: in Sect. 2 we provide descriptions about the models (BOLAM and MOLOCH) and data (ERA5-Land and observational datasets) used in the work; we also give details about the methods utilised to validate the rainfall predictions. The results are shown in Sect. 3 and discussed in Sect. 4 along with their implications and limitations.

2 Models, data and methods

2.1 Models

To produce the atmospheric hindcast for the period 1979–2019, we implemented a two one-way nested domains configuration based on two limited-area numerical weather prediction models: BOlogna Limited Area Model (BOLAM) and MOdello LOcale in Hybrid Coordinates (MOLOCH). Both codes are developed and maintained at the Institute of Atmospheric Sciences and Climate of the Italian National Research Council (CNR). They were initially developed for research purposes (Buzzi et al. 1994), but today are being used operationally by various regional meteorological

Table 1 Setup of the key characteristics of the BOLAM and MOLOCH simulations

	BOLAM	MOLOCH
Grid spacing (km)	7	2.5
Number of rows and columns	482 and 890	626 and 506
Number of vertical levels	50	50
Number of soil levels	7	7
Grid points	≃21.5 million	≃15.8 million
Time step (s)	45	30
Boundary layer scheme	1.5-order E-l closure (Zampieri et al. 2005)	
Radiation scheme	Ritter and Geleyn (1992) and	
	ECMWF radiation scheme (Morcrette et al. 2008)	
Microphysics scheme	Drofa and Malguzzi (2004)	
Turbulence scheme	1.5-order E-l closure (Trini Castelli et al. 2020)	
Convection parameterisation	Kain (2004)	None

services both in Italy and abroad (Lagouvardos et al. 2003; Corazza et al. 2018). Davolio et al. (2020) provides a general but comprehensive introduction to the BOLAM and MOLOCH models and a list of the several applications over which they are used. Here we provide a short description of the two models to make the text self-consistent; additional details and insights can be found in the cited literature. BOLAM is a primitive equations hydrostatic model with parameterised convection (using a modified version of the scheme proposed in Kain (2004)). In our work, it

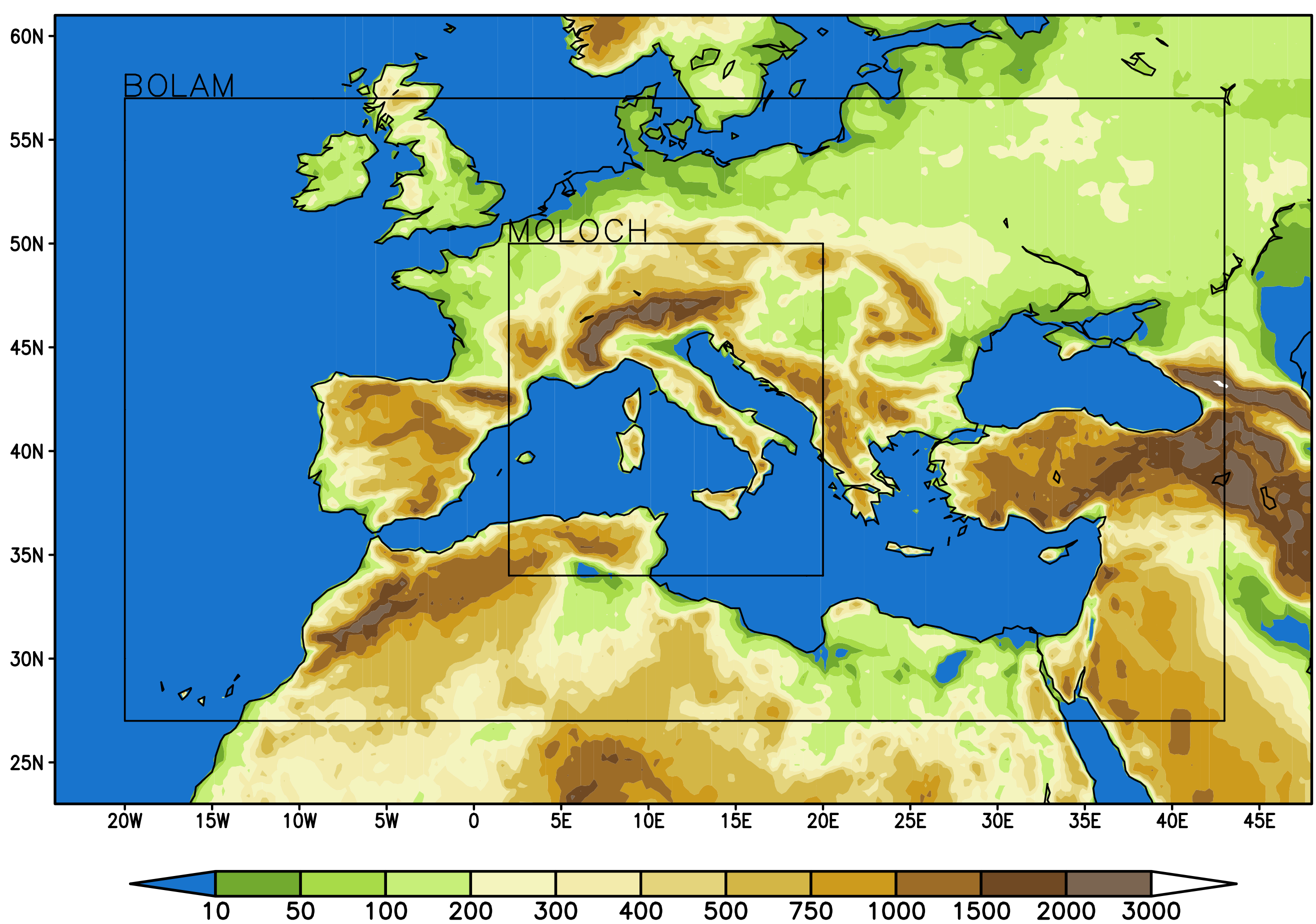

Fig. 1 Extent of the BOLAM and MOLOCH domains of integration with superimposed topography (m). The BOLAM domain, outer black box, approximately corresponds to the Med-CORDEX area; the MOLOCH domain, inner black box, covers Italy and surrounding areas

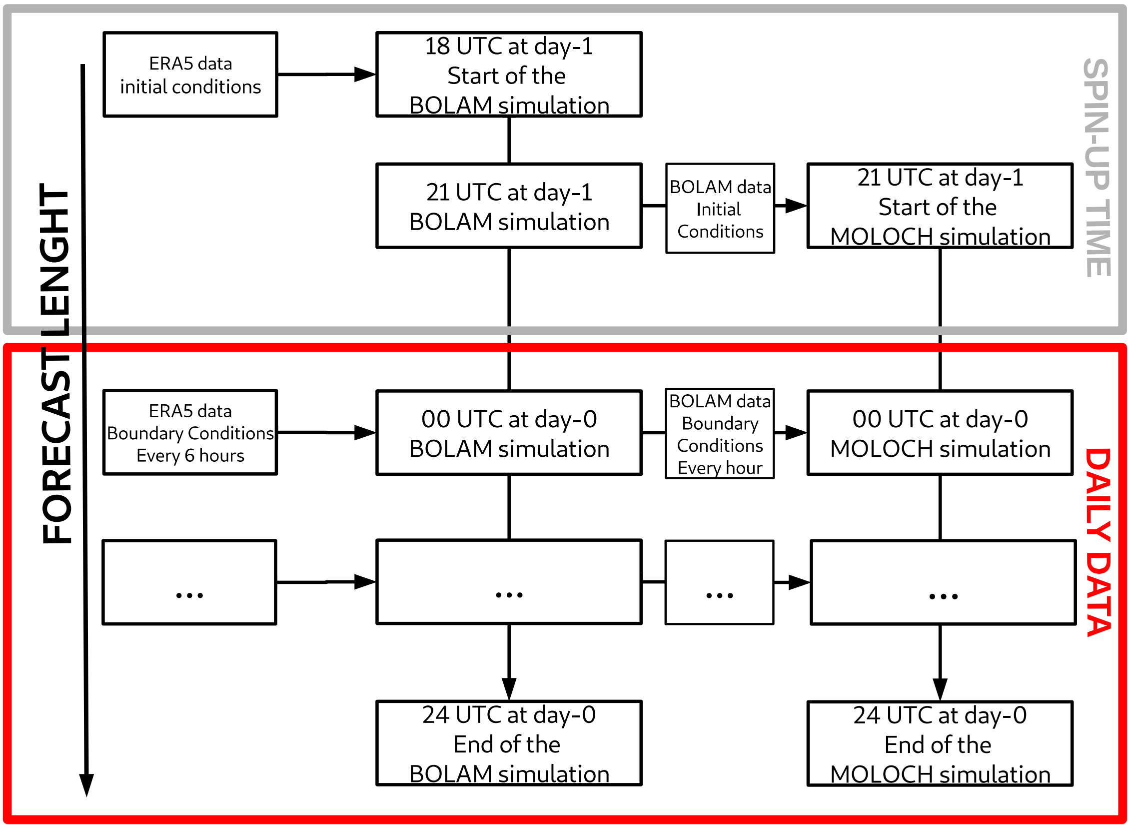

Fig. 2 Modelling setup for the production of one single day of the BOLAM/MOLOCH hindcast

was employed with a grid spacing of approximately 7 km to provide lateral boundary conditions to MOLOCH every hour. MOLOCH is a nonhydrostatic, fully compressible model that uses a hybrid terrain-following coordinate, relaxing smoothly to horizontal surfaces. The microphysical scheme is an upgrade of the parameterisation proposed by Drofa and Malguzzi (2004), which describes the interactions of cloud water, cloud ice, rain, snow and graupel. In this study, the grid spacing of the MOLOCH model is about 2.5 km, which allows the atmospheric convection to be partially resolved. We used the version released in late 2017 for both models. To define the static properties of the grid (orography, landuse, soil and vegetation type), both BOLAM and MOLOCH use the following datasets: the global digital elevation model with a grid spacing of 30 arc seconds (approximately 1 km) provided by the U.S. Geological Survey, the global soil type FAO dataset with a grid spacing of 8 km and the global vegetation/land use dataset provided by FAO with a grid spacing of approximately 1 km. Table 1 summarises some basic settings about the model implementations and references to the schemes adopted. Further and more in-depth descriptions about the physics of the BOLAM and MOLOCH models are found in the work by Buzzi et al. (2014). Daily data of the high-resolution atmospheric hindcast 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 (outer rectangle) and it approximately covers the Med-CORDEX domain (Ruti et al. 2016). 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 hindcast 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. 2. Most of upper air variables were not saved due to disk space issues, whereas the set of near surface variables that were kept includes: 2-m above ground temperature and relative humidity, u- and v-component of wind at 10-, 50-, 80- and 100-m above ground and sensible and latent heat fluxes. Numerical integrations for the period 1979–2019 were carried out thanks to the computational resources granted in the framework of the ECMWF Special

Project SPITBRAN. Using four nodes of the XC40 Cray supercomputer and 144 computing cores, a 1-day simulation is performed in approximately 47 min.

2.2 Data

Initial and boundary conditions to the BOLAM hindcast are provided by ERA5 data (Hersbach et al. 2020). ERA5 provides the new global reanalyses of the European Centre for Medium-Range Weather Forecasting, produced within the framework of the Copernicus Climate Change Service. Such dataset is based on the Integrated Forecasting System model cycle Cy41r2, introduced into operations in 2016, and on the 4D-Var assimilation scheme (Bonavita et al. 2016), which is able to ingest observations from a wide range of platforms (insitu, radiosondes, satellite). ERA5 reanalyses have a significantly enhanced horizontal resolution (approximately 31 km) with respect to the predecessor ERA-Interim (Dee et al. 2011), whose resolution is approximately 79 km. The number of vertical levels of ERA5 is 137, whereas the temporal resolution is 1 h.

To evaluate the accuracy of the BOLAM/MOLOCH hindcast, we compared rainfall modelled data against two gridded datasets: GRIPHO and ARCIS. GRIPHO is an hourly precipitation dataset, available over Italy on a horizontal grid of approximately 3 km (Fantini 2019) and its grid size is 427×358. It is built upon quality checked rain-gauge measurements and is available for the period 2001–2016. It has been used in recent papers (Ban et al. 2021; Caillaud et al. 2021; Coppola et al. 2021) to validate numerical rainfall data at the sub-daily scale at its native resolution or, as in Reder et al. (2022), at a lower resolution (grid spacing approximately 10 km). ARCIS (Climatological Archive for Central-Northern Italy, see Pavan et al. (2019) for details) is a gridded dataset which uses a large number (about 1000) of quality-controlled and homogenised rain-gauge measurements gathered from a plethora of different Italian local bodies, such as Hydrological Services, Agro-Meteorological Services, and Regional/Local Meteorological Services. Such data are interpolated on a regular grid with a resolution of approximately 5 km using a modified version of the Shepard scheme, which is detailed in Antolini et al. (2016); its grid size is 120×120. The dataset covers the north-central Italy for the period 1961–2015 and provides rainfall data on a daily time scale (i.e., 24-h accumulated precipitation).

Data from the ERA5-Land (Muñoz-Sabater et al. 2021) dataset are used as benchmark to assess any added value of the BOLAM/MOLOCH hindcast. ERA5-Land is the result of the global numerical integration of the ECMWF land surface model (namely, the CHTESSEL model cycle Cy45r1, Boussetta et al. (2013)) forced by ERA5 atmospheric data. One of its main advantage is the grid spacing, approximately 9 km, making it suitable for near-surface applications (Muñoz-Sabater et al. 2021). Output variables are focused on the water and energy cycles over land, and in this work we considered rainfall predictions. However, we stress the fact that rainfall, although part of the ERA5-Land dataset, is not a direct output of the CHTESSEL model. On the contrary it is the result of a spatial linear interpolation (Muñoz-Sabater et al. 2021) of ERA5 data onto the triangular-cubic-octahedral TCo1279 operational grid (see Malardel et al. 2016) used by the CHTESSEL model.

2.3 Performance metrics

Several metrics were taken into consideration to give a comprehensive overview of the performances of the BOLAM/MOLOCH hindcast at different time resolutions, from the yearly to the hourly time scale, and to shed a light about any potential added value of using the convection-permitting MOLOCH model.

