Spin-up time and internal variability analysis for overlapping time slices in a regional climate model

A. Lavin-Gullon[1] · J. Milovac[1] · M. García-Díez[2] · J. Fernández[1]

Abstract

Long-term regional climate simulations are computationally very costly. One way to improve their computational efficiency is to split them into overlapping time slices, which can then be run in parallel. Although this procedure reduces the cost, sufficient spin-up must be left at the start of each slice. In any case, discontinuities will occur due to internal variability where two different slices join. In this study, we explore the relative role of spin-up time and internal variability in the discontinuities of overlapping time slice simulations and their effect on the simulated climate. This analysis has implications also for non-overlapping time slices, commonly used in very high resolution climate modelling, where long transient simulations cannot be afforded. We show that discontinuities are negligible for surface and upper-air variables, but they are noticeable in variables with long response times, such as soil moisture or snow depth. For these variables, differences between the slices are mainly attributed to internal variability, but also to insufficient spin-up time, depending on the region. In general, the results show that the overlapping time slice approach is valid to accomplish long term regional climate simulations.

Keywords Regional climate model · Time slice · Spin-up time · Internal variability · CORDEX

1 Introduction

Climate model resolution has always increased hand in hand with the available computer power. As an example, 30 years ago, the computational demand of the first regional climate models (RCMs; Giorgi 2019) limited their use to 60 km grid spacing for a month-long simulation (Dickinson et al. 1989; Giorgi and Bates 1989). Currently, centennial RCM simulations at ca. 10 km grid spacing are routinely carried out at different research centers (Jacob et al. 2020). Still, the experiments with the highest spatial resolution (currently

at kilometer-scale grid spacing) can only be afforded for time slices of about a decade (Coppola et al. 2020; Pichelli et al. 2021). This approach considers a decadal simulation driven by a future scenario and a reference decade driven by historical conditions to explore changes in climate. Such practice limits the climate analyses to periods well below the minimal climate standard of 30 years (WMO 2017). In the near future, centennial simulations of a kilometer-scale RCM will be feasible, especially if the RCM community adopts the latest advances in computing (Leutwyler et al. 2016). However, time slices will still be required for the ever-increasing model resolution, complexity and coupling with other demanding model components. In this work, we consider the use of a set of overlapping time slices to accomplish multi-decadal RCM simulations and we explore the effects of this approach on the simulated climate.

The idea of splitting a climate simulation into pieces is nearly as old as regional climate modelling (Pan et al. 1999). There are different reasons for doing so, though. A common reason to re-initialize a climate simulation is to keep it close to the observations. For this purpose, a frequent re-initialization is advocated (Pan et al. 1999; Qian et al. 2003; Lo et al. 2008). The frequent cold-start from reanalysis initial conditions constrains the weather trajectory of the model to

✉ J. Fernández
jesus.fernandez@unican.es

A. Lavin-Gullon
alvaro.lavin@unican.es

J. Milovac
milovacj@unican.es

M. García-Díez
garciam@predictia.es

1 Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain

2 Predictia Intelligent Data Solutions, Santander, Spain

be close to the observed one. This approach can introduce discontinuities in the weather events and, more importantly for climate analyses, may disrupt the proper evolution of variables with long response times, such as deep soil variables. To prevent this disruption, the so-called poor man's reanalysis approach (Berg and Christensen 2008; Stahl et al. 2011; Lucas-Picher et al. 2013) keeps the soil variables across the different re-initializations, and updates only the atmospheric initial conditions from reanalysis data. In this latter approach, the simulation pieces are not independent of each other and there is no computational advantage in the re-initialization.

The computational advantage is a second reason to split an RCM simulation into pieces, which can then be run in parallel (Jimenez et al. 2010; Menendez et al. 2014). This form of parallelism can be more efficient than standard high-performance parallel computing paradigms such as Open Multi-Processing (OpenMP) or Message Passing Interface (MPI) (Jerez et al. 2009). Using these standard parallel computing approaches, computing time typically scales well with the number of processors up to a limit. This limit is usually much lower than the number of processors available. Even for a reasonable scaling, there is always a loss in using an increasing number of processors. Therefore, computing time is used more efficiently when splitting the simulation and running the pieces on a smaller amount of processors. For their use in climate studies, the initial part of each simulation piece must be disregarded as model spin-up. This spin-up period is, typically, at least one year (Christensen 1999), although a few months might suffice depending on the season when the pieces are initialized (Jerez et al. 2020). This computing time trade-off between the gain by a more efficient use of the processors and the waste due to spin up of each simulation piece, can be used to optimize the length of the pieces (Jerez et al. 2009).

RCM simulation splitting is hardly avoidable for very long simulations, such as those for the last millennium (Gómez-Navarro et al. 2011). This procedure can also alleviate the computational burden for research groups to perform centennial climate change RCM simulations. As an example, in this work, we analyze RCM simulations carried out in the last decade at Universidad de Cantabria (UCAN) as split runs and also as continuous simulation (Sect. 2.1). A form of simulation splitting is also used for the most computationally demanding RCM simulations (Coppola et al. 2020; Pichelli et al. 2021), where only a couple of decadal time slices can be afforded. Time slicing is just a simulation split into pieces, but with pieces that do not usually overlap. Here, we analyse the effect that overlapping a set of time slices would have on the simulated regional climatology. Note that this analysis cannot be carried out on standard time slice experiments, where the slices do not overlap and the proper initialization of the simulation cannot be assessed.

Due to internal variability in the RCM (Christensen et al. 2001; Lucas-Picher et al. 2008), a perfect match of the weather trajectory in two consecutive time slices is not possible. Internal variability is unavoidable and is triggered by the different initial conditions in the time slices, so there will always be a 'weather jump' at the joints. On top of the internal variability, the coarse initial conditions from the driving GCM or reanalysis need some time (spin-up time) to be assimilated by the RCM dynamics. This spin-up time depends on the variable. It is quite short for atmospheric variables, but it can extend for several months or even years for other slow-varying variables (Christensen 1999; Cosgrove et al. 2003; Jerez et al. 2020).

The objective of this study is to show the effect of the overlapping time slice approach on the regional climate simulated by an RCM. For this purpose, we used state-of-the-art simulations carried out under the COordinated Regional climate Downscaling EXperiment (CORDEX; Giorgi and Gutowski 2015) framework using the Weather Research & Forecasting (WRF) modelling system (Sect. 2.1). Analyses were carried out for variables with different response time, using split and continuous simulations for different regions, at different resolutions and with time slices in split simulations initialized in different seasons. We studied the weather jumps in split simulations, locating their occurrence (Sect. 3.1), analysing their evolution in time (Sect. 3.2) and their geographical location (Sect. 3.3). Finally, we analysed the potential effect of splitting the simulations into time slices on the simulated climate (Sect. 3.4).

2 Data and methods

2.1 Data

In this study we analyse three sets of simulations using the WRF model (Skamarock et al. 2008), with different parameterization schemes, spatial domains, horizontal resolutions and time periods. Simulations were performed over three model domains as defined within CORDEX. One set of simulations was carried out over the CORDEX South American domain at the standard 0.44° horizontal resolution (SAM-44), regular on a rotated latitude-longitude projection (Falco et al. 2019; Solman and Blázquez 2019). These simulations were carried out for the historical period 1951–2005 and for the future scenarios RCP 4.5 and RCP 8.5 for the period 2002–2100, all driven by the Canadian Earth System Model (CanESM2; Arora et al. 2011). The other two sets consisted of evaluation simulations over Europe at 0.44° (EUR-44) and ~15 km (EUR-15) horizontal grid spacing, driven by the ERA-Interim reanalysis (Dee et al. 2011). EUR-44 simulations span the period 1979–2010 (Vautard et al. 2013), while the EUR-15 simulations were generated under the CORDEX

Flagship Pilot Study on convective phenomena at high resolution over Europe and the Mediterranean (FPS-CONV), covering the period 1999–2009 (Coppola et al. 2020). An additional convection-permitting (CP) simulation at ~3 km horizontal resolution, is also analysed to evaluate the effect of the resolution. This CP simulation was nested into the EUR-15 domain and centered on the Alpine region (ALP-3). The model setup is the same as EUR-15, except for the cumulus parameterization, which was deactivated (Ban et al. 2021).

Two different model configurations were used in terms of physical parameterizations. Namely, EURO-CORDEX WRF configuration WRF341I (Manzanas et al. 2018) was used in EUR-44 and SAM-44, while an updated model version and configuration, WRF381BI (Ban et al. 2021), was used in EUR-15. The most important difference concerning our results is the different land surface model (LSM). WRF381BI used the new Noah with multi-parameterization options (Noah-MP v1.1) LSM (Niu et al. 2011), while WRF341I was coupled to its predecessor Noah LSM (Chen and Dudhia 2001). All simulations were carried out with prescribed sea surface temperature and sea ice, evolving as provided by the driving global model or reanalysis.

The three sets of simulations were performed by splitting the runs into several time slices. In order to allow for the required spin-up time, adjacent time slices were overlapped for at least one year (Fig. 1). Additionally, the EUR-15 and EUR-44 simulations were also performed continuously. The continuous runs were produced in two different ways. For EUR-44, the first time slice (S1) was extended to cover the full period. For EUR-15, the full period was simulated again, but the model was compiled with a different compiler version and linked to a different version of the parallel computing (OpenMPI) libraries. Therefore, the differences between the EUR-15 continuous run and the first time slice of the split simulation (EUR-15 S1) will be due to different numerical round-off in the model executable. These differences are expected to grow and evolve with the flow as those caused by perturbing initial conditions, i.e. internal variability (Geyer et al. 2021). The continuous runs are considered as a reference to investigate possible inhomogeneities caused by the time slicing of the split simulations.

It is worth noting that these sets of simulations were not specifically designed for this study. We use them as an ensemble of opportunity to study the spin-up length and the role of internal variability in the climate simulated by overlapping time slices. As such, we can only explore the variables available for each simulation, and the initialization seasons of the slices used in these multi-year simulations. A designed, systematic exploration of the required spin-up time has been recently carried out by Jerez et al. (2020) over Europe for a 1-year test period. Our approach extends this work by considering domains in different climates, different spatial resolutions, longer spin-up lengths, and the role of interannual and internal variability.

