Received: 5 February 2021 | Revised: 22 July 2022 | Accepted: 31 July 2022

DOI: 10.1002/met.2092

RESEARCH ARTICLE

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

Towards a convection-permitting regional reanalysis over the Italian domain

Ines Maria Luisa Cerenzia[1] | **Antonio Giordani**[1,2] | **Tiziana Paccagnella**[1] | **Andrea Montani**[1,3]

[1]HydroMeteoClimate Service, ARPAE-Emilia Romagna, Bologna, Italy

[2]Department of Physics and Astronomy (DIFA) "Augusto Righi", University of Bologna, Bologna, Italy

[3]Computing Department, European Centre for Medium-Range Weather Forecasts, Bologna, Italy

Correspondence
Ines Maria Luisa Cerenzia, HydroMeteoClimate Service, ARPAE-Emilia Romagna, Bologna, 40122, Italy.
Email: icerenzia@arpae.it

Abstract

A new convection-permitting regional reanalysis, SPHERA (High Resolution REAnalysis over Italy), has been developed over Italy and the surrounding seas using the COSMO model at 2.2 km horizontal resolution. The reanalysis system is nested in the global reanalysis ERA5; upper-air and surface observations are assimilated at the convection-permitting scale by the COSMO nudging scheme. Before the complete production of the hourly three-dimensional fields and surface/soil parameters over the period 1995–2020, general issues regarding the reanalysis set-up needed to be addressed over a shorter test period. These include the identification of the best approach to downscale the lateral boundary conditions from the global driver, and the definition of the bottom boundary condition related to deep soil temperature. With respect to the downscaling methodology, the results show a clear benefit in using lateral boundary conditions directly from the global ERA5, despite the large resolution difference between the two modes (1:15), instead of providing them from an intermediate resolution COSMO-based reanalysis. Moreover, the soil bottom boundary condition for temperature is reconstructed from the shallower ERA5 soil, using a site-dependent method based on a delayed running mean of the ERA5 temperature at the deepest soil level. Finally, an evaluation of SPHERA has been performed with respect to the skill in simulating daily precipitation over 2 years. Compared with ERA5, SPHERA shows a higher ability in simulating moderate and intense events, markedly during summer, in terms of skill scores, frequency of occurrence and bias.

KEYWORDS

convection-permitting, coupled, heavy precipitation events, hydro-meteorology, long-range, precipitation, precipitation verification, regional reanalysis, SPHERA, deep soil temperature in reanalysis, evaluation, model development, forecasting, modelling, nesting modality

1 | INTRODUCTION

Reanalysis data provide a spatio-temporal consistent description of the past atmospheric state on a regular grid. They are usually generated using a numerical weather prediction (NWP) data assimilation procedure to combine model simulations and meteorological observations. Reanalysis datasets generally have a long temporal extension, spanning from several years to several decades, and are a fundamental tool for weather related diagnostic and climate studies (Trenberth et al., 2008). The need to account for high-impact weather events, and their climatological trends, requires km-scale data assimilation procedures. For this reason, the interest towards high-resolution reanalyses has been increasing as they enable the representation of local and rapidly evolving processes like those related to severe precipitation and linked to deep convection (Keller & Wahl, 2021). Moreover, downstream modelling applications in several fields such as hydrology (Berg et al., 2018), renewable energy sector (Doddy Clarke et al., 2021) or tracer dispersion modelling (Ngan & Stein, 2017) could certainly benefit from more detailed meteorological forcings. Finally, also modern post-processing approaches (e.g. model based risk assessment) need a sound statistical basis in the form of spatial climatology, which can be provided by high-resolution regional reanalyses.

Also, global reanalyses at increased resolution and enhanced quality are now available thanks to the improvements in NWP modelling and computational capacity. The most recent generation of global reanalyses includes the Modern-Era Retrospective Analysis for Research and Applications version 2 (Gelaro et al., 2017, MERRA-2) produced by NASA (National Aeronautics and Space Administration), the Climate Forecast System Reanalysis (Saha et al., 2010, CFSR) by the National Center for Environmental Prediction (NCEP), the Japan Meteorological Agency (JMA) 55-year Reanalysis (Kobayashi et al., 2015). The European Centre for Medium-Range Weather Forecasts (ECMWF) recently produced ERA5, the first fourth-generation global reanalysis (Hersbach et al., 2020). The horizontal resolutions of global reanalyses range between 125 and 31 km, with an output frequency between 6 and 1 h. Regional reanalyses are usually based on dynamical downscalings of global datasets forced by a high-resolution limited-area NWP model and coupled with a regional data assimilation scheme.

A considerable number of regional reanalyses have been produced so far covering several continents: the North American Regional Reanalysis (Mesinger et al., 2006, NARR), the China Regional Reanalysis project (Zhang et al., 2017, CNRR), the Arctic System Reanalysis version 2 (Bromwich et al., 2018, ASRv2) or the regional reanalysis for the Australian and Indonesian regions (Su et al., 2019, BARRA). In the last decade, large efforts have been devoted to this task also in Europe. The main examples include the regional datasets covering part of the European CORDEX domain (Coordinated Regional Downscaling EXperiment, cordex.org), COSMO-REA6 (Bollmeyer et al., 2015), the HIRLAM-based reanalysis (Dahlgren et al., 2016) and MERIDA covering the Italian peninsula (Bonanno et al., 2019). Of particular interest is the European Reanalysis and Observations for Monitoring (EURO4M, www.euro4m.eu) project and its continuation UERRA (Uncertainties in Ensembles of Regional Reanalyses, www.uerra.eu), which are based on the collaborative effort of several institutions aiming to demonstrate the potential of probabilistic approaches in regionalization. Based on the developments of UERRA, a Regional ReAnalysis (RRA) for Europe is available at 11-km horizontal resolution (Schimanke et al., 2018) and an upgrade to 5.5 km is under development within the Copernicus Climate Change Service (CERRA[1]). Anyhow, deep moist convection is still parameterized in these regional reanalyses, as the computational cost to produce datasets at km-scale grid resolution is still very high for extended domains. Convection-permitting (hereafter, CP) resolution has been achieved only for national or even smaller domains. Over Europe, CP reanalyses include MERA (Met Éireann ReAnalysis), over Ireland and United Kingdom with a horizontal resolution of 2.5 km (Gleeson et al., 2017), and COSMO-REA2, a 2-km resolution dataset covering Central Europe (Wahl et al., 2017). They both highlight improvements over coarser reanalyses, particularly in the representation of local precipitation. In COSMO-REA2, enhancements relate mainly to spatial variability and frequency of precipitation, especially for high intensity events. These ameliorations are mainly linked to the enhanced spatial detail and to the assimilation of additional local observations, such as, for example, radar-derived rain rates (Wahl et al., 2017). Other studies have been performed with a pure dynamical downscaling of global reanalyses (i.e. without data assimilation) employing regional climate models at the CP resolution. Despite this limitation, they show the benefits associated with deep convection explicit representation. These include an enhanced representation of summer precipitation (Pal et al., 2019), a more realistic representation of the diurnal cycle of convective precipitation (Brisson et al., 2016; Dai et al., 1999; Fosser et al., 2015; Prein et al., 2013), a better detection of the most severe precipitation peaks (Fosser

[1]https://climate.copernicus.eu/copernicus-regional-reanalysis-europe-cerra (visited in July 2022).

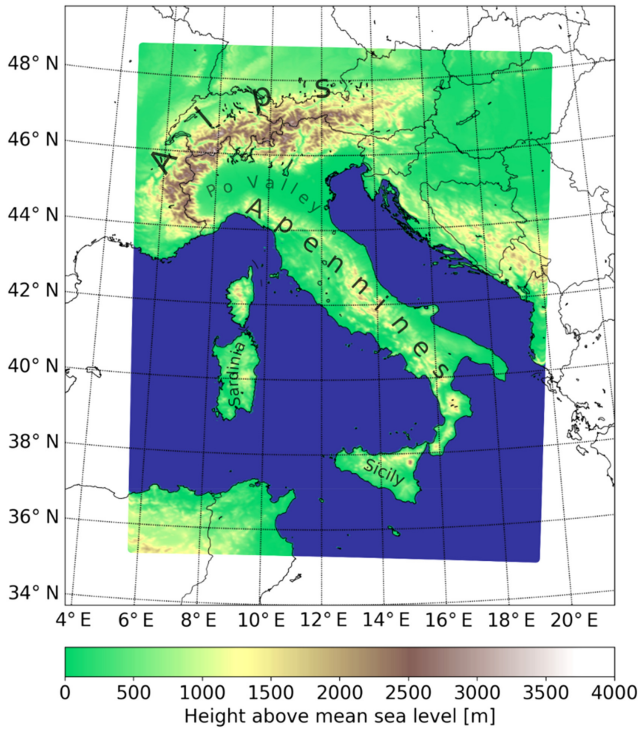

FIGURE 1 Geographical coverage of SPHERA reanalysis; the colour coding over land indicates the altitude (in metres) above the mean sea level.

et al., 2015; Kendon et al., 2012; Prein et al., 2013) and a reduction in low-precipitation event frequency bias (Berg et al., 2013). Similar results are widely highlighted in several applications of convection-permitting models (CPMs) when compared with their coarser convection-parameterizing drivers (e.g. Clark et al., 2016; Iyer et al., 2016; Klasa et al., 2018).

The main benefits of CPMs can be summarized as follows: better representation of organized convective systems, more realistic advection of microphysical species and improved ability to simulate phenomena such as the formation of daughter cells or self-regenerating storms (e.g. Clark et al., 2016; White et al., 2018). Moreover, CPMs better represent the flow interaction with small-scale surface features such as orography and land/sea contrast, hence enhancing the description of orographic convection (e.g. Kirshbaum et al., 2018) or inland rainfall penetration (Clark et al., 2016). Consequently, the predictive skill of precipitation spatial patterns and their frequency of occurrence, when simulated by CPMs, increase, especially in case of intense and localized rainfall (e.g. Clark et al., 2016; Iyer et al., 2016; Lewis et al., 2015; Weusthoff et al., 2010). CPMs typically provide a better representation of the diurnal cycle of convective precipitation, especially with respect to the timing of its associated peak (Baldauff et al., 2011; Lean et al., 2008; Prein et al., 2015; Weisman et al., 2008; Weusthoff et al., 2010; Woodhams et al., 2018). Improvements in rainfall simulation are further increased when a suitable data assimilation is coupled with the CPM (Dow and Macpherson, 2013; Clark et al., 2016; Gustafsson et al., 2018; Lewis et al., 2015; Seity et al., 2011).

This article presents SPHERA (High Resolution REAnalysis over Italy), the RRA based on the COSMO (Consortium for Small scale MOdelling) model (Schättler et al., 2018) at the CP horizontal resolution of 2.2 km. The reanalysis system is nested in the global reanalysis ERA5, after investigating the best downscaling configuration for the purpose. In SPHERA, surface and upper-air observations are assimilated at the CP scale by the COSMO nudging scheme (Schraff & Hess, 2013). SPHERA has been developed at ARPAE-SIMC, the hydro-meteo-climate service of the Emilia-Romagna Region, and its domain covers the Italian peninsula and the neighbouring countries (Figure 1). A first description of SPHERA performance in the representation of precipitation, paying particular attention to intense rainfall events, is included in the present work for a sample period of 2 years to quantify the benefits over the driving convection-parameterized reanalysis. The article is organized as follows: the experimental design is reported in Section 2, which includes the description of the main features of SPHERA, the definition of the nesting procedure, the methodology developed to impose the deep soil temperature boundary condition and the observational data used for verification. Results are reported and discussed in Section 3, while conclusions are drawn in Section 4.

2 | EXPERIMENTAL DESIGN

This section presents the configuration of SPHERA (Section 2.1), the nesting approach used to downscale ERA5 (Section 2.2), the definition of the bottom boundary condition related to deep soil temperature (Section 2.3), the observational dataset used for the validation of the reanalyses (Section 2.4) and the methodology employed to assess the performance of the datasets (Section 2.5).