Although the BOLAM/MOLOCH hindcast is available for the years 1979–2019, it is validated against observed data for those periods overlapping observations availability; that is from the 1st of January 2001 to the 31st of December 2016 as regards GRIPHO and from the 1st of January 1979 to the 31st of December 2015 as regards ARCIS. We decided to perform the validation with respect to both GRIPHO and ARCIS because the former provides data over the whole Italian domain, whereas the latter spans over a longer temporal range (1979–2015), albeit for a smaller portion of Italy (central-northern regions). We also stress the fact that in both cases the domain of verification is smaller than the MOLOCH domain of integration (compare Figs. 1 and 3). Because of the different grid size of observed and modelled data and due to computational constraints, it was necessary to reduce the total number of observed grid points. A thinning method was applied by extracting a regular grid of points out of the GRIPHO dataset and selecting one pixel every 0.25°, both in the latitude and longitude direction; see the geographical distribution in Fig. 3. The data belonging to this regular network of points represent the ground truth; in the following text, such data are often referred as pseudo weather stations. The subset of GRIPHO points belonging to the ARCIS extent represents the set of ARCIS pseudo stations. Observed data collected at locations shown in Fig. 3 were compared with corresponding modelled data, which were extracted at the nearest grid point containing the pseudo station locations. We focused our analysis at three different time scales: annual, daily and hourly time scales. To have a first clue about the results produced by the BOLAM/MOLOCH hindcast, we show the yearly average maps of accumulated rainfall and compare them with those obtained with GRIPHO/ARCIS data. To validate the outputs quantitatively, we considered some standard verification statistics for continuous forecasts, namely: root mean squared error

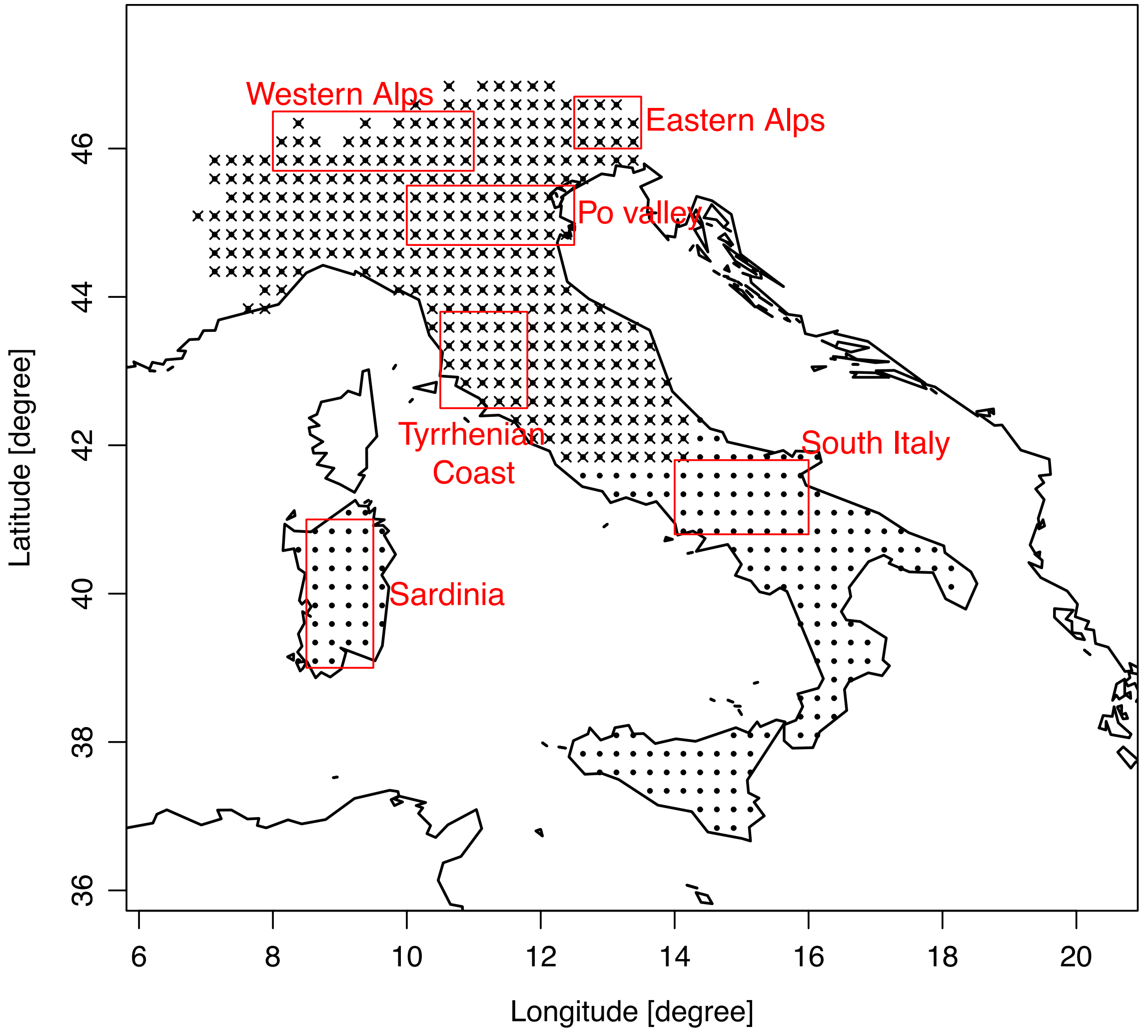

Fig. 3 Closed circles (• symbol) indicate the locations of the GRIPHO pseudo stations (total number is 524). Pseudo stations are obtained by selecting one pixel every 0.25° out of the GRIPHO grid (both in the latitude and longitude direction). Crosses (× symbol) indicate the locations of the ARCIS pseudo stations (total number is 327) and are obtained in the same fashion way. In the central and northern part of Italy the two datasets overlap, whereas in southern Italy ARCIS data are missing. The red boxes indicate the six areas used to plot Fig. 14

(*RMSE*), mean error (*ME*), multiplicative bias (m-*bias*), and Pearson correlation coefficient (*r*). See the Appendix A for details. Furthermore, to summarise multiple aspects in a single plot and have a more streamlined representation of the model performance, we realised the performance diagram (Roebber 2009) against predefined thresholds of precipitation. Such diagrams plot four measures of the dichotomous forecast: probability of detection (*POD*), success ratio (*SR*), bias and critical success index (*CSI*). See Appendix A for details about how these indexes are calculated. Performance diagrams were realised not only for annual precipitations but also for daily rainfall accumulations. The daily time scale is further evaluated by considering the spatial pattern of heavy precipitations, which are often (Ban et al. 2014) identified by considering rainfall amounts exceeding the 90th percentile of wet-days, where a wet-day is defined by a rainfall daily amount greater than 1 mm/day. To illustrate how the hindcasts perform during severe weather events, we show the results for two extreme rainfall events, among the heavier across the last 40 years: the Tanaro flooding (Buzzi et al. 1998) and the Cinque Terre flooding (Buzzi et al. 2014). We chose these two events, because they exhibit different features that characterised precipitations. The first event is a long-standing rainfall event driven by a large-scale humid flow impinging over the Alps. Precipitation was mainly determined by a strong orographic uplift with convective cells embedded in the stratiform precipitation. The second event is concentrated over a small portion of the Ligurian coast; convection processes, triggered by a sharped low-level convergence line, played a major role in determining the rainfall amounts in a relatively short time period (< 12 h). The accuracy of numerical predictions is firstly assessed qualitatively through the visual inspection of forecast against rain-gauge data. To provide a quantitative assessment, we calculated the fractional skill score (*FSS*). In Appendix A, we provide a general description on how to calculate the *FSS*. For a complete description of the method and for details about the underpinning formulas, the reader is referred to the seminal paper by Roberts and Lean (2008).

The models performance at hourly time scale is evaluated by looking at the spatial patterns of extreme hourly precipitations (i.e., amounts greater than the 90th percentile). Precipitation percentiles are calculated by using only wet-hours defined as hours with amounts greater than 0.1 mm/h. The accuracy of hourly precipitation intensities is further evaluated by considering the mean diurnal cycle of precipitations. It is computed averaging rainfall data over six selected zones, whose location is shown in Fig. 3 with the red boxes. These zones are a subjective interpretation of the climatic characterisation of the Italian regions suggested by Brunetti et al. (2001). The quality of the mean diurnal cycle of precipitations is evaluated by looking at the hour of the day when rainfall reaches its peak and the intensity of the peak, which is often referred as amplitude or magnitude (Ban et al. 2014). Furthermore, as done recently by Reder et al. (2022), to evaluate quantitatively the mean diurnal cycle of precipitations, we used the Kling-Gupta Efficiency (*KGE*, see Gupta et al. 2009 and Appendix A for details).

3 Results: performance of precipitation data

3.1 Annual precipitation

To have a first clue about the spatial representation of mean precipitation in the BOLAM/MOLOCH hindcast and compare its results with ERA5-Land data, we show in Fig. 4 the total annual precipitation, averaged over the period 2001–2016. Panel (a) shows the mean annual precipitation derived from GRIPHO data, whereas panels (b–d) illustrate

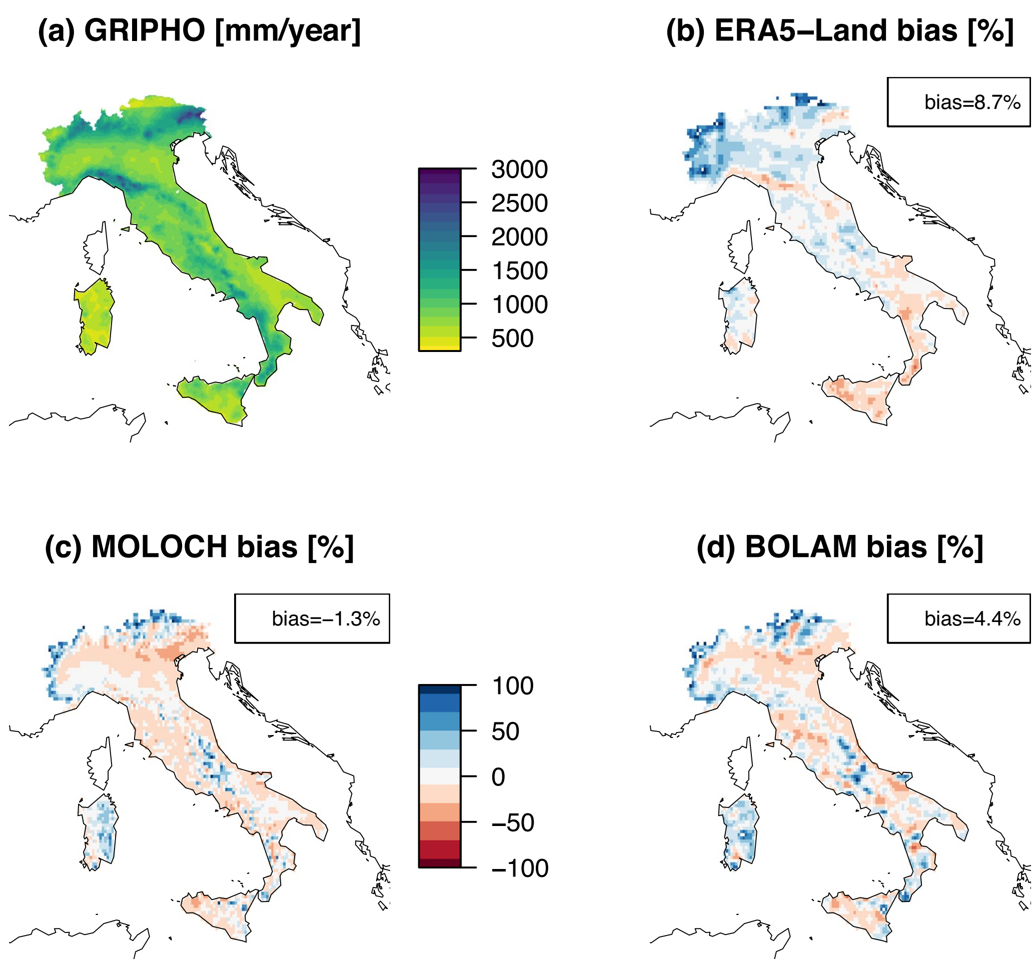

Fig. 4 Panel **a**: Mean annual precipitation (unit mm/year) for the GRIPHO dataset averaged over the 2001–2016 period. Panels **b**–**d**: Relative bias (unit %) for ERA5-Land, MOLOCH and BOLAM, respectively. The inset label in panels (**b**–**d**) indicates the average bias in the whole domain

the relative bias (i.e., the normalised difference between model and observed data, expressed in %) of ERA5-Land, MOLOCH and BOLAM, respectively; blue shades represent overestimation of modelled data, red ones underestimation. From the visual inspection of Fig. 4, we can qualitatively deduce that the three datasets provide similar outputs, although ERA5-Land and BOLAM tend to overestimate total annual amounts (average bias is approximately 9% and 4%, respectively), whereas MOLOCH tends to underestimate (average bias is approximately −1%). Overestimation occurs more frequently over the reliefs (both Alps and central-southern Apennines) and it is more sharped over the western Alps in the ERA5-Land data (see Fig. 4b). The corresponding Figure using ARCIS data (see plots in Fig. 5) confirms such preliminary assessment, although we can note a stronger wet bias of the ERA5-Land, over the Alps in particular, with an average bias value approximately equal to 18%. We underline the fact that to plot panels (b–d) in Figs. 4 and 5, we upscale both observed and BOLAM/MOLOCH data to the pixel size of ERA5-Land, which represents the coarser resolution among the four datasets. For this reason, some of the conclusions drawn so far may be artificially influenced by the algorithm (the bilinear interpolation) used to resample data onto the same grid and may not represent actual deviations from observed data.

To evaluate predictions against selected rainfall amounts, we show in Figs. 6 and 7 the performance diagrams against GRIPHO and ARCIS, respectively. To make the yes/no decision, rainfall thresholds were chosen equal to approximately the 25th, 50th, 75th, 90th, 95th and 98th percentiles of the observed mean annual precipitation for the GRIPHO dataset. These thresholds correspond to 700, 900, 1100, 1400, 1600 and 1800 mm/year, respectively. From the visual inspection of the plots, we can argue that ERA5-Land and BOLAM tend to overestimate annual rainfall amounts (since the red and blue points lie above the diagonal in all the panels but panel (a)). As we consider the higher precipitation thresholds, that is panels (d–f) corresponding to the 90th, 95th and 98th percentiles of the GRIPHO precipitations, the *POD* indexes are greater that 0.5 at the cost of low *SR* values (approximately 0.4), indicating the occurrence of false alarms. However, we note that these high precipitation