We focus our analyses on three types of variables. Initially, we consider slow-varying variables as their accurate initialization and representation are of key importance for weather and climate modelling. They require significant spin-up times as their initial conditions, taken from the driving model, usually differ greatly from the conditions generated by the RCM (Jerez et al. 2020). For this purpose, among the available variables, here we analyze total soil moisture and snow depth. These variables control energy partitioning at the land surface and,

Fig. 1 Schematic representation of the analysed simulations. Time slice simulations (S1, S2, etc.) are used to compose split simulations for each domain (switching time between slices is indicated by arrows). Continuous simulations are represented either as independent simulations (EUR-15) or by extending the S1 slice (EUR-44). Overlapping periods are shaded in grey shades: light grey for the overlap of two simulations, and dark grey for three overlapping simulations
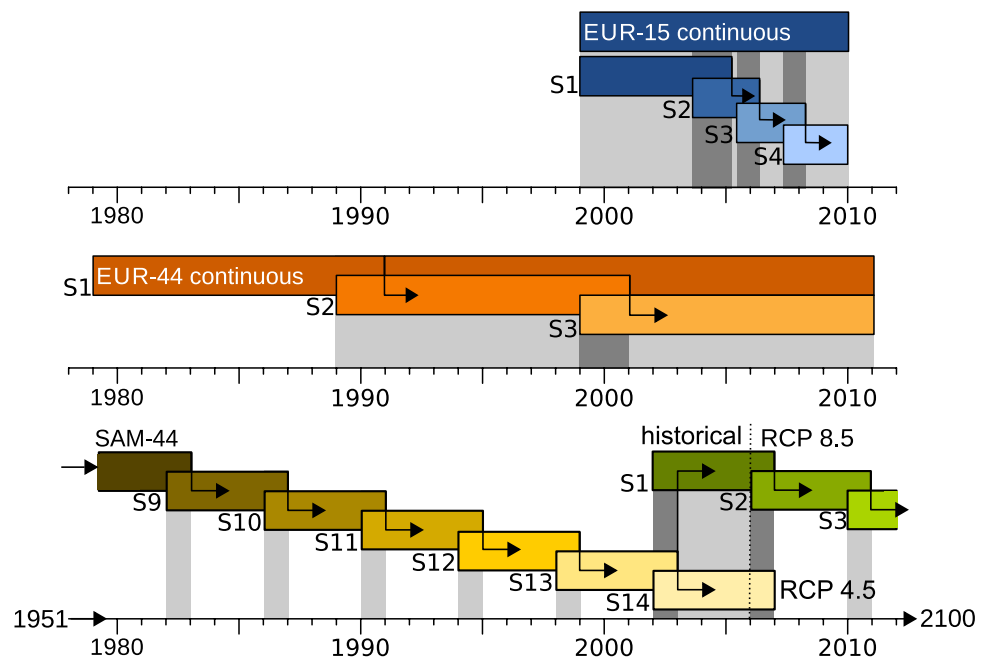
through land-atmosphere feedbacks, they influence the evolution of the atmospheric conditions in the planetary boundary layer. In such a way, these variables impact local weather and regional climate (Seneviratne et al. 2006). Furthermore, we considered near-surface temperature and precipitation, two fundamental variables that characterize the regional climate and are often considered in climate and impact studies. Both are highly variable in time and space. The atmospheric circulation was analysed as well, since it has also a strong but non-local impact on the regional climate (Zappa 2019). Atmospheric circulation shows the shortest response time, as compared to surface or subsurface fields. We characterize the circulation by means of the geopotential height at 850 hPa. The analysis was carried out for daily mean values for all variables.

2.2 Methodology

The analysis of discrepancies across the simulation slices is based on simple differences. For a given variable $X(s, n, t)$, taken from a simulation slice s at grid point n and time step t, we define the following differences:

$$D_t X(s, n, t) = X(s, n, t) - X(s, n, t - 1) \tag{1}$$

$$D_s X(s, n, t) = X(s, n, t) - X(s - 1, n, t) \tag{2}$$

$$D_{st} X(s, n, t) = X(s, n, t) - X(s - 1, n, t - 1) \tag{3}$$

for consecutive time steps $t - 1$ and t, and consecutive simulation slices $s - 1$ and s. Note that the meteorological jump (D_{st}) in variable X, occurring at the joint between time $t - 1$ in simulation slice $s - 1$ and time t in slice s, can be decomposed as the difference between slices D_s at time $t - 1$ plus the variable tendency in time (D_t) within slice s:

$$\begin{aligned} D_{st} X(s, n, t) &= D_s X(s, n, t) + D_t X(s - 1, n, t) \\ &= D_s X(s, n, t - 1) + D_t X(s, n, t), \end{aligned} \tag{4}$$

In order to have a relative measure, we consider non-dimensional differences in terms of standard deviation units, by dividing each of them by the standard deviation in time at each grid point:

$$d_{st} X(s, n, t) = \frac{D_{st} X(s, n, t)}{sdD_t X(s, n)} \tag{5}$$

where the standard deviation $sdD_t X(s, n)$ is calculated as

$$sdD_t X(s, n) = \sqrt{\frac{1}{T(\tau) - 1} \sum_{t \in \tau} \left[D_t X(s, n, t) - \overline{D_t X}^\tau (s, n) \right]^2} \tag{6}$$

from a 45-day time period (τ) prior to the target time t. The overline represents time average over a given time period τ:

$$\overline{X}^\tau (s, n) = \frac{1}{T(\tau)} \sum_{t \in \tau} X(s, n, t) \tag{7}$$

with $T(\tau)$ representing the number of time steps in τ. This is done to use intra-seasonal variability as reference, thus preventing the annual cycle variability to mask large differences for a given season. These time period averages have also been used to assess the long-term impact of time slicing on the climatology of a given variable (Sect. 3.4).

Differences (D) are spatio-temporal fields. We summarize them by means of spatial root mean squared differences (RMSD), to avoid compensation of opposite differences across the domain. For each slice, the *intra-slice daily tendency* ($RMSD_t X$) summarizes the differences between consecutive time steps:

$$RMSD_t X(s, t) = \sqrt{\frac{1}{N} \sum_{n=1}^N D_t X(s, n, t)^2} \tag{8}$$

Inter-slice differences ($RMSD_s X$) are employed to quantify the differences between two slices along the overlapped period (see Sect. 3.3):

$$RMSD_s X(s, t) = \sqrt{\frac{1}{N} \sum_{n=1}^N D_s X(s, n, t)^2} \tag{9}$$

We also consider the quadratic average of $D_{st} X$ ($RMSD_{st} X$), which arises in the context of split simulations; it is the $RMSD_t X$ at the joint of time slices. This measure quantifies the inhomogeneity introduced at the joint for different variables (see Sect. 3.1):

$$RMSD_{st} X(s, t) = \sqrt{\frac{1}{N} \sum_{n=1}^N D_{st} X(s, n, t)^2} \tag{10}$$

Finally, as a reference for natural variability we also estimate transient eddy variability (Caya and Biner 2004; Lucas-Picher et al. 2008; Lavin-Gullon et al. 2021):

$$TEV(s, \tau) = \sqrt{\frac{1}{N} \sum_{n=1}^N \overline{\left[X(s, n, t) - \overline{X}^\tau (s, n) \right]^2}^\tau} \tag{11}$$

where τ represents in this case all days in a calendar month, in order to have a monthly TEV estimate.

3 Results

3.1 Detection of meteorological inhomogeneities

Unlike continuous regional climate simulations, split simulations contain meteorological inhomogeneities, i.e.

unphysical changes in the state of the system at the joints of the time slices. This is unavoidable, given that an exact match of two climate simulations is impossible due to the chaotic nature of the climate system. For RCMs, the constraint exerted by the lateral boundary conditions make the inhomogeneities much smaller than in global models. Still, substantial internal variability develops in RCMs (Lucas-Picher et al. 2008; Bassett et al. 2020; Lavin-Gullon et al. 2021), preventing a smooth transition between simulation slices.

The ability to detect these meteorological inhomogeneities depends on the simulation sampling frequency (i.e. output frequency) used. The inhomogeneity will pass unnoticed if it is smaller than the change between consecutive output times. And this change is larger as sampling frequency decreases. Intra-slice daily tendencies ($RMSD_t$) quantify the changes between consecutive time steps. At the slice joints in a split simulation, $RMSD_t$ becomes $RMSD_{st}$ and quantifies the size of the inhomogeneity along with the variable tendency.

Of course, the relevance of the inhomogeneity depends also on the variable. For geopotential height (Fig. 2), inhomogeneities go unnoticed. Average daily geopotential tendencies in mid-latitudes (EUR-15, EUR-44) range between 20 and 100 *m*, with a prominent annual cycle. The geopotential change remains within this range when passing from one time slice on one day to the next time slice on the next day (indicated by arrows in Fig. 2), regardless of the season when joining occurs. The SAM-44 domain spans mid-latitude as well as tropical regions and, thus, geopotential heights show a smaller range of change (20–80 *m*) and a much weaker seasonal cycle. However, daily inhomogeneities go unnoticed. Geopotential height is strongly driven by the lateral boundary conditions and the pass of weather systems through the domain. Moreover, the 1-year spin-up period considered in the time slices is expected to be long enough for upper atmospheric variables, such as geopotential height, to reach physical equilibrium within the model. Therefore, these variables do not suffer from noticeable inhomogenities.

The same result applies for variables that are influenced to a greater or lesser extent by the lateral boundary forcing, such as near-surface temperature, precipitation or snow depth (not shown). For snow depth, which varies slowly, regional inhomogeneities are apparent, but the domain-wide summary in $RMSD_t$ masks the differences in the relatively small snow-covered regions in the domains considered. Soil variables (e.g. total soil moisture in Fig. 3), however, show large discontinuities at the slice joints. $RMSD_t$ shows inhomogeneities in all three sets of simulations, clearly unveiling the time when two slices join. Daily tendencies in total soil moisture range between 2 and 10 kg/m^2, except for peaks on certain days with values beyond 30 kg/m^2, corresponding to the joints of the time slices. The order of magnitude of these peaks is not sensitive to the season in which the adjacent slices join (winter in EUR-44, summer in SAM-44, winter
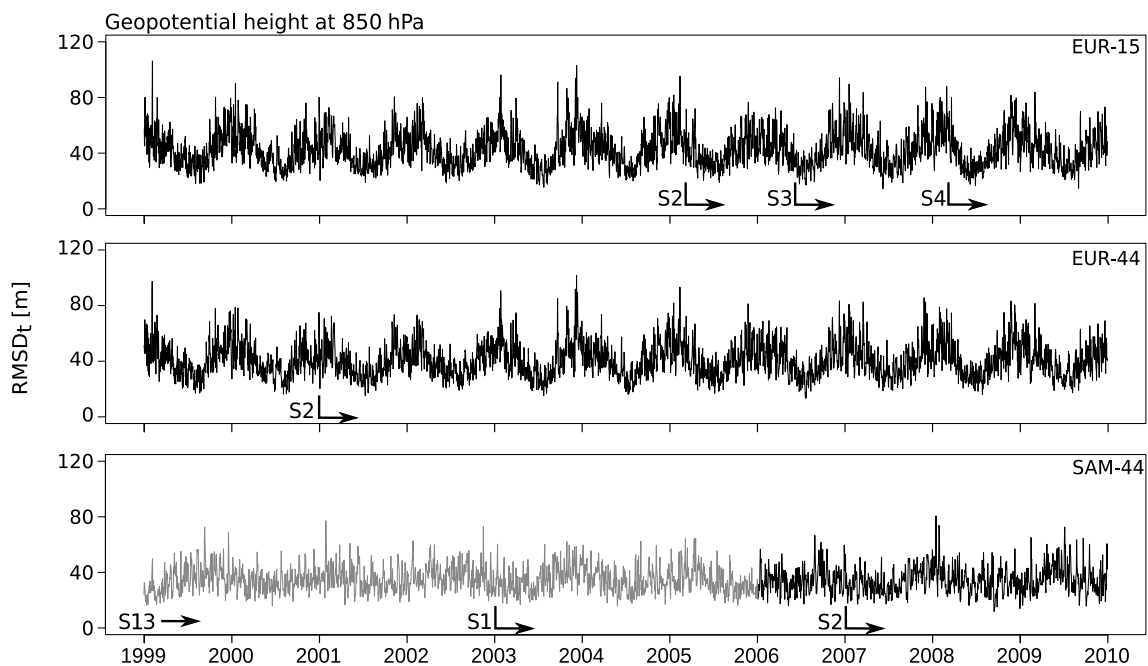