2.1 | SPHERA configuration

The SPHERA reanalysis is based on the NWP limited-area COSMO model (Schättler et al., 2018). COSMO is used in the operational NWP suites in Italy, as well as in several other countries (Switzerland, Germany, Greece, Romania, Israel, Poland and Russia). At ARPAE-SIMC, COSMO runs operationally in deterministic mode with

horizontal resolutions of 5 and 2.2 km, and in ensemble mode at 2.2 km resolution only. For the set-up of SPHERA, the model configuration is as close as possible to that applied for the higher resolution operational run. This aims at producing a long-term evaluation of the COSMO model for supporting operational forecasting, especially in cases of high-impact weather events. The COSMO model includes three-dimensional non-hydrostatic, compressible hydro-thermodynamic equations in advective form, which are solved numerically in the SPHERA set-up with a third-order Runge–Kutta split-explicit scheme (Wicker & Skamarock, 2002). The grid structure is based on the Arakawa C-grid, with Lorenz vertical grid staggering in a rotated (latitude/longitude) coordinate system. The vertical coordinate is terrain-following and in particular the Gal-Chen height coordinate (Gal-Chen & Sommerville, 1975) is used. The vertical axis extends up to 22,000 m and it is divided into 65 layers, with 19 layers in the lowest 1000 m. The horizontal grid mesh has 576 grid points in the west–east direction and 701 in the south–north direction, with 0.02 degree (∼2.2 km) horizontal resolution. The domain is defined by rotating the South Pole to W 10°, S 43°, and its coverage is reported in Figure 1. In the soil, seven vertical levels at depths of 0.005, 0.02, 0.06, 0.18, 0.54, 1.62, 4.86 and 14.58 m below the surface describe the near-surface, root-level and deep soil layers. The applied physical package includes the following: a δ-two-stream radiation scheme after Ritter and Geleyn (1992); a grid-scale cloud and precipitation scheme with prognostic cloud water, cloud ice and graupel; and a statistical scheme for sub-grid clouds (Sommeria & Deardorff, 1977), a shallow convection scheme (a reduction from Tiedtke, 1989), a turbulence scheme based on Mellor and Yamada (1982) at 2.5-level of their truncation hierarchy, a multi-layer transfer scheme (Doms et al., 2018), a multi-level soil scheme after Jacobsen and Heise (1982) based on the direct numerical solution of the heat conduction equation, and a two-layer bulk lake scheme (Mironov et al., 2010). Extensive descriptions of the parameterization schemes can be found in Schättler et al. (2018). The initial state is provided by the global reanalysis ERA5, available at a horizontal resolution of about 31 km and with 137 vertical levels (Hersbach et al., 2020), interpolated to the COSMO grid. This pre-processing of ERA5 data is performed by a separate procedure described in Schättler and Blahak (2017a). The production of SPHERA is organized in 4-year-long streams. Each stream is preceded by a 6-month-long "prerun" to account for the soil spin-up. The COSMO model is run in a sequence of 24-h-long integrations, starting from the beginning of each simulation stream. For each integration, the 24-h forecast is used as the initial condition for the following day's run in order to obtain a continuous series. The lateral boundary conditions, provided by ERA5, are updated every hour. Investigations performed to define the nesting methodology are reported in Section 2.2. Non-penetrative, free-slip boundary conditions are imposed at the model top. At the bottom boundary, free-percolation is imposed at the soil for depths below 2.43 m, hence the soil moisture drained below this depth is lost by the system. Despite this being the default configuration in the COSMO model, it may cause excessive drying of the soil for long-term integrations. This would have marked consequences on the atmospheric simulation, both at the surface (e.g. too high daily temperature and low relative humidity) and in the vertical (e.g. too low convective precipitation and reduced low level clouds). This motivates the investigation to find the optimal prescription for time-evolving deep soil temperature at the lowest soil level of SPHERA (14.58 m) as derived from the vertical profile of ERA5 soil temperature (described in Section 2.3). An investigation on the temporal trend over the production stream highlights that the amount of soil moisture lost at this level is small (about 2%–5% of the total). Moisture is mainly lost in the first 12 months of simulation (including the 6 months of spin-up). Moreover, this does not imply moisture trends in the soil layers above this depth, where the temporal evolution turns out to be coherent with the soil moisture of ERA5 (not shown). While the soil, lake and snow states are initialized from the ERA5 state at the corresponding time and can freely evolve during the simulation, the sea surface temperature (SST) needs to be prescribed. Dahlgren et al. (2016) pointed out the importance of a correct update of SST in RRA applications. In SPHERA, SST is updated every day at 00 UTC using ERA5. The data assimilation scheme implemented in SPHERA is based on a continuous nudging or Newtonian relaxation approach (Stauffer & Seaman, 1990), and directly assimilates in situ observations. The nudging consists of a relaxation of the model prognostic variables towards the prescribed observations, within a predetermined time window. It is performed by adding a small tendency term (i.e. the nudging term) to the model prognostic equation for each prognostic model variable. The nudging term determines the relaxation rate of the model variable towards the observations and depends on three main factors: the difference between model and observations, the weight associated with each observation, and the elapsed time from each observation collection. In this way, the nudging provides a continuous adaptation of the model values towards the observations during the integration of the model. For a complete description of the implementation in the COSMO model, the reader is referred to Schraff and Hess (2013). In SPHERA, the observational

nudging is applied to a set of observations included in the integration domain and available in the ECMWF archive. The main types of observations nudged in the model include SYNOP, SHIP and DRIBU sites, radiosoundings from TEMP and PILOT and aircraft observations (AIREP, AMDAR and ACARS). Meridional and zonal wind components, relative humidity, temperature and pressure are among the assimilated variables, while 2-m temperature is not nudged. This observational set corresponds to the one ingested in the operational NWP suites. It should be noted that neither radar nor satellite data are directly assimilated in SPHERA. Satellite information is indirectly included through both the lateral boundary conditions and the SST interpolated from ERA5.

2.2 | Nesting methodology

The downscaling (or nesting) methodology provides the way to convey the meteorological information from the large-scale driver to finer scale model. For multi-year simulations, the weight of the initial boundary conditions is negligible, as it is rapidly advected outside the model domain (with the exception of the soil, for which a spin-up period is used, as described later in this section). The forcing provided by the lateral boundary conditions of the driving coarser model plays a crucial role. The CP scales can be reached by one or more nested run(s) of the same model at higher and higher horizontal (and vertical) resolutions. However, in this case, the use of an intermediate resolution model would act in the so-called "grey zone" (e.g. Gerard et al., 2009), where convection is partly resolved and partly parameterized (Brisson et al., 2016; Marsigli et al., 2014). In addition, a supplementary model integration increases the overall computational cost, which is a relevant issue for decadal-long datasets. The ratio in spatial resolutions between coarse- and fine-mesh models usually ranges between 2:1 and 5:1. Larger values, implying abrupt increase in horizontal resolution, can generate strong gradients at the lateral boundaries of the nested model and gravity-inertia waves (Denis et al., 2003; Warner et al., 1997). It is worth pointing out that this ratio is sensitive to several factors, including the size of the nested domain, the temporal update of the lateral boundary conditions and the consistency among the physical parameterizations of the models (Warner et al., 1997). Furthermore, this ratio constraint has been proven to be valid when considering free dynamical downscalings (e.g. Pham et al., 2016), but may not hold when a data assimilation scheme is added to the modelling framework. In this case, further constraints are imposed over the dynamical evolution of the simulations. Therefore, beyond the common guidelines, every application needs to be tested specifically. Experiments in NWP forecasting using a higher ratio between model resolutions (e.g. from 11:1 to 18:1) have already given comparable or weakly positive results (Marsigli et al., 2014, MeteoSwiss forecasting system[2]). In order to evaluate the effect of a direct nesting of SPHERA (COSMO at 2.2 km horizontal resolution, hereafter referred to as COSMO-2.2 km) into ERA5 (at 31 km horizontal resolution) instead of employing an intermediate COSMO integration run, two options have been considered:

- SPHERA-2step (hereafter referred to as S2s): COSMO-2.2 km one-way nested in COSMO-10 km (i.e. a COSMO model configuration with horizontal resolution of 10 km, domain covering the whole Mediterranean Sea and convection-parameterized by Tiedtke scheme (Tiedtke, 1989)), which in turn is one-way nested in ERA5. The ratios of spatial resolutions between COSMO-2.2 km, COSMO-10 km and ERA5 are respectively 5:1 and 3:1. In COSMO-10 km, the soil state is self-evolved by the COSMO model after the first-day initialization from ERA5. Within the one-way nesting to COSMO-2.2 km, the soil fields from COSMO-10 km are interpolated to the 2.2 km grid mesh and updated daily at 00 UTC.
- SPHERA-1step (hereafter referred to as S1s): COSMO-2.2 km directly one-way nested in ERA5, with a resolution step of 15:1. The integration domain is enlarged by 25 grid points at the border in each direction (about 55 km) with respect to the one used in S2s, in order to mitigate any boundary effects potentially associated with a large ratio between model resolutions. The soil fields are updated daily at 00 UTC.

These two configurations have been tested on two parallel suites over 1 year (2015), preceded by 6 months of integration used to spin up the model soil fields.

2.3 | Definition of the deep soil temperature

Another crucial point concerning reanalysis set-up is the definition of the deep soil bottom boundary condition. In fact, the large soil inertia and moisture reservoir can trigger differences at the surface level and in the atmosphere at long time scales (Cheruy et al., 2017). Several studies

[2]https://www.meteoswiss.admin.ch/home/measurement-and-forecasting-systems/warning-and-forecasting-systems/cosmo-forecasting-system.html, visited in July 2022.

have focused on the issue of soil initialization for regional reanalyses, particularly the duration of spin-up time (Gleeson et al., 2017; Ridal et al., 2017), but less has been done on the soil bottom boundary condition. In general, the deep soil temperature in RRA is transmitted from the driving model through a vertical interpolation. This is reasonable when the deepest soil level of the driving model is deeper than the nested model. If not, a simple interpolation might lead to errors in the amplitude of the seasonal temperature variation (which decreases with depth) and to a temporal shift of the surface temperature signal (which is delayed by depth).

The approach used to deal with this issue is examined by assessing the simulation feedback associated with a change in the deep soil temperature at 14.58 m depth (i.e. the deepest soil level in COSMO). To define the long-term constraint of the bottom boundary condition, three different prescriptions are tested. They are all based on the simplified analytic solution of the heat transfer equation of soil (assuming sinusoidal annual wave of temperature and vertically homogeneous soil):

$$T_{z_2}(t) = \overline{T_{z_1}} + A_{Tz_1} e^{-(z_2 - z_1)/d} \sin\left[\omega(t - t_m) - \frac{z_2 - z_1}{d}\right],$$

where T_{z_2} and T_{z_1} are the soil temperatures at depths z_2 and z_1 respectively, A_{Tz_1} is the amplitude of the annual wave of temperature at depth z_1, while d indicates the damping depth of the annual wave, which depends on the thermal soil diffusivity α and on the angular velocity of Earth's rotation ω with the relation:

$$d = \sqrt{\frac{2\alpha}{\omega}}.$$

The time variable is indicated by t, and t_m is the time when T_{z_1} is equal to T_{z_2}. The overbar represents the annual average. The considered algorithms are as follows:

1. The simplified analytic solution of the temperature wave in which the thermal diffusivity is described by the amplitude algorithm (Horton et al., 1983), hereafter referred to as M1:

$$\alpha = \frac{\omega}{2}\left[\frac{z_2 - z_1}{\log(A_{Tz1}/A_{Tz2})}\right]^2;$$

2. The simplified analytic solution of the temperature wave in which the thermal diffusivity is described by the phase algorithm (Horton et al., 1983), hereafter referred to as M2:

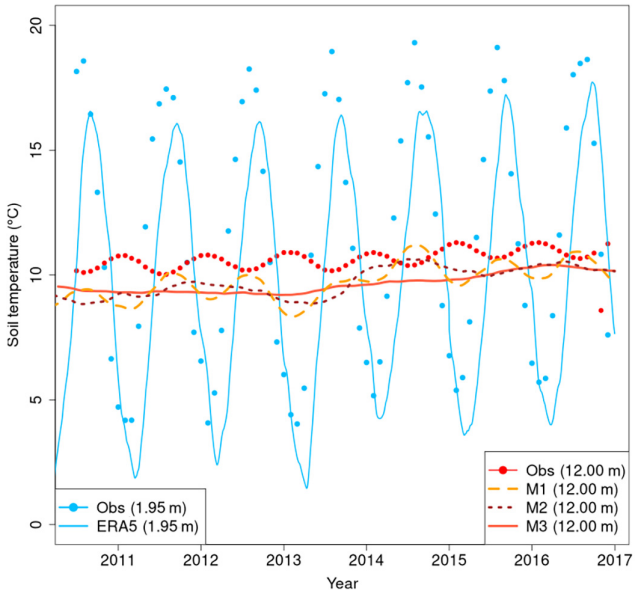

FIGURE 2 Time series of the soil temperature at 12 m depth observed at Potsdam (in red with solid dots) and reconstructed by the three parameterizations M1, M2 and M3 (in orange, dark red and red lines respectively). In light blue are plotted the time series of ERA5 at its deepest soil level (1.95 m) and the interpolation from the observations at the same level.

$$\alpha = \frac{\omega}{2}\left[\frac{z_2 - z_1}{\varphi_{z2} - \varphi_{z1}}\right]^2,$$

where φ_{z2} and φ_{z1} are the initial wave phases at depths z_2 and z_1; the three-yearly running mean of the time series of the deepest soil temperature available in ERA5 (at 1.95 m depth), with a time delay defined by the simplified analytic solution in which the thermal diffusivity is parameterized by the phase method M2. The time lag is computed as $\Delta t = z_1 - z_2/d\omega$. This method is hereafter referred to as M3

M1 and M2 methods are among the most known algorithms to estimate the thermal soil conductivity. Theoretically, the two methods are identical for vertically homogeneous dry soil. However, several studies highlighted better performance for the former or the latter, depending on the soil type, texture and moisture content (e.g. Adeniyi & Nymphas, 2012; Verhoef et al., 1996). Vice versa, the M3 option is based on the hypothesis that the amplitude of the annual thermal wave is very small at 14.58 m depth, thus the temperature can be approximated by a multi-year running mean of ERA5 soil bottom temperature at 1.95 m depth delayed by a proper lag due to soil inertia. Ideally these approaches should be compared with real soil observations in order to identify the best performing one over the domain of interest and at the required depth. However, over Europe, very few sites offer continuous multi-year time series of the soil

TABLE 1 RMSE and correlation index (R) computed for the three parameterizations of the soil temperature at 12.00 m depth against the measurements in Potsdam station, Germany

Name	Parameterization	RMSE (°C)	*R*
M1	Analytic solution using amplitude method	1.17	−0.015
M2	Analytic solution using phase method	1.04	0.39
M3	3-yearly running mean with time lag	1.07	0.38

temperature at depths lower than 1 m, and to the authors' knowledge, only the weather station at the University of Potsdam (52°24 N, 13°10 E, Germany) provides a multi-year time series at a depth comparable to the deepest soil level of the COSMO model. For this reason, the observed data at this location are used to investigate which method to employ for the long-term constraint of the bottom boundary condition related to deep soil temperature.