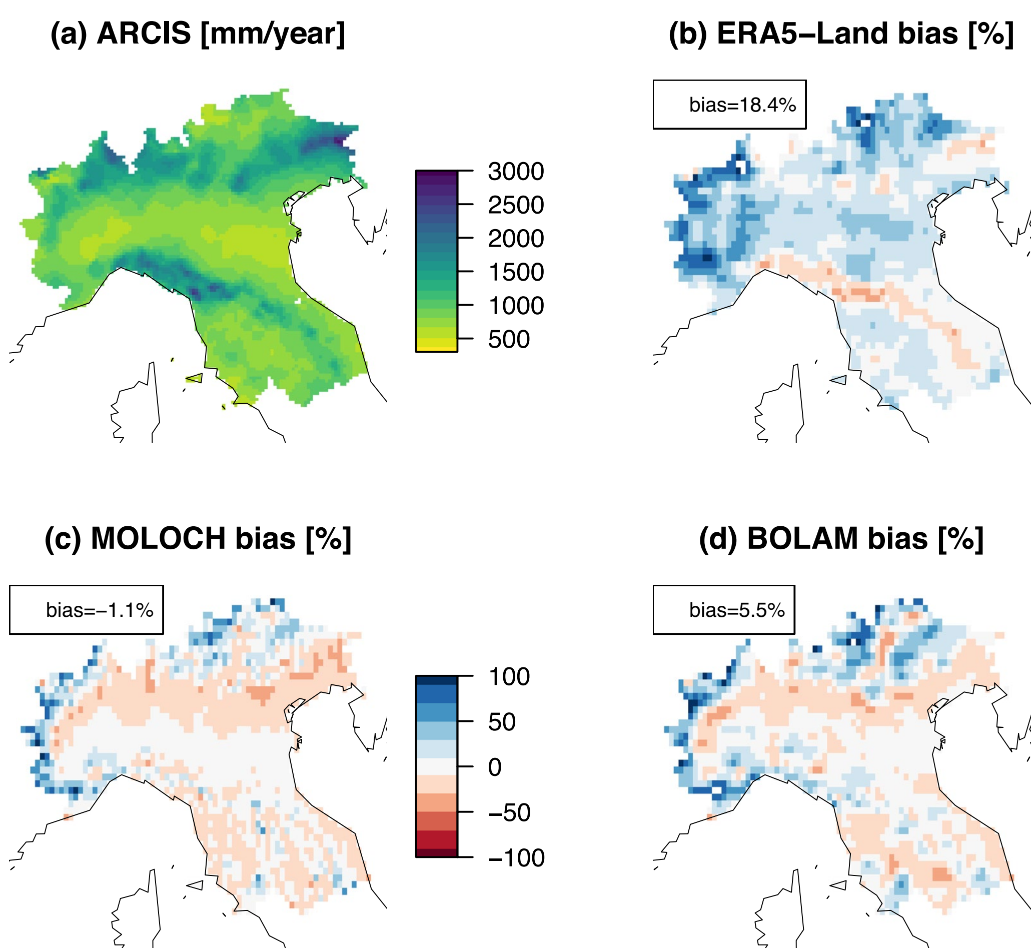

Fig. 5 As in Fig. 4, but for the ARCIS dataset and the 1979–2015 period

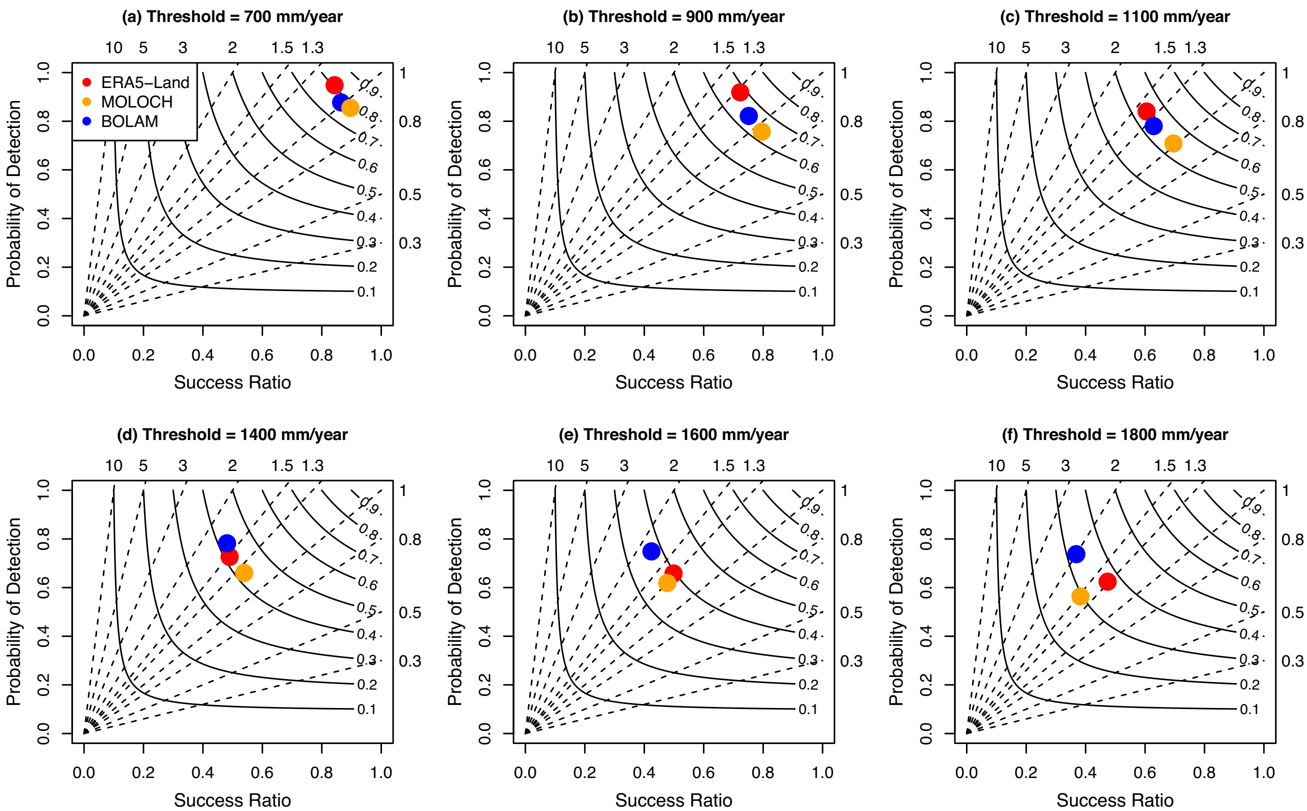

Fig. 6 Performance diagram of total annual precipitations of ERA5-Land (red), MOLOCH (orange), and BOLAM (blue) with respect to the GRIPHO observations. Skill scores are averaged over the 2001–2016 period. The x axis shows the success ratio (*SR*), the y axis shows the probability of detection (*POD*), the curved lines represent the critical success index (*CSI*), and the dashed diagonal lines represent the bias

thresholds interest only a small portion of the Italian territory (less than 7%), corresponding to the Western and Eastern Alps, the Northern Apennines and the southern Apennines (as regards the 90th percentile only).

The interannual variability of the skill scores described in Eqs. (A1)–(A4) are shown in Figs. 8 and 9 as regards GRIPHO and ARCIS, respectively. The average values of skill scores are summarised in Tables 2 and 3. As regards the GRIPHO observational dataset (Fig. 8 and Table 2), MOLOCH data provide the best overall performances with lower errors (291 mm/year and 8 mm/year for the *RMSE* and *ME*, respectively), and a m-*bias* very close to the perfect score, which is 1. The Pearson correlation score *r* is approximately 0.7 for all the models' datasets and differences are found only in the second decimal digit. No significant long-term trend is found for the skill score profiles for all the datasets. The comparison with the ARCIS observations (Fig. 9 and Table 3), demonstrates that MOLOCH provides the best result as regards *RMSE*, but also a stronger underestimation as revealed by the *ME*, which is equal to – 63 mm/year, and the m-*bias*, which is approximately 0.93. On average BOLAM data show the best trade-off considering both the biases (*ME*=26 mm/year and m-*bias*=1.01) and *RMSE*, which is equal to approximately 308 mm/year. ERA5-Land again provides the higher overestimation of rainfall, which is the consequence of overestimation in both the Alpine areas and the plains of the Po Valley (see panel (b) in Fig. 5). Correlation coefficients range from 0.69 for ERA5-Land to 0.74 for BOLAM. As in the GRIPHO case, no significant long-term trend is found in the skill score profiles.

3.2 Daily precipitation

As regards the summer (June, July and August, hereinafter JJA) daily precipitation, we first looked at the spatial pattern of the 90th percentile of wet day precipitation amounts. As in Ban et al. (2014), a wet day is defined as a day with precipitation greater than 1 mm/day. The results are shown in Fig. 10 for the GRIPHO database; panel (a) shows the magnitude of the 90th percentile of observed rainfall, which spans from 10 mm/day to approximately 50 mm/day. Ban et al. (2014) The visual examination of panels (b) and (d) in Fig. 10 suggests that a systematic underestimation of heavy daily precipitations occurs in the parameterised

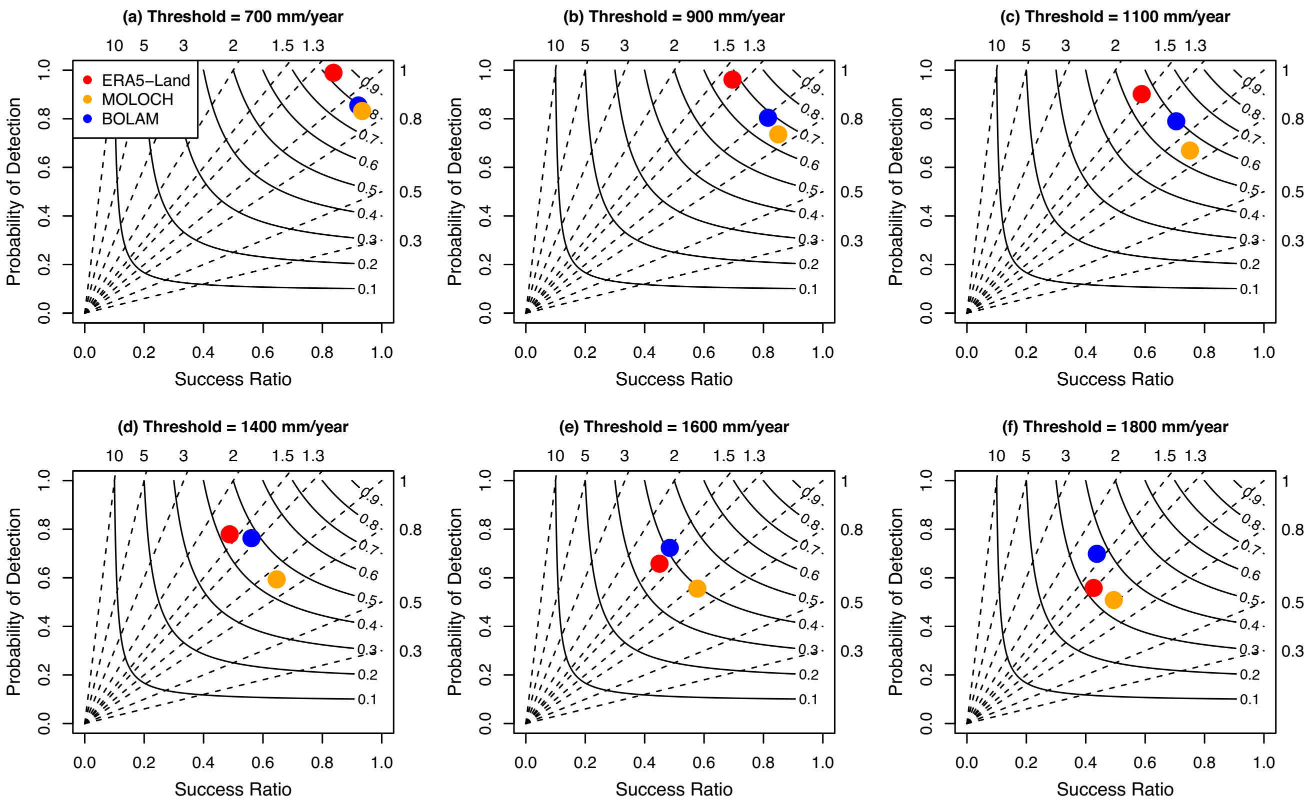

Fig. 7 As in Fig. 6 but with respect to the ARCIS observations and the 1979–2015 period. Precipitation thresholds are those used for the GRIPHO observations

models, whereas the MOLOCH model provides an overestimation, sharper in the South of Italy and over the two main islands (Sardinia and Sicily). This qualitative assessment is confirmed by the average bias (mean over the whole domain), which is −28.5% and −13.9% for the ERA5-Land and BOLAM data, respectively and 17.1% for the MOLOCH hindcast. The comparison performed against ARCIS data (plot not shown) further confirms that the MOLOCH hindcast lacks in reproducing heavy precipitations in the South of Italy and over the two main islands; in fact the overall relative bias is reduced to 5.6%, whereas ERA5-Land and BOLAM data provide approximately the same scores found using GRIPHO observations.

In Fig. 11, we show the performance diagrams of daily precipitations for GRIPHO. To make the yes/no decision, the six precipitation thresholds were subjectively chosen and correspond to 1, 2, 5, 10, 20 and 50 mm/day. Previous works, outlined that parameterised regional climate models tend to produce too much widespread light rain and lack in the description of daily maxima (Kendon et al. 2012; Ban et al. 2014). We find the confirmation of this tendency by looking at panels (a–c) of Fig. 11, which show the red points lying above the diagonal (*bias* is approximately 1.3), although providing satisfactory values as regards *POD* (>0.75) and *SR* (> 0.6). When looking to higher precipitation thresholds (panels (d), (e) and (f) of Fig. 11), ERA5-Land outperforms MOLOCH and BOLAM for the 10 mm/day threshold, then it starts to underestimate the number of cases, that is *bias* < 1, for the latter two thresholds. For BOLAM and MOLOCH data, the performances are almost similar, with both *POD* and *SR* decreasing from 0.6 to 0.5 and then 0.4 as we consider the 10 mm/day, 20 mm/day and 50 mm/day thresholds, respectively. Similar conclusions can be drawn when comparing numerical data against ARCIS observations (plot not shown). However, we report that the skill scores found for the ARCIS database are poorer than those found for the GRIPHO case for all the thresholds. This may be related to the fact that errors are larger in northern Italy because of an inadequate representation of the orography of the Alps.

3.3 Hourly precipitation

Figure 12 shows the spatial pattern of the 90th percentile of summer wet hour precipitations. The normalised bias (expressed in %) of ERA5-Land, MOLOCH and BOLAM is shown in panels (b–d), respectively, whereas panel (a) shows GRIPHO data. Figure 13 shows the corresponding results obtained for winter. Panels (b) and (d) in Fig. 12

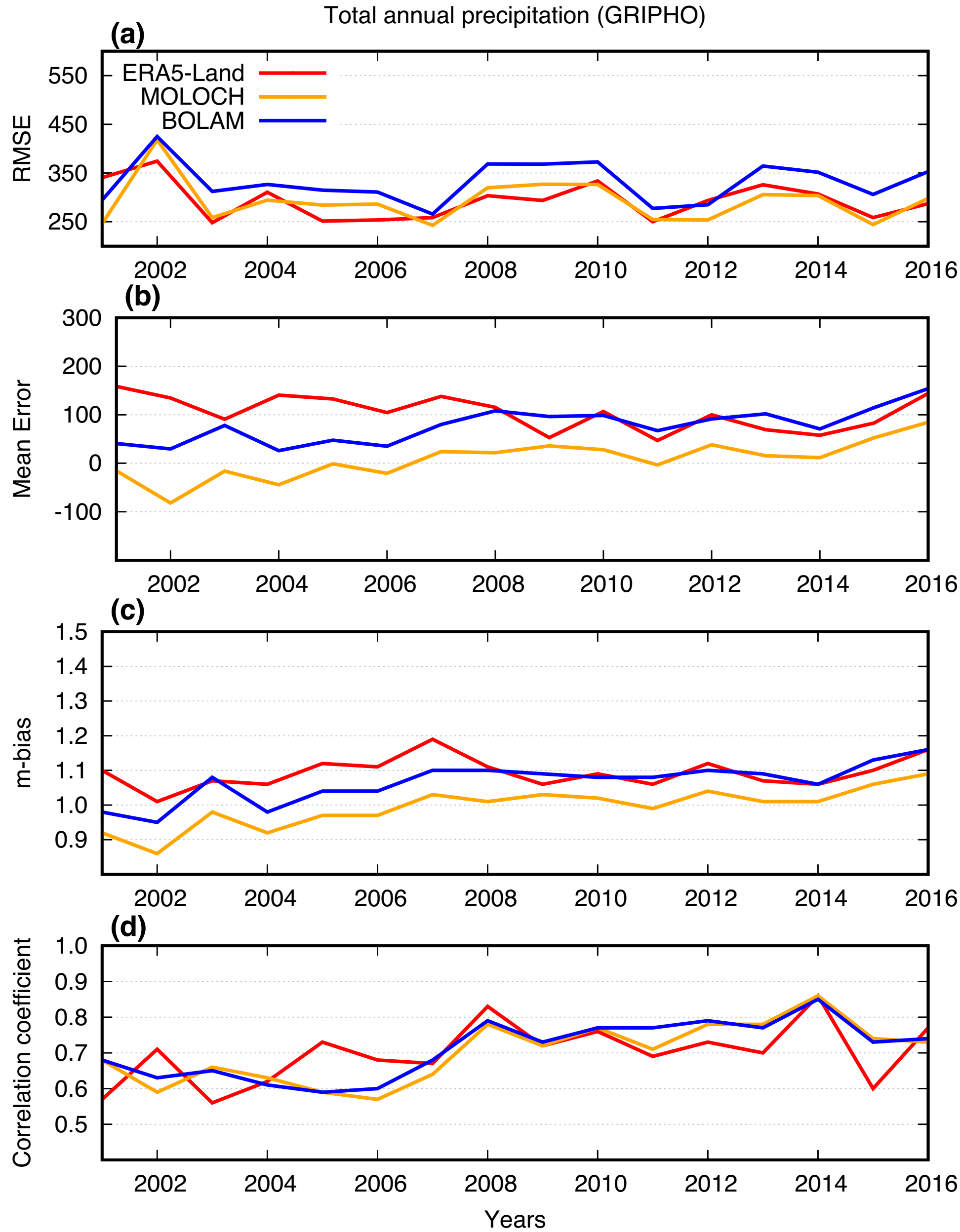

Fig. 8 2001–2016 interannual variability of the skill scores for the ERA5-Land (red line), MOLOCH (orange line) and BOLAM (blue line) datasets with respect to the GRIPHO observations. Skill scores shown are: **a** root mean square error, **b** mean error, **c** multiplicative bias and **d** correlation coefficient

demonstrate that both ERA5-Land and BOLAM data underestimate heavy hourly precipitations across the whole Italian domain, with an average relative bias of −61.1% and −49.9%, respectively. The underestimation is stronger in southern Italy, where average bias values are lower than – 50% for ERA5-Land in Sicily and Sardinia islands in particular. On the other hand, MOLOCH (panel (c) in Fig. 12), exhibits a general overestimation of the 90th percentile of hourly precipitations with an average bias of approximately 25.0% across the whole Italian domain. Such overestimation is stronger in the Po Valley and central-southern Italy, where bias values are greater than 50%. Over the Alps, MOLOCH provides reliable data, with an overestimation, in general, less than 15–20%. Underestimation is found over the coasts, in particular in the Ligurian and northern Tyrrhenian coasts, in northern Sicily and western Sardinia; such underestimation is, approximately, no less than – 30%. Looking at the winter results (shown in Fig. 13), models' skills improve. MOLOCH provides the best overall score in reproducing the 90th percentile of winter hourly precipitation; the bias is positive (i.e., indicating overestimation of rainfall) and greater than 20% in the western and central Alps, in the Sardinia island and in scattered zones across southern Italy, but in the rest of the country, the relative bias is approximately between 10 and –20% (average −4.3%). BOLAM provides a similar pattern, but a larger average underestimation (bias −