Fig. 2 RMSD between consecutive days ($RMSD_t$) for geopotential height at 850 hPa (*m*) in the EUR-15, EUR-44 and SAM-44 split simulations. Joints of time slices are indicated by arrows. In SAM-44, light gray shading refers to the historical run while dark gray shading corresponds to RCP 8.5 forced run
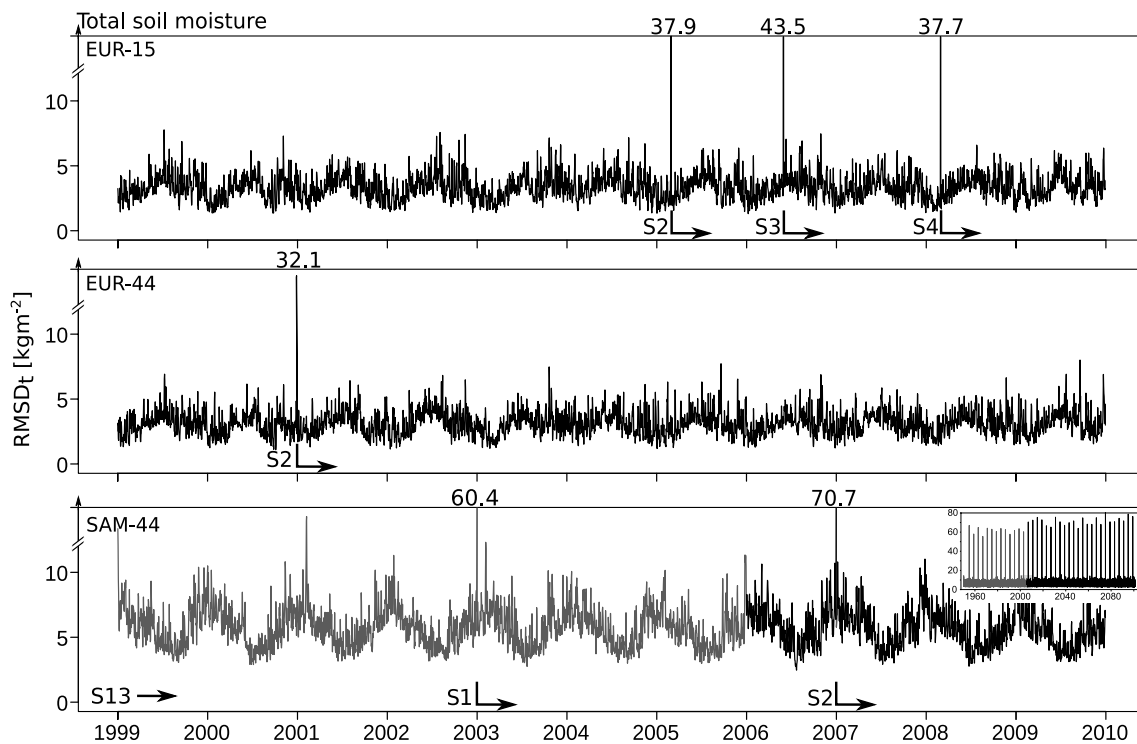
Fig. 3 As Fig. 2, but for the total soil moisture content (kg/m^2). The inset in the lower panel shows the whole 1950-2100 SAM-44 historical (grey) plus RCP 8.5 scenario period. Numbers at the top of the each panel represent maximum values (out of scale) at the joints of the slices

and spring in EUR-15). These peaks also show low interannual variability, clearly standing out from the background variable tendency for every joint in different years (see inset in Fig. 3). Despite the strong signal in $RMSD_t$, these peaks in the differences are still one order of magnitude smaller than the variable itself; e.g. the quadratic mean of total soil moisture in SAM-44 is about 600 kg/m^2.

These strong discontinuities at the joints indicate that there are high discrepancies in soil moisture between two time slices. This could be due to two reasons: (1) the spin-up period considered is not long enough for soil moisture to balance within the model or/and (2) soil moisture internal variability is larger than the daily tendency. This is investigated next.

3.2 Meteorological inhomogeneities in time

In a split simulation, a set of time-sliced simulations are concatenated after removing an initial spin-up period. In principle, the longer the spin-up period, the better. A given slice enters the split simulation just after the previous slice has finished (as depicted in Fig. 1). However, the switch between the slices can be chosen to occur at any time during the overlapping period (grey-shaded areas in Fig. 1). We can quantify the size of the discontinuity by means of inter-slice differences ($RMSD_s$, Eq. 9). As an example, Fig. 4

shows $RMSD_s$ for the 1.5-year overlap between S1 and S2 time slices for EUR-15 domain, covering the period from Sep., 2003 to Feb., 2005. The EUR-15 continuous simulation is also included (black line in the Fig. 4) as reference for $RMSD_s$. Figure 5 shows another example for a 1-year overlap period between S1 and S2 for the SAM-44 domain, covering the complete year 2006. In this case, no reference continuous run was available, but we included TEV as a reference for natural variability (black line in the Fig. 5).

3.2.1 Total soil moisture

Total soil moisture intra-slice daily tendencies (Fig. 4a, top panel) differ between the two time slices at the beginning of the overlapping period. At this time, total soil moisture in S2 is mainly provided by the initial conditions from ERA-Interim, while in S1 the soil state is generated by the RCM itself. In EUR-15, the overlap period is initiated in late summer (September) so that the difference in the daily tendencies between S1 and S2 decreases rapidly. Day-to-day variability is mainly determined by the variability in the top soil layer, which in turn depends on precipitation. Therefore, there is evident correlation between day-to-day variability of total soil moisture and precipitation (see Fig. 4a and d, top panels). Thus, under drier conditions like it is the case in late summer, daily soil moisture tendencies from the driving
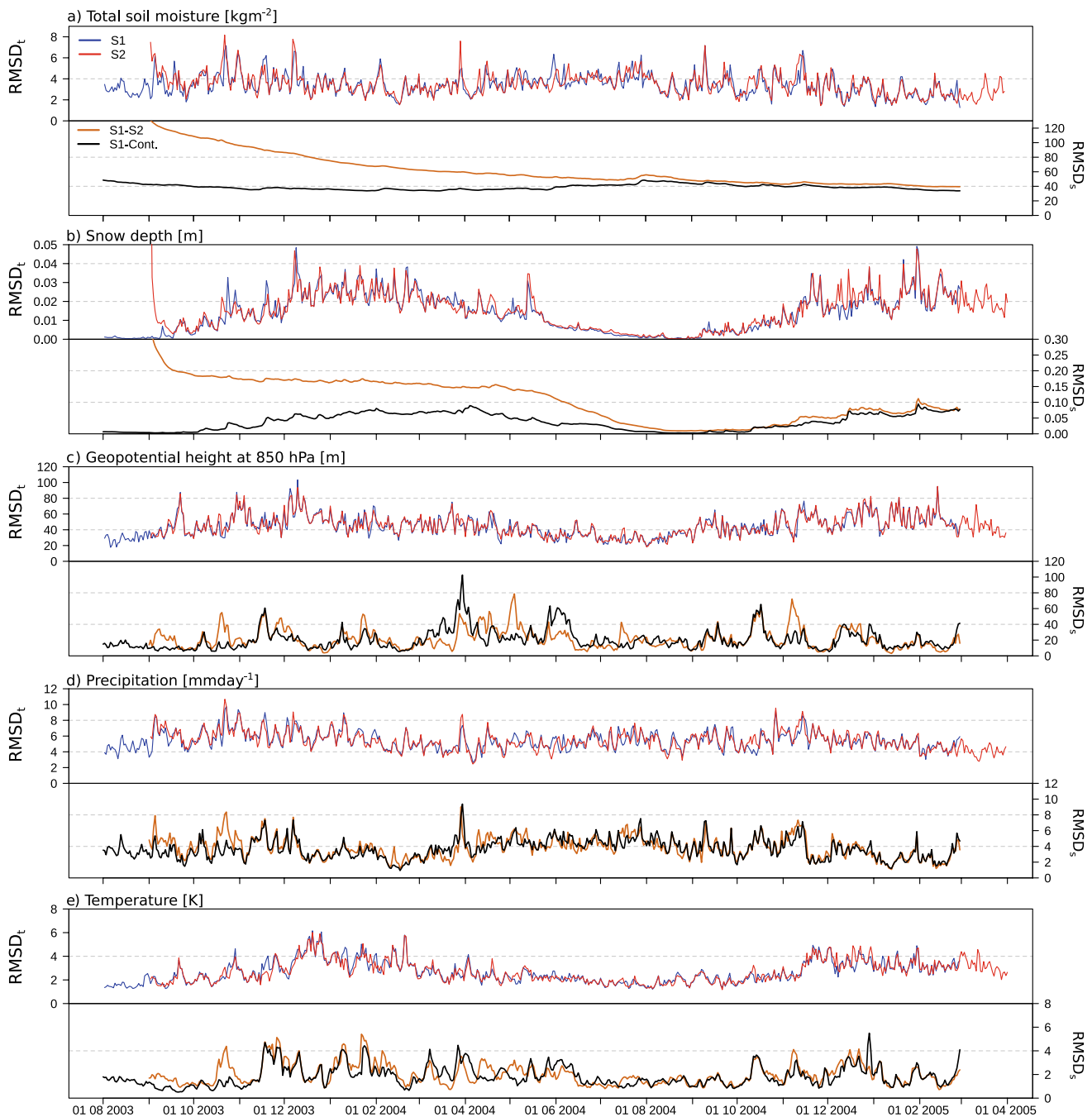
Fig. 4 RMSD for **a** total soil moisture, in kg/m^2, **b** snow depth, in m, **c** near-surface temperature, in K, **d** daily accumulated precipitation, in mm, and **e** geopotential height at 850 hPa, in m, for EUR-15 S1, S2 and continuous simulations; see Fig. 1. For each variable, intra-slice daily tendencies ($RMSD_tX$) for each simulation (top panel) and inter-slice differences ($RMSD_sX$) between S1 and S2, and between S1 and continuous (bottom panel)

model adapt faster to the soil moisture conditions in the long term RCM run (i.e. slice S1 here).