Figure 2 reports Potsdam recorded data[3] at 12.00 m depth against the time series reconstructed by the three methods at the same depth, considering the closest grid point to the observation station.

It can be noticed that ERA5 (blue solid line) compares well to the observations interpolated at the model levels (i.e. 0.04, 0.18, 0.64 and 1.95 m). For the deepest ERA5 soil level, a small time shift and a slight underestimation of the annual mean are detected, while the model amplitude is coherent with the observed. This gives a good confidence in using ERA5 data as input for the algorithms in this specific location. As for the three proposed methods, M1 shows a correct annual amplitude compared with the observations at 12.00 m depth, but with an opposite phase. On the other hand, M2 underestimates the wave amplitude and shows an inter-annual trend non-coherent with real data. The difference in the annual amplitude estimates derives from the damping depth, which results larger using the phase method (M2) than using the amplitude case (M1) (respectively, 4.89 and 2.53 m). Finally, M3 does not estimate the annual wave amplitude (it is null per definition), but the inter-annual trend appears coherent with the observed one. The quantitative scores (i.e. RMSE and correlation index, reported in Table 1) computed for the time series suggest that M2 and M3 methods are qualitatively similar, and both outperform M1. Since the M3 method is simpler to implement, it is chosen for reconstructing the deep soil temperature in SPHERA. The parameterization is applied to each grid point of the parent domain (ERA5), taking as input the time series of the soil temperatures at the model level depths of 0.64 and 1.95 m. The output is provided by the daily fields of the soil temperature at a depth of 12.00 m, then remapped on SPHERA grid.

[3]Available at: http://www.pik-potsdam.de.

To evaluate the effect of the prescribed deep soil temperature in SPHERA, two parallel simulations are run for 2 years (2015–2016), preceded by 6 months of initialization used to spin-up the model soil fields. The first simulation ('SPHERA-cTdeep') sets the SPHERA deep soil temperature to the ERA5 deepest soil level (1.95 m) at the initialization time (1 July 2014) and keeps this value constant along the model integration (for 2 years). The other simulation ('SPHERA-vTdeep') updates daily (at 00 UTC) the deep soil temperature at 14.58 m depth using the M3 method.

The summary of the tests performed for the definition of the nesting method and the deep soil temperature prescribed in SPHERA are reported in Table 2.

2.4 | The observational dataset

The performances of the different tests are evaluated through the comparison of reanalysis estimates with rainfall and temperature observations provided by the national network of meteorological stations available on Dewetra, the Italian Civil Protection database of non-GTS (Global Telecommunications System) stations (Italian Civil Protection Department, CIMA Research Foundation, 2014). Their data are independent from the reanalyses as they are not assimilated during the production of the datasets. About 5500 rain-gauges homogeneously covering the Italian territory and supplying hourly data are considered. The daily aggregated precipitations are calculated from hourly measurements after being quality checked. Furthermore, since special emphasis is given to the evaluation of precipitation extremes, additional manual quality controls are performed to verify the highest accumulations reported in the dataset (i.e. daily accumulations exceeding 500 mm). Anyhow, these data are inevitably affected by various limitations due to their nature: that is, a low pluviometric coverage in specific regions (e.g. mountainous areas), which may cause erroneous performance assessments especially in cases of localized and severe events, and a too low fraction of heated rain-gauges, which may cause an underestimated and delayed detection of precipitation in case of snowfalls (e.g. Grossi et al., 2017; Mair et al., 2016). On the other hand, surface temperature

TABLE 2 Summary of the tests performed for setting up the nesting modality and the prescription for the deep soil temperature for SPHERA

Test	Nesting method	Deep soil temperature prescription	Years of simulation
1	S1s	c-Tdeep	2015–2016
2	S2s	c-Tdeep	2015
3	S1s	v-Tdeep	2015–2016

sensors amount to about 1000 and include data from the high-resolution regional network Dewetra as well as from the SYNOP stations (that are not assimilated within SPHERA).

2.5 | Performance evaluation

The comparison among SPHERA different configurations and ERA5 is performed in terms of verification against observations for daily accumulated precipitation and 2-m temperature. As for the verification of precipitation, a neighbourhood ('fuzzy') method (Ebert, 2008, 2009) is preferred to the 'nearest-point' technique (i.e. comparison between each observation and the nearest model grid point), in order to reward closeness rather than the perfect point-to-point match between model and observation. Neighbourhood approaches relax the exact match constraint as they require the agreement of the model–observation pair in a spatial window surrounding the forecast and the observed points. For this reason, they are more suitable to assess the model skill in representing precipitation, especially for CP models. Due to the intrinsic limitations in the predictability of convection (e.g. Hohenegger & Schar, 2007), model simulations of convective processes can often be affected by small displacements and/or temporal shifts, causing a penalty in standard verification scores (Ebert, 2008; Theis et al., 2005). Among the neighbourhood techniques, this work considers an upscaling method (Marsigli et al., 2008; Weygandt et al., 2004; Yates et al., 2006), by aggregating the precipitation fields, of both observations and renalayses, over a common coarser grid to provide a single forecast–observation pair for each box of the verification domain. Because of the large differences in grid resolutions between the two reanalyses considered (i.e. $0.02 \times 0.02^{\circ}$ for SPHERA and $0.3 \times 0.3^{\circ}$ for ERA5), a fair comparison is achieved by upscaling SPHERA to ERA5 grid over the Italian domain. The comparison is performed by considering the maximum and the average of daily accumulated precipitation, as well as the 95th percentile of daily distributions in each box between model and observations. This allows us to evaluate different properties of the simulated precipitation fields (more details can be found in Marsigli et al., 2008). In fact, by considering different parameters of precipitation statistical distributions such as the maximum (or 95th percentile) or the average, it is possible to evaluate respectively the performance in the representation of the most extreme events (the tail of the distribution) as well as of the average precipitation within the time period considered. This type of comparison is useful to understand the extent to which deviations from the observed state are attributable to the method chosen to aggregate the fields. To avoid errors due to both the inevitable dishomogeneities arising from an heterogeneous spatial distribution of the rain-gauges, and from the larger domain covered by reanalyses grids compared with that of the observations, an observational mask is applied (similar to Marsigli et al., 2008) to retain only those boxes presenting an appropriate closeness between reanalysis and observed distributions. Furthermore, boxes containing less than five rain-gauges, indicative of an undersampling of the observed distributions, are withdrawn from the analysis not to incur in misrepresentations of the results due to the lack of representativeness of the used data. A categorical approach is then applied to quantify the performance of the reanalyses in representing precipitation: in each box, a 2×2 contingency table is calculated for a set of rainfall thresholds, which are used to describe the joint distribution between the binary observations and the binary outcome of the simulations. From the results of the contingency table, a series of dichotomous scores is calculated to evaluate different aspects of the reanalysis simulations, namely: the probability of detection (POD), the false alarm ratio (FAR) (that can be expressed as the success ratio [SR $= 1 -$ FAR]), the threat score (TS) and the frequency bias. TS combines POD and SR information, measuring the correspondence between observed and modelled events when correct negatives are removed from consideration. POD, SR and TS range between 0 and 1, where 1 indicates a perfect score, while the frequency bias ranges between 0 and ∞, with a perfect score of 1 for an unbiased forecast. The results are reported in the performance diagram (Roebber, 2009), which exploits the geometric relationships between these multiple verification indices. Additionally, to consider the full set of fields of the contingency tables, including the correct negatives, the Heidke skill score (HSS) is calculated (Heidke, 1926), where HSS $= 1$ represents perfect forecasts and HSS $= 0$ indicates no skill. Hence, the values of the scores are aggregated over the whole verification domain and reported for different temporal ranges related to summer and winter seasons of the year (JJA and DJF, respectively). For a detailed description of the scores, the reader is referred to Wilks (2019).

For the verification of 2-m temperature, the nearest-point method is chosen by associating each observation

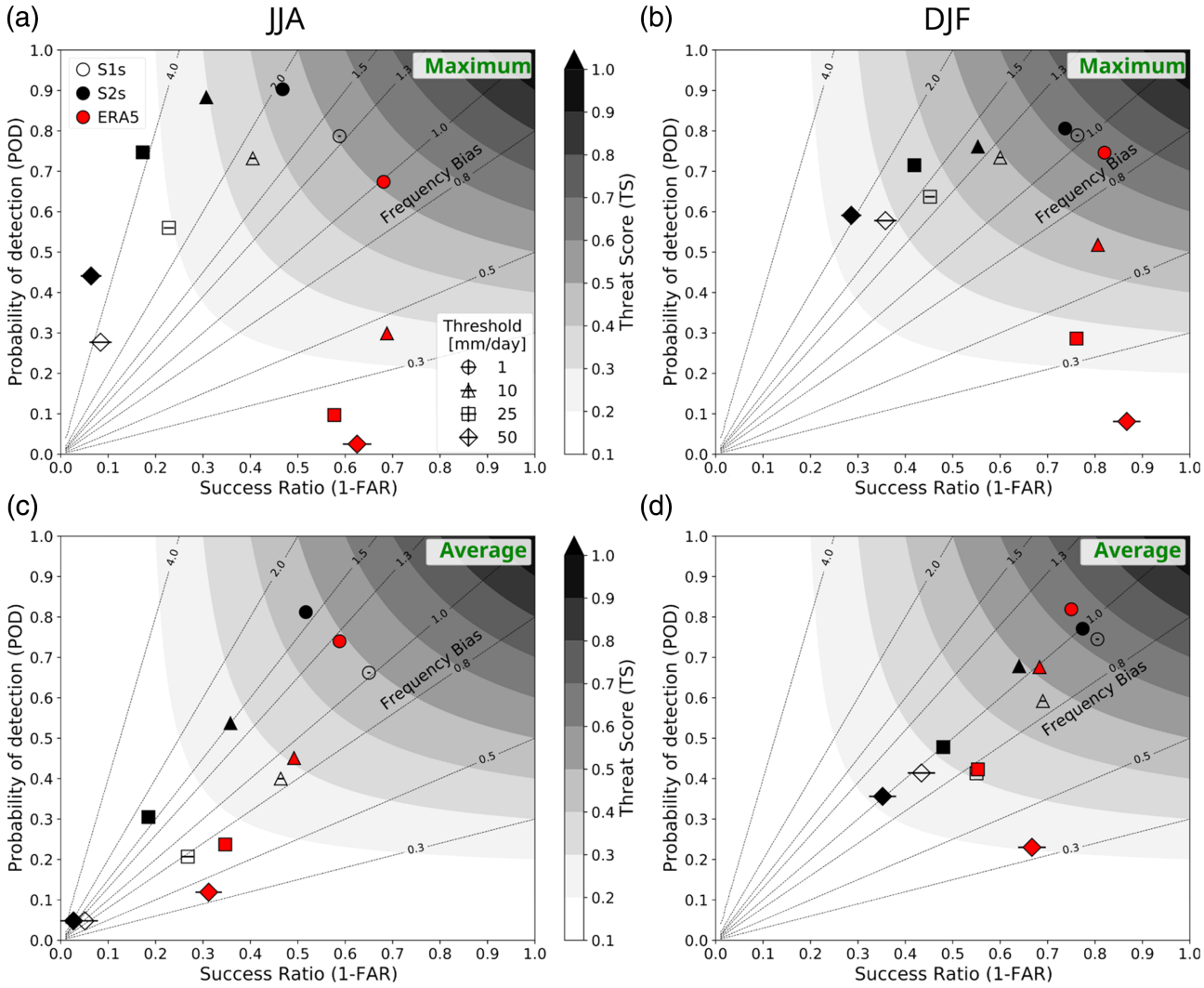

FIGURE 3 Performance diagrams for JJA (a, c) and DJF (b, d) of daily accumulated precipitation during 2015 when maximum (a, b) and average (c, d) values over boxes of the interpolated domain over ERA5 grid are considered. Different reanalysis set-ups (S1s, S2s and ERA5) are reported with different colours. The threat score is indicated in different shades of grey, and the results pertaining to various daily precipitation thresholds (ranging from 1 to 50 mm/day) are shown with different symbols. The frequency bias can be estimated from the deviations from the 45° black line, indicating an unbiased forecast (i.e. frequency bias = 1). Cross-hairs indicate the uncertainty related to the sampling variability of the data and are calculated from a resampling with bootstrapping of 1000 new samples.

to the closest grid point, which is common practice for the verification of continuous meteorological fields like temperature (conversely to highly discontinuous variables like precipitation). Then the bias and RMSE scores are calculated every 3 h. Height altitude correction is not applied, but data with a height difference between model and observation larger than 500 m are not used.