Fig. 9 As in Fig. 8 but with respect to the ARCIS observations and the 1979–2015 period

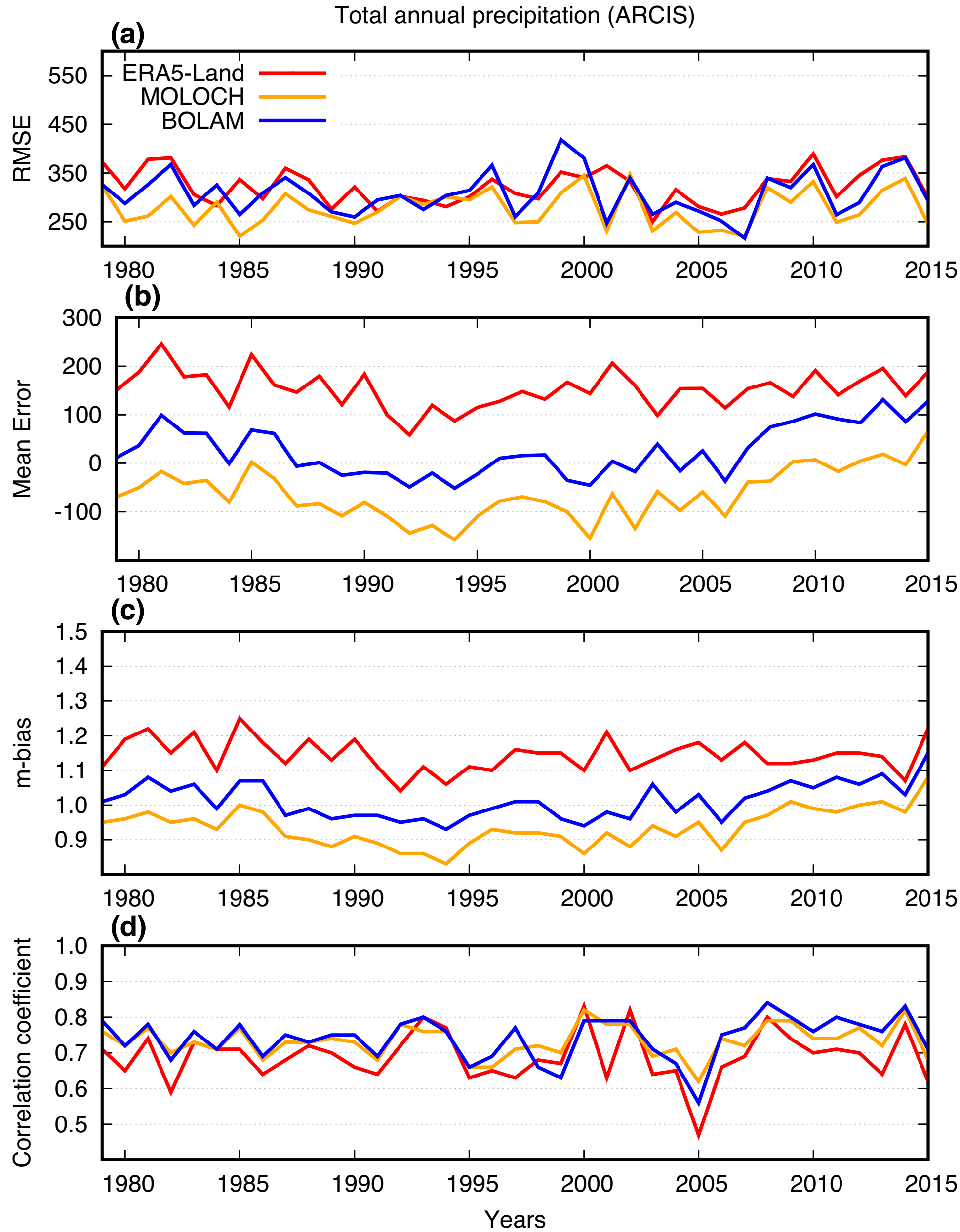

Table 2 Average skill scores for annual precipitation with respect to GRIPHO observations

Dataset	RMSE	Mean error	m-bias	Correlation coefficient
ERA5-Land	293	104	1.09	0.70
MOLOCH	**291**	**8**	**0.99**	0.70
BOLAM	331	77	1.06	**0.71**

Best scores are indicated in bold

Averages are calculated over the 2001–2016 period

Table 3 As in Table 2 but with respect to the ARCIS observations and the 1979–2015 period

Dataset	RMSE	Mean error	m-bias	Correlation coefficient
ERA5-Land	320	151	1.14	0.69
MOLOCH	**277**	– 63	0.93	0.73
BOLAM	308	**26**	**1.01**	**0.74**

Best scores are indicated in bold

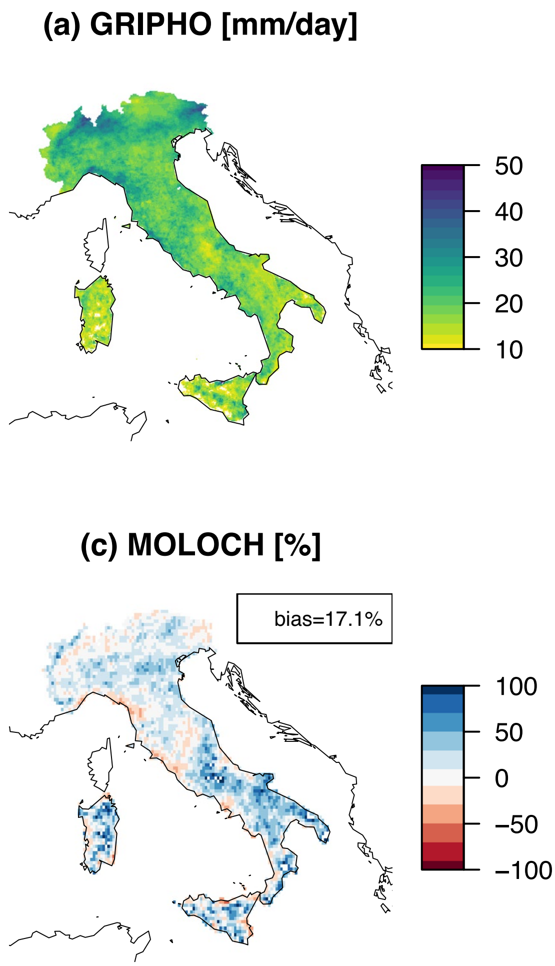

Fig. 10 **a** Spatial distribution of the 90th percentile of summer daily precipitation (unit mm/day) for the GRIPHO observations averaged over the 2001–2016 period. **b**–**d** Relative bias (unit %) for ERA5-Land, MOLOCH and BOLAM, respectively. The inset label in panels (**b**–**d**) indicates the average bias in the whole domain. The 90th percentiles are computed with respect to wet days (rainfall > 1 mm/day)

19.0%), in particular in central and southern Italy with values less than –20%. ERA5-Land data further degrade the score with a general underestimation in the whole domain and an average bias approximately equal to −35.8%.

To further evaluate how the models' hindcasts are able to represent sub-daily precipitations, in Fig. 14 we show the mean diurnal cycle of JJA precipitations over the six areas sketched in Fig. 3 with the red boxes; data are averaged over the 2001–2016 period. We show the results obtained for JJA, because during summer the convection is the leading factor that determines precipitation. For the Western and Eastern Alps (panels (a) and (b), respectively), the two parameterised models (ERA5-Land and BOLAM) provide an early onset of convection processes (approximately at 10:00 UTC for BOLAM and even earlier for ERA5-Land), which results in an anticipated peak of afternoon precipitation (approximately at 15:00 UTC for ERA5-Land). This assessment is not new in the Alps and it is associated to the use of a deep convection parameterization scheme rather than to the poor representation of orography (Prein et al. 2013). On the other hand, the convection-permitting MOLOCH model agrees better with the GRIPHO data, although its rainfall peak is anticipated by one hour for the Western Alps (17:00 UTC vs 18:00 UTC). The amplitude, that is the maximum value of the daily cycle, of the MOLOCH profile matches better observations. This is confirmed by the *KGE* index, which is the higher among the three hindcasts considered (0.21 for Western Alps and 0.60 for Eastern Alps). To summarise such results, in Table 4 we report, for each area, the hour of the rainfall peak, the amplitude and the *KGE* index. As we consider the Po Valley (see panel (c) in Fig. 14), located in the northern part of Italy and characterised by flat areas with an average topography of approximately 27 m, the mean diurnal cycle of observed rainfall do not show a sharped bell-shaped profile. Such feature is better reproduced by the MOLOCH data than the other two datasets; in fact, the *KGE* index is 0.11 (the greatest), the timing of the peak, which is 17:00 UTC, is matched and the amplitude is closer to the observed

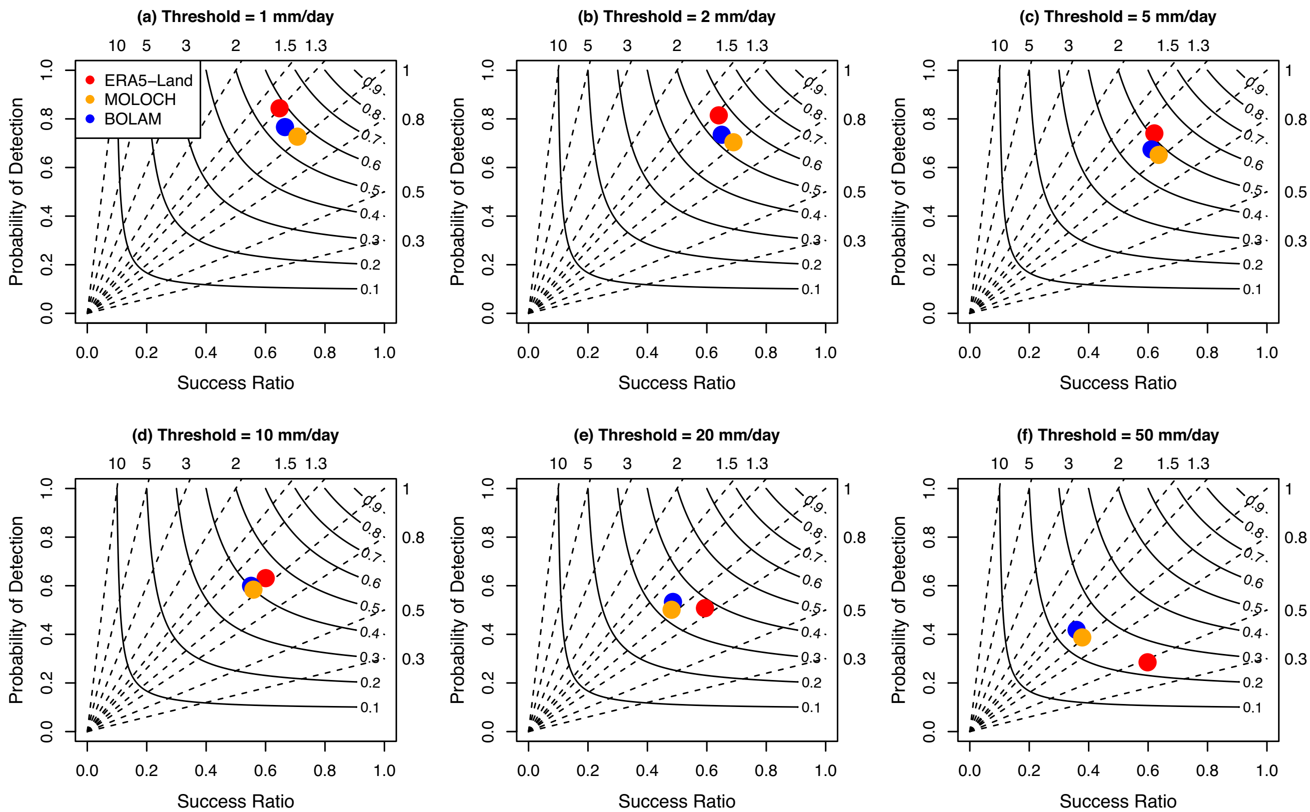

Fig. 11 Performance diagram of the daily precipitation of ERA5-Land (red), MOLOCH (orange), and BOLAM (blue) with respect to the GRIPHO observations. Skill scores are averaged over the 2001–2016 period

one (0.08 vs 0.09 mm/h) than ERA5-Land and BOLAM data. Panels (d–f) of Fig. 14 show the mean diurnal cycle for three areas located to the South of the Apennines mountainous chain: the northern Tyrrhenian Coast (d), the Sardinia Island (e) and a portion of the central-southern Italy (f). These areas are hilly, characterised by an approximate average topography of 230 m, 352 m and 412 m, respectively. The mean diurnal cycle of both observed and modelled data are bell-shaped profiles. In general, ERA5-Land provides data that agree better to observed ones, since both the hour of rainfall peak and the amplitude match observations, and the *KGE* index is the highest and greater than 0.75 overall (see Table 4). However, we underline how both BOLAM and MOLOCH data provide reliable results, in particular as regards the timing of the peak provided by MOLOCH; but the amplitude is either overestimated (0.07 vs 0.05 mm/h in Sardinia and 0.20 vs 0.15 mm/h Central-South Italy), or underestimated in the Tyrrhenian Coast (0.11 vs 0.13 mm/h).

3.4 Study cases

Once investigated the performances of annual to hourly precipitations among the three datasets, we now focus on two subjectively selected heavy precipitation events to illustrate how rainfall patterns and peaks are represented differently. We selected two heavy rainfall events among those listed in the Polaris (Popolazione a Rischio da Frana e da Inondazione in Italia) project. The Polaris database (https://polaris.irpi.cnr.it/, accessed 9 April 2022) is managed by the Research Institute for Geo-Hydrological Protection (IRPI) of the Italian National Research Council (CNR). It gathers severe weather events that caused flooding or landslides/debris flows over Italy across the period 1951–2019. The events we chose are: the Piedmont flooding occurred during the first days of November 1994 (hereinafter the PIE1994 case) in the north-western Italy (Buzzi et al. 1998), and the convective precipitation event of 25 October 2011 (Buzzi et al. 2014) that caused the flooding of the Cinque Terre (hereinafter the CT2011 case), located onshore of the Ligurian Sea.