In SAM-44, the overlapped period starts in austral summer, which in this region corresponds to the rainy season. In this case, intra-slice daily tendencies for soil moisture (Fig. 5a, top panel) show that more than 3 months are

necessary for the two adjacent slices, S1 and S2, to start evolving coherently in time. It is interesting to notice that the two slices tend to diverge again at the end of the overlapping period, which also corresponds to the austral summer. This can be associated to the discrepancies in summertime precipitation between the two slices (see Fig. 5d top panel).
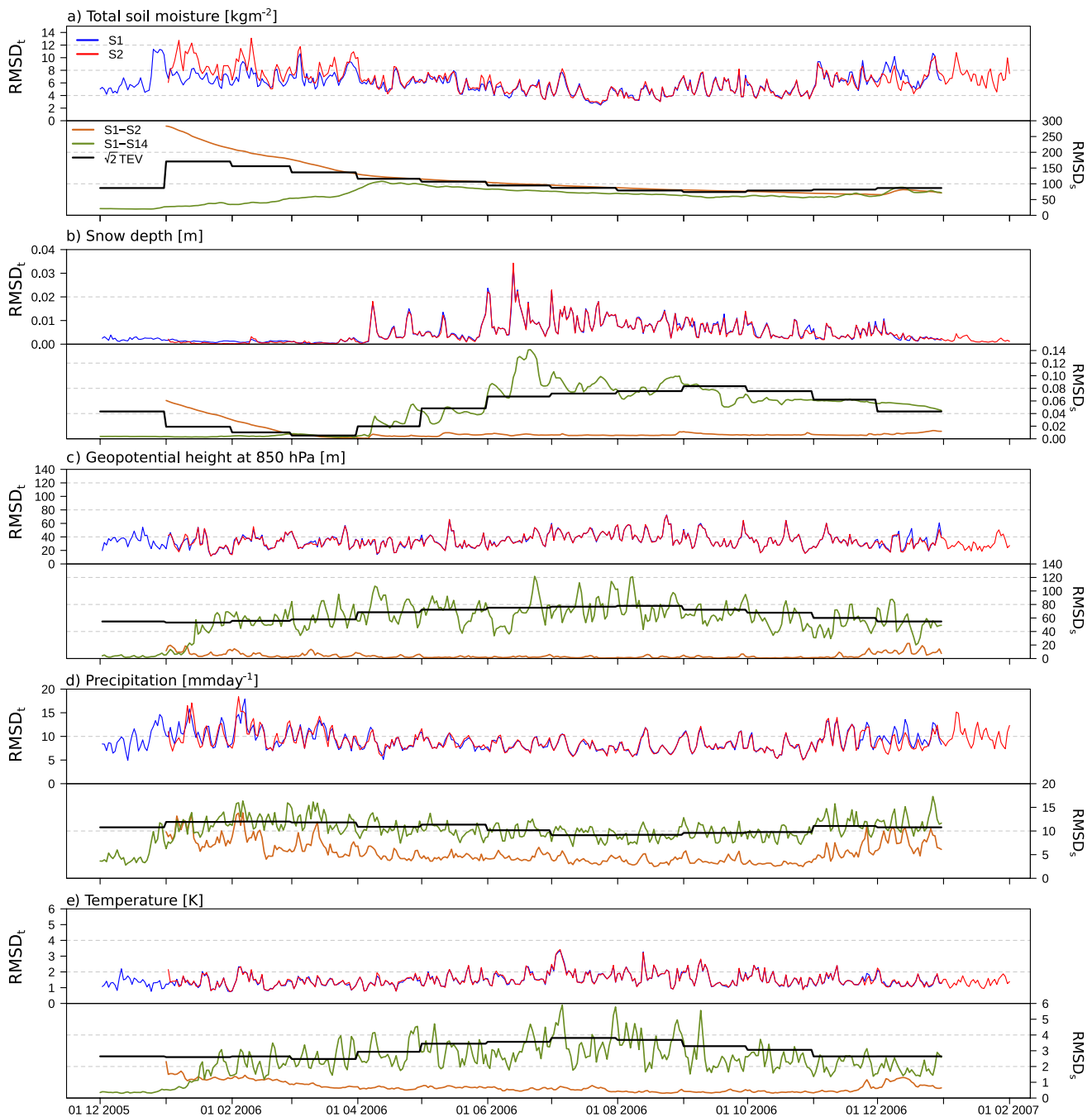
Fig. 5 As Fig. 4, but for the SAM-44 domain, and overlapping time slices S1 and S2

Daily soil moisture inter-slice differences ($RMSD_s$, bottom panels of Figs. 4a and 5a) are more controlled by the moisture state of deeper soil layers, therefore we can consider these differences to be strongly influenced by soil moisture spin-up . These differences will reach a minimum and never a zero value due to internal variability (Lavin-Gullon et al. 2021). When $RMSD_s$ is stabilized, this denotes that internal variability had overcome the initial spin-up transient. To distinguish internal variability from insufficient spin-up, we can also use the continuous run as the reference when available (i.e. EUR-15 domain). In EUR-15 case, we computed the inter-slice difference between S1 (initialized 5 years before the overlapping period, see Fig. 1) and the continuous simulation (black line in Fig. 4a). These differences are controlled by internal variability only, and they set the lower limit for inter-slice differences. $RMSD_s$ between S1 and S2 during the overlapped period are initially very high in both domains (bottom panels of Figs. 4a and 5a). In

EUR-15, $RMSD_s$ stabilizes at 40 kg/m^2 after about one year, while in SAM-44 the minimum value of 70 kg/m^2 is reached after just a few months. At these points internal variability becomes the major cause of the differences between S1 and S2, which leads to a conclusion that spin-up in SAM-44 is shorter than in EUR-15.

This highlights that not only the season determines the spin-up time, but also the synoptic regimes specific for the region. In SAM-44, the overlapped period starts in the austral summer (DJF) but, unlike in EUR-15, summer is the wet season in Central South America (Liebmann and Mechoso 2011), which largely contributes to the annual precipitation (and soil moisture) in the whole SAM-44 domain. Shorter spin-up in moist regimes (SAM-44) than in drier regimes (EUR-15) may be related to the parameterization of vertical water transport within the soil. It is based on Richards' equation in all LSMs used in the analysed simulation sets, which is more efficient under moist conditions in the wet season (Khodayar et al. 2015).

It is worth comparing the scales of inter- and intra-slice differences for soil moisture, which differ by about one order of magnitude. This scale difference causes the $RMSD_t$ peaks in split simulations denoting the soil moisture inhomogeneities at the slice joints shown in the previous section. In particular, the first peak in Fig. 3 (top panel) has a contribution from both $RMSD_t(S2)$ and $RMSD_s(S1 - S2)$ lines (Fig. 4a) at the end of the overlapping period, when slice S1 switches to S2 in the EUR-15 split simulation. This shows that an earlier switch from S1 to S2 would have not led to smaller inhomogeneities.

3.2.2 Snow depth

Snow depth shows a different behavior than soil moisture in EUR-15 (Fig. 4b, top panel). Initial $RMSD_t$ values in S2 are significantly higher than those in S1, but they get close to each other in just a few days. In SAM-44 (Fig. 5b, top panel), this initial difference is not evident as the snow depth is insignificant there during the austral summer, when the overlapped period starts. After balancing, $RMSD_t$ evolves coherently in time for both slices in both domains, with an evident seasonal cycle. The variability of day-to-day $RMSD_t$ is higher in winter and early spring, when snow depth changes due to snowfalls and snow melting, and lower values occur in summer when snow coverage is small and limited to areas with permanent snow. Inter-slice differences during the overlapped period in EUR-15 (bottom panel in Fig. 4b) show an interesting behavior. After just a few days, $RMSD_s$ for S1 and S2 slices reaches a winter minimum that remains almost constant until the end of April 2004. After that time, it starts to drop again. In summer, since the snow coverage is minimal, inter-slice differences reach the overall

minimum. Afterwards, as the snow depth starts to increase, $RMSD_s$ increases again following the seasonality pattern of internal variability. However, $RMSD_s$ never reaches the first winter minimum again. This leads to a conclusion that the occurrence of the initial drop to the first winter minimum does not mean the balancing of the snow. When the $RMSD_s$ values drop to the overall minimum, which is limited by the inter-slice differences between slice S1 and the continuous run (black line on the bottom panel in Fig. 4b), snow depth spin-up can be considered as completed. These results show that, for this specific simulation, joining the slices at the end of summer would minimize the inhomogeneity of the snow depth, while not affecting significantly the inhomogeneity of soil moisture at the S1–S2 joint. Moreover, this is valid also for shorter overlapped periods initialized in June (for EUR-15 for S2–S3 slices) or in March (for EUR-15 for S3–S4 slices), which can be seen in the Supplementary Material in Figures ESM1 and ESM2, respectively. Minimal snow depth inhomogeneities across time slices are obtained by switching slices at the end of late summer (September) in all cases, although for the S2–S3 slices when the overlapped period starts in June (just 3 months before) 1 year long overlap period is insufficient for soil moisture to spin up. In SAM-44, minimum values also appear in austral summer for $RMSD_t$ (Fig. 5a) and in late summer for $RMSD_s$ (Fig. 5b), when it would be most convenient to join the slices. Although snow in SAM-44 is scarce, covering only some areas in the southern Andes, selecting austral summer to start the overlapping in SAM-44 is beneficial.

In EUR-44 (Figs. ESM 3a and b in the Supplementary Material) a 2-year overlap period is available. This allows for better assessment of the annual cycles and variability of $RMSD_t$, inter-slice differences ($RMSD_s$) between S2 and S3, and the internal variability, estimated by $RMSD_s$ between S2 and the continuous run. This time slice overlap (S2-S3) confirms all the results previously shown for EUR-15: (1) late summer is a good season to choose to switch slices for snow cover, (2) the seasonal cycle in soil moisture internal variability, peaks also in late summer, which is clearer here. However, for this particular year, soil moisture seems not to be completely spun up. This highlights the need for an interannual assessment of soil moisture spin-up times.

3.2.3 Other variables

The other variables (Fig. 4c–e) show no initial discrepancies as a hint of spin-up period, with an evolution of day-to-day changes ($RMSD_t$) similar for both slices, and inter-slice differences ($RMSD_s$) consistent with internal variability. The order of magnitude of both RMSDs is the same, therefore there is no apparent inhomogeneities at the joints (e.g. see Fig. 2).