3 | RESULTS AND DISCUSSION

This section analyses the results concerning the nesting method and the prescription of the deep soil temperature for the optimization of SPHERA's set-up. Furthermore, a first evaluation of the performance of the new reanalysis and a comparison with its driver for the simulation of daily accumulated rainfalls over a period of 2 years (2015–2016) is presented.

3.1 | Nesting impact

The comparisons between the two SPHERA implementations with different nesting methods (i.e. S2s and S1s, respectively, corresponding to Test 2 and the first year of Test 1 in Table 2) and ERA5, for different precipitation thresholds during winter (DJF) and summer (JJA) of

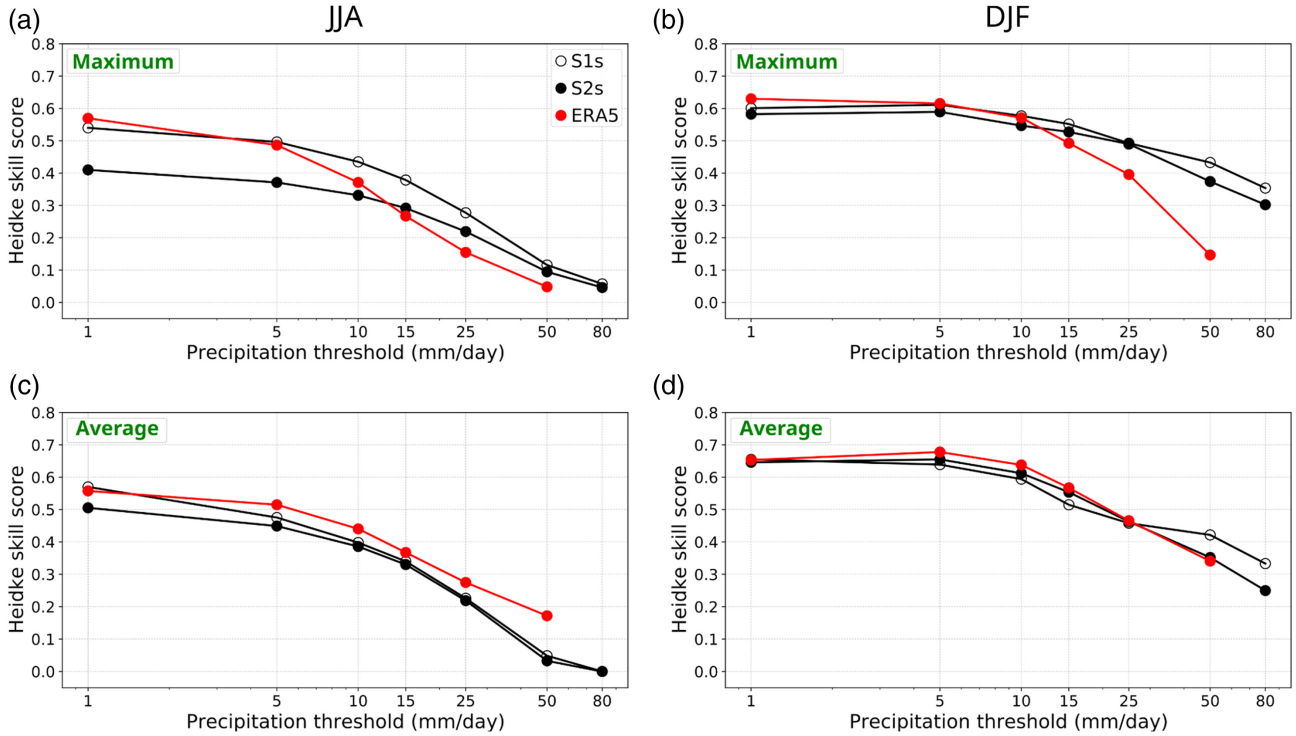

FIGURE 4 Heidke skill scores (HSS) for JJA (a, c) and DJF (b, d) of daily accumulated precipitation during 2015 when maximum (a, b) and average (c, d) values over boxes of the interpolated domain over ERA5 grid are considered. SPHERA set-ups (S1s, S2s) and ERA5 are reported with different dot fillings and colours.

2015, are reported with performance diagrams (Figure 3) and HSSs (Figure 4). Both the maximum and the average values of the boxed rainfall distributions are considered. As for maxima, both S1s and S2s perform better than ERA5 for precipitation values higher than 25 mm/day (15 mm/day considering HSS), both during summer and winter (Figures 3a,b and 4a,b). This result can be related to the improved representation of convective cells in SPHERA and the higher ability to detect the correct intensity, particularly for high thresholds. This is in agreement with what is observed in operational forecast modelling and represents one of the major advantages of the increased resolution (Baldauf et al., 2011; Kendon et al., 2012; Weusthoff et al., 2010), as will be discussed in more detail in Section 3.3. Regarding the SPHERA set-ups, the diagrams show that S1s performs better than S2s at every threshold in both seasons when considering the maximum of the precipitation distribution. The improvement, associated with a smaller bias and higher SR, TS and HSS in S1s, is larger during summer. In both seasons, but particularly during summer, S2s shows a higher POD at every threshold. This is due to an excessive amount of simulated precipitation as supported by the associated very small SR values. This is confirmed by the time series of differences in daily accumulated precipitation between S1s and S2s averaged over the domain in Figure 5. S2s is

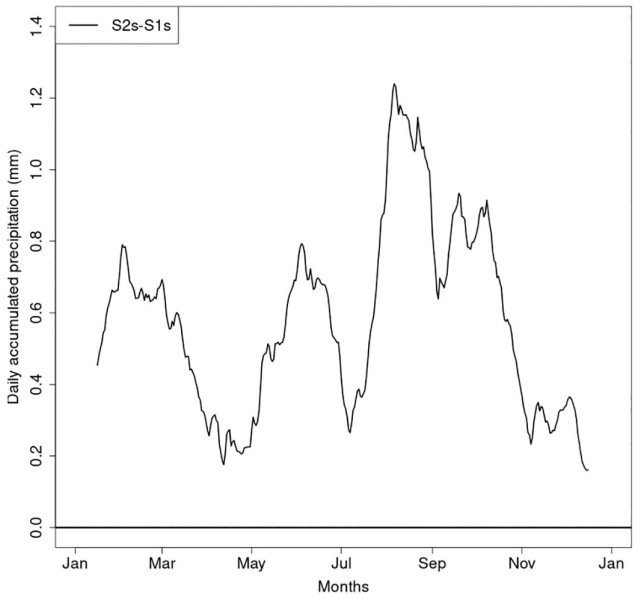

FIGURE 5 Time series of the difference between the precipitation of S2s and S1s averaged over the domain during 2015

wetter than S1s during all the year, with major divergence between the end of summer and the beginning of autumn. In this period, strong convective events over the Italian domain are favoured by the frequent occurrence

of deep troughs over the Atlantic-Mediterranean region with associated moist flows impinging over the Italian orography. The excessive precipitation (and lower skill) in S2s might be associated with several reasons: spurious gravity waves may develop using the multiple interpolation procedure in S2s compared with S1s (as they might be excited at the lateral boundaries and further contaminate part of the integration domain, Warner et al., 1997) or an overestimation of precipitation in the intermediate resolution (convection-parameterized) COSMO integration may produce a too wet soil in S2s (indeed the soil state in S2s has been imposed as a lower boundary condition from the intermediate resolution COSMO integration). On the other hand, the lower precipitation (and enhanced performance) of S1s may also be attributed to the large resolution jump from ERA5 to 2.2 km that would prevent the propagation or the spin up of convective events entering the domain, while the 10 km model would already allow for some explicit convection to propagate (Fosser et al., 2020). Independently from the reason, the comparison with the observations indicates the lower precipitation of S1s for the box maxima as a benefit with respect to S2s.

As for the average values of precipitation (Figures 3c, d and 4c,d), the benefit of SPHERA compared with the driver is reduced and the values of the scores at all thresholds are more similar. However, ERA5, besides presenting higher TS and HSS scores than both S1s and S2s, especially during JJA, underestimates the number of precipitation events. The undersampling increases with rainfall intensity, as shown by the downward deviation from the unity frequency bias line in the performance diagrams, even if less systematically than for the maximum counterparts. This highlights the greater difficulty of ERA5 in producing an adequate number of events also when considering the central part of the precipitation distributions and not only the tail of extreme events. Concerning the comparison between the two SPHERA nestings, a general decrease of POD, SR and HSS scores comparing the average to the respective maximum values is detected, especially during JJA. However, a general tendency to produce an excessive number of rainfall events is detected for S2s (frequency bias always >1), which always presents higher PODs at the expense of lower SRs compared with S1s, while regarding HSS the two nestings report very similar results. Consequently, S2s is affected by an oversampling also when considering the average of the distributions, while S1s presents a close-to-1 frequency bias in both seasons, but particularly in summer.

As for the verification of temperature with different nesting methods (Figure 6), the three-hourly bias and RMSE scores are comparable or weakly better for S1s in both summer and winter (similar results are obtained also for spring and fall, not shown). It is interesting to note that, as compared with ERA5, both the RRA set-ups improve the RMSE by about one-third and reduce the bias to about 0.5°C, probably due to a better representation of topography implied with an enhanced horizontal resolution.

On the basis of these results, it is decided to adopt the S1s nesting approach to extend the production of SPHERA and to carry out the performance analyses reported hereafter.

3.2 | Impact of deep soil temperature

To quantify the impact of different deep soil temperature prescriptions in SPHERA, a comparison is performed between the two parallel simulations described in Section 2.3, namely 'SPHERA-cTdeep' and 'SPHERA-vTdeep' (corresponding respectively to Test 1 and Test 3 in Table 2). In SPHERA-vTdeep, the deep soil temperature remains relatively constant over time, with an average seasonal variability of about 0.1°C (Figure 7). On the domain average, SPHERA-vTdeep is about 0.9°C colder than SPHERA-cTdeep at 14.58 m depth. This difference propagates upwards into the soil and after 24 months the signal is visible at 0.54 m depth, with a temperature difference of about 0.1°C. Above this soil level, the average signal has smaller amplitude and at the surface is null (Figure 7). However, when considering the spatial distribution, surface differences in specific sites can be larger than 0.5°C (Figure 8). Despite these differences at soil level, significant changes are not found in total precipitation nor in temperature and relative humidity at near-surface and upper atmospheric levels at the end of the 24-month simulation (not shown). However, it should be borne in mind that SPHERA outputs are produced in streams of 4 years; at the beginning of each stream, soil data are reinitialized, and then let to change based on atmospheric forcing. Since the length of each stream exceeds the testing period, and the soil variations amplify with time (Figure 7), the surface or the atmosphere might be affected in the second part of each simulation stream, beyond the range described in the test dataset.

On the basis of these results, the SPHERA-vTdeep approach is adopted for the production of SPHERA.