The PIE1994 case was initially analysed by Buzzi et al. (1998), who provided an exhaustive description of both the synoptic and small scale meteorological conditions that led to the disastrous flooding of a relatively vast area in north-western Italy. Recently, several authors (Capecchi 2020; Cerenzia et al. 2020; Davolio et al. 2020; Garbero and Milelli 2020; Parodi et al. 2020) produced new reforecasts of the event by using state-of-the-art and limited-area numerical

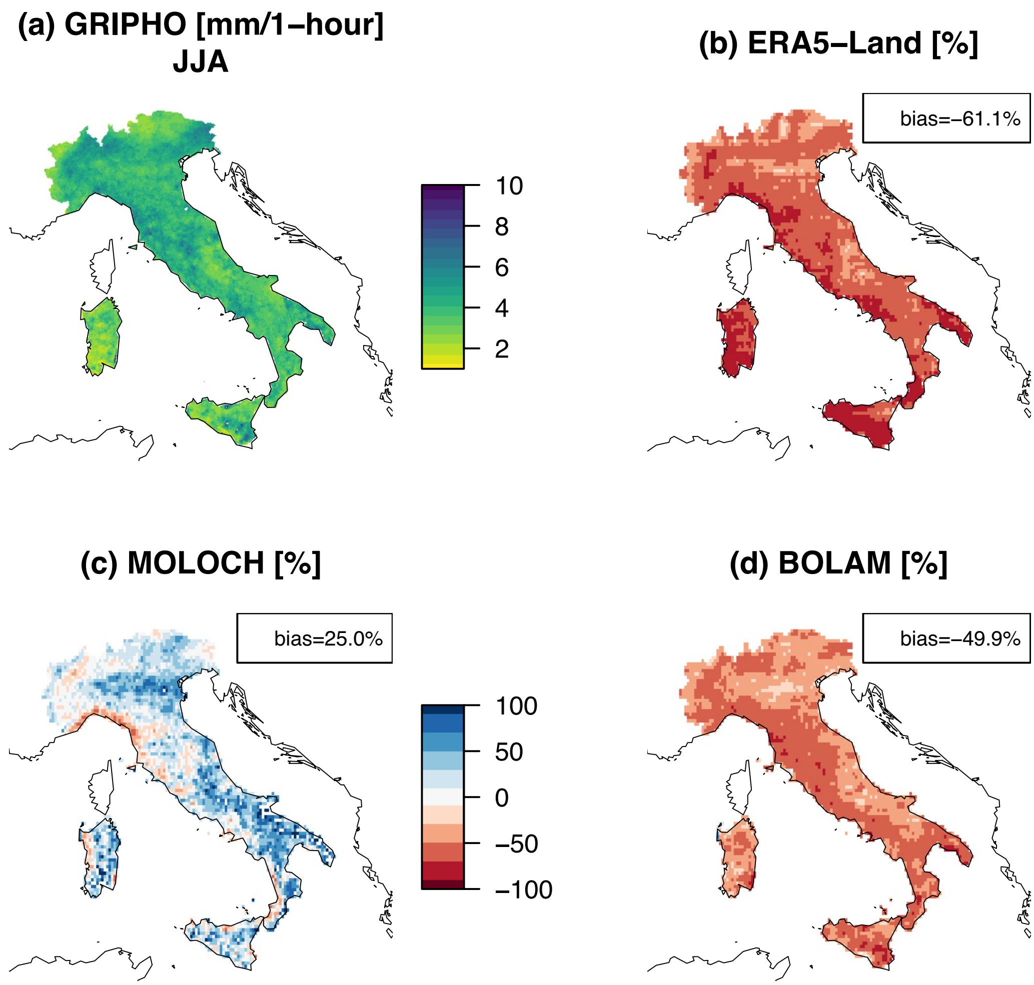

Fig. 12 **a** Spatial distribution of the 90th percentile of summer hourly precipitation (unit mm/h) for the GRIPHO observations averaged over the 2001–2016 period. **b**–**d** Relative bias (unit %) for ERA5-Land, MOLOCH and BOLAM, respectively. The inset label in panels (**b**–**d**) indicates the average bias in the whole domain

weather models (Meso-NH, COSMO, MOLOCH and WRF), all set in a convection-permitting mode. Although numerical experiments that shortly followed PIE1994 (Buzzi et al. 1998; Ferretti et al. 2000) were able to reproduce the rainfall due to the direct uplift of the flow impinging the Alps, it is by implementing higher-resolution and convection-permitting models that is possible to attain more accurate forecasts. In fact, as outlined in Davolio et al. (2020) such models are able to reconstruct both the convective activity extending to the North of the Ligurian Apennines and the activity of the convective cells embedded in the stratiform orographic precipitation in the northern Alpine part of the area (Rotunno and Houze 2007). Rainfall amounts accumulated in the 24-h period ending at 00:00 UTC 6 November 1994 are shown in Fig. 15 for: (a) observations, (b) ERA5-Land, (c) MOLOCH and (d) BOLAM data. All the three modelled datasets provide areas of maximum precipitation that agree well with rain-gauge data, with two maxima, one locate in the northern part of the region and one, lower, in the southern part. However, MOLOCH maxima (316 mm) are closer to observation peaks (323 mm) than the other two datasets in the northern part of the region. In fact, both ERA5-Land and BOLAM underestimate maximum precipitation by approximately 30% (201 mm and 218 mm, respectively). In the southern part, where convection occurred and played a major role (Capecchi 2020; Davolio et al. 2020), the MOLOCH model yields too much rain (maxima up to 294 mm against observed values up to 252 mm), possibly due to a strong convective activity, as found also by Davolio et al. (2020) in their numerical experiments with the MOLOCH model. BOLAM provides an accurate estimate of precipitation maxima (243 mm close to the observed 252 mm), whereas ERA5-Land underestimates by approximately 50% the maximum (127 mm).

As regards CT2011, a comprehensive description of the event is provided by Buzzi et al. (2014), who tested

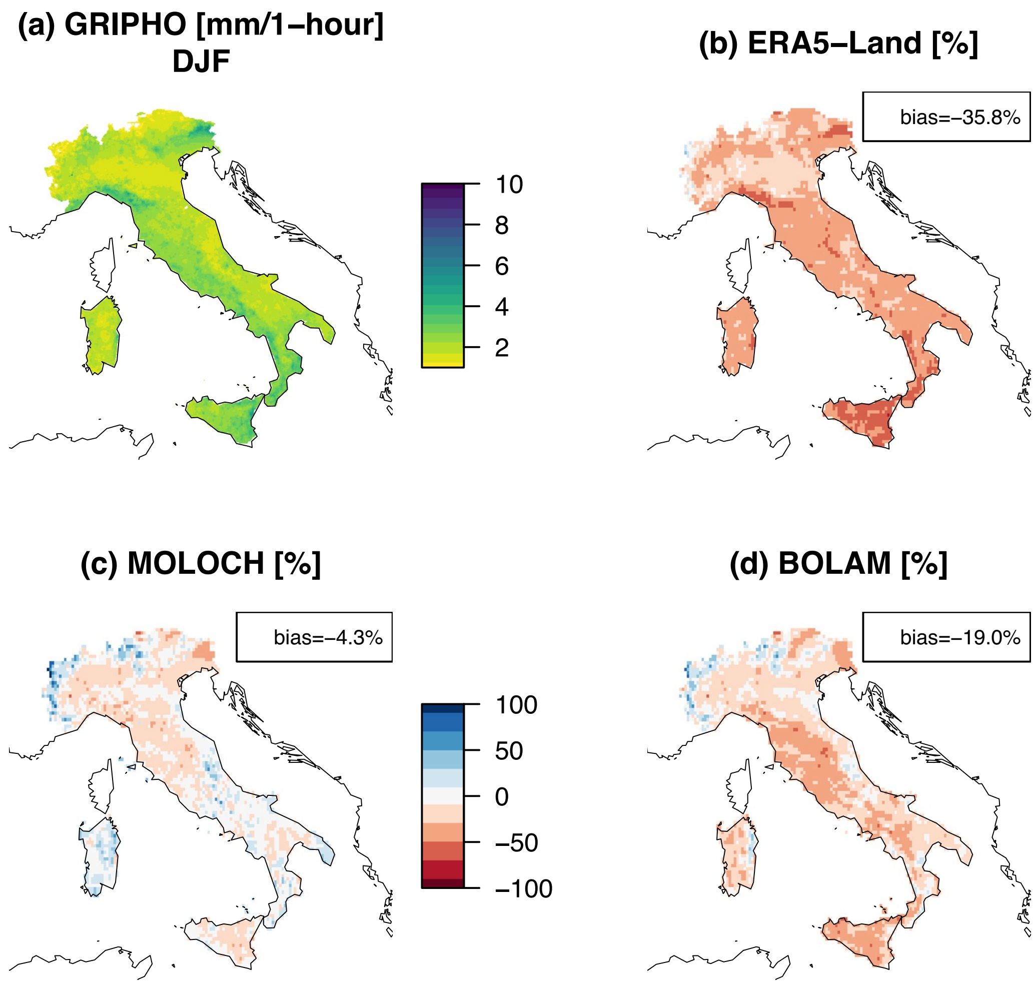

Fig. 13 As in Fig. 12 but for winter hourly precipitations (unit mm/h)

the predictability capabilities of the MOLOCH model at increasing horizontal resolution. We present the results of the BOLAM/MOLOCH models for the CT2011 case to enable comparisons with previous works (Buzzi et al. 2014; Capecchi 2021) and to evaluate the added value carried by the use of ERA5 data as initial and boundary conditions. The precipitation accumulated in the 24-h period ending at 00:00 UTC 26 October 2011 is shown in Fig. 16. Panel (a) shows available rain-gauge measurements and panels (b–d) show predictions produced by the ERA5-Land, MOLOCH and BOLAM datasets, respectively. In all the four panels, inset numbers indicate precipitations maxima in the surrounding area. In the relatively small area of interest (extent is approximately 10^4 km^2) two maxima are present in the observations: 538 mm in the western part and 218 mm in the south-eastern part, separated by an area of relatively low precipitation values (< 50 mm). Such pattern is reconstructed by the convection-permitting MOLOCH model, but is missed by both ERA5-Land and BOLAM. Furthermore, precipitation maxima of the MOLOCH model (196 mm) are closer to rain-gauge data (538 mm) than the other two models, although largely underestimate observations. This is not surprising, in fact Buzzi et al. (2014) demonstrated that even with very high-resolution numerical experiments (grid spacing up to 1.5 km) simulated precipitation maxima of the CT2011 case can reach no more than 335 mm in 24 h. On the other hand Capecchi (2021) proved that with a multi-model ensemble initialised with the 50 members of ECMWF ensemble data at different lead times (up to 600 members) the highest precipitation peaks achievable is approximately 400 mm.

For both events, the higher the resolution, the better the agreement of rainfall peaks between observations and numerical data. This is well known (Buzzi et al. 2014); however we also found that the MOLOCH short-range simulation (forecast's length is 30-h) fed by the ERA5 reanalyses does not consistently outperform in terms of rainfall peaks previous experiments, which used operational analysis and forecasts as initial and boundary conditions. As regards PIE1994, the precipitation maxima we found (316 mm/

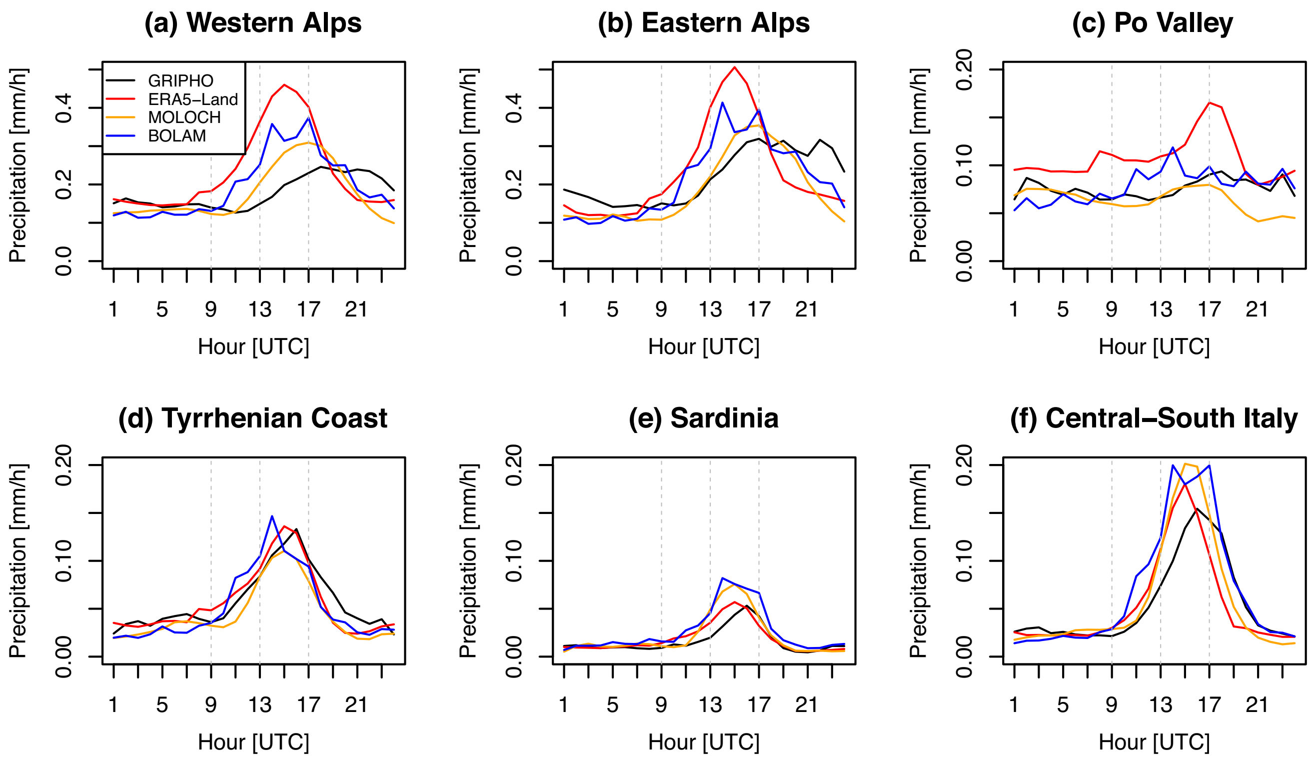

Fig. 14 Mean diurnal cycle of summer precipitations (unit mm/h) obtained for GRIPHO (black line), ERA5-Land (red line), MOLOCH (orange line) and BOLAM (blue line) data and averaged over **a** Western Alps, **b** Eastern Alps, **c** Po Valley, **d** Tyrrhenian Coast, **e** Sardinia and **f** Central-South Italy. Note the different y-axis range between the Alps (panels (**a**) and (**b**)) and the remaining zones

Table 4 For each dataset and for each area (shown in Fig. 3) it is reported: the maximum of the daily cycle (namely, 'Amplitude') and the hour at which the maximum is reached (namely, 'H max')

Dataset	H max	Amplitude	KGE	H max	Amplitude	KGE	H max	Amplitude	KGE
	Western Alps			Eastern Alps			Po Valley		
GRIPHO	18	0.24	/	17	0.31	/	18	0.09	/
ERA5-Land	15	0.46	−0.80	15	0.50	0.07	**17**	0.16	−0.57
MOLOCH	**17**	**0.30**	**0.21**	**17**	**0.35**	**0.60**	**17**	**0.08**	**0.11**
BOLAM	**17**	0.37	−0.11	14	0.41	0.47	14	0.11	−0.08
	Tyrrhenian Coast			Sardinia			Central-South Italy		
GRIPHO	16	0.13	/	16	0.05	/	16	0.15	/
ERA5-Land	**15**	**0.13**	**0.88**	**15**	**0.05**	**0.75**	**15**	**0.18**	**0.81**
MOLOCH	**15**	0.11	0.76	**15**	0.07	0.27	**15**	0.20	0.62
BOLAM	14	0.14	0.75	14	0.08	−0.02	17	0.20	0.46

The *KGE* index is shown in the last column of each area. The best score among the three numerical datasets is highlighted in bold

day and 294 mm/day in northern and southern Piedmont, respectively) agree well with those found by Davolio et al. (2020). As regards CT2011, the precipitation maxima we found when running MOLOCH at 2.5 km grid spacing (196 mm/day) is between what Davolio et al. (2015) found running two MOLOCH resolutions: the first one at 3 km grid spacing (175 mm/day) and second one at 2 km grid spacing (215 mm/day).

