Climate Dynamics manuscript No.
(will be inserted by the editor)

Pan-European climate at convection-permitting scale: a model intercomparison study

Ségolène Berthou · Elizabeth Kendon · Steven Chan · Nikolina Ban · David Leutwyler · Christoph Schär · Giorgia Fosser

Received: date / Accepted: date

Abstract We investigate the effect of using convection-permitting models (CPMs) spanning a pan-European domain on the representation of precipitation distribution at a climatic scale. In particular we compare two 2.2km models with two 12km models run by ETH Zürich (ETH-12 km and ETH-2.2 km) and the Met-Office (UKMO-12 km and UKMO-2.2 km).

The two CPMs yield qualitatively similar differences to the precipitation climatology compared to the 12 km models, despite using different dynamical cores and different parameterization packages. A quantitative analysis confirms that the CPMs give the largest differences compared to 12 km models in the hourly precipitation distribution in regions and seasons where convection is a key process: in summer across the whole of Europe and in autumn over the Mediterranean Sea and coasts. Mean precipitation is increased over high orography, with an increased amplitude of the diurnal cycle. We highlight that both CPMs show an

S. Berthou
Met Office Hadley Centre, FitzRoy Rd, Exeter, Devon EX1 3PB, UK
Tel.: +44 1392 884967
E-mail: segolene.berthou@metoffice.gov.uk

E. Kendon
Met Office Hadley Centre, FitzRoy Rd, Exeter, Devon EX1 3PB, UK

S. Chan
University of Newcastle, UK

N. Ban
Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland

D. Leutwyler
Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland

C. Schär
Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland

G. Fosser
Met Office Hadley Centre, FitzRoy Rd, Exeter, Devon EX1 3PB, UK

increased number of moderate to intense short-lasting events and a decreased number of longer-lasting low-intensity events everywhere, correcting (and often over-correcting) biases in the 12 km models. The overall hourly distribution and the intensity of the most intense events is improved in Switzerland and to a lesser extent in the UK but deteriorates in Germany. The timing of the peak in the diurnal cycle of precipitation is improved. At the daily time-scale, differences in the precipitation distribution are less clear but the greater Alpine region stands out with the largest differences. Also, Mediterranean autumnal intense events are better represented at the daily time-scale in both 2.2 km models, due to improved representation of mesoscale processes.

Keywords convection-permitting models · Europe · Mediterranean · diurnal cycle · convection

1 Introduction

Global climate models (GCMs) are our primary tool for understanding how climate may change in the future with increasing greenhouse gases. These typically have coarse resolutions with grid spacings of 60-300 km (Taylor et al 2012). To provide regional detail, higher resolution regional climate models (RCMs; 12-50 km grid spacing) are often used, which only span a limited area (Jacob et al 2014). These give a better representation of mountains and coastlines and fine-scale (order 10-100km) physical and dynamical processes. In general, RCMs are able to capture the average statistics of daily precipitation on scales of a few grid boxes, with greatest agreement for moderate intensities and model biases increasing for heavier events (Boberg et al 2009; Kjellström et al 2010).

Both GCMs and RCMs with typical grid spacings (>10km) rely on a convection parameterisation scheme to represent the average effects of convection. This simplification is a known source of model errors and leads to deficiencies in the diurnal cycle of convection (Brockhaus et al 2008) and the inability by design to produce hourly precipitation extremes (Hanel and Buishand 2010; Gregersen et al 2013). Very high resolution models (order 1 km grid spacing), can represent deep convection explicitly without the need for such parameterisation schemes (Kendon et al 2012; Hohenegger et al 2008). Such models are termed ‘convection-permitting’ (or for simplicity sometimes ‘convection-resolving’ but this is not stricly true): larger storms and mesoscale convective organisation are permitted (largely resolved) but most turbulent kinetic motions are not represented (Wyngaard 2004). More specifically, while there is some evidence that km-scale resolution represents convection in some bulk sense (Langhans et al 2013), resolving convective updrafts requires about ten times higher resolutions (Dauhut et al 2015).

Convection-permitting models (CPMs) are commonly used in short-range weather forecasting, where they have been shown to give a much more realistic representation of convection and can be used to forecast the possibility of localised high-impact rainfall not captured at coarser resolutions (Done et al 2004; Richard et al 2007; Lean et al 2008; Weisman et al 2008; Weusthoff et al 2010; Schwartz 2014). However, due to their high computational cost, they have not commonly been applied at climate-time scales. Studies to date show that convection-permitting models do not necessarily better represent daily mean precipitation (Chan et al

2013), but have significantly better sub-daily rainfall characteristics with improved representation of the diurnal cycle of convection (Ban et al 2014), the spatial structure of rainfall and its duration-intensity characteristics (Kendon et al 2012), the intensity of hourly precipitation extremes (Chan et al 2014; Ban et al 2014; Fosser et al 2015), orographic precipitation and snowpack (Liu et al 2016), which are typically poorly represented in climate models.

Convection-permitting models provide a step change in our ability to represent convection, but there are still remaining issues. Smaller showers are not properly resolved, which results in a tendency for heavy rain to be too intense and for cell sizes to be too large. CPMs are also sensitive to sub-grid scale process representation (turbulence, microphysics), associated with many unknowns. The use of ever higher resolution does not necessarily result in convergence in terms of the representation of convection. For example, showers tend to become smaller (more speckly) with finer resolution rather than upscale on to the correct meteorological scale (Hanley et al 2015) and improvement with resolution can depend on use of appropriate parameterisation (Bryan and Morrison 2012).

Although convection-permitting simulations have been used at climate-scales on small domains in several regions of Europe and North America (see Prein et al (2015) for a review), Mediterranean intense precipitation events occuring in autumn have not yet been studied with such high-resolution on long time-scales. These events have been widely studied mainly on single cases with convection-permitting models within the the Mesoscale Alpine Programme (MAP, Richard et al (2007) and the HyMeX project (Drobinski et al 2014) and climatologically with convection-parameterised models within Med-CORDEX framework (Berthou et al 2016; Cavicchia et al 2016; Vaittinada Ayar et al 2016; Ruti et al 2016). Khodayar et al (2016) compared various convection-permitting models and convection-parameterised models on a single case study and showed that the former better represent the short-intense convective events whereas the convection-parametrized models tend to produce a large number of weak and long-lasting events. Although convection-parameterised models at scales of 10-40 km are able to capture the role of orography, blocking and convergence lines in shaping heavy-precipitation events, organised convection only represented at convection-permitting scales and interaction of this convection with the orography can be important in the triggering, propagation and life-time of some heavy precipitation in the Mediterranean (Ducrocq et al 2008a; Bresson et al 2012; Manzato et al 2015; Meredith et al 2015; Barthlott and Davolio 2016).

Following the work of Leutwyler et al (2017), who provided an analysis of the performance of the 10-year long ETH-2.2 km simulation in comparison with the driving 12 km simulation, we compare 9 years of simulations for a pan-european domain from the UKMO and the ETH 2.2 km models with 12 km models and with observations. The main added value of the article is to provide the first model-intercomparison study of convection-permitting climate simulations across the wide variety of European climates and to objectively assess in which regions and seasons they differ most with coarser resolution models in terms of precipitation.

After presenting the models and datasets in Sect. 2 and the methods in Sect. 3, we identify regions and seasons where the 2.2 km models differ most from the 12 km models in terms of distribution shape and mean of hourly precipitation in Sect. 4 and evaluate if this is an improvement against observations. We then gain more

country/region	name	reference	native resolution	years
France	SAFRAN	Quintana-Segui et al (2008)	8 km	1999-2007
Germany	HYRAS	Rauthe et al (2013)	5 km	1998-2006
Spain	Spain02	Herrera et al (2012)	12 km	1999-2007
United Kigdom	UKCPOBS	Perry et al (2009)	5 km	1999-2007
Alps	APGD_EURO4M	Isotta et al (2014)	5 km	1999-2007
Switzerland	RdisaggH	Perry et al (2009)	1 km	2003-2010
Germany	GERMANY	Paulat et al (2008)	7 km	2001-2008
United Kingdom	NIMROD	Golding (1998)	5 km	2003-2011

Table 1 Datasets used in this study: daily datasets in the top part of the table, hourly datasets in the bottom part of the table. Years indicate the years of the datasets used in this study. Most datasets span a longer period.

insight as to how the distribution changes in summer in Sect. 5. Finally, we focus on the representation of Mediterranean heavy precipitation in Sect. 6 with the use of high percentiles and an illustrative case study. We provide conclusions and a discussion in Sect. 7.

2 Datasets and simulations

2.1 Datasets

2.1.1 Daily precipitation

For the analysis of daily precipitation we use the regional gridded datasets presented in the top section of Table 1 covering the UK, France, Germany, the Netherlands, the Alps and Spain. Regional datasets were chosen for the comparison, as advised by Prein and Gobiet (2017): their native resolution are higher than the european-wide EOBS dataset (Haylock et al 2008) and include higher densities of raingauges (up to 44 times more). Furthermore, EOBS is not advised to be used for coastal areas and mountainous regions of Southern Europe (Flaounas et al 2012) and can be biased over regions with a low density of stations, especially regarding the extremes (Hofstra et al 2010; Lenderink 2010; Prein and Gobiet 2017). Further information about how each dataset was computed can be found in Sect. 8.1. CMORPH (NOAA Climate Prediction Center morphing method) was also used to evaluate the representation of heavy precipitation events in the Mediterranean in autumn in Sect. 6. It was not included in the rest of the analysis as it is not representative of the whole precipitation spectrum in northern Europe (Kidd et al 2012).

2.1.2 Hourly precipitation datasets

For the analysis of hourly precipitation we use the datasets presented in the bottom section of Table 1 covering the UK, Germany and Switzerland, which were all the gridded hourly datasets available to the authors. 8 or 9 years were used to compare with the models, they are not necessarily the same as the model years due to data availability (see Table 1). However, we are interested in the multi-year climatology of hourly precipitation, and this is not expected to depend strongly on the exact choice of years providing a sufficient number of years are chosen.

The percentage of missing values for the hourly datasets for 2003-2008 in summer are shown as a map in Figure 1b. The German dataset shows between 10 and 20% of days with missing data all over Germany, the Swiss dataset about 10% of missing data in the southeast of the country and the UK dataset does not cover some regions in the southeast and northeast of England, and has variable coverage in Scotland with about 40% of missing data. For this dataset, only grid-points with less than 30% of missing data are used and the same points are used in the models to avoid inconsistencies.

It should be kept in mind that possible uncertainties in the datasets arise from rain-gauge undercatch, gridding procedures (Frei et al 2003), and weather radar measurements (Wuest et al 2010). The rain-gauge undercatch implies that rainfall intensities may well be underestimated with an amplitude that is difficult to assess. Prein and Gobiet (2017) mention that it can reach up to 80% in mountainous region for snowfall at exposed locations.

All the datasets were conservatively regridded to the 12 km UKMO grid with the Python interface to the Earth System Modeling Framework (ESMF) regridding utility interface before the calculation of indices. The first-order conservative regridding is a variant of a constant method which compares the proportions of overlapping source and destination cells to determine appropriate weights.

2.2 Models

Both CPMs use the same pan-european domain as shown in Fig. 1a defined on a 2.2 km regular grid with a rotated pole located at (43N, 190E). The grid has 1536x1536 points and 70 vertical levels for the UKMO model and 60 for the ETH model. Both models are forced at their boundaries with 6-hourly ERA-interim reanalyses. ETH-2.2 km uses a 12 km-simulation as an intermediate step for the downscaling (dashed domain in Fig. 1a), whereas the UKMO-2.2 km is directly forced by ERA-interim. This large resolution jump (factor 34) for the UKMO configuration implies that the spin-up zone for small-scale transient eddies to develop is larger than for the ETH model. In fact, Matte et al (2017) suggest that spin-up effects for small-scale transient eddies in the vorticity field are present on a 3xL zone, where L is the e-folding distance on which the asymptotic value is reached. According to their findings, we get a spin-up zone of 3x2x75km/2.2km $\simeq$ 205 grid points. Comparing maps of mean precipitation between the UKMO-12km and UKMO-2.2km (not shown), we removed 220 points from the domain on each side for our analysis (zone depicted in Fig. 1) to prevent contamination from the downscaling method.

The simulations are starting in March 1998 for UKMO-2.2 km and in November 1998 for ETH-2.2 km. The soil moisture initial conditions in UKMO comes from ERA-interim from the start of the run. The ETH-2.2 km initialisation is based on the soil moisture fields of ETH-12 km after 5 years of simulation initialised with the CCLM EURO-CORDEX simulation (Kotlarski et al 2014). The UKMO-12 km simulation was set on a wider domain (in yellow in Fig. 1a) and started in January 1998. The article is based on 9 years of simulation from January 1999 to December 2007.

2.2.1 UKMO 12 km and 2.2 km

The Met Office Unified Model (UM) can be run in climate mode (Walters et al 2016), seasonal forecasting mode (Scaife et al 2014) or at convection-permitting scales for numerical weather prediction (NWP) (Clark et al 2016). The UKMO 2.2 km (UM version 10.1) model is based on the UKV Met Office regional model which has been in use for operational numerical weather prediction since 2012 (Clark et al 2016). The UKMO 12 km (UM version 10.3) is based on the climate version (Williams et al in rev.).

The UM is a non-hydrostatic model with a deep-atmosphere formulation based on a semi-implicit semi-Lagrangian dynamical core: ENDGame (Even Newer Dynamics for General atmospheric modelling of the environment) (Wood et al 2014). The prognostic fields are discretised horizontally onto a rotated-pole grid with Arakawa C-grid staggering (Arakawa and Lamb 1977) whilst vertical decomposition is done via CharneyPhillips staggering (Charney and Phillips 1953) using terrain-following hybrid height coordinates on 70 levels for the 2.2 km model and 63 levels for the 12 km model. Both models have a 40km top, but different spacing of levels in the lower troposphere. The lowest grid level is 2.5 m above the ground and the grid spacing increases quadratically with height. The model time-step is 1 mn at 2.2 km and 4 mn at 12 km.

The 2.2 km model does not include any convection parametrization and relies on the model dynamics to explictly represent convective clouds. Although it is acknowledged that not all types of convection are represented with such grid-spacing, this choice was made in the current absence of a scale-aware convection scheme which correctly parametrizes sub-grid convective motion and hands over to the model dynamics for clouds larger than the model filter scale. The UKMO 12 km model uses a mass flux convection scheme based on Gregory and Rowntree (1990) with various extensions which include downdrafts (Gregory and Allen 1991) and convective momentum transport.