3.3 | Evaluation of daily precipitation

Before proceeding with the production of the complete 25 years dataset, an evaluation of the performance of

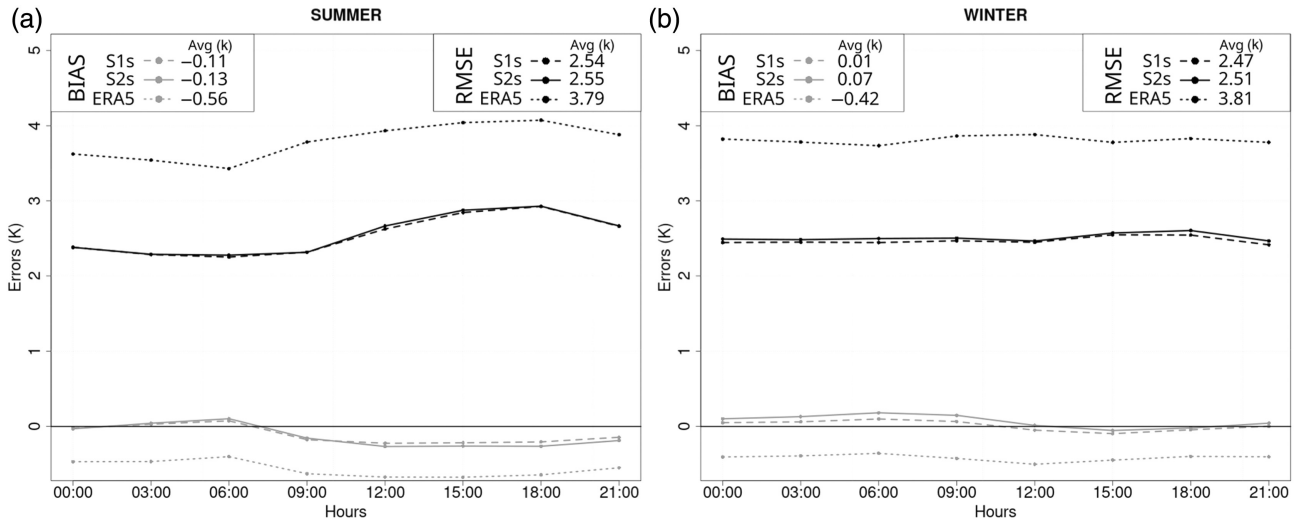

FIGURE 6 Bias (grey lines) and RMSE (black lines) of temperature at 2 m in S1s (dashed), S2s (solid) and ERA5 (dotted) averaged on the day hours for summer 2015 (a) and winter 2015 (b). In the legends the respective seasonal averages of the scores are reported.

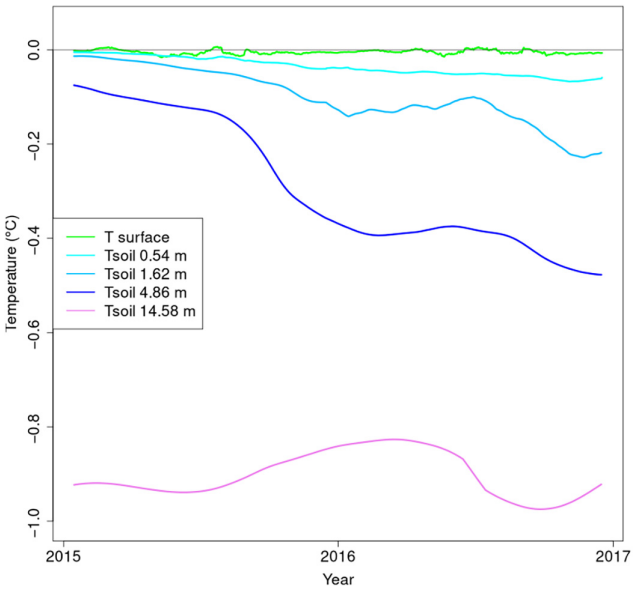

FIGURE 7 Temporal evolution of the difference in soil temperature at different depths (as indicated in the legend) averaged over the domain between the simulations performed by SPHERA-vTdeep and SPHERA-cTdeep

SPHERA is carried out for 2 years of integration (i.e. 2015 and 2016) with the selected setting (corresponding to Test 3 in Table 2). Past climate and climate projections are often evaluated in terms of Essential Climate Variables. Since climate change is expected to increase the severity and number of high-impact weather events, past and future climate should be evaluated also with respect to the change in frequency and intensity of these events. The evaluation is carried out focusing on 24-h accumulated precipitation whose maxima are usually connected with major flood events (e.g. Luino, 1999, Fiori et al., 2011). The analysis is performed using the dense observational dataset from the Italian Civil Protection described in Section 2.4. SPHERA is assessed against the driver model ERA5 to evaluate the impact of the increased horizontal resolution, which is linked especially to the possibility of switching off the parameterization of deep convection. As regards intense and localized precipitation, km-scale models should be verified by considering the loss of predictability at the pixel scale (e.g. Hohenegger & Schar, 2007). To reduce the penalty that would arise for high-resolution models if point-to-point matches between observation and reanalysis are requested (Ebert 2008, 2009), the neighbourhood technique described in Section 2.5 is adopted. The model pixel scale is aggregated both for the reanalyses and for the observation dataset at the scale of the coarser resolution model ERA5, and considering the maximum, the 95th percentile and the average of daily precipitation values within each grid box (Marsigli et al., 2008). This choice is consistent with what is done at ARPAE-SIMC for the verification of operational numerical forecasts.

3.3.1 | Spatial distribution

As a first investigation, the spatial distribution of precipitation is analysed over the years 2015–2016. The choice of the distribution parameter used for aggregating the rainfall distributions in the upscaled domain significantly influences the results. In fact, when coming to spatially distributed fields, considering exclusively the maximum would give a non-robust estimation tailed towards the

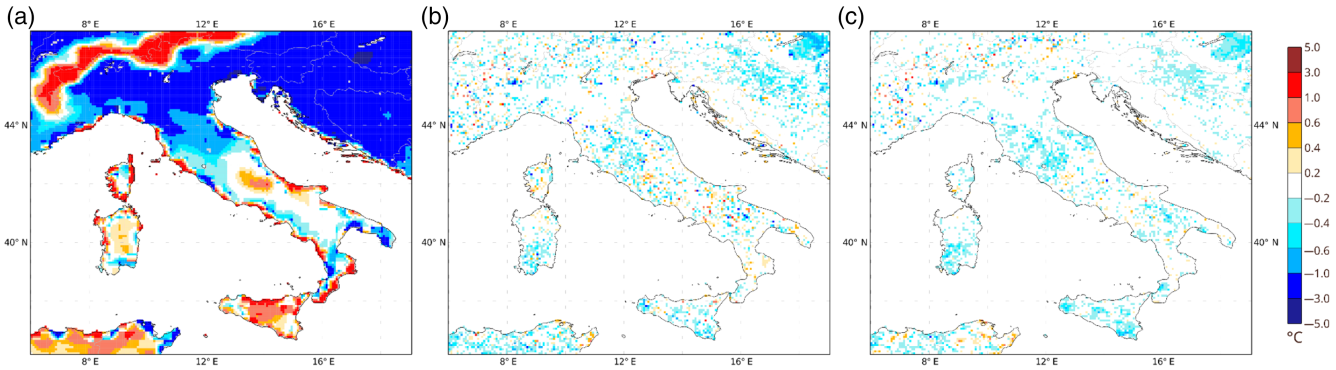

FIGURE 8 Difference in soil temperature at depth 14.58 m (a), 0.18 m (b) and 0.00 m (c) between the simulations performed by SPHERA-vTdeep and SPHERA-cTdeep averaged over the last month of simulation considered (December 2016) after 24 months from the initialization

most extremal events (as it derives from a single parameter of the entire distribution in a box). On the other hand, the mean value assesses the average skill of the reanalyses, implying an abrupt cutting of the tails related to severe precipitation events (for which we are seeking the benefits of the high resolution). For these reasons, the 95th percentile of the distributions is chosen to assess the spatial behaviour of the simulated precipitation fields, with a focus towards intense rainfall events. The results (Figure 9), on the annual term, show the largest values of observed daily accumulated precipitation (Figure 9a), between 6 and 9 mm/day, over the Alpine and Apennine regions. Conversely, in the Po Valley in Northern Italy, in some regions in Sardinia and Sicily, and in southeastern Italy, lower amounts of about 3 mm/day and below are recorded (the reader is referred to Figure 1 for geographical references). Compared with the observational dataset, ERA5 presents a dry negative bias over almost the whole Italian territory (Figure 9c), particularly marked over specific spots in mountainous areas, while it performs well over the plains (well visible in the Po Valley). The ERA5 daily bias calculated for each cell of the upscaled grid and averaged over the whole Italian domain (hereafter referred to as mean bias per box) on the annual term (i.e. the spatial average of Figure 9c) amounts to -1.3 mm. SPHERA for the annual aggregation shows a wet bias (with a mean bias per box of 1.6 mm) mainly over the Po Valley and the southern regions (Figure 9b). The opposite tendencies of the two reanalyses biases are confirmed also by the density histograms reported in Figure 9d.

To better understand the behaviours of reanalyses on different temporal aggregation scales, the seasonal biases are also reported in Figure 9 for summer (JJA) and winter (DJF). During summer, ERA5 underestimates precipitation over the whole Italian peninsula (Figure 9g), with a mean bias at the box of −1.3 mm, and high spatial coherence with the annual bias. Conversely, SPHERA presents a marked wet bias over the western Po Valley, and a less pronounced wet bias along the western sectors of the peninsula and over the Islands (Figure 9f). The daily average bias over summer per box is 2.8 mm, while the bias density histogram reports a weakly drier median centred around 2.5 mm/day (Figure 9h). It has to be pointed out that the overestimation in summer precipitation when considering the extreme tail of the rainfall distribution is, at least partly, attributable to the limits in the ability to sample the observed state of the pluviometers network: if observations are not sampled with the sufficient spatial frequency (as it could be the case, see Section 2.4), it is likely for the high-resolution system to overestimate the rainfall fields in the bias calculation between the 95th percentiles of SPHERA and the sampled real-state. Consequently, the wet bias is at least in part expected, and it may suggest the potential of SPHERA in simulating very high rainfall rates. During wintertime, SPHERA reports significantly smaller bias in precipitation (Figure 9j), with a mean bias at box of 0.4 mm, indicating an overall quasi-unbiased spatial distribution. On the other hand, ERA5 in DJF shows a less spatially systematic dry bias (Figure 9k), presenting a mean value per box of −1.1 mm, with important 'hot spots' of rainfall underestimation in the northern Apennines and over the southern Apennines in southern Italy. This spatial distribution is explained by the seasonality of orographic precipitation over these mountainous areas (as visible from the observed pattern in Figure 9i). Here precipitation is particularly enhanced during winter as a consequence of the predominant characteristics of mesoscale humid flows impinging over the orography (e.g. Krichak et al., 2015). In this case, the difficulties in adequately representing orographic precipitations are also shown by SPHERA, which, however, reports a less accentuated and more localized dry bias over the