To add more insights into the dynamical aspects of the PIE1994 and CT2011 events, in Fig. 17 we show the temporal evolution of hourly rainfall accumulations for two selected rain-gauges: Oropa as regards PIE1994 (upper panel) and Calice al Cornoviglio as regards CT2011 (lower panel). Observed data for PIE1994, shown with the black line, indicate that rainfall was mainly due to the orographic uplift of moist air (Buzzi et al. 1998), in fact the Oropa

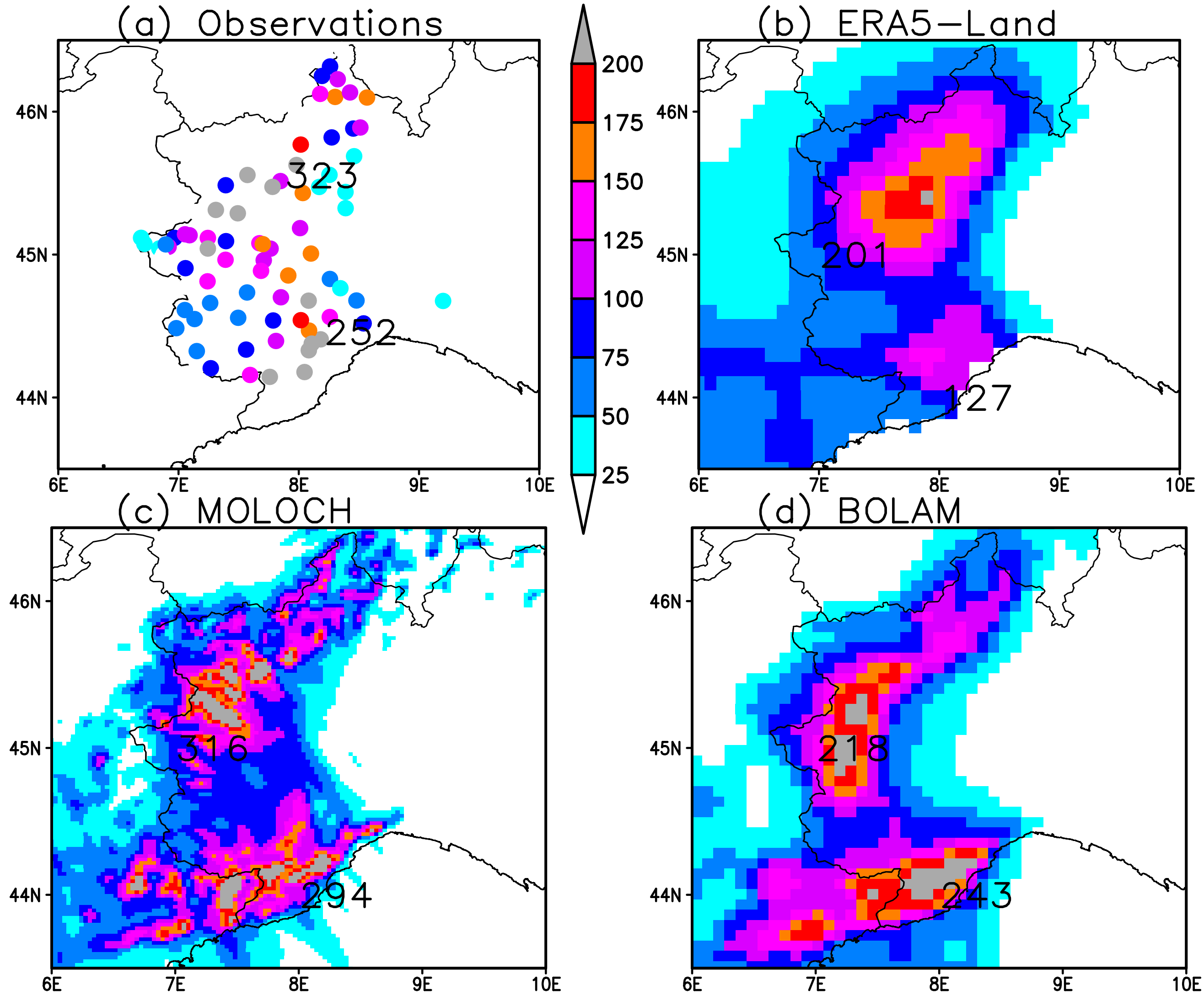

Fig. 15 Rainfall amounts accumulated in the 24-h period ending at 00:00 UTC 6 November 1994 for: **a** Observations, **b** ERA5-Land, **c** MOLOCH and **d** BOLAM. In each panel, inset numbers indicate precipitation maxima in the surrounding area

rain-gauge is located at 1186 m above sea level and the maximum rainfall rate is approximately 27 mm/1-h. All the three numerical datasets, reconstruct such dynamics, with a steady increment of the rainfall profile across the day; however, MOLOCH almost equals the observed daily accumulated precipitation (292 mm against 298 mm) and outperforms ERA5-Land and BOLAM, which underestimate the rainfall peak by approximately 47% and 29%, respectively. Further, MOLOCH rainfall rate is up to 22 mm/1-h and agrees well with observations. Lower panel of Fig. 17 reveals the convective nature of precipitation during the CT2011 event, as recorded at the Calice al Cornoviglio rain-gauge. Maximum rainfall rate is 121 mm/1-h, average rainfall rate is 18 mm/1-hour and standard deviation of 30 mm/1-h; in the second part of the day (last 12 h) the average is 31 mm/1-h and the standard deviation is 39 mm/1-h. None of the modelled data are able to reconstruct such dynamics in terms of rainfall rate, and the daily precipitation amount is underestimated by approximately 78%, 57% and 62% for ERA5-Land, MOLOCH and BOLAM, respectively.

To evaluate more quantitatively rainfall predictions for the PIE94 and CT2011 cases, we show the *FSS* in Fig. 18. Since we kept the neighbourhood size constant and equal to 18 km, this assessment allows for a precipitation intensity analysis. By looking at the precipitation thresholds below which the *FSS* corresponds to a *POD* less than 0.5 and *CSI* less than 0.33 (this is commonly referred as the FSS_u value, which is indicated by the horizontal dashed line in Fig. 18), we claim that, the BOLAM/MOLOCH simulations outperform the ERA5-Land accuracy for both events. In fact, as regards PIE1994 (plot on the left), the MOLOCH profile (orange line) is greater than

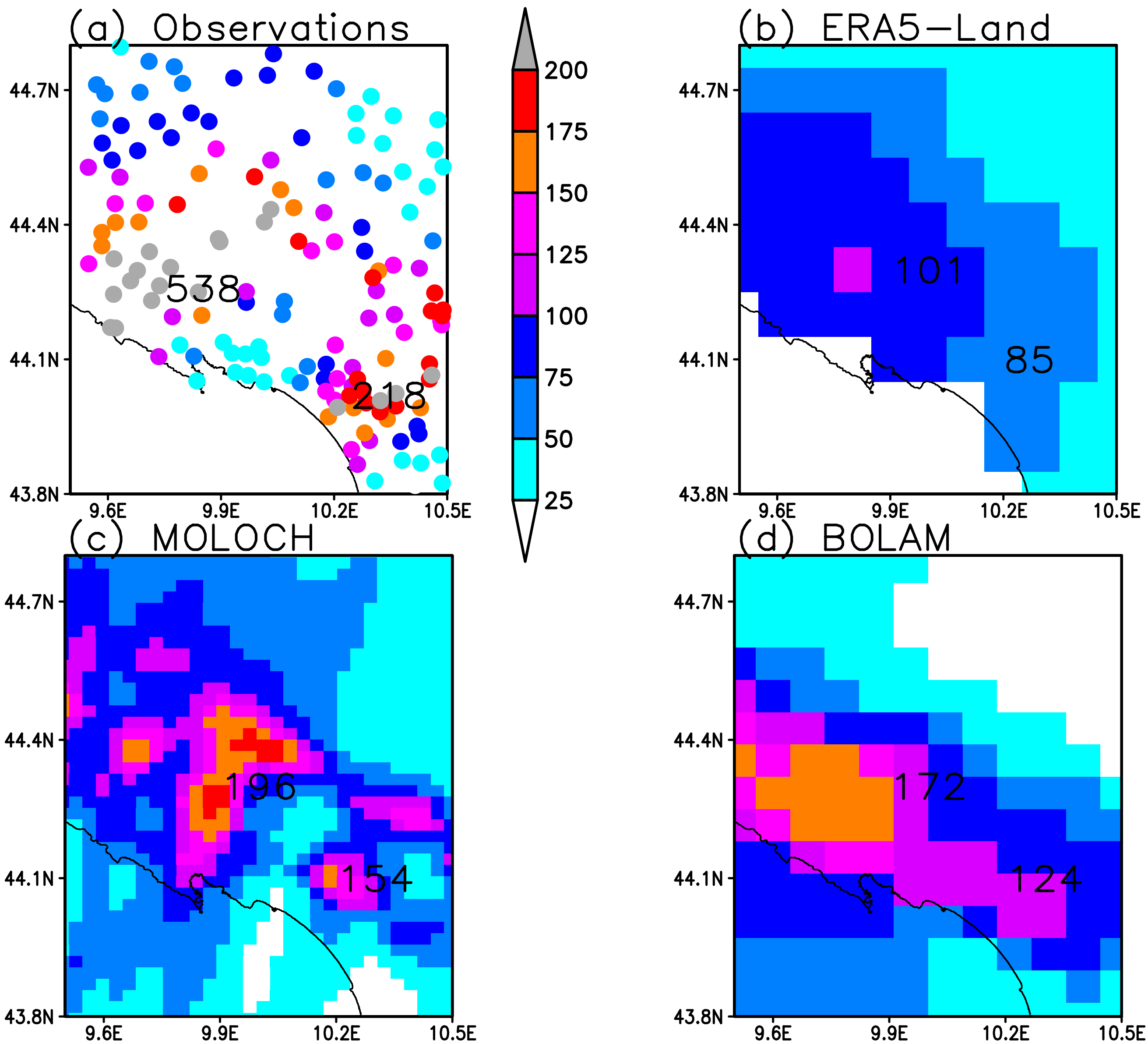

Fig. 16 As in Fig. 15 but for the 24-h period ending at 00:00 UTC 26 October 2011

FSS_u for all the precipitation thresholds up to 250 mm, the BOLAM profile (blue line) drops below the dashed line at approximately the 200 mm threshold, and the red line (corresponding to ERA5-Land data) is below the dashed line for each precipitation threshold greater than 180 mm, approximately. As regards the CT2011 case (plot on the right), we can draw similar conclusions, except for the fact that MOLOCH is producing useful predictions up to 175 mm, BOLAM up to 150 mm and ERA5-Land up to 100 mm, approximately. We stress that the FSS_u values shown in Fig. 18 are calculated for the 250 mm and 200 mm precipitation thresholds for the PIE1994 and CT2011 case, respectively.

4 Discussions

We realised a limited-area, high-resolution and long-term hindcast based on the BOLAM and MOLOCH models fed by ERA5 data as initial and boundary conditions. Precipitation performances were verified against two gridded observational datasets (GRIPHO and ARCIS) and ERA5-Land data were used as benchmark to assess the added value of running the convection-permitting MOLOCH model.

Considering annual precipitations, results demonstrate that the MOLOCH model provides the best overall scores, reducing the wet bias found in the Alps for ERA5-Land. The wet bias of ERA5/ERA5-Land is a known issue and it was recently suggested (Lavers et al. 2022) that it is due

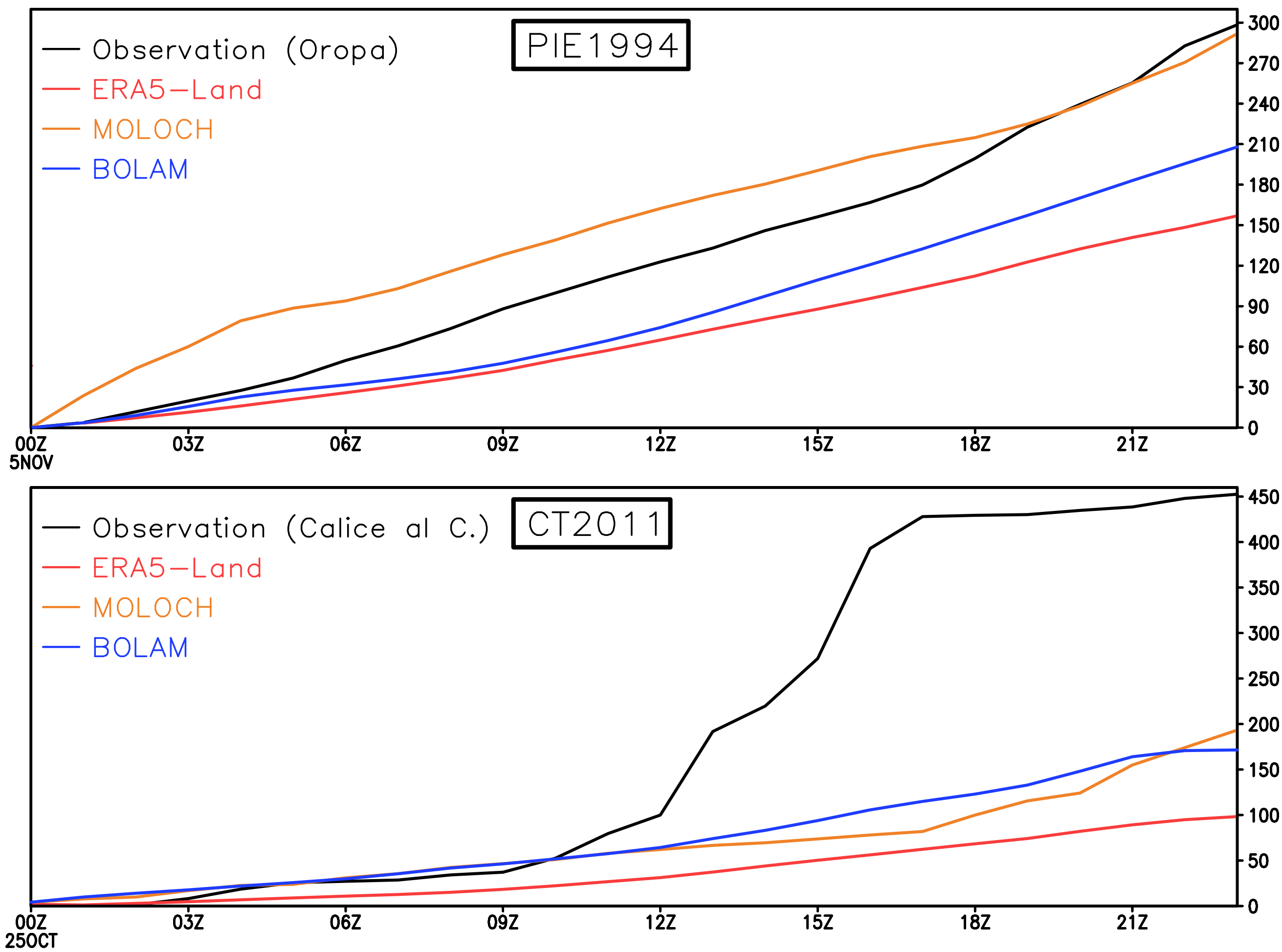

Fig. 17 PIE1994 (upper panel): accumulated precipitation from 00:00 UTC to 23:00 UTC on the 5th of November registered at Oropa rain-gauge. CT2011 (lower panel): accumulated precipitation from 00:00 UTC to 23:00 UTC on the 25th of October registered at Calice al Cornoviglio rain-gauge. In both panels the black line indicates observed data, the red, orange and blue lines indicate the corresponding ERA5-Land, MOLOCH and BOLAM data, respectively