The UKMO 12 km model uses a prognostic cloud fraction and prognostic condensate scheme (PC2; Wilson et al (2008)) whereas the UKMO 2.2 km model, like other convection-permitting UM formulations, uses the diagnostic Smith (1990) scheme.

Both models use the radiative transfer scheme of Edwards and Slingo (1996) with a similar configuration as described by Walters et al (2011), with several upgrades (more details in Stratton et al (in rev.)). Aerosol absorption and scattering assumes climatological aerosol properties. Full radiation calculations are made every 15 minutes, with sub-stepped corrections due to cloud evolution performed every 5 minutes. The treatment of cloud microphysical processes is based on Wilson and Ballard (1999), with extensive modifications described in Williams et al (in rev.). The UKMO 2.2 km model includes graupel as a prognostic variable in addition to the moist variables of water vapour, cloud liquid, cloud ice and rain used by the 12 km model. This allows the inclusion of a lightning flash rate prediction scheme (McCaul et al 2009). The UKMO 2.2 km model uses the blended boundary-layer parametrization (Boutle et al 2014). This scheme transitions from the one-dimensional vertical scheme of Lock et al (2000), used for lower resolution simulations such as UKMO 12 km, to a three-dimensional turbulent mixing scheme based on Smagorinsky (1963) and is suitable for high-resolution simulations, with a weighting which is a function of the ratio of the grid-length to a turbulent length scale. The UM uses the JULES (Best et al 2011; Clark et al 2011) land surface scheme with the default four soil layers with thicknesses of 0.1, 0.25, 0.65 and 1.0 m, giving a total depth of 3 m. The tiles share a common soil water reservoir, with the van Genuchten et al (1991) relationship describing soil hydraulic conductivity and soil moisture. Note, however, it has recently been discovered that there may be an inconsistency between the Van-Genuchten hydrology and the soil properties provided in the ancillary, such that soil moisture infiltration rates may be too low. Initial tests using Brooks-Corey hydraulic equations, which are consistent with the soil properties, show that this impacts the soil moisture content but appears to have only limited impact on surface temperature and precipitation. The 12km and 2.2km models also have a different set up in the treatment of saturated soil layers: in the 2.2km model excess water moves upward, whilst in the 12km model it moves downward. The sensitivity of the results to this setting are discussed in Sect. 7 (see supplementary material for more detail).

The sub-grid hydrology model is also different: the UKMO 2.2 km configurations use the Probability Distributed Model (PDM Moore (1985)) and the 12 km follows the climate configuration of the TOPMODEL (Beven and Kirkby 1979).

More details can be found in Walters et al (2016), Williams et al (in rev.) and Stratton et al (in rev.). The latter article provides a more detailed description of a similar model set-up over Africa.

Note that unlike flux formulated schemes, semi Lagrangian advection schemes are typically not designed to locally conserve the advected quantities. Correctors are applied in the global UM, but in regional configurations the issue is complicated by the need to account for fluxes through the lateral boundaries in the calculation of the error and no correction scheme is implemented in these versions of the model. Stratton et al (in rev.) showed that it is likely causing enhanced mean precipitation (by $\simeq$ 20% in Africa), especially due to increased intense rainfall events.

2.2.2 ETH 12 km and 2.2 km

The simulation setup has been introduced in Leutwyler et al (2016) and verification was performed in Leutwyler et al (2017). Therefore we here only briefly summarize the most important aspects.

The 12 km and 2.2 km ETH simulations have been performed with version 4.19 of the Consortium for Small-scale Modeling weather and climate model (COSMO) (Böhm et al 2006; Rockel et al 2008). COSMO is a non-hydrostatic limited-area model solving the fully compressible governing equations with finite-difference methods in a rotated coordinate system, projected on a regular structured grid (Steppeler et al 2003; Förstner and Doms 2004). To integrate the prognostic variables forward in time, a split-explicit 3-stage Runge-Kutta integrator is used (Wicker and Skamarock 2002). For horizontal advection a fifth-order upwind scheme and in the vertical an implicit Crank-Nicholson scheme are used (Baldauf et al 2011). Multi-dimensional advection of scalar fields is implemented using the one-dimensional Bott scheme (Bott 1989; Schneider and Bott 2014). The model time-step is 90 s for the 12 km model and 20 s for the 2.2 km model, considerably shorter than for the UKMO equivalent model.

Depending upon resolution, sub-grid convection is parameterized using an adapted version of the Tiedtke mass-flux scheme with moisture-convergence closure (Tiedtke 1989). Cloud-microphysics are parameterized with a single-moment bulk scheme using five species (cloud water, cloud ice, rain, snow, and graupel) (Reinhardt and Seifert 2005), radiative transfer is based on the δ-two-stream approach (Ritter and Geleyn 1992), and a turbulent-kinetic-energy-based parametrization is used in the planetary boundary layer (PBL) as well as for surface transfer (Mellor and Yamada 1982; Raschendorfer 2001). The ten-layer soil modeel TERRA_ML has a total soil depth of 15.24 m (Heise et al 2006) and the aerosol climatology has been changed from the default climatology (Tanré et al 1984) to the AeroCom climatology (Kinne et al 2006).

The model configuration follows a two-step one-way nesting approach with the outer nest consisting of a simulation with parameterized convection (ETH 12km) and the inner nest of a simulation with the parameterization of deep-convection switched off (ETH2km, Fig. 1a). It should be noted that the parameterization of shallow convection remains active in the ETH 2.2km model, which is an important difference compared to the UKMO configuration (which has no convective parameterization). The outer nest has a grid spacing of 12 km and the inner nest follows the same setup as the UKMO 2.2 km simulation. In both ETH simulations, the vertical direction is discretized using 60 stretched model levels, ranging from the first model level at 20 m to the model top at 23.5 km. To provide adequately spun-up soil moisture fields, the soil layers in ETH 12 km have been initialized on 1 November 1993 based on the soil-moisture fields from the CCLM EURO-CORDEX simulation (Kotlarski et al 2014), and thereafter integrated for 5 years. Subsequently ETH 2.2 km was initialized on 1 November 1998 with the soil moisture fields of ETH 12 km, leaving two months of integration for soil spinup.

The simulations have been performed with a version of COSMO capable of using GPU accelerators (Fuhrer et al 2014). The new COSMO version enables execution of the time stepping algorithm entirely on accelerators, which is essential to minimize expensive data movements between the CPU and the GPU. To this end the dynamical core has been rewritten in C++, using the domain-specific Stencil Loop Language (STELLA) (Gysi et al 2015; Osuna et al 2015), and the physical parametrization have been ported using OpenACC (2011) compiler directives (Lapillonne and Fuhrer 2014). Data exchange at the sub-domain boundaries (i.e. halo exchange) is handled using a re-usable communication framework. On 144 compute nodes of a hybrid Cray XC30 system, the time-to-solution for a 10-year-long integration is about 1.7 months (Leutwyler et al 2016).

3 Methods

All the models and datasets are regridded to the UKMO 12 km grid before the computation of all diagnostics in order to show a fair comparison between models. Therefore, scales smaller than 12 km are not evaluated. However, most of the regional datasets are also not necessarily accurate enough to evaluate scales smaller than 12 km. It should be stressed that this approach is not entirely fair for the 12 km models, as they are not supposed to represent the 12 km scale properly, but rather a 25 km scale or larger (Skamarock 2004).

We use the average of values above the 99^{th} percentile of all days to evaluate the representation of moderate to intense events ($p99avg$) to compare fairly the model extremes, independently from the wet-day/wet-hour frequency, as recommended by Schär et al (2016).

To gain insight into the distribution of precipitation, we use the ASoP method ("Analyzing Scales of Precipitation", version 1.0 ASoP1) presented in Klingaman et al (2017), which gives a spectrum of the precipitation intensities contributing to the mean precipitation rate. This allows a comparison of the contribution of different intensities to the mean across different time-scales and grid-point by grid-point to better understand the underlying model physics. It provides a view of differences in the distribution in its entirety and also allows differences coming from a pure shift to higher/lower intensities to be distinguished from an increase/decrease of precipitation in all the bins.

Fig. 2 shows the steps of calculation for the ASoP method and illustrates the differences with a probability density function. The example uses the distribution of daily precipitation in the southern UK from the UKCPOBS dataset and the UKMO-12 km model for 1999-2007. The bins used to calculate precipitation frequency in the ASoP method are designed such that the number of events per bin is rather similar across bins (except in the largest bins so that their signal is not lost in one single bin). This is illustrated in panel b compared to panel a, where the vertical bars representing each bin are spaced differently. The function defining the bins is given by Eq. (1).

$$b_n = e^{\left(ln(0.005) + \left[n \frac{(ln(120) - ln(0.005))^2}{59} \right]^{\frac{1}{2}} \right)} \tag{1}$$

The frequency of events f_i in each i^{th} bin is multiplied by the mean precipitation rate of the bin p_i: $C_i = f_i \, p_i$. This provides the actual contribution C_i of the bin to the mean precipitation rate. The sum across all bins (area under the curve) gives the mean precipitation rate. The resulting spectrum is shown in Figure 2c. It provides information about the relative contribution of each bin to the mean.

Further dividing each bin's actual contribution by the mean precipitation rate (sum across all bins of the actual contribution spectrum), as shown in Eq. (2) gives a spectrum which area under the curve is unity (fractional contribution, panel d), providing information mostly about the shape of the distribution, independently from the mean precipitation.

$$FC_i = \frac{C_i}{\sum C_i} = \frac{C_i}{mean} \tag{2}$$

This method provides a quantitative visualisation of model differences or biases against observations in all parts of the precipitation distribution, and not only in the head or tail of the distribution like more traditional approaches (probability distribution function, cumulative distribution function). As an example, two spectrums are plotted: a reference spectrum and a model spectrum for the actual contribution (Fig. 2e) and the fractional contribution (Fig. 2f). Their difference is plotted in panels g and h, respectively. These figures illustrate two facts: first of all, the model shows a dry bias compared to the reference: the area under the red curve in panel e is smaller than the area under the blue curve, which is more easily seen in panel g, where the negative area between the curve and the zero line is larger than the positive area. More importantly, it illustrates which bins contribute to the mean bias: the model shows mainly too much precipitation from the intensities below 8 mm/day but a stronger underestimation of the contribution from events between 8 and 100 mm/day. This latter contribution has the largest effect on the mean as the sum of all bins is negative. Panels e and g therefore mix the information between the mean bias and the shape of the distribution.

Looking at the fractional contributions in panels f and h, they mainly illustrate the differences in the shape of the distribution between the model and the data. By construction, the integral of the difference between the two curves is zero: the positive and negative grey areas in panel h compensate each other. These figures mainly show that the lower intensity bins contribute too much to the mean compared to the higher intensity means. It loses the information about the differences in the means of the models. In this case, actual contribution and fractional contributions are not very different, but it is easy to think about a model which would have the right shape of fractional distribution but too much precipitation coming from all the bins: the actual contributions would be larger and the mean bias positive, but the fractional contribution would be similar as in the observations. These contributions are calculated at each grid-point and can then be averaged over a given region or maps can be shown by aggregating the contributions over several bin categories.

On top of this method, we build indices to summarize information about the shape of the distribution and the mean precipitation differences between datasets to serve two purposes:

- identify the regions, seasons and timescales where the mean precipitation and the shape of the precipitation distribution are most different between the 12 km and the 2.2 km
- in these cases, identify whether the 2.2 km models provide an overall better or worse representation of the contributions of different intensities to mean precipitation where observations are available.

The first index gives information about how much the fractional contributions differ between a model (mod) and a reference dataset (ref): the index FC (Fractional Contribution Index) is given in Eq. (3).

$$FC(mod, ref) = \sum_i |FC_i^{mod} - FC_i^{ref}| \tag{3}$$

FC represents how different the shapes of the two distributions are independently from the differences in the means. It has no units either but is the area between the two fractional contribution spectrums (in grey in (Figure 2f and h). Its minimum and best value is zero while the maximum is two and means no overlap between the distributions.

The second index assesses which model (between $m1$ and $m2$) performs best in terms of fractional contributions to the mean (Eq. (4)).

$$FC_{best}(m1, m2) = \frac{FC(m1, obs) - FC(m2, obs)}{FC(m2, obs)} x100 \tag{4}$$

This index measures the percentage of improvement or worsening of the fractional spectrum of model $m1$ over model $m2$ with regards to the observations. If $m2$ agrees better with the observations, the index is positive (the area between $m1$

and *obs* is larger than the area between $m2$ and obs) and the index is negative if $m1$ agrees better. The index gives some credit to a model which has a better fractional contribution but a worse bias, meaning that it would potentially reproduce well the underlying physical processes but just do too much of all of them.

The indices are calculated at each grid-point and then averaged over regions or presented as maps. With these score we require the models not only to capture the area mean, but also each grid point accurately.

4 Regions and seasons of largest difference across resolution

4.1 Comparing 2.2 km with 12 km models

In this section, we identify where and for which season the 2.2 km hourly precipitation statistics differ most from the 12 km ones. We use a combination of the absolute mean difference ratio: AMD=(|mean(2.2 km) - mean(12 km)|)/mean(12km) and the fractional contribution index (FC) presented in Sect. 3, calculated between the 2.2 km and the 12 km models. Calculated at each grid point, these measures are then averaged over the different domains presented in Fig. 1a. Only points with average precipitation above 0.03mm $h^{-1}$for hourly precipitation and 0.5mm $day^{-1}$for daily precipitation are taken into account, to avoid regions with too little precipitation which do not have a robust precipitation spectrum. The results are not very sensitive to these chosen thresholds (not shown).

Each domain can be quite vast but this region definition is chosen as a first order description of the variability of climates using a limited number of regions, inspired by the Köppen-Geiger map of climates (Peel et al 2007). Figure 3 shows a plot of FC(2.2 km, 12 km) as a function of AMD for hourly precipitation. The higher the value of FC, the larger the differences in shape of the distribution between the 2.2km and 12km models and the larger the AMD, the larger the difference in means of the two models.