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

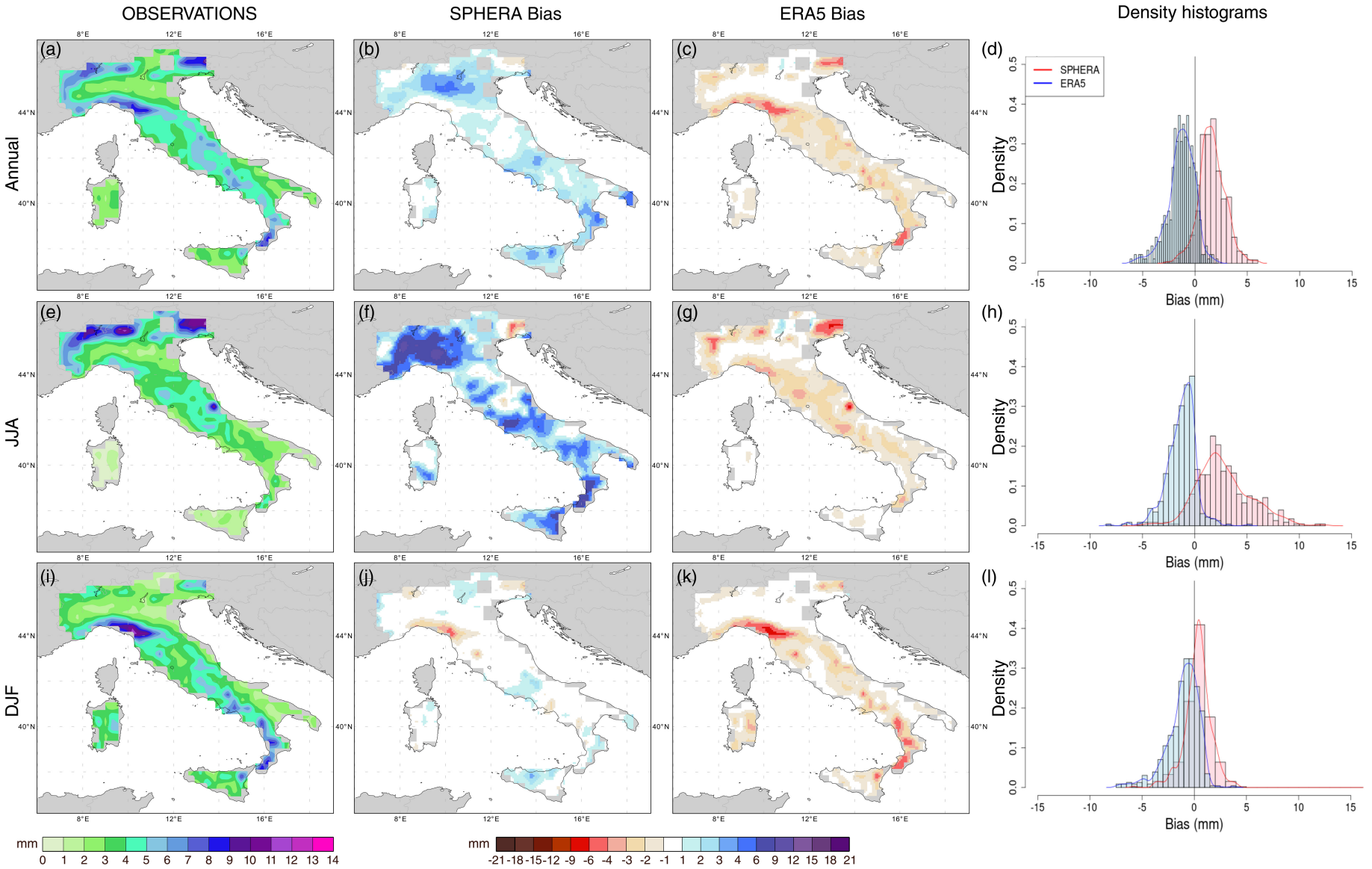

FIGURE 9 Observed (left column—panels a,e,i) average of the 95th percentile of daily precipitation distributions for different temporal aggregations over the years 2015–2016: Annual (top row - panels a through d), summer (JJA, central row—panels e through h) and winter (DJF, bottom row—panels i through l). On centre-left and centre-right columns are reported the deviations from the observed spatial distributions for SPHERA and ERA5 respectively. Both SPHERA and observations datasets are upscaled through an interpolation over ERA5 grid. The grey colour in the maps indicates areas over land. On the right column (panels d,h,l) are reported the respective density histograms of the annual, summer and winter daily biases relative to the same rainfall distributions of SPHERA (in red) and ERA5 (in blue).

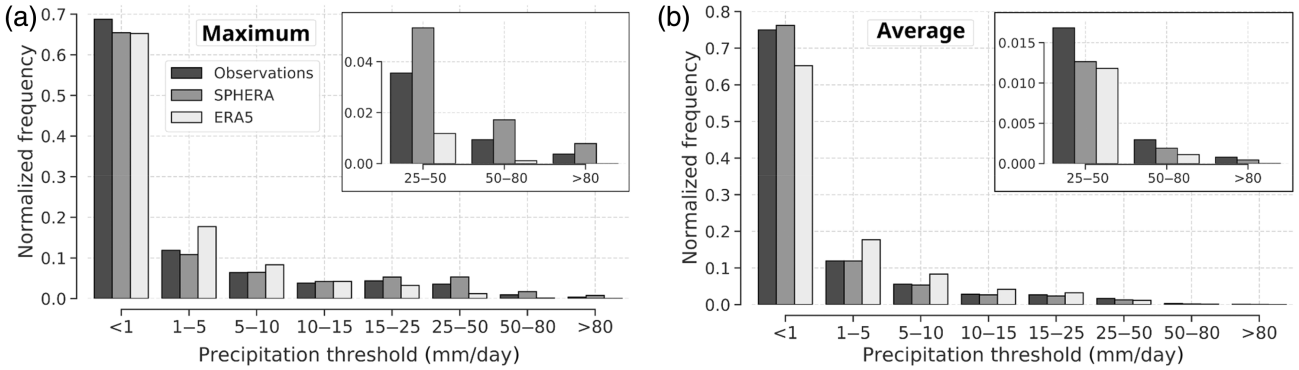

FIGURE 10 Comparison of the normalized frequency histograms of the distributions of daily rainfall occurrences over 2015–2016 among observations, SPHERA and ERA5 datasets, when maximum (a) and average (b) values of the distributions over boxes obtained with the interpolation over ERA5 native grid are considered. For a better visualization the sub-distributions for the highest thresholds are highlighted in black-framed subplots.

northern Apennines (Figure 9j), and no predominant biases over the southern Italian Apennines. These features are confirmed by the comparison of the density histograms of the two reanalyses during winter (Figure 9l): in this case, SPHERA shows a sharper and slightly positively biased distribution, in contrast to a more blunted and negatively centred distribution related to ERA5.

3.3.2 | Rainfall relative frequencies

In the following, the normalized relative frequencies of daily precipitation intensities for SPHERA and ERA5 against the observational dataset, in terms of maximum and average values of the boxed distributions, are considered (Figure 10) to add further details to the validation exercise. The histogram related to the maxima (left plot) reveals relevant differences among the three datasets: while ERA5 slightly overestimates the number of low-precipitation events (i.e. with intensities lower than 10 mm/day), excluding the range of dry days (<1 mm/day), SPHERA slightly underestimates the number of events in the range <5 mm/day. On the contrary, for larger precipitations, ERA5 shows an undersampling for accumulations >10 mm/day, worsening as the amount increases. This result can be attributed to the difficulty in simulating severe rainfall events, likely to be associated with deep moist convection processes, which characterizes coarse and convection-parameterized models (e.g. Clark et al., 2016; Iyer et al., 2016). On the other hand, SPHERA clearly overpredicts the number of precipitation occurrences, with the maximum overestimation in the range of 25–50 mm/day. A similar analysis of frequency distribution was performed with the other CP reanalysis based on the COSMO model and produced over the German territory, that is, COSMO-REA2, does not report any tendency to overestimate the number of intense rainfall events (Wahl et al., 2017). Hence, the behaviour exhibited by SPHERA should not be attributable to an intrinsic bias of the COSMO model. However, we must point out that COSMO-REA2, conversely to SPHERA, makes use of a latent heat nudging scheme for assimilating radar data, which puts further constraints to the model simulation, compared with a nudging approach. Furthermore, in contrast to Wahl et al. (2017), in the present work, the frequency analysis is based on the distribution of precipitation after the box aggregation, which is a crucial point. In fact, the choice to consider the box maximum can be, at least in part, responsible of the frequency overestimation for intense rainfall events: due to the larger amount of SPHERA grid points compared with rain-gauges per box, it is more likely for SPHERA to catch a higher precipitation peak than it is for the less dense observational network. In the same way, part of the large ERA5 underestimate can be attributed to the choice to consider the maximum per box, which penalizes its coarser resolution. To better understand the impact of the statistical parameters used for the aggregation, the relative frequency histogram for the average values in the boxes is considered (Figure 10b). In this case, SPHERA tends to simulate always a more adequate number of events than ERA5 at all intensities, with a very good agreement with the observations in the range <25 mm/day, and an undersampling for larger intensities. ERA5 is characterized by larger errors than SPHERA in sampling events at all intensities, with a significant oversampling for accumulations in the range of 1–25 mm/day, and an undersampling always stronger than SPHERA for events >25 mm/day. This further confirms the dry bias affecting ERA5 in this range of rainfall intensities, previously detected when considering the maximum values,

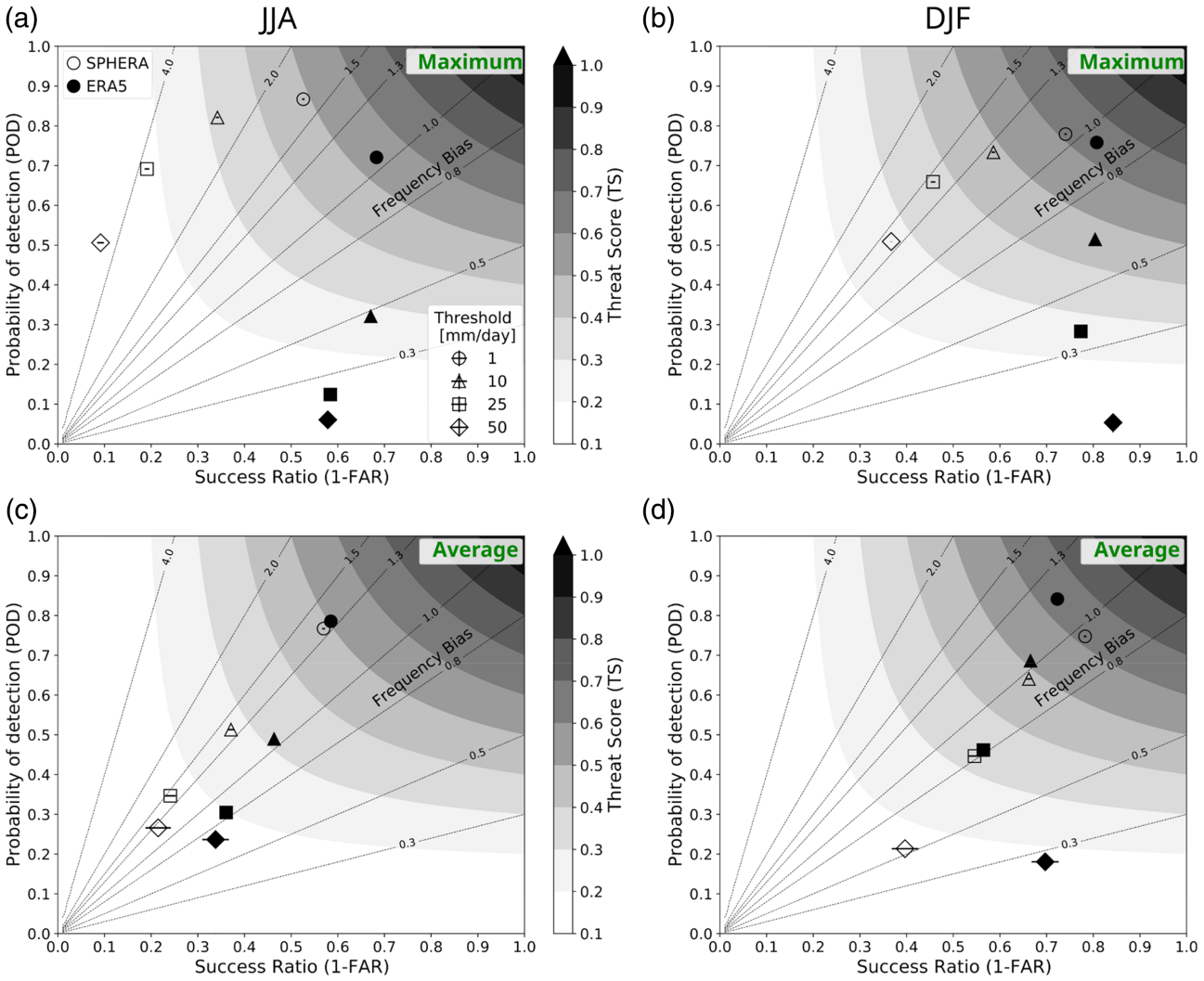

FIGURE 11 As Figure 3 but for the aggregation of the data over 2015–2016 and relative to SPHERA (empty symbols) and ERA5 (black symbols)

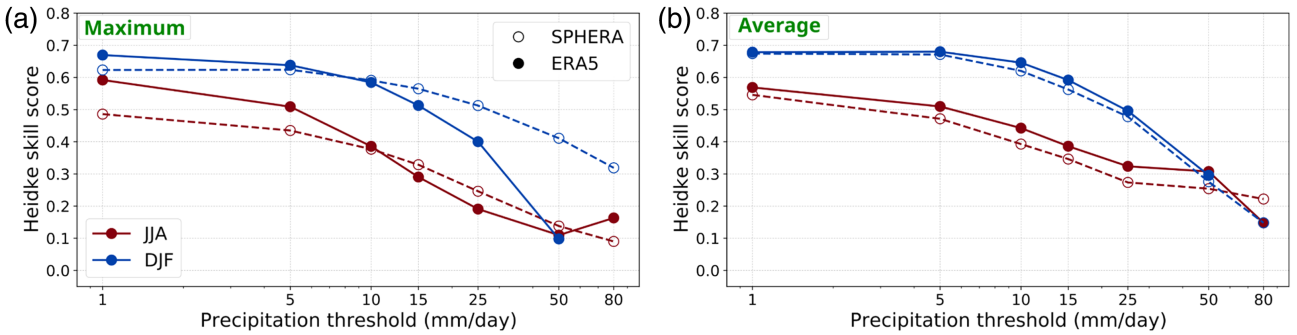

FIGURE 12 Heidke skill scores (HSS) for JJA and DJF (in red and blue colours respectively) of daily accumulated precipitation during 2015–2016 when maximum (a) and average (b) values over boxes of the interpolated domain over ERA5 grid are considered, for SPHERA (in empty dots and dashed lines) and ERA5 (in solid dots and continuous lines)

indicating the general tendency of the global reanalysis to underestimate intense rainfall regardless of the method employed for the statistical aggregation (even if a smaller gap with SPHERA is shown when considering the average values).