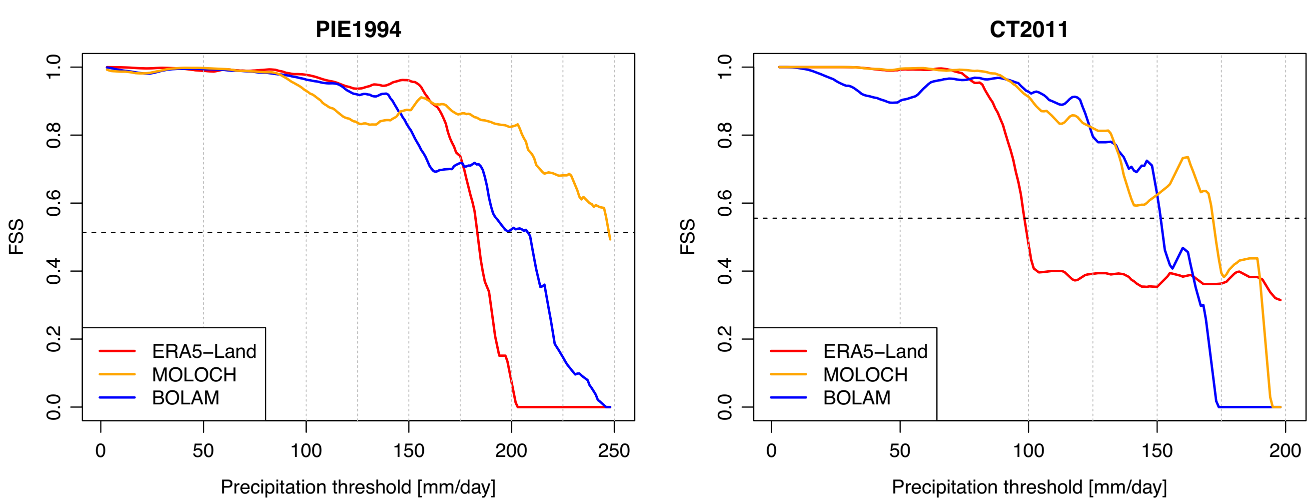

Fig. 18 Fractional Skill Score (*FSS*) as a function of the precipitation threshold for the PIE1994 (left) and CT2011 (right) cases. The neighbourhood length is constant and equal to 18 km. The dashed line represents FSS_u (see the text for explanations). The ARCIS database was used to calculate the scores

to the ERA5 point elevations which are, on average, higher than actual stations and potentially can cause the orographic enhancement of precipitation. The MOLOCH model, which has a better representation of orography, was proved to reduce the wet bias, in particular in complex orography areas and for all the precipitation thresholds up to the 75th percentile of observed data (see panels (a)–(c) in Fig. 6 in particular). We also focused on the interannual variability of the accuracy of total annual precipitations (see Figs. 8 and 9) and found that the profiles of the skill scores do not present any significant increase/decrease over time. This assessment is relevant because it is known that changes in the amounts and types of observational data that is assimilated by ERA5 may produce artificial trends or variability along the time series.

At the daily time scale, beside the underestimate of ERA5-Land, which was expected since the relatively coarse resolution of ERA5-Land data tend to generate a smoothing of extreme precipitation at the daily time scale (Lavers et al. 2022; Reder et al. 2022), what deserves a discussion is the average wet bias of MOLOCH. Looking at panel (c) in Fig. 10, it is evident how the wet bias is prominently present in the southern part of the country. This area is narrow (longest distance from coast to coast is approximately 270 km) and characterised by complex orography with the Apennines chain crossing the area from North to South (highest peak 2912 m above sea level) and with steep orography between mountainous and coastal areas. Here, the MOLOCH provides too many rainfall events exceeding the 90th percentiles of summer precipitations, possibly because of too high temperatures triggering rainfall induced by thermals.

The performance diagram shown in Fig. 11 confirms (Lavers et al. 2022) that ERA5-Land data exhibit weak wet biases as regards light rains (all red points lie above the diagonal in panels (a)–(c)), although providing satisfactory scores as regards POD and SR (both greater than 0.6). Performances in reproducing medium to heavy daily rainfall accumulations (20 and 50 mm/day) are worst, whereas a daily rainfall accumulation equal to 10 mm/day appears the optimal threshold over which ERA5-Land data perform better, coherently to what was found also by Bonanno et al. (2019). The decreasing agreement between the ERA5 data and observations with increasing daily precipitation intensity was already highlighted at the global (Rivoire et al. 2021) and European scale (Bandhauer et al. 2022). The BOLAM/MOLOCH data are able to reduce such disagreement, in particular for the two higher thresholds (20 and 50 mm/day). We stress the fact that the scores obtained when using the GRIPHO dataset as reference are, for almost all thresholds, better than those obtained when using ARCIS. This holds possibly because (i) a higher incidence of correctly simulated cases for the GRIPHO database, which is shorter in time, than ARCIS or (ii) a higher incidence of extreme events in the ARCIS database, which are more difficult to simulate, because the Alps cover the larger part of the domain. However, any possible cause raises questions and concerns about the coherence between GRIPHO and ARCIS in the overlapping period 2001–2015. We didn't perform any evaluation of the coherence between the two databases because it would required a consistent effort, and would be out of the scope of the present work.

As regards the analysis on heavy hourly precipitation corresponding to the 90th percentile of wet hours, results show that MOLOCH provides the best average bias (i.e., closer to zero) than the other two datasets. This holds for both summer and winter, although in summer (Fig. 12) the bias is positive and equal to approximately 25.0% and in winter (Fig. 13) it is negative and equal to approximately −4.3%. As stated previously, the reason for such a behaviour possibly relies on too high temperatures during summer triggering convective precipitations. To assess whether such hypothesis is correct or not, a verification on low-level temperatures is currently ongoing. The analysis on the mean diurnal cycle of precipitations shown in Fig. 14 and Table 4 suggests that the use of a convection-permitting model improves the representation of summer daily rainfall in regions where orography is complex (namely Western and Eastern Alps) and in the flat plains of the Po Valley in northern Italy. In particular the timing of the peak agrees well with observations, with an anticipation no more than one hour, and the MOLOCH amplitude is closer to the observations than BOLAM and ERA5-Land. The fact that lower resolution models lack in the representation of the mean diurnal cycle is due to the use of a parameterisation scheme for precipitation and not to the coarse representation of the orography. This assessment is not new and agrees with what is found in the literature (Ban et al. 2021). On the other hand, in the three hilly areas of the Tyrrhenian Coast, Sardinia Island and central-southern Italy, ERA5-Land outperforms BOLAM and MOLOCH and no added value is found when using the convection-permitting model. However, we underline that both BOLAM and MOLOCH are reliable as confirmed by the *KGE* values.

We finally considered two heavy precipitations events to investigate the capability of BOLAM/MOLOCH data to capture extreme precipitations and assess the added value with respect to ERA5-Land predictions. We argue that two single case studies cannot provide statistically robust results. However, a more rigorous verification of the long-term BOLAM/MOLOCH hindcast for selected heavy precipitation events can be feasible considering the availability of the Polaris database, which counts heavy rainfall events since the early '60s. We underline the fact that when calculating the *FSS* score shown in Fig. 18, we set a constant value for the neighbourhood size, which is equal to 18 km. We chose this value because it is known (Skamarock 2004; Ricard et al. 2013) that the effective resolution of numerical weather models is a multiple (say K) of the grid spacing Δx. Some studies found that this inflating factor K ranges from 4 for the Meso-NH model (Lac et al. 2018) run at 2.5 km grid spacing to 7 for the WRF model (Skamarock et al. 2008) run at 4 km grid spacing. Since we want to highlight any potential added value of the convection-permitting MOLOCH model, we calculated the *FSS* for a neighbourhood size which approximately matches the MOLOCH effective resolution. Since no studies exist as regards MOLOCH and its K value, we chose a conservative setting for K equal to 7 to get 17.5 km as effective resolution (in fact, in our study the grid spacing Δx of the MOLOCH model is 2.5 km). For this reason, the profiles shown in Fig. 18 allow for an intensity analysis, and not for an investigation on the scales over which the simulations improve their skill. What we can conclude from the plots in Fig. 18 is that MOLOCH provides valuable information on the intensity of precipitation up to 250 mm/day as regards PIE1994 and up to 200 mm as regards CT2011; it outperforms both BOLAM and ERA-Land.

All the assessments we made regarding the performances of the BOLAM/MOLOCH hindcast rely on the quality of observations. It has to be always kept in mind that rain-gauge undercatch, measurement errors, and interpolation procedures may lead to uncertainties in observed gridded dataset. This holds in particular in mountainous areas, where the rainfall underestimation may be induced by rain-gauge positioning, strong wind and/or snowfall. Some papers report an underestimation in the range 5–40% in such areas (Frei et al. 2003), and as a benchmark, Pichelli et al. (2021) consider precipitation biases in the range between −5 and 25% as an acceptable range in some of their analyses. For this reason, we should use caution when interpreting the results. For example, the conclusions we drew about the accuracy of the MOLOCH hindcast can be better evaluated considering the possible underestimation in the observations. We can argue that the systematic wet bias found in the spatial pattern of the 90th percentile of daily and hourly precipitations is more likely lower than the one shown in panel (c) of Figs. 10 and 12. Several previous studies (Ban et al. 2014; Pichelli et al. 2021) used the E-OBS dataset (Cornes et al. 2018) as observational dataset. E-OBS is a gridded dataset over Europe at a resolution of about 0.1 degree both in the longitude and latitude direction. Using E-OBS would have triggered comparisons between our results and those found in similar studies. However, it is known (Prein and Gobiet 2017; Bandhauer et al. 2022), that in regions where the rain-gauge density is poor, E-OBS suffers from a systematic underestimation as regards extreme events and average values over mountainous areas. For these reason, as done by many authors (Ban et al. 2014; Berthou et al. 2020; Reder et al. 2022), we preferred to validate modelled data against high-resolution observational datasets, although they are available over a shorter time-period, as in the case of GRIPHO, or for a limited portion of the area of interest, as in the case of ARCIS which lacks data over southern Italy. This holds, besides the fact that validating the mean daily cycle of precipitations and the 90th percentile of hourly rainfall rate imposes the constraint about the need for hourly records. To our knowledge, GRIPHO is the only long-term, hourly and gridded database freely available over the whole Italian domain.

As discussed in the Introduction, a key factor to create an accurate picture of the past climate through reanalyses is the implementation of a data assimilation scheme; this to nudge the model towards the observed states. Otherwise than similar studies (Bollmeyer et al. 2015; Whelan et al. 2018; Bonanno et al. 2019; Reder et al. 2022), the BOLAM/MOLOCH hindcast we presented, didn't built upon any assimilation of observed data at the local scale and we acknowledge that this might represent a shortcoming of the dataset. Nevertheless, we claim that this is a pragmatic approach to take full advantage of the the improved quality of global reanalyses (ERA5) and to path the way for a fast downscaling of future global reanalyses, such as ERA6 data which are expected to even better resolve small scale features thanks to advances in assimilation techniques. Further, regional data assimilation requires a robust quality-checked database of observations, not ingested into global reanalyses, and which are required to cover long temporal and spatial scales. Observations gathering and preprocessing would slow down any future implementation of new limited-area reanalyses based on convection-permitting models. This holds beyond the technical and scientific challenges to be tackled when assimilating data at such high-resolution scale.

5 Conclusions

The dynamical downscaling of climate projections at the convection-permitting scale is at the cutting-edge of numerical climate experiments (Prein et al. 2015). To evaluate the uncertainty at the regional scale of such simulations, several projects developed multi-model climate ensembles (Coppola et al. 2020; Ban et al. 2021; Pichelli et al. 2021). Thus, our study aims at assessing the reliability of the BOLAM/MOLOCH numerical chain over a long past period to promote its use in climate projections downscaling, for which the impact of the driving boundary conditions is prominent respect to initial conditions. However, we acknowledge that BOLAM and MOLOCH are numerical weather prediction models and not climate models, thus before the implementation in climate studies, efforts are required to adapt them. Our work is a first assessment about the reliability, affordability and ease of implementation of the BOLAM/MOLOCH suite for downscaling reanalyses and thus, potentially, climate projections. As regards the affordability, we

Table 5 The 2×2 contingency table

		Event observed	
		Yes	No
Event forecast	Yes	A	B
	No	C	D

highlight that, assuming the same computational power as in Raffa et al. (2021b), that is 2160 computing cores, a 1-year BOLAM/MOLOCH data can be delivered in less than 19 h. This holds if we assume that the deployment of the BOLAM/MOLOCH setup on 60 nodes is an embarrassingly-parallel job, which is a reasonable hypothesis.

Thanks to the numerous ongoing activities regarding the refinement of ERA5 data at the local scale, the BOLAM/MOLOCH hindcast can complement similar datasets (Bonanno et al. 2019; Raffa et al. 2021b; Cerenzia et al. 2022) to create a multi-model and convection-permitting ensemble, aimed at assessing the uncertainty of past rainfall regimes in Italy.

Appendix A

Let F_i and O_i be the modelled and observed values, respectively, at location i, then *RMSE*, *ME*, m-*bias* and r are defined as follows (see also Wilks 2011):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(F_i - O_i)^2}, \tag{A1}$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}(F_i - O_i), \tag{A2}$$

$$\text{m-}bias = \frac{\frac{1}{n}\sum_{i=1}^{n} F_i}{\frac{1}{n}\sum_{i=1}^{n} O_i} = \frac{\overline{F}}{\overline{O}}, \tag{A3}$$

$$r = \frac{\sum_{i=1}^{n}(F_i - \overline{F})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n}(F_i - \overline{F})^2}\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}}, \tag{A4}$$

where $\overline{F}$ and $\overline{O}$ are the average values (over the index i) of forecasts and observations, respectively, and n is the total number of weather stations considered.

Using the 2×2 contingency table for the dichotomous (yes/no) forecast shown in Table 5, four skill measures are defined as follows:

$$SR = 1 - \frac{B}{A + B}, \tag{A5}$$

$$POD = \frac{A}{A + C}, \tag{A6}$$

$$CSI = \frac{A}{A + B + C}, \tag{A7}$$

$$bias = \frac{A + B}{A + C}. \tag{A8}$$

The *FSS* method (Roberts and Lean 2008) belongs to the object-based verification methods and it was conceived to analyse the spatial agreement between predicted and observed gridded values. Let p be a rainfall amount in a time period T (for example 24-h), firstly the method compares the fraction of grid points exceeding p in a window of side length n (often referred as the neighbourhood length) in both the modelled and observed grids. This is done by converting the precipitation values to a binary field with the purpose to remove any bias, and transform the continuous forecast to a dichotomous one (for the particular event "the rainfall is greater than p"). In a second step, the method calculates, for each square of side length equal to n, the square root of the *RMSE* (defined $MSE_{(n)}$) between forecast and observed fractions and compares this value with a reference value $MSE_{(n)ref}$, which can be thought of as the largest possible *MSE* that can be obtained from the forecast and observed fractions. Finally the *FSS* is computed by using the formula:

$$FSS_{(n)} = 1 - \frac{MSE_{(n)}}{MSE_{(n)ref}}. \tag{A9}$$

FSS spans in the interval $[0, 1] \in \mathbb{R}$, with 1 as a perfect score. For each value p, Roberts and Lean (2008) identified a thresholds below which *FSS* values indicate a forecast with a *POD* less than 0.5 and a *CSI* less than 0.33; this threshold is denoted as "uniform" *FSS* (denoted with FSS_u using the authors' notation) and is considered to be a suitable value for the "target skill".