For both model sets and for each individual region, the 2.2 km models differ most from the 12 km models in summer in terms of precipitation distribution (FC). In terms of differences in the mean (AMD), it is largest over high orography in all seasons (>1500 m), and it is high in the Mediterranean in summer too. Note however that the 'Med Sea' and 'Med coast' points for summer are not as reliable since they are based on a limited number of points because mean precipitation is very low.

A second conclusion which can be drawn is that the UKMO is overall more sensitive to the changes in resolution than the ETH model for most regions and seasons, comparing the right and left panels. This may partly be due to the fact that the UKMO 12 km model and 2.2 km model do not have exactly the same model physics.

Another point is that in all seasons, the largest differences between resolutions in terms of shape of the distribution (FC) are found over the Mediterranean sea or coasts, these points especially stand out compared to the other regions in autumn and summer.

In most seasons and models, the smallest sensitivity to resolution is found in flat lands in Northern Europe and Central Europe except in summer in the UKMO where these regions show large differences in the shape of the distribution. Most of the differences occur in the shape of the distribution and not in the mean state except in summer in the UKMO. In the ETH model, CEurM (orography between 500 m and 1500 m in Central Europe) shows more sensitivity to resolution than flat lands in all seasons in both the mean and shape of the distribution.

Guided by these findings, we will mainly focus the rest of the study on the summer season in all regions and the Mediterranean coasts and sea in autumn, where the differences between models are largest.

4.2 Model performances against available observations

We show the mean bias compared to observations and a comparison of the fractional distribution differences (FC best (2.2 km, 12 km)) at each grid-point on a map for daily (Fig. 4)) and hourly precipitation (Fig. 5).

Regarding daily precipitation differences, Fig. 4 first highlights that the Alpine region stands out as a region of large increase in the mean precipitation in the two models, as highlighted in the previous part. The bias increases with height above 800 m in both of the 2.2 km models (Fig. 5 of the supplementary materials) and areas above 1500 m in this region show a wet bias of around 30-70%. Although this region tends to be more biased in the 2.2 km mean, it shows a better performance for the distribution in the western part of the mountain range and worse in the northeastern part (panels c and f). The wet bias partly comes from an overestimation of wet days for all intensities, which was quantified as an increase by 10-30% of the wet-day frequency (not shown). It should be stressed that the observations over high ground may underestimate precipitation by at least 10% as discussed in Sect. 8.1.

A second point which can be made is the overall improvement in the shape of the fractional contribution to the mean in both 2.2km models south of the Alps and to a lesser extent north of the range. The improvement is about 30-50% compared to the 12 km performances. This is associated with a smaller mean bias in the ETH 2.2 km in this region. The UKMO is however dominated by a dry bias in this region, although there are improvements in the mean bias in Liguria.

Northern Germany, the Netherland and the UK coasts are also regions of improvement in both 2.2 km models. The other regions do not show any clear improvement in the distribution between the 2.2 km and the 12 km models.

The mean bias is not very different across resolutions in the ETH model, and is worse in the UKMO 2.2 km model with an overall dry bias of 20-50% in northern Italy, northern Spain, France and western Germany. The fact that there is not much of a resolution-dependence in the model skill in capturing the shape of the distribution, but a large dependence in the mean indicates that the dry bias in the UKMO 2.2 km mostly comes from a reduction in the overall wet day frequency, which was quantified as being around 20%.

Regarding the mean and shape of the distribution for hourly precipitation, Fig. 5 shows similar mean biases as for daily precipitation, which is reassuring given that the reference datasets are different and the time-period of comparison is not the same (Table 1). The signal in FC_{best}, showing which model shows the best overlap with the observation in terms of fractional contributions is now much stronger than for daily data. There is a clear improvement by the 2.2 km models in terms of which intensities contribute to the mean for Switzerland in both models, especially at higher altitudes for the UKMO. In Germany, the overall tendency is to a worsening of the distribution in the 2.2 km models, especially strong on a southwest-northeast diagonal. A common improvement is however found in north and northwestern Germany. In the UK, the model performance is very spatially dependent and there is mostly a tendency of improvement along the coasts of the Irish Sea and of a deterioration inland.

Overall, the 2.2 km models improve the daily and hourly distribution shape over the western Alps but show a tendency of having too many wet-days in the high-grounds, although the raingauge under-catch is hard to evaluate in this region. They seem to deteriorate the hourly distribution on flat land away from the coasts in the UK and Germany. The UKMO has also an overall dry bias linked with too few wet hours and days in France, Spain and northern Italy.

5 Shift to shorter and more intense wet-spell intensities in 2.2 km models

5.1 Shift to larger contributions from moderate and intense precipitation

We now look further into the distributions to evaluate which parts are most affected by the changes in resolution. We focus on hourly distributions, since the differences are clearer at this scale.

Both the differences in fractional and actual contributions against observations are shown in Figure 6. They illustrate the very different behaviour of the 2.2 km models compared to the 12 km models on the hourly time-scales. The 12 km models tend to show a too large contribution to total precipitation from low-intensity events (below 2-3 mm/h) by 5 to 40% depending on countries, which is can be over-corrected in the 2.2 km, which tend to have too much rainfall contributed by moderate and intense (3-30mm/h) events by 10 to 40%. There is a significant improvement in Switzerland and to a lesser extend in the UK in terms of fractional contributions but the 2.2 km ETH model overestimates precipitation in all bins in terms of actual contributions. In Germany, the 12 km models already have too large a contribution from intense events (>8 mm/h) by around 40% and the 2.2 km models have even larger contributions from events above 2 mm/h, resulting in a 40% increase in contributions from intensities above 2 mm/h. It increases the distribution biases against observations in this country. In both models, this is due both to a decrease in the actual contribution of low-intensity events and an increase in the moderate events. The decrease in actual contribution from low-intensity events is larger in the UKMO than in ETH and results in biases of -8 to -34% in the UKMO 2.2 km model to -12 to -15% in ETH (in the UK and Germany only) for this range of intensities.

Although the number of hourly datasets available is limited, the 12 km and 2.2 km model contributions to total rainfall can be compared on the whole domain by plotting maps of the fractional contribution to total rainfall from low intensity events (<2 mm/h), moderate events (2-8 mm/h) and intense events (>8 mm/h), as shown in Figure 7. These maps show that the shift in contribution of precipitation from low to moderate and intense precipitation in both 2.2 km models is present everywhere on land and is much larger than differences between the 12 km models.

This leads to improvement in Switzerland and to a lesser extend in the UK but to larger biases in Germany.

5.2 Analysis of wet-spell durations and intensities

Figure 8 presents the distribution of hourly wet-spell frequencies by duration (in hours) and mean intensity over the wet-spell for the available observation datasets and the four models. A wet spell is defined as consecutive hours with precipitation rates larger than 0.1 mm/h at a single grid-point. For the observations, the wet spell frequency is shown and for the models we show the difference in the number of wet spells per year in each intensity/duration bin between the model and the observations normalised by the number of wet spells per year in the observations. This way, a positive difference between model and observations in a given bin reflects an overestimation of wet spells in this bin, not just a larger share of this bin in the wet-spell distribution. We also show percentage differences in the number of wet-spells against the observations in each panel title.

In all three countries, the 2.2 km models increase the frequency of short-lasting (<10 h) moderate to intense (average intensity of 1-20 mm/h) events and decrease the share of long-lasting (>5 h) low-intensity (<1 mm/h) wet spells compared to 12 km models. The latter effect is especially strong in the UKMO 2.2 km model. As a result, the 2.2 km tend to underestimate the long-lasting weak wet-spells contrary to the 12 km models which overestimate them: the 2.2 km models yield better results for these events in all countries for the ETH2.2 km and only in Switzerland and to a lesser extent the UK for the UKMO2.2 km. The total number of wet-spells generally decreases from 12 to 2.2 km, the effect is more pronounced in the UKMO 2.2 km due to the former point. The short-lasting moderate to intense wet-spells tend to be underestimated in the 12 km models and overestimated by the 2.2 km models (except in the UK). Improvement for these high-impact events occur for the UK and Switzerland (only for the ETH model).

The ETH 2.2 km model also decreases the occurrence of short-lasting wet-spells whereas the UKMO 2.2 km increases these occurrences compared to the 12 km model: in this model, low-intensity wet spells become shorter.

Note that the UKMO 12 km model shows intense and very short-lasting (¡3 h) wet spells, in disagreement with the German and Swiss datasets but not the British one, this is probably due to grid-point storms.

5.3 Changes in the tail of the precipitation distribution

Looking at the representation of intense events, Figure 7 shows a larger contribution from intense events (>8 mm/h) to total precipitation in the 2.2 km models, especially in the ETH 2.2 km where these events can represent up to 20% of the mean, as also shown in Figure 6. The average top 1% of all hours shown in Figure 9 shows that the increase in contribution from the moderate and intense events in the 2.2 km models is partly due to more intense hourly rainfall in both models and not only linked with a decrease in number of low-intensity hours. This is again an improvement for Switzerland and the UK and a deterioration for Germany, where this index is overestimated by 10-30% in the UKMO2.2 km and 10-50% in the

ETH 2.2 km. This is not the case for daily precipitation on flat land where this diagnostic does not show a large intensification (not shown).

5.4 Diurnal cycle in summer

Finally, Figure 10 and Figure 11 respectively show the amplitude and the phase (hour of the maximum precipitation in local time) of the mean diurnal cycle at each grid-point. They show stronger amplitudes in the 2.2 km models over high orography (>1500m) compared to the 12 km models, especially in the Swiss, Austrian and north-italian Alps. According to the Swiss and German datasets over the Alps, this is an improvement, although the amplitude may tend to be too strong in the convection-permitting models. Both models also generally show larger amplitudes of the diurnal cycle on lower level topography (Massif Central, Appenines, Dinaric Alps), where MCSs are often triggered (Morel and Senesi 2002). The UKMO and to a lesser extent the ETH 2.2 km models also reproduce the larger amplitude of the diurnal cycle in southern Germany along the Alpine foothills, where MCSs are observed (Hagen and Finke 1999; Kaltenböck 2004).

Figure 11 shows the better timing of the peak precipitation in the 2.2 km models, the peak being shifted from late morning-early afternoon in parameterised models to mid-late afternoon in the convection-permitting models, which is more realistic, in line with Fosser et al (2015); Ban et al (2014). It is worth noting that the UKMO-2.2 km still produces precipitation too early in the day in the Swiss Alps (around 2-4 pm), whereas the ETH-2.2 km model is in better agreement with the observations with a peak between 4 pm and 8 pm. Nisi et al (2016) observations are also more in line with the late peak of ETH 2.2 km in the Po valley. Generally, the UKMO 2.2 km modeL tends to produce earlier afternoon peaks by about 2 h than the ETH 2.2 km model, further away from the observations. Both models reproduce well the spatial gradients of the hour of maximum precipitation in the UK on the southwestern coasts.

6 Mediterranean heavy precipitation events

In autumn, the heaviest precipitation events in Europe occur on the Mediterranean coasts, as illustrated by the average of rainfall on the top 1% of all days shown in Fig. 12e. In this figure, we use daily CMORPH observations (2001-2008) (see Sect. 8.1) as a complement to the daily precipitation datasets for this metric. This satellite-derived product is not as reliable as daily observation products and not as high resolution (0.25°), but it provides some estimate of convection over the sea and in the regions not covered by high resolution datasets, although it was shown to underestimate coastal heavy precipitation events in this region (Stampoulis et al 2013). This can also probably be seen in the sharp transition between high values in Italy in the Appenines in the Alpine dataset and lower values in CMORPH.

Regions particularly hit by heavy precipitation events are the Valencian country in Spain, the southern part of the Massif Central (Cévennes) and the Alps in France, the Ligurian region in Italy, the whole southern edge of the Alps and the Dinaric Alps. Intense convection also occurs in the Gulf of Lions and the

Tyrrhenian Sea. Liguria, most of Italy and the Dinaric Alps were identified as regions with rather large convection-parameterised model biases in the extremes in convection-parameterised models (Berthou et al 2016; Cavicchia et al 2016; Fantini et al 2016).

6.1 Contribution of intense events to mean precipitation in autumn

Fig. 12 shows the $p99_{avg}$ metric for all the models. The two 2.2 km models seem to actually converge to a solution closer to the observations compared to the 12 km models which differ from each other. The convection-parameterised models have very different biases: the UKMO-12 km model shows very intense wet biases on the upslope side of all mountain ridges and on the coasts, while the ETH 12 km model underestimates this metric by around 30-50%. The ETH 2.2 km is in better agreement with the observations and the UKMO 2.2 km mostly shows stronger intensities in northern Italy. All models show stronger precipitation in the coastal Pyrenees compared to the observations. The 2.2 km show stronger precipitation in the Valencian country, in better agreement with the observations.

Over the sea, precipitation maximum in CMORPH occurs in the Gulf of Lions and the Thyrrenian Sea whereas it is maximum in the Ionian Sea in the UKMO 2.2 km and in the Thyrrenian Sea in ETH 2.2 km. Precipitation is more intense over the sea in each 2.2 km model compared to its 12 km counterpart. This suggests that convection is more easily triggered over the sea away from the influence of the orography or the coasts in the 2.2 km models.

6.2 Case study: 8-9 Sept. 2002 in Southern France

Having examined the climatological differences between the 12km and the 2.2km models, we now focus on a single case study to illustrate how processes are represented differently across resolution. The chosen case is a Mediterranean heavy precipitation event which occurred on the 8^{th} and 9^{th} Sept. 2002 in the Gard region in Southern France. This case was chosen for three main reasons: first, it is well documented (Delrieu et al 2005; Anquetin et al 2005; Nuissier et al 2008; Ducrocq et al 2008b). Second, it was strongly forced synoptically (Nuissier et al 2008) so we can expect it to be present in the climate models (which only receive atmospheric information on the observed state at the lateral boundary) and third, cold pool interactions with the mesoscale environment played an important role in setting the location and intensity of the event, so we may expect the 2.2km models to behave differently from the 12km models (Ducrocq et al 2008b).

Over the two days of the event, maximum rainfall of 600-700 mm was recorded (Fig. 13e). The meteorological environment of the heavy rainfall event was characterized by an upper-level trough centred over Ireland and extending meridionally to the Iberian peninsula, progressively veering to a northwest, southeast axis. It generated a south-westerly diffluent flow over south-eastern France. An associated surface cold front, first located over western France, moved progressively eastward. Convection first formed well ahead of the front in the warm sector, where a low-level south-easterly flow prevailed and was later reinforced by embedded convection in the front. Fig. 13 shows that for both 12km models maximum precipitation falls on the southeast facing slopes of the Cévennes. In both 2.2km models, precipitation occurs both on the slopes of the Cévennes and in the Rhone valley, the latter being where the maximum in the observations is found. All models underestimate the precipitation in the Rhone Valley, but the 2.2km models have smaller negative biases.