Geopotential height (Fig. 4c top panel) exhibits seasonal cycle with larger day-to-day changes ($RMSD_t$) in winter than summer, which is typical mid-latitude variability in Europe (Caya and Biner 2004). On the other hand, this seasonality is not evident in SAM-44 (Fig. 5c top panel). The SAM-44 domain is not limited only to mid-latitudes, but it also covers large tropical areas showing no variable seasonality, which smooths the final results. Nevertheless, the larger summer internal variability is apparent in the $RMSD_s$. Inter-slice differences ($RMSD_s$) between S1 and S2 slices do not show strong deviations from internal variability (i.e. $RMSD_s$ between the continuous simulation and S1 in EUR-15), therefore most of the differences between time slices can be explained by the internal variability itself. For precipitation intra-slice daily tendencies evolve more coherently in SAM-44 than in EUR-15, especially in the austral winter. In both domains, differences in $RMSD_t$ between the two slices does not change along all the overlapped period, regardless of the season. On the other hand, a seasonal cycle appears in inter-slice differences, with higher (lower) differences between the slices in the austral summer (winter), following the seasonal pattern observed for geopotential height at 850 hPa. This seasonal behaviour is more apparent in SAM-44 than in EUR-15.

The results for near-surface temperature follow the similar pattern as those for the geopotential height at 850 hPa.

3.2.4 Horizontal resolution

We bilinearly remapped all EUR-15 variables to the ALP-3 domain (not shown). Apart from a slightly higher initial state of the soil moisture in ALP-3, all $RMSD_t$ and $RMSD_s$ time series during the overlap periods are virtually identical in the remapped EUR-15 and ALP-3 resolutions. Horizontal resolution does not seem to play any major role on the model spin-up and inhomogeneities of split simulations. This small sensitivity to the change in horizontal resolution can be seen also in Figs. 2 and 3, with very similar $RMSD_t$ evolution in both EUR-44 and EUR-15, regardless of the time slice considered and despite the differences in model version and configuration.

3.2.5 Another view on spin-up time

The SAM-44 simulation setup does not allow to estimate the internal variability limit, since the only year with two long-term overlapping simulations (S14 and S1) is 2006. In this year, the GCM boundary conditions driving S14 and S1 bifurcate into two different global climate realizations forced by the RCP 4.5 (S14) and 8.5 (S1 and S2) concentration scenarios. As a result, differences between S14 and S1 in 2006 do not represent RCM internal variability, but different global driving fields. In particular, the changes in

the forcing between the RCP 4.5 and 8.5 scenarios are so small in 2006 that the two global climate realizations can be considered as resulting from the GCM internal variability. At least, regarding the atmospheric fields fed to the RCM. In a few weeks, the slight forcing differences make the GCM circulation diverge and the synoptic situation of corresponding days in the two global model realizations will be as different as two random days in the corresponding season. Note that no inhomogeneity occurs, since both GCM realizations are started from the same final state of the historical run at the end of 2005. Synoptic conditions depart smoothly as slight changes in the forcing introduce small perturbations which are amplified by the model dynamics to become finite perturbations.

From the point of view of the RCM, soil variables will evolve smoothly, with the land surface model responding in a physically consistent manner to the new atmospheric conditions. Snow cover should also adjust smoothly to the new synoptic conditions fed through the boundaries. This adjustment process is similar to the spin-up, since the RCM internal fields need to adjust to the new driving fields. Unlike the spin-up process, model states are physically consistent during the whole process; no tendencies develop in the model to account for the mismatch between the initial conditions and a balanced model state. The expected $RMSD_s(S1 - S14)$ value after the adjustment is not the RCM internal variability limit in this case (since the driving fields differ), but the GCM internal variability. This can be estimated from the transient eddy variability (TEV, Eq. 11). In particular, for uncorrelated fields from two GCM realizations, $RMSD_s$ should reach $\sqrt{2}$TEV (Caya and Biner 2004).

This decorrelation time to reach the GCM internal variability level sets a minimum response time for the spin-up time. The RCM starts the adjustment process from an internally consistent state, unlike in the spin-up process, which needs to bring the initial state into line with the RCM dynamics. Therefore, spin-up times should be longer than the smooth adjustment time to decorrelated synoptic situations. This is illustrated in Fig. 5, where the monthly $\sqrt{2}$ TEV lines have been included as reference. As expected, surface and upper air variables adjust in a few weeks. In fact, geopotential height will adjust almost immediately, since the 2-week delay shown in Fig. 5e is likely the time taken by the GCM circulation to decorrelate. Soil moisture takes about 3 months to reach decorrelation ($\sqrt{2}$TEV line in Fig. 5a). Note that the Figure includes Dec-2005, which still represents RCM internal variability levels.

Snow depth takes longer to decorrelate, since the adjustment starts in austral summer, with no snow, and $RMSD_s$ keeps low until April (mid-autumn). Then, decorrelates relatively fast, growing along with the TEV line. This is different from the spin-up process, which efficiently uses the summer months to reset the snow cover fields.

Interestingly, for soil moisture, inter-slice differences between S2 and S1 stabilize at the $\sqrt{2}$TEV level. This means that soil moisture differences between time slices subject to the same boundary conditions are as different as those in two random days in this month (note that TEV was computed considering interannual variability). This may indicate a generally low departure from average conditions in this variable. It could also be result of insufficient spin-up. Since the low internal variability level ($RMSD_s(S1 - S14)$) from Dec-2005 is not reach by the end of the overlapping period, this suggest that the soil moisture spin-up for $RMSD_s(S1 - S2)$ may have not finished in Dec-2006, despite the apparent stabilization.

Precipitation also gets close to the decorrelation limit during the austral summer months. In this case, summer convective precipitation is weakly forced by the boundaries and precipitation centers are likely mislocated between slices, even if forced by the same boundary conditions.

3.3 Meteorological inhomogeneities: spatial distribution

In the previous sections we quantified day-to-day changes and changes between time slices with spatial root mean squared differences, which masked the spatial distribution of these changes. Therefore, in this section, we analyse how the inhomogeneities in split simulations are spatially distributed. As an example we show the differences in EUR15 (Fig. 6) that cause the $RMSD_t$ peak on March 1st, 2005, shown in Fig. 3 (top panel). The results for other joints are qualitatively similar (not shown). The first column shows the observed change at the time-slice joint ($D_{st}X$) for the different variables. The changes shown will stand out as a noticeable inhomogeneity (i.e. peaks in Figs. 2 or 3) if (and where) D_sX (second column) is significantly larger (for an order of magnitude) than D_tX (third column). In order to have a relative measure to compare different variables, in the fourth column we show the differences in standard deviation units ($d_{st}X$).

For soil moisture content (Fig. 6, first row), as already inferred from the $RMSD_t$ time series, the change between consecutive days (D_tX) is negligible as compared to change between slices (D_sX). The relatively high differences at joints are spread across the whole domain (Fig. 6d), with highest values (i.e above three standard deviations) in eastern and northeastern parts of Europe, as well as in northern Africa. In the European regions, these high discrepancies can be explained only by changes between the slices (D_sX) at the specific joint, since soil moisture has very long response time so the changes between consecutive days (D_tX) are negligible. The two slices are joined within the transitional time during which soil moisture reaches the

full equilibrium and internal variability starts to dominate (Fig. 4). This can vary with the region, since the length of the spin-up depends on the soil characteristic and the climate conditions. Therefore, these inhomogeneities can be contributed to the internal variability for sure, but regionally they may contain also the effects of the spin-up due to the different climate conditions (Yang et al. 2011; Lim et al. 2012). On the other hand, high differences in northern Africa are the consequence of our standardization - this area is very dry and absolute changes of soil moisture are very small, which leads to large (i.e. above three standard deviation) relative changes in standard deviation units. The results for snow depth are qualitatively similar. Due to its long response time, most of the differences occur between the slices, except for an elongated area north of the Black Sea. Since the spin-up was sufficient (Fig. 4b), we may attribute the discrepancies between both slices to the internal variability. These snow accumulation differences between slices (Fig. 6f) have typical depths from individual misplaced snowfall events, such as the one north of the Black Sea, occurring on the analysed day (Fig. 6g). As such, unlike soil moisture, which is relatively stable at this joint (Fig. 4), the specific spatial pattern can fully differ from one joint to another for snow depth, even considering the same season. As an example, on March 1st, 2008 (Fig. ESM 5), the synoptic conditions over Europe barely provided any snow and, as a result, inhomogenities in snow depth are very small. This emphasizes the role of interannual variability at the joint, which in turn may increase or decrease the inhomogeneities. Four areas with differences up to two standard deviations stand out in the 850 hPa geopotential height (Fig. 6i, l). Unlike for the previous slow-varying variables, these differences are almost exclusively due to changes in synoptic conditions between consecutive days (Fig. 6k). Along the day, two lows develop north of France and Scandinavia, and a third low moves and deepens from the west to the north of the Black Sea. Only west of the Black Sea the inhomogeneities have a slightly larger contribution of the changes between between the two slices (Fig. 6j), which weakens and slightly shifts the low northwards. The season is again an important factor. In winter (Figs. 6 and ESM 5), when boundary forcing at mid latitudes is dominant, the differences are mainly attributed to the day-to-day natural variability of the atmosphere (e.g. low pressure systems entering or moving across the domain). However, by the end of the spring, the strength of the boundary forcing decreases and internal variability increases, allowing for larger discrepancies between time slices. Thus, the discrepancies between the two slices are larger at the joint on June 1st, 2006 (Fig. ESM 4), as it can be observed over northern Europe.