The *KGE* index is a standard measure, often used in hydrology, to evaluate the goodness of a discrete modelled time-series F_j with respect to observations O_j, where $j \in \{1, \dots, n\} \subset \mathbb{N}$. In formula:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_F}{\sigma_O} - 1\right)^2 + \left(\frac{\overline{F}}{\overline{O}} - 1\right)^2}, \tag{A10}$$

where σ_F and σ_O are the standard deviations of predictions and observations, respectively, over time, $\overline{F}$ and $\overline{O}$ are the time averages and r is the Pearson correlation coefficient defined in Eq. A4. The *KGE* index belongs to the interval $(-\infty, 1] \in \mathbb{R}$, with 1 as perfect score. Knoben et al. (2019) argued that *KGE* values less than −0.41 indicate not skilful predictions when using the mean of the observations as benchmark.

Acknowledgements The Copernicus Climate Change Service is acknowledged for providing ERA5 data. ISAC/CNR owns the intellectual property rights on the BOLAM and MOLOCH models. Their use in the present study is authorised as part of the scientific agreement among the ISAC/CNR, ARPAL, ISPRA and LaMMA institutes. Acknowledgement is made for the use of ECMWF's computing and archive facilities within the framework of the SPITBRAN Special Project. The authors wish also to thank Adriano Fantini, Fabio Di Sante and Erika Coppola (ICTP) for providing access to the Italian database of precipitation GRIPHO. The help of Tommaso Torrigiani Malaspina (LaMMA) is acknowledged for providing access to ARCIS.

Author contributions VC and CB designed the study, VC and FP produced the BOLAM/MOLOCH hindcast, VC analysed the output data, produced the figures/tables and wrote most of the paper. All authors proofread the manuscript, provided a critical review and approved the final version of the manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability Sample hourly rainfall data from the MOLOCH simulations and for the years 1994 and 2011 are available on Zenodo with DOI 10.5281/zenodo.6990599. The whole dataset for the period 1979–2019 is available on request from the corresponding author (VC).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Adinolfi M, Raffa M, Reder A et al (2020) Evaluation and expected changes of summer precipitation at convection permitting scale with COSMO-CLM over alpine space. Atmosphere 12(1):54

Antolini G, Auteri L, Pavan V et al (2016) A daily high-resolution gridded climatic data set for emilia-romagna, italy, during 1961–2010. Int J Climatol 36(4):1970–1986

Ban N, Schmidli J, Schär C (2014) Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. Journal of Geophysical Research: Atmospheres 119(13):7889–7907

Ban N, Schmidli J, Schär C (2015) Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? Geophys Res Lett 42(4):1165–1172

Ban N, Caillaud C, Coppola E et al (2021) The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation. Climate Dynamics pp 1–28

Bandhauer M, Isotta F, Lakatos M et al (2022) Evaluation of daily precipitation analyses in E-OBS (v19. 0e) and era5 by comparison to regional high-resolution datasets in European regions. Int J Climatol 42(2):727–747

Berthou S, Kendon EJ, Chan SC et al (2020) Pan-European climate at convection-permitting scale: a model intercomparison study. Clim Dyn 55(1):35–59

Bollmeyer C, Keller J, Ohlwein C et al (2015) Towards a high-resolution regional reanalysis for the European CORDEX domain. Q J R Meteorol Soc 141(686):1–15

Bonanno R, Lacavalla M, Sperati S (2019) A new high-resolution meteorological reanalysis Italian dataset: MERIDA. Q J R Meteorol Soc 145(721):1756–1779

Bonavita M, Hólm E, Isaksen L et al (2016) The evolution of the ECMWF hybrid data assimilation system. Q J R Meteorol Soc 142(694):287–303

Boussetta S, Balsamo G, Beljaars A et al (2013) Natural land carbon dioxide exchanges in the ecmwf integrated forecasting system: implementation and offline validation. J Geophys Res 118(12):5923–5946

Brunetti M, Maugeri M, Nanni T (2001) Changes in total precipitation, rainy days and extreme events in northeastern Italy. Int J Climatol 21(7):861–871

Bubnová R, Hello G, Bénard P et al (1995) Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the arpege/aladin nwp system. Mon Weather Rev 123(2):515–535

Buizza R, Poli P, Rixen M et al (2018) Advancing global and regional reanalyses. Bull Am Meteorol Soc 99(8):ES139–ES144

Buzzi A, Fantini M, Malguzzi P et al (1994) Validation of a limited area model in cases of Mediterranean cyclogenesis: surface fields and precipitation scores. Meteorol Atmos Phys 53(3):137–153

Buzzi A, Tartaglione N, Malguzzi P (1998) Numerical simulations of the 1994 Piedmont flood: Role of orography and moist processes. Mon Weather Rev 126(9):2369–2383

Buzzi A, Davolio S, Malguzzi P et al (2014) Heavy rainfall episodes over Liguria in autumn 2011: numerical forecasting experiments. Nat Hazard 14(5):1325

Caillaud C, Somot S, Alias A et al (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(5):1717–1752

Capecchi V (2020) Reforecasting the November 1994 flooding of piedmont with a convection-permitting model. Bull Atmos Sci Technol 1(3):355–372

Capecchi V (2021) Reforecasting two heavy-precipitation events with three convection-permitting ensembles. Weather Forecast 36(3):769–790

Cerenzia IML, Giordani A, Montani A (2022) Towards a convection-permitting regional reanalysis over the Italian domain. Meteorol Appl 29(5):e2092

Cerenzia IML, Pincini G, Paccagnella T et al (2020) Forecast of precipitation for the 1994 flood in Piedmont: performance of an ensemble system at convection-permitting resolution. Bull Atmos Sci Technol 1(3):319–338

Coppola E, Sobolowski S, Pichelli E et al (2020) A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. Clim Dyn 55(1):3–34

Coppola E, Fantini A, Giorgi F et al (2021) A new spatially distributed added value index for regional climate models: the

EURO-CORDEX and the CORDEX-CORE highest resolution ensembles. Clim Dyn 57(5):1403–1424
Corazza M, Sacchetti D, Antonelli M et al (2018) The ARPAL operational high resolution poor man's ensemble, description and validation. Atmos Res 203:1–15
Cornes RC, van der Schrier G, van den Besselaar EJ et al (2018) An ensemble version of the E-OBS temperature and precipitation data sets. J Geophys Res 123(17):9391–9409
Davolio S, Silvestro F, Malguzzi P (2015) Effects of increasing horizontal resolution in a convection-permitting model on flood forecasting: the 2011 dramatic events in Liguria, Italy. J Hydrometeorol 16(4):1843–1856
Davolio S, Malguzzi P, Drofa O, Mastrangelo D, Buzzi A (2020) The Piedmont flood of November 1994: a testbed of forecasting capabilities of the CNR-ISAC meteorological model suite. Bull Atmos Sci Technol 1(3):263–282
Dee DP, Uppala SM, Simmons AJ et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597. https://doi.org/10.1002/qj.828
Doddy Clarke E, Griffin S, McDermott F et al (2021) Which reanalysis dataset should we use for renewable energy analysis in Ireland? Atmosphere 12(5):624
Drofa O, Malguzzi P (2004) Parameterization of microphysical processes in a non hydrostatic prediction model. In: Proceedings of 14th Intern. Conf. on Clouds and Precipitation (ICCP), Bologna, Italy, pp 19–23
Fantini A (2019) Climate change impact on flood hazard over Italy. PhD thesis, Università degli Studi di Trieste, 2019
Ferrari F, Besio G, Cassola F et al (2020) Optimized wind and wave energy resource assessment and offshore exploitability in the mediterranean sea. Energy 190(116):447
Ferretti R, Low-Nan S, Rotunno R (2000) Numerical simulations of the Piedmont flood of 4–6 November 1994. Tellus A 52(2):162–180
Fischer C, Montmerle T, Berre L et al (2005) An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. Quart J R Meteorol Soc 131(613):3477–3492
Fosser G, Khodayar S, Berg P (2015) Benefit of convection permitting climate model simulations in the representation of convective precipitation. Clim Dyn 44(1):45–60
Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. J Geophys Res 108:D3. https://doi.org/10.1029/2002JD002287
Garbero V, Milelli M (2020) Reforecast of the november 1994 flood in piedmont using ERA5 and COSMO model: an operational point of view. Bull Atmos Sci Technol 1(3):339–354
Giard D, Bazile E (2000) Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. Mon Weather Rev 128(4):997–1015
Giorgi F, Gutowski WJ Jr (2015) Regional dynamical downscaling and the CORDEX initiative. Annu Rev Environ Resour 40:467–490
Giorgi F, Jones C, Asrar GR et al (2009) Addressing climate information needs at the regional level: the CORDEX framework. World Meteorol Organ Bull 58(3):175
Gupta HV, Kling H, Yilmaz KK et al (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J Hydrol 377(1–2):80–91
Gustafsson N, Berre L, Hörnquist S et al (2001) Three-dimensional variational data assimilation for a limited area model: part I: general formulation and the background error constraint. Tellus A 53(4):425–446
Hersbach H, Bell B, Berrisford P et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049
Jiang Y, Yang K, Shao C et al (2021) A downscaling approach for constructing high-resolution precipitation dataset over the Tibetan Plateau from ERA5 reanalysis. Atmos Res 256(105):574
Kain JS (2004) The Kain-Fritsch convective parameterization: an update. J Appl Meteorol 43(1):170–181
Kendon EJ, Roberts NM, Senior CA et al (2012) Realism of rainfall in a very high-resolution regional climate model. J Clim 25(17):5791–5806. https://doi.org/10.1175/JCLI-D-11-00562.1
Knoben WJ, Freer JE, Woods RA (2019) Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrol Earth Syst Sci 23(10):4323–4331
Lac C, Chaboureau JP, Masson V et al (2018) Overview of the Meso-NH model version 5.4 and its applications. Geosci Model Dev 11(5):1929
Lagouvardos K, Kotroni V, Koussis A et al (2003) The meteorological model bolam at the national observatory of athens: assessment of two-year operational use. J Appl Meteorol 42(11):1667–1678
Lavers DA, Simmons A, Vamborg F, Rodwell MJ (2022) An evaluation of ERA5 precipitation for climate monitoring. Q J R Meteorol Soc 148(748):3152–3165. https://doi.org/10.1002/qj.4351
Lind P, Lindstedt D, Kjellström E et al (2016) Spatial and temporal characteristics of summer precipitation over central Europe in a suite of high-resolution climate models. J Clim 29(10):3501–3518
Malardel S, Wedi N, Deconinck W et al (2016) A new grid for the IFS. ECMWF Newsl 146:23–28
Malguzzi P, Tartagione N (1999) An economical second-order advection scheme for numerical weather prediction. Q J R Meteorol Soc 125(558):2291–2303
Masson V, Le Moigne P, Martin E et al (2013) The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geosci Model Dev 6(4):929–960
Mentaschi L, Besio G, Cassola F et al (2013) Developing and validating a forecast/hindcast system for the Mediterranean Sea. J Coastal Res 65(10065):1551–1556
Morcrette J, Barker HW, Cole J et al (2008) Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. Mon Weather Rev 136(12):4773–4798
Muñoz-Sabater J, Dutra E, Agustí-Panareda A et al (2021) Era5-land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst Sci Data 13(9):4349–4383
Osinski RD, Radtke H (2020) Ensemble hindcasting of wind and wave conditions with WRF and WAVEWATCH III® driven by ERA5. Ocean Sci 16(2):355–371
Parodi A, Lagasio M, Meroni AN et al (2020) A hindcast study of the piedmont 1994 flood: the CIMA research foundation hydro-meteorological forecasting chain. Bull Atmos Sci Technol 1(3):297–318
Pavan V, Antolini G, Barbiero R et al (2019) High resolution climate precipitation analysis for north-central Italy, 1961–2015. Clim Dyn 52(5):3435–3453
Pichelli E, Coppola E, Sobolowski S et al (2021) The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation. Clim Dyn 56(11):3581–3602
Prein AF, Gobiet A (2017) Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. Int J Climatol 37(1):305–327
Prein A, Gobiet A, Suklitsch M et al (2013) Added value of convection permitting seasonal simulations. Clim Dyn 41(9):2655–2677
Prein AF, Langhans W, Fosser G et al (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. Rev Geophys 53(2):323–361
Raffa M, Reder A, Adinolfi M et al (2021) A comparison between one-step and two-step nesting strategy in the dynamical

downscaling of regional climate model COSMO-CLM at 2.2 km driven by ERA5 reanalysis. Atmosphere 12(2):260
Raffa M, Reder A, Marras GF et al (2021) VHR-REA_IT dataset: very high resolution dynamical downscaling of ERA5 reanalysis over Italy by COSMO-CLM. Data 6(8):88
Reder A, Raffa M, Padulano R, Rianna G, Mercogliano P (2022) Characterizing extreme values of precipitation at very high resolution: an experiment over twenty European cities. Weather Clim Extremes 35:100407
Ricard D, Lac C, Riette S et al (2013) Kinetic energy spectra characteristics of two convection-permitting limited-area models AROME and Meso-NH. Q J R Meteorol Soc 139(674):1327–1341
Ritter B, Geleyn JF (1992) A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. Mon Weather Rev 120(2):303–325
Rivoire P, Martius O, Naveau P (2021) A comparison of moderate and extreme ERA-5 daily precipitation with two observational data sets. Earth Space Sci 8(4):e2020EA001,633
Roberts NM, Lean HW (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Mon Weather Rev 136(1):78–97
Rockel B, Will A, Hense A (2008) The regional climate model COSMO-CLM (CCLM). Meteorol Z 17(4):347–348
Roebber PJ (2009) Visualizing multiple measures of forecast quality. Weather Forecast 24(2):601–608
Rotunno R, Houze RA (2007) Lessons on orographic precipitation from the Mesoscale Alpine Programme. Quart J R Meteorol Soc 133(625):811–830
Ruti PM, Somot S, Giorgi F et al (2016) Med-cordex initiative for mediterranean climate studies. Bull Am Meteor Soc 97(7):1187–1208
Schraff C, Hess R (2003) A description of the non-hydrostatic regional model LM - Part III: Data assimilation. Technical report, Deutscher Wetterdienst, P.O. Box 100465, 63004 Offenbach, Germany
Seity Y, Brousseau P, Malardel S et al (2011) The AROME-France convective-scale operational model. Mon Weather Rev 139(3):976–991
Singh T, Saha U, Prasad V et al (2021) Assessment of newly-developed high resolution reanalyses (IMDAA, NGFS and ERA5) against rainfall observations for indian region. Atmos Res 259(105):679
Skamarock WC, Klemp JB, Dudhia J et al (2008) A description of the Advanced Research WRF version 3. In: NCAR Tech. Note NCAR/TN-475+ STR, 10.5065/D68S4MVH
Skamarock WC (2004) Evaluating mesoscale NWP models using kinetic energy spectra. Mon Weather Rev 132(12):3019–3032
Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon Weather Rev 117(8):1779–1800
Trini Castelli S, Bisignano A, Donateo A et al (2020) Evaluation of the turbulence parametrization in the MOLOCH meteorological model. Q J R Meteorol Soc 146(726):124–140
Uboldi F, Lussana C, Salvati M (2008) Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks. Meteorol Appl 15(3):331–345
Vannucchi V, Taddei S, Capecchi V et al (2021) Dynamical downscaling of ERA5 data on the North-Western Mediterranean Sea: from atmosphere to high-resolution coastal wave climate. J Mar Sci Eng 9(2):208
Wahl S, Bollmeyer C, Crewell S et al (2017) A novel convective-scale regional reanalysis COSMO-REA2: improving the representation of precipitation. Meteorol Z 26(4):345–361
Whelan E, Gleeson E, Hanley J (2018) An evaluation of MERA, a high-resolution mesoscale regional reanalysis. J Appl Meteorol Climatol 57(9):2179–2196
Wilks DS (2011) Statistical methods in the atmospheric sciences, vol 100. Academic press
Zampieri M, Malguzzi P, Buzzi A (2005) Sensitivity of quantitative precipitation forecasts to boundary layer parameterization: a flash flood case study in the Western Mediterranean. Nat Hazards Earth Syst Sci 5(4):603–612
Zhang W, Villarini G, Scoccimarro E et al (2021) Examining the precipitation associated with Medicanes in the high-resolution ERA-5 reanalysis data. Int J Climatol 41:E126–E132