The UKMO climate models show different time-evolutions of the surface cold front and first generate precipitation over orography, in association with a strong temperature gradient, on the afternoon of the 8th (this differs from the real event which already shows cold pools and precipitation in the valley by the afternoon of 8^{th}, not shown). The 500 hPa synoptic situation is closer to ERA-interim in the 2.2 km model than in the 12 km model, probably as a result of domain size (not shown). The UKMO 12 km model mostly shows orographic precipitation and convection embedded in the cold front during the whole event. In the UKMO 2.2 km model, following the triggering of precipitation over orography, convection-induced cold air accumulates in the Rhone valley, leading to the formation of a mesoscale cold front. By the morning of the 9^{th}, convective cells are triggered on the edge of the cold pool (Fig. 14) which gradually propagates upstream of a 50-60 knot southerly flow, maintaining convective cells in the valley in the 2.2 km model. There is no hint of interaction of the flow with a cold pool at any stage of the event in the UKMO 12 km model (not shown). The more realistic positioning of the rainfall maximum, and higher rainfall totals, in the 2.2km models therefore seems to be related to their ability to represent cold pools and some form of organised convection. Given this is just a single case study, and we would not expect the timing or position of rainfall to be exactly captured across models, it is not possible to make any definite conclusions. However, the results are illustrative of the potential for improved representation of mesoscale processes and associated extreme precipitation events at convection-permitting resolution.

7 Discussion and conclusion

This first intercomparison pan-European CPMs confirms and builds on previous studies on smaller domains or with single models. Quantitatively we find that the largest precipitation differences between CPMs and 12 km parameterised models occur at hourly time-scales in summer in most regions. Regions of high topography show the largest differences in mean precipitation at the convection-permitting scales and the Mediterranean coasts and sea are most affected in terms of precipitation distribution, especially in summer and autumn.

The two pan-european CPMs behave similarly in terms of differences in precipitation distribution at the hourly timescale in summer compared to 12 km models. Mean precipitation comes from an increased contribution of short-lasting moderate and intense events and a decreased contribution of longer lasting low-intensity events everywhere. This leads to an overall improvement compared with the 12km models in Switzerland (also found in Ban et al (2014); Lind et al (2016)) and parts of the UK (also in Kendon et al (2012)) but deteriorates the distribution in most of Germany with too much moderate and intense precipitation, unlike the findings of Fosser et al (2015) who evaluated their model against hourly raingauges in Southwestern Germany. The lack of low-intensity events in both models is especially large in the UKMO 2.2 km model and is responsible for a 10-30% dry bias in France, Spain and Italy in this model.

The daily precipitation distribution is mostly affected by resolution changes in the Alps, in northern Italy and near the coasts (UK/Germany). The Austrian Alps show a deterioration of the distribution while the southwestern Alps and northern Italy benefit from higher resolution. Mean precipitation is increased over the Alps and becomes larger than in the observations. This bias increases with height above 800 m in both 2.2 km models and it is unclear which part is due to observation uncertainties or model deficiencies ((Lind et al 2016) yield similar results). Mediterranean intense events in autumn at the daily scale are better represented by the 2.2 km models, which converge to a solution closer to the observations in terms of location and intensity than their 12 km counterparts.

The phase of the diurnal cycle is better represented in the CPMs but the UKMO-2.2 km has still too early a peak over orography. This is a well-known improvement in CPMs due to the fact that convective instability takes more time to build-up as it is not consumed by parameterised convection which tends to start convection around midday (Kendon et al 2012; Prein et al 2013; Ban et al 2014; Fosser et al 2015). Both CPMs have an enhanced amplitude over orography compared to the 12 km models, which is an improvement.

Regarding model differences, the UKMO-2.2 km has a much reduced wet-day frequency compared to the UKMO-12 km, which is a clear bias compared to the observations; this is not the case in the ETH model. It is not clear whether it comes directly from resolution changes. One of the model differences that we investigated is the the way saturated layers of soil are treated. At higher resolution, when the top layer of soil is saturated, excess water disappears into the surface run-off whereas it is drained into the second layer in the 12km. Initial sensitivity tests have shown that modifying the treatment of saturated layers moistens the lower soil layers slightly, but has negligible impact on the surface soil moisture (not shown) and the surface climate (supplementary material). We note that the impact of soil moisture infiltration rates being too low in the UKMO models, due to the use of Van-Genuchten hydraulic equations, may impact the 2.2 km model differently to the 12 km model, given in the former rainfall is more intense and hence the surface layer is more likely to become saturated. Initial tests, however, suggest the impact of changing the hydraulic equations on the surface temperature is small, with warm/dry biases in the UKMO-2.2km persisting. Thus it is possible that the intense/intermittent nature of rainfall in the 2.2km model is responsible for dry soil conditions and associated warm temperature bias over Eastern Europe but further work looking at more variables such as the work of Brisson et al (2016) on clouds is needed. In the ETH model such an effect is less apparent, possibly due to the use of a shallow convection parameterisation in this model. Other regions such as the UK are less sensitive, as soil moisture is not close to critical value for limiting evaporation. It should be noted that Liu et al (2016) using ERA-interim driven WRF 4 km simulations also show a warm and dry bias in the Central US in 13-year long simulations over the US.

In this study we have shown that two 2.2km convection-permitting models yield qualitatively similar differences to the precipitation climatology compared to 12 km models, despite using different dynamical cores and different parameterization packages. Its also highlights that both convection-permitting models will need to address how to better balance the increased number of moderate to intense events and the decreased number of low-intensity events, which are needed to improve the 12 km model hourly distributions but are overcompensated in both models. Work is on-going to introduce a scale-aware convection parameterisation in future model versions of the UKMO, which would enable some sub-grid convection. Work on the boundary layer scheme and its coupling with convection is also on-going.

This intercomparison study would benefit from the availability of new generations of hourly precipitation datasets. Future work will examine whether there are similarly robust signals of future precipitation change across different CPMs, reducing uncertainty in projections of intense events at hourly and km-scales. To this end, the CORDEX-Flagship pilot study on CPMs is a promising initiative, allowing comparison of more CPMs beyond the two available for analysis here.

8 Appendix

8.1 Daily datasets

FRANCE: SAFRAN (8km) Systeme d'Analyse Fournissant des Renseignements Atmospheriques á la Neige (SAFRAN) is a precipitation analysis for continental France that uses an optimal interpolation method. One of the main features of SAFRAN is that the analyses are performed over climatically homogeneous zones, which are areas of irregular shape covering a surface usually smaller than 1000 km and where the horizontal climatic gradients (especially for precipitation) are weak. SAFRAN estimates one value of each parameter for each zone at several altitude levels. Within the zone, analyzed parameters depend only on elevation and aspect. First, SAFRAN performs a quality control of the observations. This is an iterative procedure based on the comparison between observed and analyzed quantities at the observation location. There were 3675 measurement stations for 2004/2005. The precipitation analysis is performed daily at 0600 UTC, to include in the analysis the numerous rain-gauges that measure precipitation on a daily basis (in particular in the climatological and snow networks). The first guess is a very simple and constant field. An hourly separation is then performed, but in this study we use the daily precipitation amount. Further description can be found in Quintana-Segui et al (2008).

ALPS: APGD_EURO4M (5km) The Alpine rain-gauge dataset typically comprises 5500 observations on any day of the period 1971 - 2008. The analysis is based on a first guess for a day that is the long-term mean precipitation (period 1971 - 1990) of the relevant calendar month. The precipitation-elevation relationship is calculated locally and taken into account in this first guess. Then an anomaly is computed for every grid point using the stations located within a radius that depends on the station density. It can be up to 60 km from the grid point. The dataset has a 5km resolution, but its effective resolution is closer to 10-15km. The dataset is provided by the Federal Office of Meteorology and Climatology MeteoSwiss. The dataset incorporates local precipitation topography relationships at the climatological time-scale, which aims at reducing the risk of systematic underestimates at high elevations but does not correct for any gauge undercatch, which is comparatively larger during episodes with strong wind and during weather with low rainfall intensity or with snowfall. Sevruk and Zahlavova (1994) and Richter (1995) estimated measurement errors ranging from 7% (5%)

over the flatland regions in winter (summer) to 30% (10%) above 1500m in winter (summer). Further description can be found in Isotta et al (2014).

SPAIN: Spain02 (0.11°) Daily precipitation gridded dataset developed for peninsular Spain and the Balearic Islands using 2756 quality-controlled stations over the time period from 1971 - 2010 (Herrera et al 2012). The grid was produced by applying the kriging method in a two-step process. First, the occurrence was interpolated using a binary kriging and, in a second step, the amounts were interpolated by applying ordinary kriging to the occurrence outcomes. The elevation is not explicitly included in the development of the dataset because the available dense gauge network represents the orography corresponding to the 0.11° grid appropriately. Explicit comparison of Spain02 with the E-OBS dataset shows better performance of Spain02 to represent extreme events of daily precipitation in the region of Valencia regarding the amount and spatial distribution of precipitation (Herrera et al 2012).

UK: UKCPOBS (5km) The National Climate Information Centre daily UK gridded precipitation dataset (Perry et al 2009) spans the period 1958-present day, and from 1990 uses approximately 2500-3500 surface gauge observations. Quality control is performed through computerized and manual comparisons of individual daily station values against the daily all-station average and daily values from nearby stations. Any stations that have failed quality control are excluded from the computation of the gridded values. The gridding of the gauge data to a 5km×5km grid uses a cubic inverse-distance weighting interpolation using stations within 50km radius of the grid box.

CMORPH 1.0 (0.25°) The CMORPH (NOAA Climate Prediction Center morphing method, Joyce et al (2004)) algorithm uses the relatively high-resolution IR information to infer the hydrometeorological position between two consecutive PMW estimates. IR maps are used to derive cloud system advection vectors (CSAVs) to propagate PMW rainfall estimates. Such propagation is performed forward and backward for each time step using information provided by the CSAVs. Final values are achieved by averaging forward and backward rainfall analyses proportionally to step distance.

8.2 Hourly datasets

Nimrod (UK): Gridded hourly radar data for the UK at 5km resolution are available from the Nimrod database (Golding 1998) for the period 2003-present-day. There are many issues with radar (clutter, anaprop, bright band, beam attenuation), and in particular radar data are known to systematically underestimate heavy rainfall amounts. The Met Office calibrates radar against rain gauges and employs algorithms to take account of known issues but some problems cannot be fully rectified. One of these is that the hourly gauges used in the calibration are relatively sparse, and thus are not able to fully correct for locally-varying effects such as attenuation.

Germany: The hourly precipitation data set assembled by Paulat et al (2008) is used. It features a horizontal grid spacing of 7 km and an effective horizontal resolution of 14-28 km. The time period of the dataset is 2001-2008 (8 years). To assemble this dataset, measurements from rain gauges have been gridded as daily sums, following the procedure by Frei and Schär (1998). Afterwards, the daily sums were disaggregated into hourly values using rain rate retrievals from radar (Wuest et al 2010). Beyond uncertainties arising from rain-gauge undercatch, gridding procedures (Frei et al 2003), and weather radar measurements (Wuest et al 2010), possible inconsistencies between gauge observations and radar restricts the data set to 92% of the possible days, at the respective grid points (Paulat et al 2008).

Switzerland: RdisaggH is an experimental precipitation data set for Switzerland which provides gridded, radar-disaggregated rain-gauge observations (Wuest et al 2010). In order to obtain hourly data, a gridded daily product was disaggregated into hourly sums, using information from weather radar fields. The resulting dataset has a grid-spacing of $0.01^{\circ} \times 0.01^{\circ}$ covers Switzerland and is available for the time period May 2003-2010.

Acknowledgements The authors acknowledge the HyMeX data-base teams (ESPRI/IPSL and SEDOO/Observatoire Midi-Pyrns) for their help in accessing the SAFRAN dataset, H. Wernli for accessing the German hourly dataset and MeteoSwiss for the Swiss dataset. The authors also thank AEMET and UC for providing access to the Spain02 dataset, http://www.meteo.unican.es/datasets/spain02. The work of the ETH group was supported by the Swiss National Sciences Foundation through the Sinergia grant $CRSII2_154486$ ‘crCLIM. E.J. Kendon gratefully acknowledges funding from the Joint Department of Energy and Climate Change (DECC) and Department for Environment Food and Rural Affairs (Defra) Met Office Hadley Centre Climate Programme (GA01101). This work also forms part of the European Research Council funded INTENSE project (ERC-2013-CoG-617329; grant holder and PI: Hayley J Fowler, Newcastle University). We thank the reviewers for helping to clarify the paper.