Even though differences found in the geopotential height are small, they may affect the results in other variables, especially those that are dependent on the synoptic
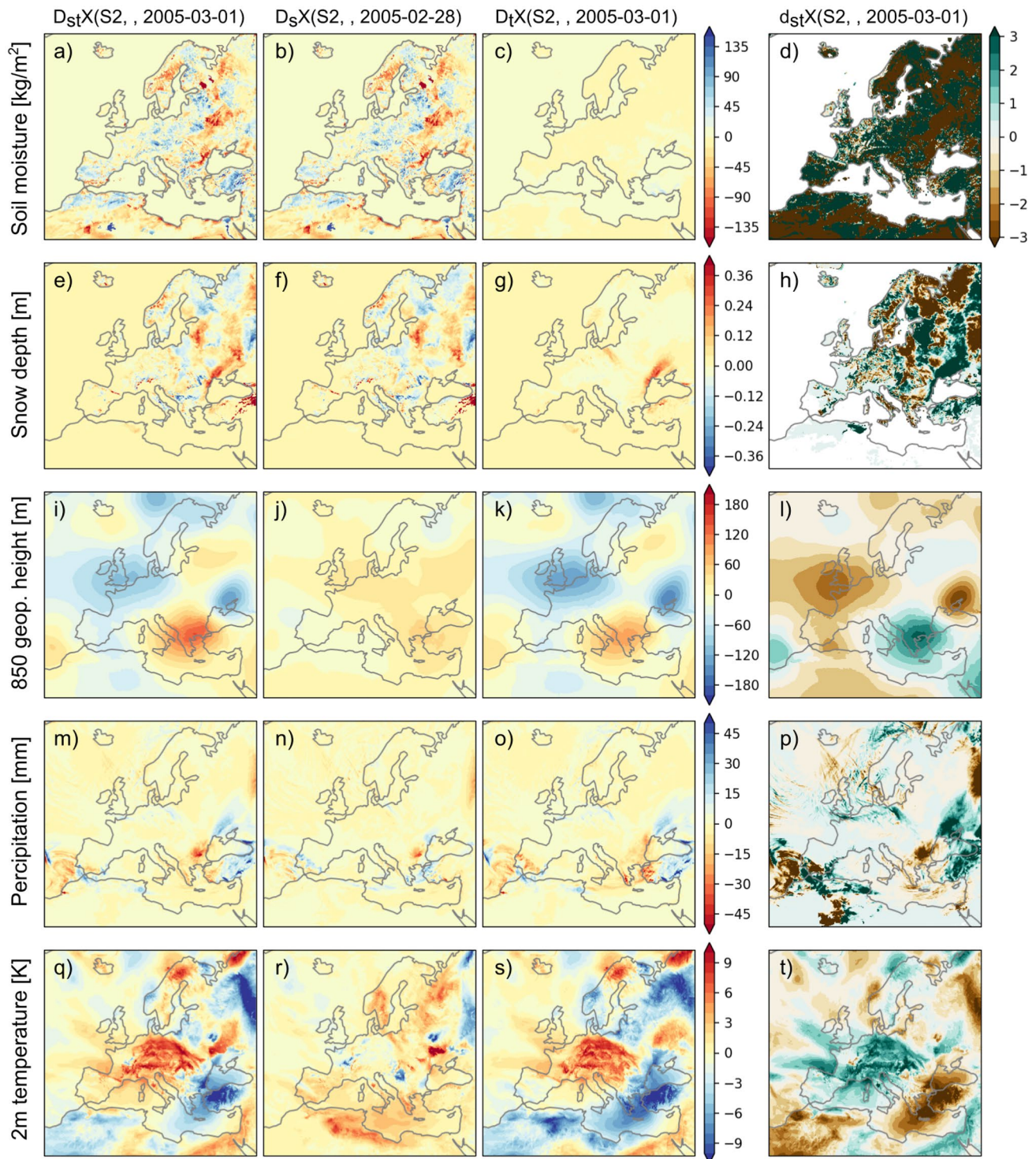
Fig. 6 Spatial distribution of differences (Equations 1–3 and 5) for EUR-15 at the joint from 28th February $(t-1)$ to 1st March, 2005 (t) between slices S2 (s) and S1 $(s-1)$. From left to right: $D_{st}X(s, n, t)$, $D_s X(s, n, t-1)$, $D_t X(s, n, t)$, and $d_{st}X(s, n, t)$. Note that the latter is non-dimensional and the same colorbar is used for all variables (from top to bottom): total soil moisture (kg/m^2), snow depth (m), geopotential height at 850 hPa (m), daily accumulated precipitation (mm) and near-surface temperature (K)

circulation, such as precipitation. Discrepancies between the slices (D_sX) in the low pressure area, west of the Black Sea (Fig. 6f), drive precipitation changes there, causing inhomogeneities to increase up to 50 *mm* in northern Bulgaria. This is due to a northward shift in precipitation which accompanies the corresponding shift in the low pressure system. This is observed in the other domains, as well. For example, in SAM-44 (not shown), the joint on January 1st, 2015 exhibits extended precipitation inhomogeneities east of Brazil, driven by time slice differences in simulating a low pressure system over the Atlantic.

Inter-slice differences for near-surface temperature (Fig. 6r) are generally small over the domain. There are only isolated locations with notable differences, especially in Ukraine, but still these are within one standard deviation (Fig. 6t). Therefore, inhomogeneities that appear at the joint can be mostly explained with the day-to-day natural change of the variable (Fig. 6s).

3.4 Simulated climate

In previous sections we have shown that discontinuities in split simulations can be relatively large, especially for slow-varying variables. Since these discontinuities occur on individual time steps, it is expected that the simulated climate should not be affected by the splitting procedure, especially when the differences are within the range of model's internal variability. In this section, we compare the climatology of the split and continuous simulations. For this purpose, we used the EUR-44 set of simulations, which has the longest simulated overlapping period (20 years) among all of our simulation sets (see Fig. 1).

Figure 7 shows the winter (DJF) and summer (JJA) differences in the seasonal climatology for the full overlapping period 1991–2010 (first and fourth columns). All variables considered are shown in different rows and differences use a common non-dimensional scale of seasonal standard deviation units. In winter, soil moisture and snow depth show significant differences, while upper air and surface variables show much lower, non-significant differences. Differences are spatially smooth for all variables except snow depth, which shows patchy differences over snow-dominated areas in the domain. In summer (Fig. 7, fourth column), snow depth differences vanish due to lack of snow. Somewhat larger differences arise in upper air and surface variables, especially in the summer season. Significant precipitation differences (Fig. 7, row d, column 4) are concentrated mostly over Mediterranean, having a very patchy pattern which is consistent with a weak mid-latitude lateral boundary forcing in summer months, and the mainly convective origin of precipitation. A relative low geopotential height develops in the split simulation over northern Africa/central Mediterranean Sea in summer (row c and column 4 in Fig. 7), which

is consistent with the relatively colder region over northern Africa (row e and column 4 in Fig. 7).

In order to check whether these differences are consistent with low frequency internal variability or a side effect of the splitting, we computed the seasonal climatology differences for two different 10-year periods. The years 1991–1995 and 2001–2005 (columns 2 and 5 in Fig. 7) are considered as years during which the differences could potentially be affected by a long term spin-up transient, since they were initialized less than 6 years before the simulation start (see Fig. 1). On the other hand, we consider the years 1996–2000 and 2006–2010 (columns 3 and 6 in Fig. 7) as periods when the differences are dominated by the model internal variability, since the RCM initialization occurred at least 6 years after the simulation start. Note that the 20-year seasonal climatology differences (columns 1 and 4 in Fig. 7) are the average of these two 10-year climatology differences. In this way, differences in the long-term climatology can be ascribed to periods dominated either by potential spin-up or internal variability.

In general, no significant difference can be ascribed particularly to the period with a potential spin-up transient regime. Seasonal differences in the two 10-year periods show similar magnitudes for all variables. Moreover, some features in the 20-year climatology differences, such as summer differences in geopotential height or temperature, correspond to stronger differences during the internal variability dominated 10-year period. More or less co-located differences during the first 10-year period lead to significant differences during the full 20-year period. Patchy summer precipitation differences in the 20-year period also correspond to constructive averages with even patchier differences in the 10-year periods. The same is true for winter snow depth. This points to no spin-up transient effect in the first 10-year period, and to differences compatible with internal variability, even if statistically significant.

No systematic effect is apparent in both periods, except for the soil moisture differences over northern Africa, which are quite robust in all periods and seasons and reach several standard deviations. This is likely related to the extremely dry soils and low variability there, which would need further research. Unfortunately, for this simulations soil moisture at different depths is not available to investigate properly the source of this difference between the continuous and split runs. Over continental Europe these differences are also significant, reaching about half the standard deviation. However, they are compatible with internal variability since there is no systematic location of differences across the 10-year periods. The differences found in the full 20-year period are the result of partly overlapping positive and negative differences in the other two time periods.
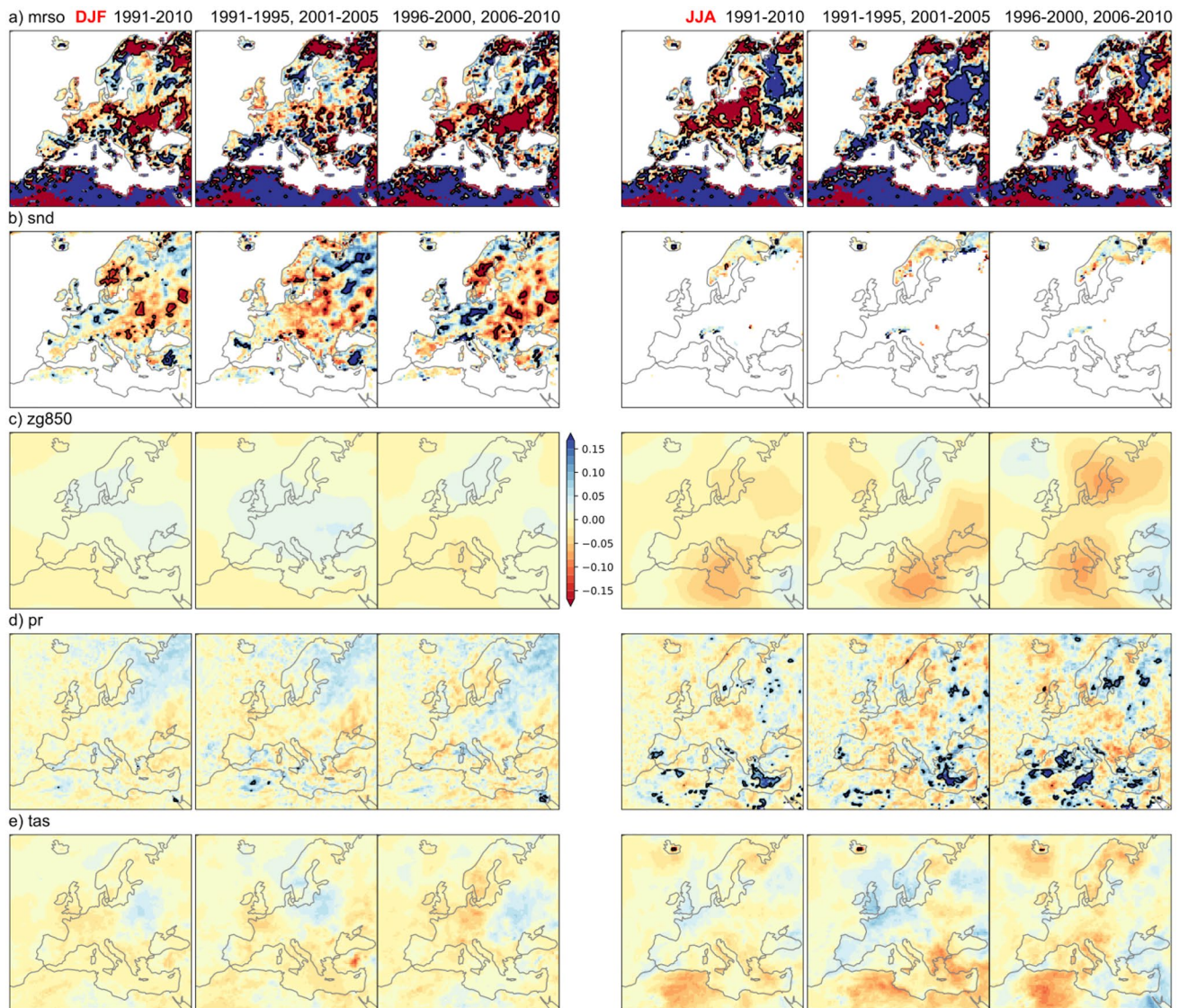
Fig. 7 Seasonal climatology differences for different periods (in columns) and variables (in rows) between the EUR-44 split and EUR-44 continuous. Differences are in non-dimensional standard deviation units. Black contours show statistically significant differences according to a two-sample t-test with 95% confidence. Snow depth (snd) differences are masked out where the variability is below 0.001 m

4 Summary and conclusions

We presented a *post hoc* analysis of several regional climate simulation experiments carried out with the WRF RCM as a set of overlapping time slices. The simulations span different domains, boundary conditions, horizontal grid spacing and overlapping periods. Overlapping time slices were joined to build split simulations. Continuous simulations are available as reference to evaluate split simulations for some of these experiments. We evaluated the discontinuities in time and space introduced by this procedure at the joints of the time slices, devising a methodology to discern between insufficient spin-up and RCM internal variability effects. Finally, the effect of overlapping time slices on the regional climate was assessed.