3.3.3 | Quantitative evaluation by statistical indices

To quantitatively assess the performance of the reanalyses in simulating precipitation, the performance

diagrams and the HSSs for daily maximum and average values for JJA and DJF during 2015–2016 are reported (Figures 11 and 12 respectively). The results concerning SPHERA show similar behaviours to those of the test version S1s[4] previously investigated in Figure 3.

Considering the maximum (Figures 11a,b and 12a), SPHERA presents for every precipitation intensity a higher or similar POD and a lower SR compared with those of ERA5 for both seasons. This indicates a higher performance in detecting the actually occurred rainfall events in SPHERA, especially for significant accumulations, coming with a worse ability to avoid false alarms. This recalls the tendency of SPHERA to overpredict the number of precipitation events, particularly for intense rainfall events, already shown in the frequency analysis (Figure 10a), and further confirmed by the greater-than-1 frequency bias for all thresholds (Figure 11a,b). ERA5 reports the opposite behaviour, with a great ability to avoid false alarms, almost constantly for all the intensities, at the expense of too few detected rainfall cases, markedly during summer and for moderate and intense events. This picture is associated with the tendency of ERA5 to underestimate the rainfall frequencies for thresholds larger than 15 mm/day, as shown by the frequency bias lower than 1 and by the frequency distribution (Figure 10a). TS and HSS result to be higher for SPHERA for both seasons for moderate and intense daily precipitation events (i.e. >10 mm/day), while they are both higher for ERA5 for low-intensity rainfall cases. This indicates that the added value of SPHERA over its driver is in the representation of precipitation cases with intensity from medium to heavy, in terms of detection of occurrence and intensity, in agreement with the results of CP models against their convection-parametrizing drivers (e.g. Clark et al., 2016; Iyer et al., 2016; Klasa et al., 2018). Moreover, the added value of SPHERA is slightly more evident in summer where a better performance than ERA5 is shown starting from events with lower intensities compared with winter.

On the other hand, when considering the average values (Figures 11c,d and 12b), we notice similar skills of the verification scores for the two reanalyses, with a slightly better performance of ERA5 in predicting the average amounts of precipitation, as indicated by the superior HSS and TS compared with SPHERA at all thresholds and seasons. However, in terms of the frequency bias, ERA5 undersamples the number of high-precipitation events in both seasons more evidently than SPHERA. This highlights that the two reanalyses have a similar skill in representing the average daily precipitation over Italy, which is justified by the substantial removal from the sample of the occurrences related to tailed events presenting large and extreme daily accumulations.

[4]The only difference between the currently considered SPHERA simulation and S1s is the assigned deep soil temperature (see Test 1 and Test 3 in Table 2).

4 | SUMMARY AND CONCLUSIONS

In this article, a new convection-permitting regional reanalysis SPHERA (High Resolution REAnalysis over Italy) is presented. One main focus of this work is the definition of the best model set-up ahead of the reanalysis integration. Two points are considered: the downscaling procedure and the deep soil temperature boundary condition strategy. With regard to the first issue, results suggest to nest SPHERA directly in ERA5, although the considerable leap in resolution, as the usage of an intermediate step does not produce an improvement but rather a slight detriment of the simulation quality. Concerning the bottom boundary condition, important to avoid spurious drifts during the multi-year integration, three different deep soil temperature prescriptions are considered to transfer the information from ERA5 bottom soil level (1.95 m depth) to SPHERA's counterpart (14.58 m depth). The three-yearly running mean of ERA5 bottom soil level time series with a time delay is the one showing the best accordance with the observed deep soil temperature state and the simplest to be implemented in the model, therefore it is chosen for the extension of the dataset. The second purpose of this work is to evaluate and compare the performance of SPHERA with ERA5 in representing precipitation over a testing period of 2 years (2015–2016), in preparation for the multi-decadal model integration. Various aspects of daily rainfall simulations are assessed, considering their spatial distributions, the relative frequencies of the events and employing dichotomous methods for quantifying different attributes of the simulations through statistical scores. The main improvement of the CP reanalysis over the global counterpart concerns moderate and heavy daily rainfalls. SPHERA performs significantly better both regarding the frequencies of occurrence and in terms of more skilful representations of the actually occurred events. On the other hand, ERA5 is more reliable in simulating low to moderate accumulations with an increasing negative frequency bias when moving to higher intensities. This is expected when comparing high-resolution CP reanalyses with lower resolution convection-parametrizing datasets, resulting from the enhanced representation of localized convective processes and of topography (especially in presence of very

complex terrains) that is gained through finer grid meshes. The results are in accordance with the performance shown by other CP reanalyses for similar precipitation intensities (e.g. Fosser et al., 2015; Prein et al., 2013; Wahl et al., 2017). Even if the quality of ERA5 simulations is already high, particularly when coming to avoiding non-observed events, the possibility to improve the simulation of cases having relevant intensity is undoubtedly a pivotal element for the definition of the past, present and future climate. The clear improvement in the representation of precipitation shown by SPHERA was essential to proceed with the production of the dataset for the entire 25-year period (1995–2020). It is anyway important to stress that the sample of data used in the present analysis is appropriate to get a first insight into the potential of the new dataset, but it is certainly too small to permit a robust and comprehensive assessment of SPHERA. Additionally, given the high temporal resolution of 1 h of SPHERA simulations, a sub-daily analysis of precipitation is expected to highlight even more the added value in the description of high-impact weather events (Prein et al., 2015). A follow-up paper focused on an extensive and more detailed validation of the multi-decadal dataset of SPHERA has already been submitted.

AUTHOR CONTRIBUTIONS

Ines Maria Luisa Cerenzia: Data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); project administration (lead); resources (lead); software (lead); validation (lead); visualization (equal); writing – original draft (lead); writing – review and editing (supporting). **Antonio Giordani:** Formal analysis (supporting); methodology (supporting); validation (supporting); visualization (supporting); writing – review and editing (lead). **Tiziana Paccagnella:** Conceptualization (lead); supervision (lead); writing – original draft (supporting). **Andrea Montani:** Conceptualization (supporting); methodology (supporting); supervision (supporting); writing – review and editing (equal).

ACKNOWLEDGEMENTS

The whole dataset of SPHERA and the preliminary tests have been performed on the ECMWF supercomputer system thanks to resources assigned to the special project "SPHERA" and "SPHERA-PRE". Additional reporting and documentation is available at https://www.ecmwf.int/en/research/special-projects/spitcere-2018 and https://www.ecmwf.int/en/research/special-projects/spitcere-2017. The dataset SPHERA has already been produced for the period 1995-2018 and it is archived in the ECMWF file storage system (ECFS). Selected variables will be publicly available from ARPAE-SIMC, through the MISTRAL portal. Contact the corresponding author for more information and data access. The authors thank Valentina Pavan for the useful discussion, and two anonymous reviewers for the suggestions leading to a substantial improvement of the manuscript.

ORCID

Ines Maria Luisa Cerenzia https://orcid.org/0000-0001-9902-3701

REFERENCES

Adeniyi, M.O. & Nymphas, E.F. (2012) Validation of analytical algorithms for the estimation of soil thermal properties using de vries model. *American Journal of Scientific Research*, 3, 103–114. https://doi.org/10.5251/ajsir.2012.3.2.103.114

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

Berg, P., Donnelly, C. & Gustafsson, D. (2018) Near-real-time adjusted reanalysis forcing data for hydrology. *Hydrology and Earth System Sciences*, 22, 989–1000. https://doi.org/10.5194/hess-22-989-2018

Berg, P., Wagner, S., Kunstmann, H. & Schädler, G. (2013) High resolution regional climate model simula- tions for Germany: part i—validation. *Climate Dynamics*, 40(1), 401–414. https://doi.org/10.1007/s00382-012-1508-8

Bollmeyer, C., Keller, J.D., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P. et al. (2015) Towards a high-resolution regional reanalysis for the European cordex domain. *Quarterly Journal of the Royal Meteorological Society*, 141(686), 1–15. https://doi.org/10.1002/qj.2486

Bonanno, R., Lacavalla, M. & Sperati, S. (2019) A new high-resolution meteorological reanalysis Italian dataset: MERIDA. *Journal of the Royal Meteorological Society*, 145(721), 1756–1779. https://doi.org/10.1002/qj.3530

Brisson, E., Demuzere, M. & van Lipzig, N.P. (2016) Modelling strategies for performing convection- permitting climate simulations. *Meteorologische Zeitschrift*, 25(2), 149–163. https://doi.org/10.1127/metz/2015/0598

Bromwich, D.H., Wilson, A.B., Bai, L., Liu, Z., Barlage, M., Shih, C.-F. et al. (2018) The arctic system reanalysis, version 2. *Bulletin of the American Meteorological Society*, 99(4), 805–828. https://doi.org/10.1175/BAMS-D-16-0215.1

Cheruy, F., Dufresne, J.L., Mesbah, S.A., Grandpeix, J.Y. & Wang, F. (2017) Role of soil thermal inertia in surface temperature and soil moisture-temperature feedback. *Journal of Advances in Modeling Earth Systems*, 9(8), 2906–2919.

Clark, P., Roberts, N., Lean, H., Ballard, S.P. & Charlton-Perez, C. (2016) Convection-permitting models: a step-change in rainfall forecasting. *Meeting Apps*, 23, 165–181. https://doi.org/10.1002/met.1538

Dahlgren, P., Landelius, T., Kållberg, P. & Gollvik, S. (2016) A high-resolution regional reanalysis for europe. Part 1: three-dimensional reanalysis with the regional high-resolution limited-area model (hirlam). *Quar-terly Journal of the Royal Meteorological Society*, 142(698), 2119–2131. https://doi.org/10.1002/qj.2807

Dai, A., Giorgi, F. & Trenberth, K.E. (1999) Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *Journal of Geophysical Research: Atmospheres*, 104(D6), 6377–6402. https://doi.org/10.1029/98JD02720

Denis, B., Laprise, R. & Caya, D. (2003) Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. *Climate Dynamics*, 20(2), 107–126. https://doi.org/10.1007/s00382-002-0264-6

Doddy Clarke, E., Griffin, S., McDermott, F., Monteiro Correia, J. & Sweeney, C. (2021) 2021: which reanalysis dataset should WeUse for renewable energy analysis in Ireland? *Atmosphere*, 12, 624. https://doi.org/10.3390/atmos12050624

Doms, G., Förstner, J., Heise, E., Herzog, H.-J., Mironov, D., Raschendorfer, M. et al. (2018) *A Description of the Nonhydrostatic Regional COSMO-Model. part II. Physical Parameterizations, Consortium for small scale modelling*. Offenbach: Deutscher Wetterdienst, p. 167. https://doi.org/10.5676/dwd_pub/nwv/cosmo-doc_5.05_II

Dow, G., & Macpherson, B. (2013). Benefit of convective-scale data assimilation and observing systems in the UK models. Met Office.

Ebert, E.E. (2008) Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications: A Journal of Forecasting, Practical Applications, Training Techniques and Modelling*, 15(1), 51–64.

Ebert, E.E. (2009) Neighborhood verification: A strategy for rewarding close forecasts. *Weather and Forecasting*, 24(6), 1498–1510.

Fiori, E., Comellas, A., Molini, L., Rebora, N., Siccardi, F., Gochis, D.J. et al. (2011) Analysis and hindcast simulations of an extreme rainfall event in the Mediterranean area: The Genoa 2011 case. *Atmospheric Research*, 138, 13–29.

Fosser, G., Khodayar, S. & Berg, P. (2015) Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Climate Dynamics*, 44(1), 45–60. https://doi.org/10.1007/s00382-014-2242-1

Fosser, G., Kendon, E.J., Stephenson, D. & Tucker, S. (2020) Convection-permitting models offer promise of more certain extreme rainfall projections. *Geophysical Research Letters*, 47 (13), e2020GL088151.

Gal-Chen, T. & Sommerville, R. (1975) On the use of a coordinate transformation for the solution of the navier–stokes equations. *Journal of Computational Physics*, 17, 209–228.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L. et al. (2017) The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, 30(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Gerard, L., Piriou, J.-M., Brožková, R., Geleyn, J.-F. & Banciu, D. (2009) Cloud and precipitation parameterization in a mesogamma-scale operational weather prediction model. *Monthly Weather Review*, 137(11), 3960–3977.

Gleeson, E., Whelan, E. & Hanley, J. (2017) Met Éireann high resolution reanalysis for Ireland. *Advances in Science and Research*, 1, 1–13. https://doi.org/10.5194/asr-1-1-2017

Grossi, G., Lendvai, A., Peretti, G. & Ranzi, R. (2017) Snow precipitation measured by gauges: systematic error estimation and data series correction in the central Italian Alps. *Water*, 9(7), 461.

Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H. et al. (2018) Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, 144, 1218–1256. https://doi.org/10.1002/qj.3179

Heidke, P. (1926) Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst. *Geografiska Annaler*, 8(4), 301–349.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049. https://doi.org/10.1002/qj.3803

Hohenegger, C. & Schar, C. (2007) Atmospheric predictability at synoptic versus cloud-resolving scales. *Bulletin of the American Meteorological Society*, 88, 1783–1794. https://doi.org/10.1175/BAMS-88-11-1783

Horton, R.A., Wierenga, P.J. & Nielsen, D.R. (1983) Evaluation of methods for determining the apparent thermal diffusivity of soil near the surface. *Social Science Society of America Journal*, 47, 52.

Italian Civil Protection Department & CIMA Research Foundation. (2014). The Dewetra Platform: a multi-perspective architecture for risk management during emergencies. In Information Systems for Crisis Response and Management in Mediterranean Countries: First International Conference, ISCRAM-med 2014, Toulouse, France, October 15-17, 2014. Proceedings 1 (pp. 165-177). Springer International Publishing.

Iyer, E.R., Clark, A.J., Xue, M. & Kong, F. (2016) A comparison of 36–60-h precipitation forecasts from convection-allowing and convection-parameterizing ensembles. *Weather and Forecasting*, 31, 647–661. https://doi.org/10.1175/WAF-D-15-0143.1

Jacobsen, I. & Heise, E. (1982) A new economic method for the computation of the surface temperature in numerical models. *Contributions to Atmospheric Physics*, 55, 128–141.

Keller, J.D. & Wahl, S. (2021) Representation of climate in reanalyses: an intercomparison for Europe and North America. *Journal of Climate*, 34(5), 1667–1684. Retrieved from https://journals.ametsoc.org/view/journals/clim/34/5/JCLI-D-20-0609.1.xml

Kendon, E.J., Roberts, N.M., Senior, C.A. & Roberts, M.J. (2012) Realism of rainfall in a very high-resolution regional climate model. *Journal of Climate*, 25(17), 5791–5806.

Kirshbaum, D.J., Adler, B., Kalthoff, N., Barthlott, C. & Serafin, S. (2018) Moist orographic convection: physical mechanisms and links to surface-exchange processes. *Atmosphere*, 9, 80. https://doi.org/10.3390/atmos9030080

Klasa, C., Arpagaus, M., Walser, A. & Wernli, H. (2018) An evaluation of the convection-permitting ensemble COSMO-E for three contrasting precipitation events in Switzerland. *Quarterly Journal of the Royal Meteorological Society*, 144, 744–764. https://doi.org/10.1002/qj.3245

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M. & Onogi, K. (2015) The JRA-55 reanalysis: general specifications and basic characteristics. *Journal of the meteorological Society of Japan Series II*, 93(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Krichak, S.O., Barkan, J., Breitgand, J.S., Gualdi, S. & Feldstein, S. B. (2015) The role of the export of tropical moisture into midlatitudes for extreme precipitation events in the Mediterranean region. *Theoretical and Applied Climatology*, 121(3), 499–515.

Lean, H.W., Clark, P.A., Dixon, M., Roberts, N.M., Fitch, A., Forbes, R. et al. (2008) Characteristics of high-resolution

versions of the met Office unified model for forecasting convection over the United Kingdom. *Monthly Weather Review*, 136, 3408–3424. https://doi.org/10.1175/2008MWR2332.1

Lewis, H., Mittermaier, M., Mylne, K., Norman, K., Scaife, A., Neal, R. et al. (2015) From months to minutes – exploring the value of high-resolution rainfall observation and prediction during the UKwinter storms of 2013/2014. *Meeting Apps*, 22, 90–104. https://doi.org/10.1002/met.1493

Luino, F. (1999) The flood and landslide event of November 4–6 1994 in Piedmont Region (Northwestern Italy): Causes and related effects in Tanaro Valley. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*, 24(2), 123–129, ISSN 1464-1895. https://doi.org/10.1016/S1464-1895(99)00007-1

Mair, E., Leitinger, G., Della Chiesa, S., Niedrist, G., Tappeiner, U. & Bertoldi, G. (2016) A simple method to combine snow height and meteorological observations to estimate winter precipitation at sub-daily resolution. *Hydrological Sciences Journal*, 61(11), 2050–2060.

Marsigli, C., Montani, A. & Paccagnella, T. (2014) Provision of boundary conditions for a convection- permitting ensemble: comparison of two different approaches. *Nonlinear Processes in Geophysics*, 21(2), 393–403. https://doi.org/10.5194/npg-21-393-2014

Marsigli, C., Montani, A. & Paccangnella, T. (2008) A spatial verification method applied to the evaluation of high-resolution ensemble forecasts. *Meetting Apps*, 15, 125–143. https://doi.org/10.1002/met.65

Mellor, G.L. & Yamada, T. (1982) Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20, 851–875.

Mesinger, F., Dimego, G., Kalnay, E., Mitchell, K., Shafran, P.C. & Ebisuzaki, W. (2006) North american regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3), 343–360. https://doi.org/10.1175/BAMS-87-3-343

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. & Terzhevik, A. (2010) Implementation of the lake parameterisation scheme flake into the numerical weather prediction model cosmo. *Boreal Environment Research*, 15, 218–230.

Ngan, F. & Stein, A.F. (2017) A long-term WRF meteorological archive for dispersion simulations: application to controlled tracer experiments. *Journal of Applied Meteorology and Climatology*, 56(8), 2203–2220. https://doi.org/10.1175/JAMC-D-16-0345.1

Pal, S., Chang, H.-I., Castro, C.L. & Dominguez, F. (2019) Credibility of convection-permitting modeling to improve seasonal precipitation forecasting in the southwestern United States. *Frontiers in Earth Science*, 7, 11. https://doi.org/10.3389/feart.2019.00011

Pham, V.S., Hwang, J.H. & Ku, H. (2016) Optimizing dynamic downscaling in one-way nesting using a regional ocean model. *Ocean Modelling*, 106, 104–120.

Prein, A., Gobiet, A.F.a., Suklitsch, M., Truhetz, H., Awan, N.K., Keuler, G. et al. (2013) Added value of convection permitting seasonal simulations. *Climate Dynamics*, 41(9), 2655–2677. https://doi.org/10.1007/s00382-013-1744-6

Prein, A.F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K. et al. (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53(2), 323–361. https://doi.org/10.1002/2014RG000475

Ridal, M., Undén, P., Olsson, E., Bojarova, J., Landelius, T., Zimmermann, K. et al. (2017) Harmonie reanalysis: overview. In: *5th UERRA Generaly assembly, Tarragona*.

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

Roebber, P.J. (2009) Visualizing multiple measures of forecast quality. *Weather and Forecasting*, 24(2), 601–608. https://doi.org/10.1175/2008WAF2222159.1

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S. et al. (2010) The ncep climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8), 1015–1058. https://doi.org/10.1175/2010BAMS3001.1

Schättler, U., & U. Blahak 2017a: A Description of the Nonhydrostatic Regional COSMO-Model. Part V: Preprocessing:Initial and Boundary Data for the COSMO-Model, Consortium for small scale modelling, Deutscher Wetterdienst, Offenbach, 86pp. https://doi.org/10.5676/dwd_pub/nwv/cosmo-doc_5.05_V.

Schättler, U., Doms, G. & Schraff, C. (2018) *A description of the nonhydrostatic regional COSMO-model. Part VII. User's guide, consortium for small scale modelling*. Offenbach, 195pp: Deutscher Wetterdienst. https://doi.org/10.5676/dwd_pub/nwv/cosmo-doc_5.05_VII

Schimanke, S. & the C3S_322_Lot1 Team. (2018) *Copernicus regional reanalysis for Europe*. Toulouse: User workshop.

Schraff, C., & Hess, R. (2013) A description of the nonhydrostatic regional cosmo-model. Part iii: data assimilation. User's guide, consortium for small scale modelling, Deutscher Wetterdienst, Offenbach, 93 pp. https://doi.org/10.5676/dwd_pub/nwv/cosmo-doc_5.00_III.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F. et al. (2011) The AROME-France convective-scale operational model. *Monthly Weather Review*, 139(3), 976–991.

Sommeria, G. & Deardorff, J.W. (1977) Subgrid-scale condensation in models of non-precipitating clouds. *Journal of the Atmospheric Sciences*, 34, 344–355.

Stauffer, D.R. & Seaman, N.L. (1990) Use of four-dimensional data assimilation in a limited-area mesoscale model. Part i: experiments with synoptic-scale data. *Monthly Weather Review*, 118(6), 1250–1277. https://doi.org/10.1175/1520-0493(1990)118<1250:UOFDDA>2.0.CO;2

Su, C.-H., Eizenberg, N., Steinle, P., Jakob, D., Fox-Hughes, P., White, C.J. et al. (2019) Barra v1.0: the bureau of meteorology atmospheric high-resolution regional reanalysis for Australia. *Geoscientific Model Development*, 12(5), 2049–2068. https://doi.org/10.5194/gmd-12-2049-2019

Tiedtke, M. (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, 117, 1179–1799.

Trenberth, K., Koike, T. & Onogi, K. (2008) Progress and prospects for reanalysis for weather and climate. *EOS. Transactions of the American Geophysical Union*, 89, 234–235. https://doi.org/10.1029/2008EO260002

Theis, S.E., Hense, A. & Damrath, U. (2005) Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteorological Applications*, 12(3), 257–268.

Verhoef, A., van den Hurk, B.J., Jacobs, A.F. & Heusinkveld, B.G. (1996) Thermal soil properties for vineyard (efeda-i) and

savanna (hapex-Sahel) sites. *Agricultural and Forest Meteorology*, 78(1), 1–18. https://doi.org/10.1016/0168-1923(95)02254-6

Wahl, S., Bollmeyer, C., Crewell, S., Figura, C., Friederichs, P., Hense, A. et al. (2017) A novel convective-scale regional reanalysis cosmo-rea2: improving the representation of precipitation. *Meteorologische Zeitschrift*, 26(4), 345–361. https://doi.org/10.1127/metz/2017/0824

Warner, T.T., Peterson, R.A. & Treadon, R.E. (1997) A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bulletin of the American Meteorological Society*, 78(11), 2599–2618. https://doi.org/10.1175/1520-0477(1997)078<2599:ATOLBC>2.0.CO;2

Weisman, M.L., Davis, C., Wang, W., Manning, K.W. & Klemp, J.B. (2008) Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Weather Forecasting*, 23, 407–437. https://doi.org/10.1175/2007WAF2007005.1

Weusthoff, T., Ament, F., Arpagaus, M. & Rotach, M.W. (2010) Assessing the benefits of convection-permitting models by neighborhood verification: examples from MAP D-PHASE. *Monthly Weather Review*, 138, 3418–3433. https://doi.org/10.1175/2010MWR3380.1

Weygandt, S.S., Loughe A.F., Benjamin S.G., & Mahoney J.L. (2004) Scale sensitivities in model precipitation skill scores during IHOP. 22nd Conference Severe Local Storms. American Meteorological Society: Hyannis, MA, 4–8 October 2004.

White, B., Buchanan, A., Birch, C., Stier, P. & Pearson, K. (2018) Quantifying the effects of horizontal grid length and parameterized convection on the degree of convective organization using a metric of the potential for convective interaction. *Journal of Atmospheric Sciences*, 75, 425–450. https://doi.org/10.1175/JAS-D-16-0307.1

Wicker, L. & Skamarock, W. (2002) Time–splitting methods for elastic models using forward time schemes. *Monthly Weather Review*, 130, 2088–2097.

Wilks, D.S. (2019) *Statistical methods in the atmospheric sciences*, 4th edition. Oxford: Elsevier Science.

Woodhams, B.J., Birch, C.E., Marsham, J.H., Bain, C.L., Roberts, N. M. & Boyd, D.F.A. (2018) What is the added value of a convection-permitting model for forecasting extreme rainfall over tropical East Africa? *Monthly Weather Review*, 146, 2757–2780. https://doi.org/10.1175/MWR-D-17-0396.1

Yates, E., Anquetin, S., Ducrocq, V., Creutin, J.-D., Ricard, D. & Chancibault, K. (2006) Point and areal validation of forecast precipitation fields. *Meteorological Applications*, 13, 1–20.

Zhang, Q., Pan, Y., Wang, S., Xu, J. & Tang, J. (2017) High-resolution regional reanalysis in China: Evalua- tion of 1 year period experiments. *Journal of Geophysical Research: Atmospheres*, 122(20), 10801–10819. https://doi.org/10.1002/2017JD027476

How to cite this article: Cerenzia, I. M. L., Giordani, A., Paccagnella, T., & Montani, A. (2022). Towards a convection-permitting regional reanalysis over the Italian domain. *Meteorological Applications*, *29*(5), e2092. https://doi.org/10.1002/met.2092