References

Anquetin S, Yates E, Ducrocq V, Samouillan S, Chancibault K, Davolio S, Accadia C, Casaioli M, Mariani S, Ficca G (2005) The 8 and 9 september 2002 flash flood event in France: a model intercomparison. Nat Hazards Earth Syst Sci 5(5):741–754

Arakawa A, Lamb VR (1977) Computational design of the basic dynamical processes of the UCLA general circulation model. Methods Comput Phys 17:173–265

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. Mon Weather Rev 139(12):3887–3905, DOI 10.1175/MWR-D-10-05013.1

Ban N, Schmidli J, Schär C (2014) Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. J Geophys Res 119(13):7889–7907, DOI 10.1002/2014JD021478

Barthlott C, Davolio S (2016) Mechanisms initiating heavy precipitation over italy during hymex special observation period 1: a numerical case study using two mesoscale models. Q J R Meteorol Soc 142:238–258, DOI 10.1002/qj.2630

Berthou S, Mailler S, Drobinski P, Arsouze T, Bastin S, Béranger K, Flaounas E, Lebeaupin Brossier C, Stéfanon M (2016) Influence of submonthly air-sea coupling on heavy precipitation events in the western mediterranean basin. Q J R Meteorol Soc DOI 10.1002/qj.2717

Best MJ, Pryor M, Clark DB, Rooney GG, Essery RLH, Ménard CB, Edwards JM, Hendry MA, Porson A, Gedney N, Mercado LM, Sitch S, Blyth E, Boucher O, Cox PM, Grimmond CSB, Harding RJ (2011) The joint uk land environment simulator (jules), model description - part 1: Energy and water fluxes. Geosci Model Devel 4:677–699, DOI 10.5194/gmd-4-677-2011

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrol Sci B 24:43–69

Boberg F, Berg P, Thejll P, Gutowski WJ, Christensen JH (2009) Improved confidence in climate change projections of precipitation evaluated using daily statistics from the PRUDENCE ensemble. Clim Dyn 32:1097–1106

Böhm U, Kücken M, Ahrens W, Block A, Hauffe D, Keuler K, Rockel B, Will A (2006) Clm– the climate version of lm: brief description and long-term applications. COSMO Newsl 6:225–235

Bott A (1989) A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes. Mon Weather Rev 117(5):1006–1015, DOI 10.1175/1520-0493(1989)117¿1006:APDASO¿2.0.CO;2

Boutle IA, Eyre JEJ, Lock AP (2014) Seamless stratocumulus simulation across the turbulent gray zone. Mon Weather Rev 142:1655–1668, DOI 10.1175/MWR-D-13-00229.1

Bresson E, Ducrocq V, Nuissier O, Ricard D, de Saint-Aubin C (2012) Idealized numerical simulations of quasi-stationary convective systems over the northwestern Mediterranean complex terrain. Q J R Meteorol Soc 138(668):1751–1763, DOI 10.1002/qj.1911

Brisson E, Van Weverberg K, Demuzere M, Devis A, Saeed S, Stengel M, van Lipzig NPM (2016) How well can a convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics? Clim Dyn 47(9):3043–3061, DOI 10.1007/s00382-016-3012-z

Brockhaus P, Lüthi D, Schär C (2008) Aspects of the diurnal cycle in a regional climate model. Meteorol Z 17:433–443, DOI 10.1127/0941-2948/2008/0316

Bryan GH, Morrison H (2012) Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. Mon Weather Rev 140(1):202–225, DOI 10.1175/MWR-D-11-00046.1

Cavicchia L, Scoccimarro E, Gualdi S, Marson P, Ahrens B, Berthou S, Conte D, Dell'Aquila A, Drobinski P, Djurdjevic V, Dubois C, Gallardo C, Li L, Oddo P, Sanna A, Torma C (2016) Mediterranean extreme precipitation: a multi-model assessment. Clim Dyn DOI 10.1007/s00382-016-3245-x

Chan SC, Kendon EJ, Fowler HJ, Blenkinsop S, Ferro CAT, Stephenson DB (2013) Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? Clim Dyn 41(5-6):1475–1495, DOI 10.1007/s00382-012-1568-9

Chan SC, Kendon EJ, Fowler HJ, Blenkinsop S, Roberts NM, Ferro CAT (2014) The value of high-resolution Met Office regional climate models in the simulation of multi-hourly precipitation extremes. J Clim 27(16):6155–6174, DOI 10.1175/JCLI-D-13-00723.1

Charney JG, Phillips NA (1953) Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. J Meteorol 10:71–99, DOI 10.1175/1520-0469(1953)010¡0071:NIOTQG¿2.0.CO;2

Clark DB, Mercado LM, Sitch S, Jones CD, Gedney N, Best MJ, Pryor M, Rooney GG, Essery RLH, Blyth E, Boucher O, Harding RJ, Huntingford C, Cox PM (2011) The joint uk land environment simulator (jules), model description - part 2: Carbon fluxes and vegetation dynamics. Geosci Model Devel 4:701–722, DOI 10.5194/gmd-4-701-2011

Clark P, Roberts N, Lean H, Ballard SP, Charlton-Perez C (2016) Convection-permitting models: a step-change in rainfall forecasting. Meteorol Appl 23:165–181, DOI 10.1002/met.1538

Dauhut T, Chaboureau JP, Escobar J, Mascart P (2015) Large-eddy simulations of hector the convector making the stratosphere wetter. Atmos Sci Lett 16(2):135–140, DOI 10.1002/asl2.534

Delrieu G, Nicol J, Yates E, Kirstetter PE, Creutin JD, Anquetin S, Obled C, Saulnier GM, Ducrocq V, Gaume E, Payrastre O, Andrieu H, Ayral PA, Bouvier C, Neppel L, Livet M, Lang M, Du-Châtelet JP, Walpersdorf A, Wobrock W (2005) The catastrophic flash-flood event of 89 september 2002 in the gard region, france: A first case study for the cévennesvivarais mediterranean hydrometeorological observatory. J Hydrometeorol 6(1):34–52, DOI 10.1175/JHM-400.1

Done J, Davis CA, Weisman ML (2004) The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model. Atmos Sci Lett 5:110–117

Drobinski P, Ducrocq V, Alpert P, Anagnostou E, Béranger K, Borga M, Braud I, Chanzy A, Davolio S, Delrieu G, Estournel C, Boubrahmi NF, Font J, Grubisic V, Gualdi S, Homar V, Ivancan-Picek B, Kottmeier C, Kotroni V, Lagouvardos K, Lionello P, Llasat MC, Ludwig W, Lutoff C, Mariotti A, Richard E, Romero R, Rotunno R, Roussot O, Ruin I, Somot S, Taupier-Letage I, Tintore J, Uijlenhoet R, Wernli H (2014) HyMeX, a 10-year multidisciplinary program on the Mediterranean water cycle. Bull Am Meteorol Soc 95:1063–1082, DOI 10.1175/BAMS-D-12-00242.1

Ducrocq V, Nuissier O, Ricard D, Lebeaupin C, Thouvenin T (2008a) A numerical study of three catastrophic precipitating events over southern France. II: mesoscale triggering and stationarity factors. Q J R Meteorol Soc 134(630):131–145, DOI 10.1002/qj.199

Ducrocq V, Nuissier O, Ricard D, Lebeaupin C, Thouvenin T (2008b) A numerical study of three catastrophic precipitating events over southern France. II: Mesoscale triggering and

stationarity factors. Q J R Meteorol Soc 134(630):131–145, DOI 10.1002/qj.199
Edwards JM, Slingo A (1996) Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. Q J R Meteorol Soc 122:689–720
Fantini A, Raffaele F, Torma C, Bacer S, Coppola E, Giorgi F, Ahrens B, Dubois C, Sanchez E, Verdecchia M (2016) Assessment of multiple daily precipitation statistics in era-interim driven med-cordex and euro-cordex experiments against high resolution observations. Clim Dyn pp 1–24, DOI 10.1007/s00382-016-3453-4
Flaounas E, Drobinski P, Borga M, Calvet JC, Delrieu G, Morin E, Tartari G, Toffolon R (2012) Assessment of gridded observations used for climate model validation in the Mediterranean region: the HyMeX and MED-CORDEX framework. Environ Res Lett 7(2):24,017, DOI 10.1088/1748-9326/7/2/024017
Förstner J, Doms G (2004) Runge-Kutta time integration and high-order spatial discretization of advection - a new dynamical core for the LMK
Fosser G, Khodayar S, Berg P (2015) Benefit of convection permitting climate model simulations in the representation of convective precipitation. Clim Dyn 44(1-2):45–60, DOI DOI 10.1007/s00382-014-2242-1
Frei C, Schär C (1998) A precipitation climatology of the Alps from high-resolution rain-guage observations. Int J Clim 18:873–900
Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: evaluation and intercomparison for the european alps. J Geophys Res 108(D3):4124, DOI 10.1029/2002JD002287
Fuhrer O, Osuna C, Lapillonne X, Gysi T, Cumming B, Arteaga A, Schulthess TC (2014) Towards a performance portable, architecture agnostic implementation strategy for weather and climate models. Supercomp Front Innov 1(1), DOI 10.14529/jsfi140103
van Genuchten MT, Leij FJ, Yates SR (1991) The RETC code for quantifying the hydraulic functions of unsaturated soils. Tech. Rep. EPA/600/2-91/065, U.S. Environmental Protection Agency
Golding BW (1998) Nimrod: A system for generating automated very short range forecasts. Meteorol Appl 5:1–16, DOI 10.1017/S1350482798000577
Gregersen IB, Sorup HJD, Madsen H, Rosbjerg D, Mikkelsen PS, Arnbjerg-Nielsen K (2013) Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. Clim Chang 118(3-4):783–797, DOI 10.1007/s10584-012-0669-0
Gregory D, Allen S (1991) The effect of convective downdraughts upon nwp and climate simulations. In: Ninth Conf. Numer. Weather Predict. Denver, Color., pp 122–123
Gregory D, Rowntree PR (1990) A mass-flux convection scheme with representation of cloud ensemble characteristics and stability dependent closure. Mon Weather Rev 118:1483–1506, DOI 10.1175/1520-0493(1990)118¡1483:AMFCSW¿2.0.CO;2
Gysi T, Osuna C, Fuhrer O, Bianco M, Schulthess TC (2015) Stella: A domain-specific tool for structured grid methods in weather and climate models. In: Proc. Int. Conf. HPC, Networking, Storage Anal., SC '15, pp 41:1—-41:12, DOI 10.1145/2807591.2807627
Hagen M, Finke U (1999) Motion characteristics of thunderstorms in southern germany. Meteorol Appl 6(3):227–239
Hanel M, Buishand TA (2010) On the value of hourly precipitation extremes in regional climate model simulations. J Hydrol 393:265–273, DOI 10.1016/j.jhydrol.2010.08.024
Hanley KE, Plant RS, Stein THM, Hogan RJ, Nicol JC, Lean HW, Halliwell C, Clark PA (2015) Mixing length controls on high resolution simulations of convective storms. Q J R Meteorol Soc 141(686):272–284, DOI 10.1002/qj.2356
Haylock MR, Hofstra N, KleinTank AMG, Klok EJ, Jones PD, New M (2008) A european daily high-resolution gridded dataset of surface temperature and precipitation. J Geophys Res 113:D20,119, DOI 10.1029/2008JD10201
Heise E, Ritter B, Schrodin R (2006) Operational implementation of the multilayer soil model, COSMO tech. rep., no. 9. Tech. Rep. 5, COSMO
Herrera S, Gutierrez JM, Ancell R, Pons MR, Frias MD, Fernandez J (2012) Development and analysis of a 50 year high-resolution daily gridded precipitation dataset over Spain (SPAIN02). Int J Clim 32:74–85, DOI 10.1002/joc.2256
Hofstra N, New M, McSweeney C (2010) The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. Clim Dyn 35(5):841–858, DOI 10.1007/s00382-009-0698-1
Hohenegger C, Brockhaus P, Schär C (2008) Towards climate simulations at cloud-resolving scales. Meteorol Z 17(4):383–394, DOI 10.1127/0941-2948/2008/0303

Isotta F, Frei C, Weilguni V, Perec Tadi M, Lasssegues P, Rudolf B, Pavan V, Cacciamani C, Antolini G, Ratto SM, Munari M, Micheletti S, Bonati V, Lussana C, Ronchi C, Panettieri E, Marigo G, Vertani G (2014) The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. Int J Clim 34:1657–1675, DOI 10.1002/joc.3794

Jacob D, Petersen J, Eggert B, Alias A, Christensen OB, Bouwer LM, Braun A, Colette A, Déqué M, Georgievski G, Georgopoulou E, Gobiet A, Menut L, Nikulin G, Haensler A, Hempelmann N, Jones C, Keuler K, Kovats S, Kröner N, Kotlarski S, Kriegsmann A, Martin E, van Meijgaard E, Moseley C, Pfeifer S, Preuschmann S, Radermacher C, Radtke K, Rechid D, Rounsevell M, Samuelsson P, Somot S, Soussana JF, Teichmann C, Valentini R, Vautard R, Weber B, Yiou P (2014) Euro-cordex: new high-resolution climate change projections for european impact research. Reg Environ Chang 14(2):563–578, DOI 10.1007/s10113-013-0499-2

Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) Cmorph: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J Hydrometeorol 5(3):487–503, DOI 10.1175/1525-7541(2004)005¡0487:CAMTPG¿2.0.CO;2

Kaltenböck R (2004) The outbreak of severe storms along convergence lines northeast of the alps. case study of the 3 august 2001 mesoscale convective system with a pronounced bow echo. Atmos Res 70(1):55–75, DOI https://doi.org/10.1016/j.atmosres.2003.11.003

Kendon EJ, Roberts NM, Senior CA, Roberts MJ (2012) Realism of rainfall in a very high-resolution regional climate model. J Clim 25(17):5791–5806, DOI 10.1175/JCLI-D-11-00562.1

Khodayar S, Fosser G, Berthou S, Davolio S, Drobinski P, Ducrocq V, Ferretti R, Nuret M, Pichelli E, Richard E, Bock O (2016) A seamless weatherclimate multi-model intercomparison on the representation of a high impact weather event in the western mediterranean: Hymex iop12. Q J R Meteorol Soc 142, DOI 10.1002/qj.2700

Kidd C, Bauer P, Turk J, Huffman GJ, Joyce R, Hsu KL, Braithwaite D (2012) Intercomparison of high-resolution precipitation products over northwest europe. J Hydrometeorol 13(1):67–83, DOI 10.1175/JHM-D-11-042.1

Kinne S, Schulz M, Textor C, Guibert S, Balkanski Y, Bauer SE, Berntsen T, Berglen TF, Boucher O, Chin M, Collins W, Dentener F, Diehl T, Easter R, Feichter J, Fillmore D, Ghan S, Ginoux P, Gong S, Grini A, Hendricks J, Herzog M, Horowitz L, Isaksen I, Iversen T, Kirkevåg A, Kloster S, Koch D, Kristjansson JE, Krol M, Lauer A, Lamarque JF, Lesins G, Liu X, Lohmann U, Montanaro V, Myhre G, Penner J, Pitari G, Reddy S, Seland O, Stier P, Takemura T, Tie X (2006) An AeroCom initial assessment optical properties in aerosol component modules of global models. Atmos Chem Phys 6(7):1815–1834, DOI 10.5194/acp-6-1815-2006

Kjellström E, Boberg F, Castro M, Christensen JH, Nikulin G, Sánchez E (2010) Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. Clim Res 44:135–150, DOI 10.3354/cr00932

Klingaman NP, Martin GM, Moise A (2017) Asop (v1.0): a set of methods for analyzing scales of precipitation in general circulation models. Geosci Model Dev 10(1):57–83, DOI 10.5194/gmd-10-57-2017