The analysis was carried out on variables with different response times and, at the same time, variables typically saved in coordinated dynamical downscaling initiatives, such as CORDEX. For this purpose, we also focused on daily model output, commonly available in public repositories. An analysis at higher temporal frequencies would likely ease the location of meteorological discontinuities for the variables with the shortest response times (geopotential height, temperature, precipitation), which went unnoticed in our study. However, the output frequency does not affect spin-up times or seasonal climatology analyses.

We avoided spatial averaging and used root mean squared differences to highlight any mismatch between time slice simulations. The relative size of time slice switch differences with respect to daily variable tendencies only led to noticeable inhomogeneities in soil moisture. Locally, snow depth can also be used to reveal slice joints. Surface and upper air variables show larger day-to-day variations than across time slices.

Jerez et al. (2020) showed that the optimal spin-up period is not always the longest, recommending an initialization in the warm season. We found that this depends on the region, though. As an example, unlike in Europe, the warm season in South America is also the wet season, due to deep convective events which lead to greater precipitation internal variability and, thus, more uncertain initial soil moisture conditions. The optimal starting point should therefore be found for each region to minimize the contents of slow-varying reservoirs (e.g. snow and soil water), thus avoiding gross errors in their initial levels. We also found that mismatches at the slice joints are also minimized during the warm/dry season. Therefore, a minimal overlapping time slice setup could be a 1-year spin-up period initialized at the end of the warm/dry season and entering the split simulation one year later.

The largest and more spatially extended differences were found in total soil moisture, both regarding meteorological inhomogeneities and long term climatologies. This agrees with previous studies warning on the very long spin-up required by soil variables, and especially soil moisture (Christensen 1999; Cosgrove et al. 2003; Yang et al. 2011; Jerez et al. 2020). We found that significant differences in the climatology over continental Europe might be ascribed to internal model variability. However, differences over low-precipitation, non-vegetated areas (northern Africa) present systematic differences which persist along the simulated period. This is likely due to initial soil moisture inconsistencies between the forcing reanalysis and the RCM equilibrium soil state over these areas. The land surface model takes a very long time to restore the equilibrium, especially from extremely dry initial conditions (Cosgrove et al. 2003). The situation is likely exacerbated due to the lack of precipitation and deep roots. A dedicated study with long-term, continuous simulations would be necessary to properly disentangle the internal variability and spin-up of soil variables in coupled simulations. Most studies on soil spin-up rely on off-line land surface model simulations (Cosgrove et al. 2003; Yang et al. 2011), where there is a target equilibrium soil moisture that is consistent with the prescribed atmospheric forcing. Coupled simulations, with an active soil-atmosphere feedback, could develop greater internal variability with persistent anomalies in the slowest components. There are examples, though, of the opposite effect (reduced internal variability) for simulations coupled to other slow interactive components such as a regional ocean (Ho-Hagemann et al. 2020).

For other variables, the discrepancies between the climatology in split simulations and continuous simulations can be ascribed to internal variability, even if statistically significant (e.g. for snow depth). No special effect on the climatology was found for periods closer to the initialization with respect to those farther away.

Our results are robust to interannual variability regarding the detection of meteorological inhomogeneities. Spin-up times for slow-varying variables can depend on the specific conditions of the initialization year. We showed examples with a long plateau in snow depth differences across time slices in EUR-15 in 2003 (Fig. 4), which did not occur in other overlapping years (Figs. ESM1 or ESM2). Also, the soil moisture spin-up time, as represented by the time to reach the internal variability limit, differs from year to year. While one year is usually enough, there are instances (e.g. 1999 in EUR-44, Figure ESM3) when over two years seem necessary to reach equilibrium.

All in all, this work shows that the use of overlapping time slices to accomplish long term regional climate simulations is a valid approach. This procedure can largely improve the efficiency of regional climate simulations, both for computationally heavy simulation (e.g. kilometer-scale simulation) or for a faster accomplishment of lower resolution runs, which do not scale efficiently to a large number of processors. Modelling workflow managers, such as WRF4G (Fernández-Quiruelas et al. 2015) —used in our simulations—, can help in the extra design, job submission and monitoring burden of this approach.

As a final cautionary remark, note that we considered RCM simulations with prescribed sea surface temperatures. RCM simulations coupled to a regional ocean pose an even greater spin-up challenge due to the strong thermal and dynamic inertia of the ocean surface. The results shown could also be sensitive to the model parameterizations and other model components (e.g. land surface model) used. Multi-model uncertainty was also not explored in this study and would require coordination of different modelling groups. Given the uncertainties, in order to be on the safe side of soil spin-up, longer slices and spin-up times could be considered. For example, centennial scenario simulations could be safely split into 30-year slices (e.g. near-, mid-, and far-future) with a 5-year spin-up each, especially if soil variables are initialized from an RCM soil climatology (Cosgrove et al. 2003; Rodell et al. 2005; Jerez et al. 2020).

Availability of data and material The simulation data used is fully available upon request to the corresponding author. Sample data for the analyzed overlapping periods over the CORDEX-SAM domain along with the model configuration files are available at https://doi.org/10.5281/zenodo.5012560. Split simulation data for SAM-44 and EUR-44 are freely available through the Earth System Grid Federation (ESGF). EUR-15 and ALP-3 data will also be made available through ESGF in the future.

Code availability The WRF source code is freely available through GitHub at https://github.com/wrf-model/WRF.

Declarations

References

Arora VK, Scinocca JF, Boer GJ, Christian JR, Denman KL, Flato GM, Kharin VV, Lee WG, Merryfield WJ (2011) Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. Geophys Res Lett. https://doi.org/10.1029/2010GL046270

Ban N, Caillaud C, Coppola E, Pichelli E, Sobolowski S, Adinolfi M, Ahrens B, Alias A, Anders I, Bastin S, Belusic D, Berthou S, Brisson E, Cardoso R, Chan S, Christensen OB, Fernández J, Fita L, Frisius T, Gašparac G, Giorgi F, Goergen K, Haugen JE, Hodnebrog O, Kartsios S, Katragkou E, Kendon EJ, Keuler K, Lavin-Gullon A, Lenderink G, Leutwyler D, Lorenz T, Maraun D, Mercogliano P, Milovac J, Panitz HJ, Raffa M, Remedio AR, Schär C, Soares PMM, Srnec L, Steensen BM, Stocchi P, Tölle MH, Truhetz H, Vergara-Temprado J, de Vries H, Warrach-Sagi K, Wulfmeyer V, Zander M (2021) The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation. Clim Dyn 57:275–302. https://doi.org/10.1007/s00382-021-05708-w

Bassett R, Young PJ, Blair GS, Samreen F, Simm W (2020) A large ensemble approach to quantifying internal model variability within the WRF numerical model. J Geophys Res: Atmos 125(7):e2019JD031286. https://doi.org/10.1029/2019JD031286

Berg P, Christensen J (2008) Poor man's re-analysis over Europe. Tech Rep, WATCH Technical 5 Report No. 2

Caya D, Biner S (2004) Internal variability of RCM simulations over an annual cycle. Clim Dyn 22(1):33–46. https://doi.org/10.1007/s00382-003-0360-2

Chen F, Dudhia J (2001) Coupling an advanced land surface-hydrology model with the penn state-NCAR MM5 modeling system. Part i: Model implementation and sensitivity. Mon Weather Rev 129(4):569–585. https://doi.org/10.1175/1520-0493(2001)129<0569:caalsh>2.0.co;2

Christensen OB (1999) Relaxation of soil variables in a regional climate model. Tellus A 51:674–685. https://doi.org/10.1034/j.1600-0870.1999.00010.x

Christensen OB, Gaertner MA, Prego JA, Polcher J (2001) Internal variability of regional climate models. Clim Dyn 17:875–887. https://doi.org/10.1007/s003820100154

Coppola E, Sobolowski S, Pichelli E, Raffaele F, Ahrens B, Anders I, Ban N, Bastin S, Belda M, Belusic D, Caldas-Alvarez A, Cardoso RM, Davolio S, Dobler A, Fernandez J, Fita L, Fumiere Q, Giorgi F, Goergen K, Guttler I, Halenka T, Heinzeller D, Hodnebrog O, Jacob D, Kartsios S, Katragkou E, Kendon E, Khodayar S, Kunstmann H, Knist S, Lavin-Gullon A, Lind P, Lorenz T, Maraun D, Marelle L, van Meijgaard E, Milovac J, Myhre G, Panitz HJ, Piazza M, Raffa M, Raub T, Rockel B, Schär C, Sieck K, Soares PMM, Somot S, Srnec L, Stocchi P, Tolle MH, Truhetz H, Vautard R, de Vries H, Warrach-Sagi K (2020) A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. Clim Dyn 55:3–34. https://doi.org/10.1007/s00382-018-4521-8

Cosgrove BA, Lohmann D, Mitchell KE, Houser PR, Wood EF, Schaake JC, Robock A, Sheffield J, Duan Q, Luo L, Higgins RW, Pinker RT, Tarpley JD (2003) Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS). J Geophys Res Atmos. https://doi.org/10.1029/2002JD003316

Dee DP, de Rosnay P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Van de Berg L, Bidlot J, Bormann N, Dragani R, Fuentes M, Vitart F (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. Q J Royal Meteorol Soc 137:553–597. https://doi.org/10.1002/qj.828