Kotlarski S, Keuler K, Christensen OB, Colette A, Déqué M, Gobiet A, Goergen K, Jacob D, Lüthi D, van Meijgaard E, Nikulin G, Schär C, Teichmann C, Vautard R, Warrach-Sagi K, Wulfmeyer V (2014) Regional climate modeling on european scales: a joint standard evaluation of the euro-cordex rcm ensemble. Geosci Model Dev 7(4):1297–1333, DOI 10.5194/gmd-7-1297-2014

Langhans W, Schmidli J, Fuhrer O, Bieri S, Schär C (2013) Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. J Appl Meteorol Clim 52:1490–1510, DOI 10.1175/JAMC-D-12-0167.1

Lapillonne X, Fuhrer O (2014) Using compiler directives to port large scientific applications to GPUs: An example from atmospheric science. Parallel Process Lett 24(1):1450,003 (18 pp.), DOI 10.1142/S0129626414500030

Lean HW, Clark PA, Dixon M, Roberts NM, Fitch A, Forbes R, Halliwell C (2008) Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. Mon Weather Rev 136:3408–3424, DOI 10.1175/2008MWR2332.1

Lenderink G (2010) Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. Clim Res 44(2-3):151–166, DOI 10.3354/cr00946

Leutwyler D, Fuhrer O, Lapillonne X, Lüthi D, Schär C (2016) Towards european-scale convection-resolving climate simulations. Geosci Model Devel DOI 10.5194/gmd-2016-119

Leutwyler D, Lüthi D, Ban N, Fuhrer O, Schär C (2017) Evaluation of the convection-resolving climate modeling approach on continental scales. J Geophys Res Atmos pp n/a—-n/a, DOI 10.1002/2016JD026013

Lind P, Lindstedt D, Kjellström E, Jones C (2016) Spatial and temporal characteristics of summer precipitation over central europe in a suite of high-resolution climate models. J Clim 29(10):3501–3518, DOI 10.1175/JCLI-D-15-0463.1

Liu C, Ikeda K, Rasmussen R, Barlage M, Newman AJ, Prein AF, Chen F, Chen L, Clark M, Dai A, Dudhia J, Eidhammer T, Gochis D, Gutmann E, Kurkute S, Li Y, Thompson G, Yates D (2016) Continental-scale convection-permitting modeling of the current and future climate of north america. Clim Dyn pp 1–25, DOI 10.1007/s00382-016-3327-9

Lock AP, Brown AR, Bush MR, Martin GM, Smith RNB (2000) A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. Mon Weather Rev 128(9):3187–3199

Manzato A, Davolio S, Miglietta MM, Pucillo A, Setvák M (2015) 12 september 2012: A supercell outbreak in NE italy? Atmos Res 153:98–118, DOI http://doi.org/10.1016/j.atmosres.2014.07.019

Matte D, Laprise R, Thériault JM, Lucas-Picher P (2017) Spatial spin-up of fine scales in a regional climate model simulation driven by low-resolution boundary conditions. Clim Dyn 49(1):563–574, DOI 10.1007/s00382-016-3358-2

McCaul EW, Goodman SJ, LaCasse KM, Cecil DJ (2009) Forecasting lightning threat using cloud-resolving model simulations. Weather Forecast 24:709–729, DOI 10.1175/2008WAF2222152.1

Mellor G, Yamada T (1982) Development of a turbulence closure model for geophysical fluid problems. Rev Geophys 20(4):851–875, DOI 10.1029/RG020i004p00851

Meredith EP, Maraun D, Semenov VA, Park W (2015) Evidence for added value of convection-permitting models for studying changes in extreme precipitation. J Geophys Res Atmos 120(24):12,500–12,513, DOI 10.1002/2015JD024238

Moore RJ (1985) The probability distributed principle and runoff production at point and basin scales. Hydrol Sci J 30:273–297

Morel C, Senesi S (2002) A climatology of mesoscale convective systems over europe using satellite infrared imagery. ii: Characteristics of european mesoscale convective systems. Q J R Meteorol Soc 128(584):1973–1995, DOI 10.1256/003590002320603494

Nisi L, Martius O, Hering A, Kunz M, Germann U (2016) Spatial and temporal distribution of hailstorms in the alpine region: a long-term, high resolution, radar-based analysis. Q J R Meteorol Soc pp n/a—-n/a, DOI 10.1002/qj.2771

Nuissier O, Ducrocq V, Ricard D, Lebeaupin C, Anquetin S (2008) A numerical study of three catastrophic precipitating events over southern France. i: Numerical framework and synoptic ingredients. Q J R Meteorol Soc 134(630):111–130, DOI 10.1002/qj.200

OpenACC (2011) The OpenACC application programing interface, 2011. URL https://www.openacc.org

Osuna C, Fuhrer C, Gysi T, Schulthess TC (2015) Porting the COSMO dynamical core to heterogeneous platforms using STELLA library. In: Parallel Comput. Road to Exascale, Proc. Int. Conf. Parallel Comput. ParCo 2015, 1-4 Sept. 2015, Edinburgh, Scotland, UK, pp 637–646, DOI 10.3233/978-1-61499-621-7-637

Paulat M, Frei C, Hagen M, Wernli H (2008) A gridded dataset of hourly precipitation in Germany: Its construction, climatology and application. Meteorol Z 17(6):719–732, DOI 10.1127/0941-2948/2008/0332

Peel MC, Finlayson BL, McMahon TA (2007) Updated world map of the köppen-geiger climate classification. Hydrol earth Syst Sci Discuss 4(2):439–473

Perry M, Hollis D, Elms M (2009) The generation of daily gridded datasets of temperature and rainfall for the uk

Prein AF, Gobiet A (2017) Impacts of uncertainties in european gridded precipitation observations on regional climate analysis. Int J Climatol 37(1):305–327, DOI 10.1002/joc.4706

Prein AF, Gobiet A, Suklitsch M, Truhetz H, Awan NK, Keuler K, Georgievski G (2013) Added value of convection permitting seasonal simulations. Clim Dyn 41:2655–2677, DOI 10.1007/s00382-013-1744-6

Prein AF, Langhans W, Fosser G, Ferrone A, Ban N, Goergen K, Keller M, Tölle M, Gutjahr O, Feser F, Brisson E, Kollet S, Schmidli J, van Lipzig NPM, Leung R (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges: Convection-permitting climate modeling. Rev Geophys DOI 10.1002/2014RG000475

Quintana-Segui P, Le Moigne P, Durand Y, Martin E, Habets F, Baillon M, Canellas C, Franchisteguy L, Morel S (2008) Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France. J Appl Meteorol 47:92–107, DOI 10.1175/2007JAMC1636.1

Raschendorfer M (2001) The new turbulence parameterization of LM

Rauthe M, Steiner H, Riediger U, Mazurkiewicz A, Gratzki A (2013) A central european precipitation climatology - part i: Generation and validation of a high-resolution gridded daily data set (hyras). Meteorol Zeitschrift 22(3):235–256, DOI 10.1127/0941-2948/2013/0436

Reinhardt T, Seifert A (2005) A three-category ice-scheme for LMK

Richard E, Buzzi A, Zängl G (2007) Quantitative precipitation forecasting in the alps: The advances achieved by the mesoscale alpine programme. Q J R Meteorol Soc 133(625):831–846, DOI 10.1002/qj.65

Richter D (1995) Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Messfehlers des Hellmann-Niederschlagsmessers. Selbstverl. des Dt. Wetterdienstes Offenbach aM

Ritter B, Geleyn JF (1992) A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. Mon Weather Rev 120(2):303–325

Rockel B, Will A, Hense A (2008) Special issue: Regional climate modelling with cosmo-clm (cclm). Meteorol Z 17:477–485

Ruti PM, Somot S, Giorgi F, Dubois C, Flaounas E, Obermann A, Dell'Aquila A, Pisacane G, Harzallah A, Lombardi E, Ahrens B, Akhtar N, Alias A, Arsouze T, Aznar R, Bastin S, Bartholy J, B??ranger K, Beuvier J, Bouffies-Cloch?? S, Brauch J, Cabos W, Calmanti S, Calvet JC, Carillo A, Conte D, Coppola E, Djurdjevic V, Drobinski P, Elizalde-Arellano A, Gaertner M, Gal??n P, Gallardo C, Gualdi S, Goncalves M, Jorba O, Jord?? G, L'Heveder B, Lebeaupin-Brossier C, Li L, Liguori G, Lionello P, Maci??s D, Nabat P, ??nol B, Raikovic B, Ramage K, Sevault F, Sannino G, Struglia MV, Sanna A, Torma C, Vervatis V (2016) Med-cordex initiative for mediterranean climate studies. Bull Am Meteorol Soc 97(7):1187–1208, DOI 10.1175/BAMS-D-14-00176.1

Scaife AA, Arribas A, Blockley E, Brookshaw A, Clark RT, Dunstone N, Eade R, Fereday D, Folland CK, Gordon M, Hermanson L, Knight JR, Lea DJ, MacLachlan C, Maidens A, Martin M, Peterson AK, Smith D, Vellinga M, Wallace E, Waters J, Williams A (2014) Skillful long-range prediction of european and north american winters. Geophys Res Lett 41:2514–2519, DOI 10.1002/2014GL059637

Schär C, Ban N, Fischer EM, Rajczak J, Schmidli J, Frei C, Giorgi F, Karl TR, Kendon EJ, Tank AMGK, OGorman PA, Sillmann J, Zhang X, Zwiers FW (2016) Percentile indices for assessing changes in heavy precipitation events. Clim Chang 137(1-2):201–216, DOI 10.1007/s10584-016-1669-2

Schneider W, Bott A (2014) On the time-splitting errors of one-dimensional advection schemes in numerical weather prediction models; a comparative study. Q J Roy Meteor Soc 140(684):2321–2329, DOI 10.1002/qj.2301

Schwartz CS (2014) Reproducing the September 2013 record-breaking rainfall over the Colorado Front Range with high-resolution WRF forecasts. Wea Forecast 29:393–402, DOI http://dx.doi.org/10.1175/WAF-D-13-00136.1

Sevruk B, Zahlavova L (1994) Classification system of precipitation gauge site exposure: Evaluation and application. Int J Climatol 14(6):681–689, DOI 10.1002/joc.3370140607

Skamarock WC (2004) Evaluating mesoscale nwp models using kinetic energy spectra. Mon Weather Rev 132(12):3019–3032, DOI 10.1175/MWR2830.1

Smagorinsky J (1963) General circulation experiments with the primitive equations, Part 1: The basic experiments. Mon Weather Rev 91:99–164

Smith RNB (1990) A scheme for predicting layer clouds and their water content in a general circulation model. Q J R Meteorol Soc 116:435–460

Stampoulis D, Anagnostou EN, Nikolopoulos EI (2013) Assessment of high-resolution satellite-based rainfall estimates over the mediterranean during heavy precipitation events. J Hydrometeorol 14(5):1500–1514, DOI 10.1175/JHM-D-12-0167.1

Steppeler J, Doms G, Schattler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G (2003) Meso-gamma scale forecasts using the nonhydrostatic model LM. Meteorol Atmos Phys 82:75–96, DOI 10.1007/s00703-001-0592-9

Stratton RA, Senior CA, Vosper SB, Folwell SS, Boutle IA, Earnshaw D, Kendon E, Lock AP, Malcolm A, Manners J, Morcrette J, Short C, Stirling AJ, Taylor CM, Tucker S, Webster S, Wilkinson JM (in rev.) A pan-africa convection-permitting regional climate simulation with the met office unified model : Cp4-africa. J Clim pp n–a

Tanré D, Geleyn J, Slingo J (1984) Aerosols and Their Climatic Effects, A. Deepak, Hampton, Va, chap First resu, pp 133–177

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498, DOI 10.1175/BAMS-D-11-00094.1

Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parametrization in large-scale models. Mon Weather Rev 117:1779–1800

Vaittinada Ayar P, Vrac M, Bastin S, Carreau J, Déqué M, Gallardo C (2016) Intercomparison of statistical and dynamical downscaling models under the euro- and med-cordex initiative framework: present climate evaluations. Clim Dyn 46(3):1301–1329, DOI 10.1007/s00382-015-2647-5

Walters D, Brooks M, Boutle I, Melvin T, Stratton R, Vosper S, Wells H, Williams K, Wood N, Allen T, Bushell A, Copsey D, Earnshaw P, Edwards J, Gross M, Hardiman S, Harris C, Heming J, Klingaman N, Levine R, Manners J, Martin G, Milton S, Mittermaier M, Morcrette C, Riddick T, Roberts M, Sanchez C, Selwood P, Stirling A, Smith C, Suri D, Tennant W, Vidale PL, Wilkinson J, Willett M, Woolnough S, Xavier P (2016) The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. Geosci Model Devel DOI 10.5194/gmd-2016-194

Walters DN, Best MJ, Bushell AC, Copsey D, Edwards JM, Falloon PD, Harris CM, Lock AP, Manners JC, Morcrette CJ, Roberts MJ, Stratton RA, Webster S, Wilkinson JM, Willett MR, Boutle IA, Earnshaw PD, Hill PG, MacLachlan C, Martin GM, Moufouma-Okia W, Palmer MD, Petch JC, Rooney GG, Scaife AA, Williams KD (2011) The Met Office Unified Model global atmosphere 3.0/3.1 and JULES global land 3.0/3.1 configurations. Geosci Model Devel 4:919–941, DOI 10.5194/gmd-4-919-2011

Weisman ML, Davis C, Wang W, Manning KW, Klemp JB (2008) Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. Weather Forecast 23(3):407–437

Weusthoff T, Ament F, Arpagaus M, Rotach MW (2010) Assessing the benefits of convection-permitting models by neighbourhood vertification: Examples from MAP D-PHASE. Mon Weather Rev 138:3418–3433

Wicker L, Skamarock W (2002) Time-splitting methods for elastic models using forward time schemes. Mon Weather Rev 130(8):2088–2097

Williams KD, Copsey D, Blockley EW, Calvert D, Comer R, Davis P, Graham T, Hewitt HT, Hill R, Hyder P, Ineson S, Johns TC, Keen A, Lee RW, Megann A, Milton SF, Rae JGL, Roberts MJ, Scaife AA, Schiemann R, Storkey D, Thorpe L, Watterson IG, Walters DN, West A, Wood R, Woollings T, Xavier PK (in rev.) The met office global coupled model 3.0 and 3.1 (gc3 & gc3.1) configuration. J Adv Model Earth Syst 1(0)

Wilson DR, Ballard SP (1999) A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. Q J R Meteorol Soc 125:1607–1636, DOI 10.1002/qj.49712555707

Wilson DR, Bushell AC, Kerr-Munslow AM, Price JD, Morcrette CJ (2008) PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description. Q J R Meteorol Soc 134:2093–2107, DOI 10.1002/qj.333

Wood N, Staniforth A, White A, Allen T, Diamantakis M, Gross M, Melvin T, Smith C, Vosper S, Zerroukat M, Thuburn J (2014) An inherently mass-conserving semi-implicit semi-Lagrangian discretisation of the deep-atmosphere global nonhydrostatic equations. Q J R Meteorol Soc 140:1505–1520, DOI 10.1002/qj.2235

Wuest M, Frei C, Altenhoff A, Hagen M, Litschi M, Schär C (2010) A gridded hourly precipitation dataset for Switzerland using rain-gauge analysis and radar-based disaggregation. Int J Clim 30:1764–1775

Wyngaard JC (2004) Toward numerical modeling in the terra incognita. J Atmos Sci 61:1816–1826, DOI 10.1175/1520-0469

List of Figures

1 a. Domains of the different models and subregion definitions: Neur: Northern Europe, CEurL: Central Europe (low land below 500 m); CEurM: Central Europe medium height (above 500 m and below 1500 m), > 1500m: high lands above 1500 m (Alps, Atlas and Pyrenees), Med coast: Mediterranean coasts, Med sea: Mediterranean sea. b. Percentage of missing values in the hourly precipitation datasets. 30

2 Explanation of the ASoP spectral method. a: Probability density function with regular 1 mm bins (frequency of events as a function of event intensity), vertical lines represent the bin widths; b: Probability density function with the ASoP spectral bins defined in Eq. (1); c: Actual contributions to mean precipitation (mm day^{-1}) as a function of event intensity (from b to c, each frequency was multiplied by the average intensity of the bin), the area under the curve is the mean precipitation; d: Fractional contributions (percentage of the mean that each bin represents): from c to d, each bin was divided by the mean precipitation. The area under the curve is 1; e: Actual contributions for a reference (blue) and a model (red), the grey area is the mean absolute difference (mm day^{-1}); f: Fractional contributions for a reference (blue) and a model (red), the grey area is the absolute difference (%); g: Difference in actual contributions between the model (red in panel e) and the reference (blue in panel e); f: Same as g for the fractional contrbutions: difference between the red and blue curves of panel f. 31

3 Fractional contribution index between the 2.2 km and 12 km simulation (FC(2.2 km, 12 km)) as a function of the absolute mean difference (|mean(2.2 km) - mean(12 km)|/mean(12 km)) averaged over the regions defined in Fig. 1a for a) ETH and b) UKMO models. Red is summer, blue winter, cyan spring and black autumn. 32

4 Mean daily precipitation bias in percentage of the observation values for the (top) UKMO and (bottom) ETH models in summer at (a, d) 12km and (b, e) 2.2km resolution. The best daily fractional index between the 2.2km and the 12km for (c) UKMO and (f) ETH model (as described by Eq. (4)). Blue means the 12 km model is closest to the observations, red means the 2.2 km is closest. Values indicate percentage of improvement compare to FC(12 km, obs). Regions with means smaller than 0.5mm day^{-1}in the observations are masked out. 33

5 Same as Fig. 4 for hourly precipitation. 34

6 Differences in the fractional and actual contribution of hourly precipitation between models and the observations (JJA) for different countries (left: actual contribution, right: fractional contribution). See Sect. 3 and Figure 2 for details about the method. a. Germany, b. Switzerland, c. United Kingdom (only points where less than 30% of data is missing in the observations are taken into account). . . . 35

7 Fractional contribution (ratio of actual contribution on total precipitation) of three bin categories in summer: top ($<$2mm/h), middle (from 2 to 8 mm/h), bottom: above 8 mm/h. From left to right: observations, UKMO-12 km, ETH-12 km, UKMO-2.2 km, ETH-2.2 km. 36

8 Frequency of wet spells in summer in different duration and intensity bins for the a. UK, b. Switzerland c. Germany. In each panel, observational datasets are shown as reference and model differences with the observations are shown as indicated in the panel titles (see Sect. 5.2 for details). The number written above the observation plots is the average number of wet spells per grid point per season and the percentage indicated above each model panel is the percentage difference in number of wet spells between models and observations. 37

9 Average of values above the 99^{th} percentile of all hours in summer. Top row: UKMO 12 km (left) and 2.2 km (right) models (percentage difference with the observations), bottom row: ETH 12 km (left) and 2.2 km (right) models (percentage difference with the observations); right column: observations (mm day^{-1}). 38

10 Amplitude of mean diurnal cycle in summer (maximum - minimum) (mm). Top row: UKMO 12 km (left) and 2.2 km (right) models, bottom row: ETH 12 km (left) and 2.2 km (right) models; right column: observations. 39

11 Hour of maximum precipitation of the mean diurnal cycle in summer (local time). Top row: UKMO 12 km (left) and 2.2 km (right) models, bottom row: ETH 12 km (left) and 2.2 km (right) models; right column: observations. 40

12 Average of values above the 99th percentile of all days in autumn (SON). a. UKMO 12 km, b. UKMO 2.2 km, d. ETH 12 km, e. ETH 2.2 km; f. available observations (composite of CMORPH and gridded regional products, as shown in panel c). Yellow area in panel c shows the domain of the case study in Fig. 13 and Fig. 14 (mm/day). 41

13 2-day total precipitation between 08/09/2002 and 09/09/2002. The 12 km models, 2.2 km models and SAFRAN observations are respectively on the left, centre and right. Upper and lower row are for UKMO and ETH simulations. Green lines outline surface height above 500 and 1000 m for the UKMO 12-km simulation on which all models and observations are regridded. Maximum and spatial mean are also given. The domain corresponds to the box in Fig. 12. 42

14 UKMO 2.2 km (upper panel) and 12 km (lower panel) and model-simulated snapshots of 3h-accumulated precipitation (thick black lines; 10, 20, 50 mm/3h), 925 hPa wind (barbs; knots) and virtual temperature (colour shading). White space mask when 925 hPa isobar is below ground. 43

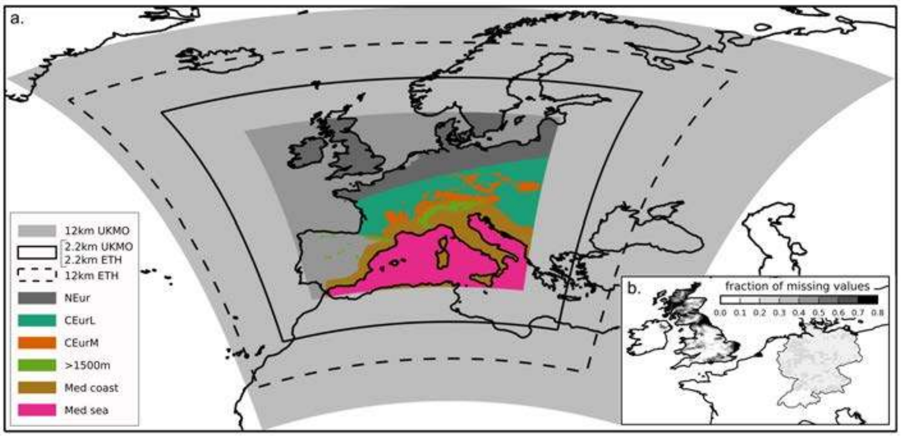


Fig. 1 a. Domains of the different models and subregion definitions: Neur: Northern Europe, CEurL: Central Europe (low land below 500 m); CEurM: Central Europe medium height (above 500 m and below 1500 m), > 1500m: high lands above 1500 m (Alps, Atlas and Pyrenees), Med coast: Mediterranean coasts, Med sea: Mediterranean sea. b. Percentage of missing values in the hourly precipitation datasets.

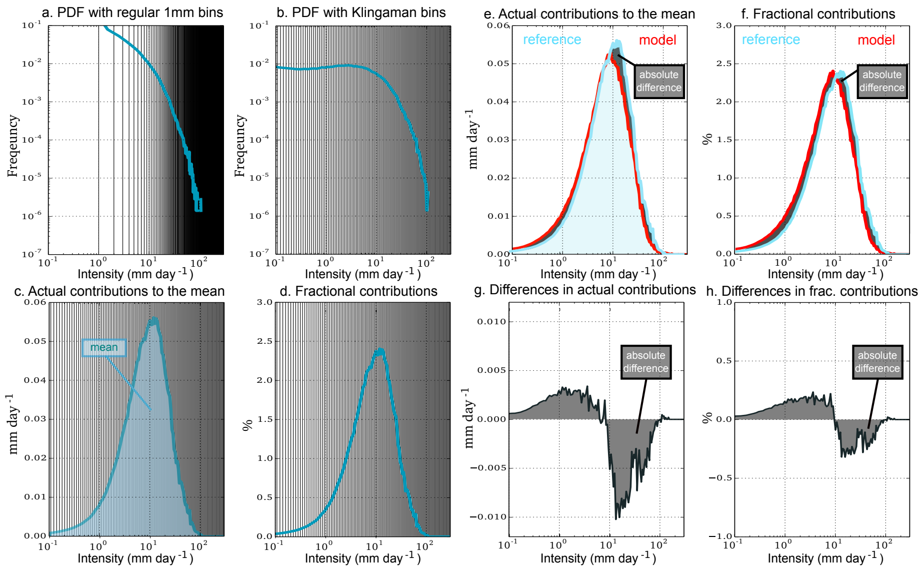


Fig. 2 Explanation of the ASoP spectral method. a: Probability density function with regular 1 mm bins (frequency of events as a function of event intensity), vertical lines represent the bin widths; b: Probability density function with the ASoP spectral bins defined in Eq. (1); c: Actual contributions to mean precipitation (mm day^{-1}) as a function of event intensity (from b to c, each frequency was multiplied by the average intensity of the bin), the area under the curve is the mean precipitation; d: Fractional contributions (percentage of the mean that each bin represents): from c to d, each bin was divided by the mean precipitation. The area under the curve is 1; e: Actual contributions for a reference (blue) and a model (red), the grey area is the mean absolute difference (mm day^{-1}); f: Fractional contributions for a reference (blue) and a model (red), the grey area is the absolute difference (%); g: Difference in actual contributions between the model (red in panel e) and the reference (blue in panel e); f: Same as g for the fractional contrbutions: difference between the red and blue curves of panel f.

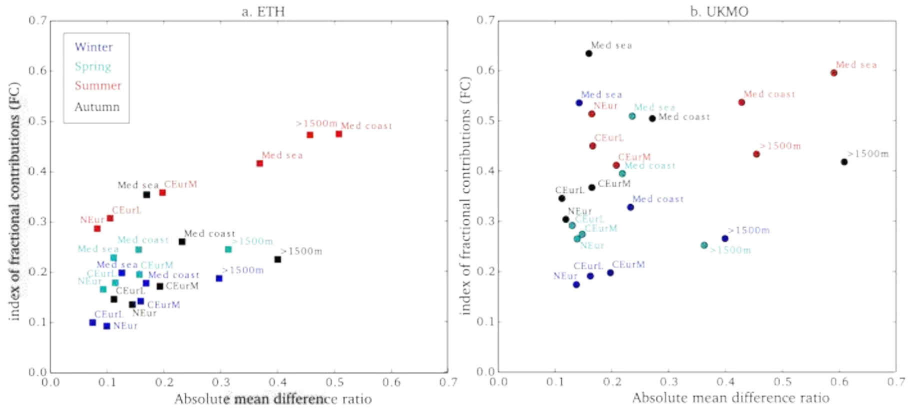


Fig. 3 Fractional contribution index between the 2.2 km and 12 km simulation (FC(2.2 km, 12 km)) as a function of the absolute mean difference (|mean(2.2 km) - mean(12 km)|/mean(12 km)) averaged over the regions defined in Fig. 1a for a) ETH and b) UKMO models. Red is summer, blue winter, cyan spring and black autumn.

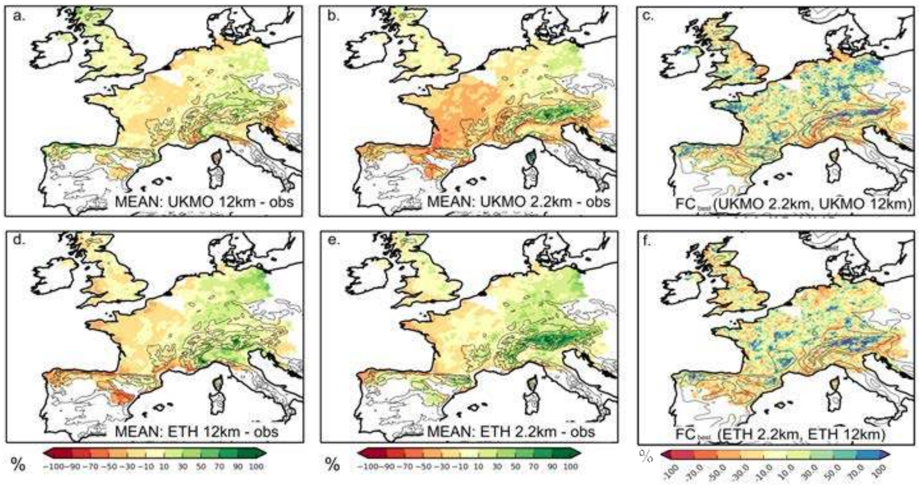


Fig. 4 Mean daily precipitation bias in percentage of the observation values for the (top) UKMO and (bottom) ETH models in summer at (a, d) 12km and (b, e) 2.2km resolution. The best daily fractional index between the 2.2km and the 12km for (c) UKMO and (f) ETH model (as described by Eq. (4)). Blue means the 12 km model is closest to the observations, red means the 2.2 km is closest. Values indicate percentage of improvement compare to FC(12 km, obs). Regions with means smaller than 0.5mm day^{-1}in the observations are masked out.