Dickinson RE, Errico RM, Giorgi F, Bates GT (1989) A regional climate model for the western United States. Clim Chang 15:383–422. https://doi.org/10.1007/BF00240465

Falco M, Carril AF, Menéndez CG, Zaninelli PG, Li LZX (2019) Assessment of CORDEX simulations over South America: added value on seasonal climatology and resolution considerations. Clim Dyn 52:4771–4786. https://doi.org/10.1007/s00382-018-4412-z

Fernández-Quiruelas V, Blanco C, Cofiño A, Fernández J (2015) Large-scale climate simulations harnessing clusters, grid and cloud infrastructures. Future Gener Comput Syst 51:36–44. https://doi.org/10.1016/j.future.2015.04.009

Geyer B, Ludwig T, von Storch H (2021) Limits of reproducibility and hydrodynamic noise in atmospheric regional modelling. Commun Earth Environ. https://doi.org/10.1038/s43247-020-00085-4

Giorgi F (2019) Thirty years of regional climate modeling: Where are we and where are we going next? J Geophys Res: Atmos 124(11):5696–5723. https://doi.org/10.1029/2018JD030094

Giorgi F, Bates GT (1989) The climatological skill of a regional model over complex terrain. Mon Weather Rev 117(11):2325–2347. https://doi.org/10.1175/1520-0493(1989)117<2325:TCSOAR>2.0.CO;2

Giorgi F, Gutowski WJ (2015) Regional dynamical downscaling and the CORDEX initiative. Annu Rev Environ Resour 40(1):467–490. https://doi.org/10.1146/annurev-environ-102014-021217

Gómez-Navarro JJ, Montávez JP, Jerez S, Jiménez-Guerrero P, Lorente-Plazas R, González-Rouco JF, Zorita E (2011) A regional climate simulation over the Iberian peninsula for the last millennium. Clim Past 7(2):451–472. https://doi.org/10.5194/cp-7-451-2011

Ho-Hagemann HTM, Hagemann S, Grayek S, Petrik R, Rockel B, Staneva J, Feser F, Schrum C (2020) Internal model variability of the regional coupled system model GCOAST-AHOI. Atmosphere. https://doi.org/10.3390/atmos11030227

Jacob D, Teichmann C, Sobolowski S, Katragkou E, Anders I, Belda M, Benestad R, Boberg F, Buonomo E, Cardoso RM, Casanueva A, Christensen OB, Christensen JH, Coppola E, De Cruz L, Davin EL, Dobler A, Dominguez M, Fealy R, Fernandez J, Gaertner MA, Garcia-Diez M, Giorgi F, Gobiet A, Goergen K, Gomez-Navarro JJ, Aleman JJG, Gutierrez C, Gutierrez JM, Guettler I, Haensler A, Halenka T, Jerez S, Jimenez-Guerrero P, Jones RG, Keuler K, Kjellstrom E, Knist S, Kotlarski S, Maraun D, van Meijgaard E, Mercogliano P, Montavez JP, Navarra A, Nikulin G, de Noblet-Ducoudre N, Panitz HJ, Pfeifer S, Piazza M, Pichelli E, Pietikainen JP, Prein AF, Preuschmann S, Rechid D, Rockel B, Romera R, Sanchez E, Sieck K, Soares PMM, Somot S, Srnec L, Sorland SL, Termonia P, Truhetz H, Vautard R, Warrach-Sagi K, Wulfmeyer V (2020) Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community. Reg Environ Change 20:51. https://doi.org/10.1007/s10113-020-01606-9

Jerez S, Gimenez D, Montávez JP (2009) Optimizing the execution of a parallel meteorology simulation code. IPDPS 2009 - Proceedings of the 2009 IEEE International Parallel and Distributed Processing Symposium https://doi.org/10.1109/IPDPS.2009.5161154

Jerez S, López-Romero JM, Turco M, Lorente-Plazas R, Gómez-Navarro JJ, Jiménez-Guerrero P, Montávez JP (2020) On the spin-up period in WRF simulations over Europe: trade-offs between length and seasonality. J Adv Model Earth Syst 12(4):e2019MS001945. https://doi.org/10.1029/2019ms001945

Jimenez PA, Gonzalez-Rouco JF, Garcia-Bustamante E, Navarro J, Montavez JP, de Arellano JVG, Dudhia J, Muñoz-Roldan A (2010) Surface wind regionalization over complex terrain: evaluation and analysis of a high-resolution WRF simulation. J Appl Meteorol Climatol 49:268–287. https://doi.org/10.1175/2009JAMC2175.1

Khodayar S, Sehlinger A, Feldmann H, Kottmeier C (2015) Sensitivity of soil moisture initialization for decadal predictions under different regional climatic conditions in Europe. Int J Climatol 35(8):1899–1915. https://doi.org/10.1002/joc.4096

Lavin-Gullon A, Fernandez J, Bastin S, Cardoso RM, Fita L, Giannaros TM, Goergen K, Gutierrez JM, Kartsios S, Katragkou E, Lorenz T, Milovac J, Soares PMM, Sobolowski S, Warrach-Sagi K (2021) Internal variability versus multi-physics uncertainty in a regional climate model. Int J Climatol 41(S1):E656–E671. https://doi.org/10.1002/joc.6717

Leutwyler D, Fuhrer O, Lapillonne X, Lüthi D, Schär C (2016) Towards European-scale convection-resolving climate simulations with GPUs: a study with COSMO 4.19. Geosci Model Dev 9(9):3393–3412. https://doi.org/10.5194/gmd-9-3393-2016

Liebmann B, Mechoso CR (2011) The South American Monsoon System. In: Chang CP, Johnson RH, Lau NC, Wang B, Yasunari T (eds) The Global Monsoon System: Research andForecast. 2nd edn. World Scientific, Singapore, pp 137–157. https://doi.org/10.1142/9789814343411_0009

Lim YJ, Hong J, Lee TY (2012) Spin-up behavior of soil moisture content over east Asia in a land surface model. Meteorol Atmos Phys 118(3–4):151–161. https://doi.org/10.1007/s00703-012-0212-x

Lo JCF, Yang ZL, Pielke RA Sr (2008) Assessment of three dynamical climate downscaling methods using the weather research and forecasting (wrf) model. J Geophys Res: Atmos. https://doi.org/10.1029/2007JD009216

Lucas-Picher P, Caya D, de Elía R, Laprise R (2008) Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain. Clim Dyn 31:927–940

Lucas-Picher P, Boberg F, Christensen JH, Berg P (2013) Dynamical downscaling with reinitializations: A method to generate finescale climate datasets suitable for impact studies. J Hydrometeorol 14(4):1159–1174. https://doi.org/10.1175/JHM-D-12-063.1

Manzanas R, Gutiérrez J, Fernández J, van Meijgaard E, Calmanti S, Magariño M, Cofiño A, Herrera S (2018) Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: added value for user applications. Clim Serv 9:44–56. https://doi.org/10.1016/j.cliser.2017.06.004

Menendez M, Garcia-Diez M, Fita L, Fernandez J, Mendez FJ, Gutierrez JM (2014) High-resolution sea wind hindcasts over the Mediterranean area. Clim Dyn 42:1857–1872. https://doi.org/10.1007/s00382-013-1912-8

Niu GY, Yang ZL, Mitchell KE, Chen F, Ek MB, Barlage M, Kumar A, Manning K, Niyogi D, Rosero E, Tewari M, Xia Y (2011) The community Noah land surface model with multiparameterization options (Noah-MP): model description and evaluation with local-scale measurements. J Geophys Res. https://doi.org/10.1029/2010jd015139

Pan Z, Takle E, Gutowski W, Turner R (1999) Long simulation of regional climate as a sequence of short segments. Month Weather Rev 127(3):308–321. https://doi.org/10.1175/1520-0493(1999)127<0308:LSORCA>2.0.CO;2

Pichelli E, Coppola E, Sobolowski S, Ban N, Giorgi F, Stocchi P, Alias A, Belušić D, Berthou S, Caillaud C, Cardoso RM, Chan S, Christensen OB, Dobler A, de Vries H, Goergen K, Kendon EJ, Keuler K, Lenderink G, Lorenz T, Mishra AN, Panitz HJ, Schär C, Soares PMM, Truhetz H, Vergara-Temprado J (2021) The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation. Clim Dyn 56:3581–3602. https://doi.org/10.1007/s00382-021-05657-4

Qian JH, Seth A, Zebiak S (2003) Reinitialized versus continuous simulations for regional climate downscaling. Month Weather Rev 131(11):2857–2874. https://doi.org/10.1175/1520-0493(2003)131<2857:RVCSFR>2.0.CO;2

Rodell M, Houser PR, Berg AA, Famiglietti JS (2005) Evaluation of 10 methods for initializing a land surface model. J Hydrometeorol 6:146–155. https://doi.org/10.1175/JHM414.1

Seneviratne SI, Lüthi D, Litschi M, Schär C (2006) Land-atmosphere coupling and climate change in Europe. Nature 443(7108):205–209. https://doi.org/10.1038/nature05095

Skamarock C, Klemp B, Dudhia J, Gill O, Barker D, Duda G, Huang Xy, Wang W, Powers G (2008) A description of the advanced research WRF version 3. University Corporation for Atmospheric Research No. NCAR/TN-475+STR, https://doi.org/10.5065/D68S4MVH

Solman S, Blázquez J (2019) Multiscale precipitation variability over South America: analysis of the added value of CORDEX RCM simulations. Clim Dyn 53:1547–1565. https://doi.org/10.1007/s00382-019-04689-1

Stahl K, Tallaksen LM, Gudmundsson L, Christensen JH (2011) Streamflow data from small basins: a challenging test to high-resolution regional climate modeling. J Hydrometeorol 12(5):900–912. https://doi.org/10.1175/2011JHM1356.1

Vautard R, Gobiet A, Jacob D, Belda M, Colette A, Deque M, Fernandez J, Garcia-Diez M, Goergen K, Guettler I, Halenka T, Karacostas T, Katragkou E, Keuler K, Kotlarski S, Mayer S, van Meijgaard E, Nikulin G, Patarcic M, Scinocca J, Sobolowski S, Suklitsch M, Teichmann C, Warrach-Sagi K, Wulfmeyer V, Yiou P (2013) The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Clim Dyn 41:2555–2575. https://doi.org/10.1007/s00382-013-1714-z

WMO (2017) WMO guidelines on the calculation of climate normals. Tech. Rep. WMO-No. 12031, World Meteorological Organization1

Yang Y, Uddstrom M, Duncan M (2011) Effects of short spin-up periods on soil moisture simulation and the causes over New Zealand. J Geophys Res 116(D24):1–12. https://doi.org/10.1029/2011JD016121

Zappa G (2019) Regional climate impacts of future changes in the mid-latitude atmospheric circulation: a storyline view. Curr Clim Change Rep 5(4):358–371. https://doi.org/10.1007/s40641-019-00146-7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.