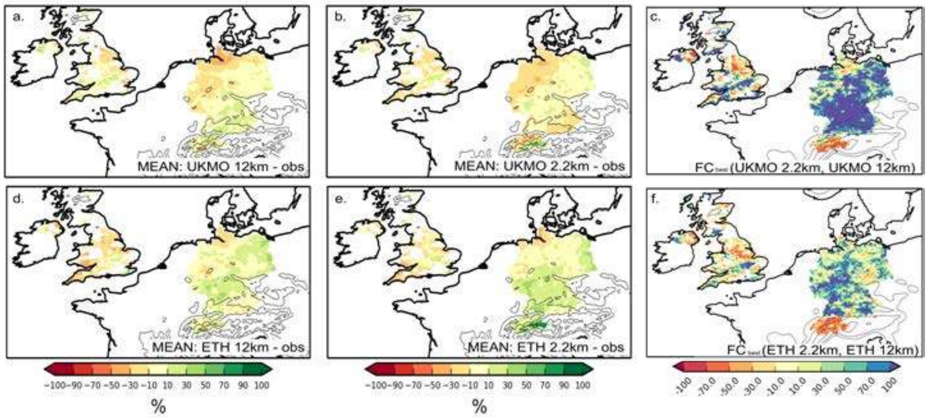


Fig. 5 Same as Fig. 4 for hourly precipitation.

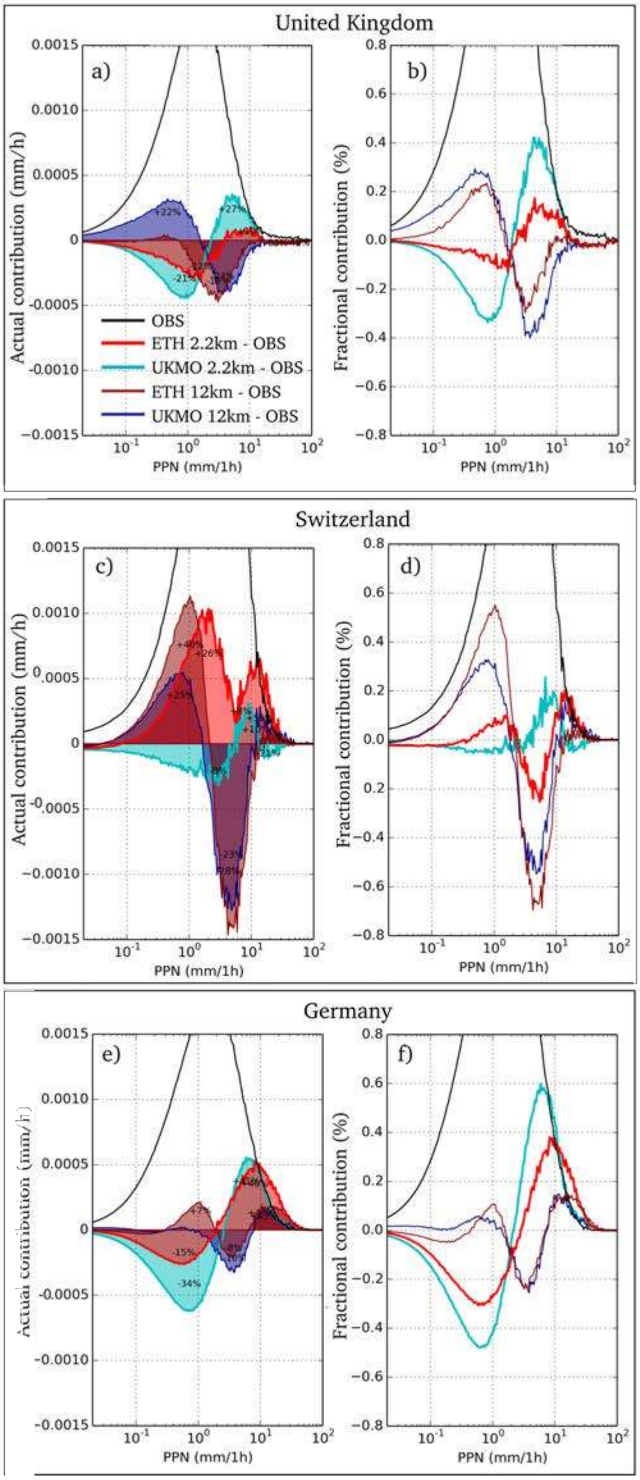


Fig. 6 Differences in the fractional and actual contribution of hourly precipitation between models and the observations (JJA) for different countries (left: actual contribution, right: fractional contribution). See Sect. 3 and Figure 2 for details about the method. a. Germany, b. Switzerland, c. United Kingdom (only points where less than 30% of data is missing in the observations are taken into account).

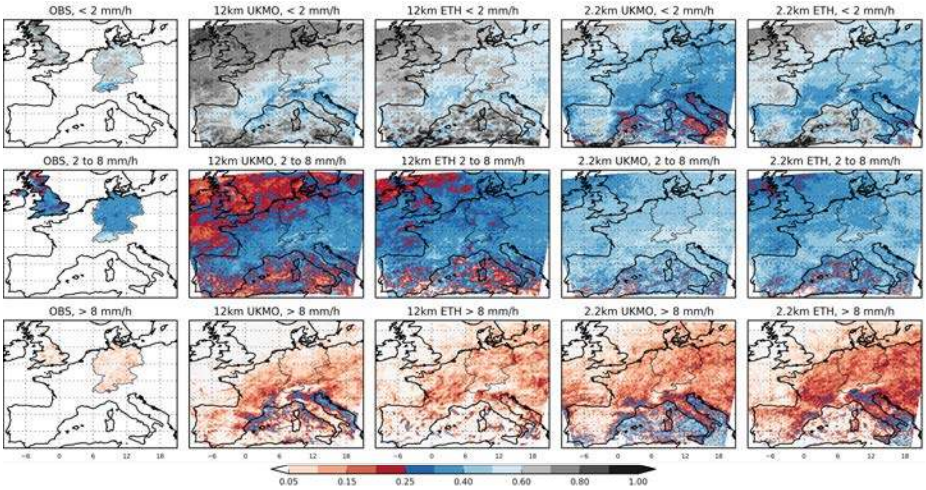


Fig. 7 Fractional contribution (ratio of actual contribution on total precipitation) of three bin categories in summer: top (<2mm/h), middle (from 2 to 8 mm/h), bottom: above 8 mm/h. From left to right: observations, UKMO-12 km, ETH-12 km, UKMO-2.2 km, ETH-2.2 km.

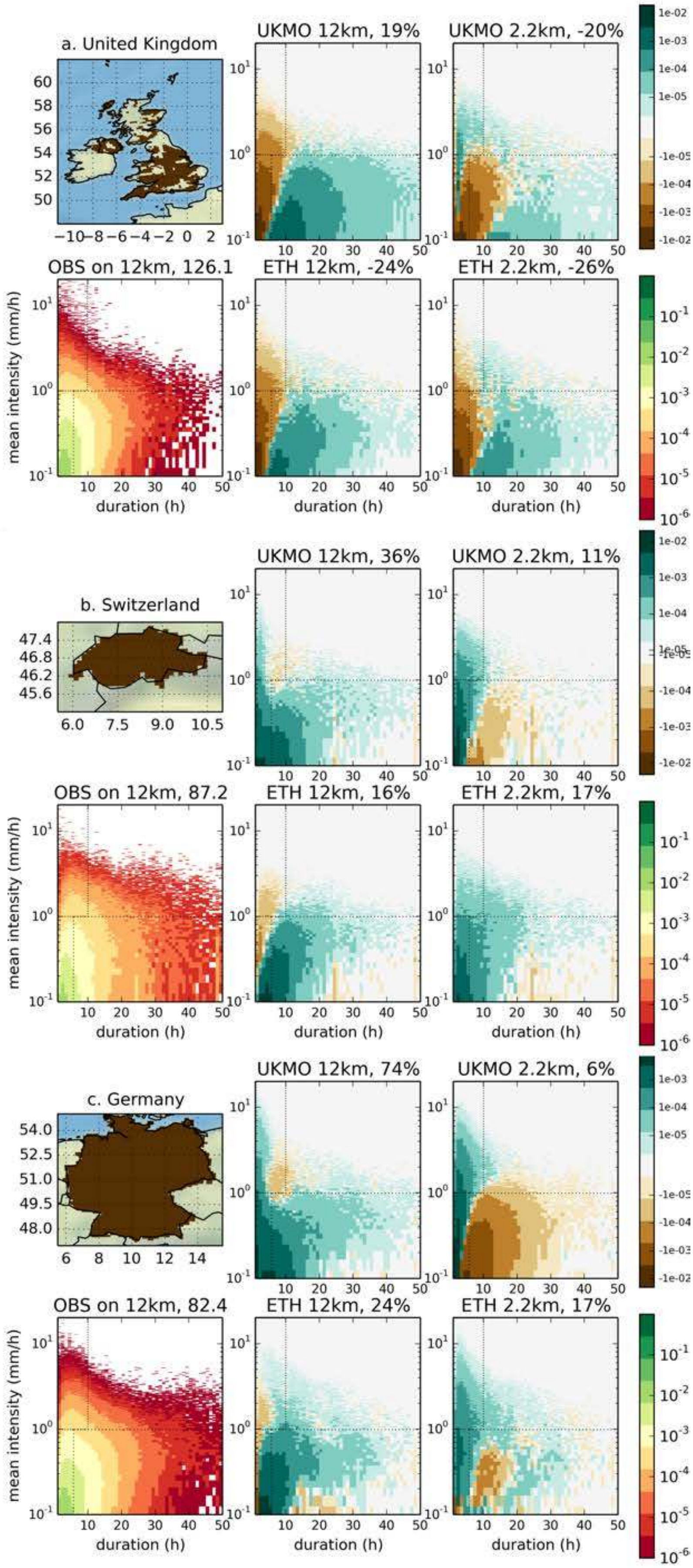


Fig. 8 Frequency of wet spells in summer in different duration and intensity bins for the a. UK, b. Switzerland c. Germany. In each panel, observational datasets are shown as reference and model differences with the observations are shown as indicated in the panel titles (see Sect. 5.2 for details). The number written above the observation plots is the average number of wet spells per grid point per season and the percentage indicated above each model panel is the percentage difference in number of wet spells between models and observations.

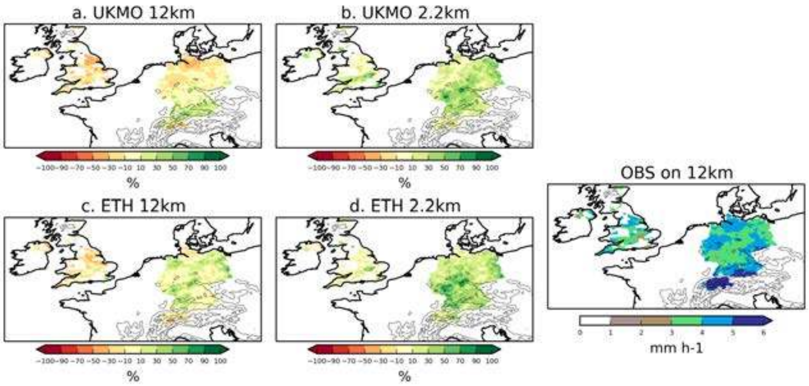


Fig. 9 Average of values above the 99^{th} percentile of all hours in summer. Top row: UKMO 12 km (left) and 2.2 km (right) models (percentage difference with the observations), bottom row: ETH 12 km (left) and 2.2 km (right) models (percentage difference with the observations); right column: observations (mm day^{-1}).

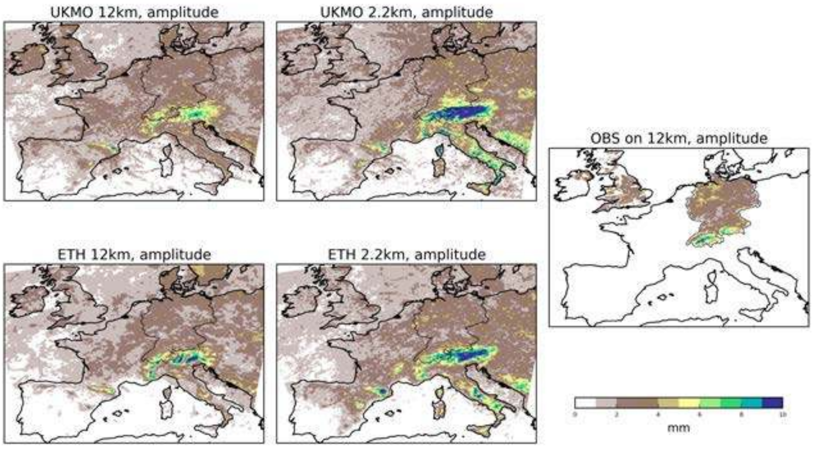


Fig. 10 Amplitude of mean diurnal cycle in summer (maximum - minimum) (mm). Top row: UKMO 12 km (left) and 2.2 km (right) models, bottom row: ETH 12 km (left) and 2.2 km (right) models; right column: observations.

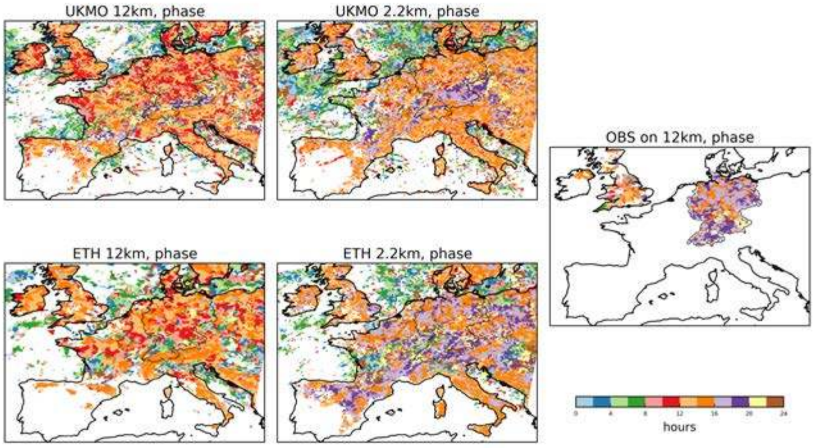


Fig. 11 Hour of maximum precipitation of the mean diurnal cycle in summer (local time). Top row: UKMO 12 km (left) and 2.2 km (right) models, bottom row: ETH 12 km (left) and 2.2 km (right) models; right column: observations.

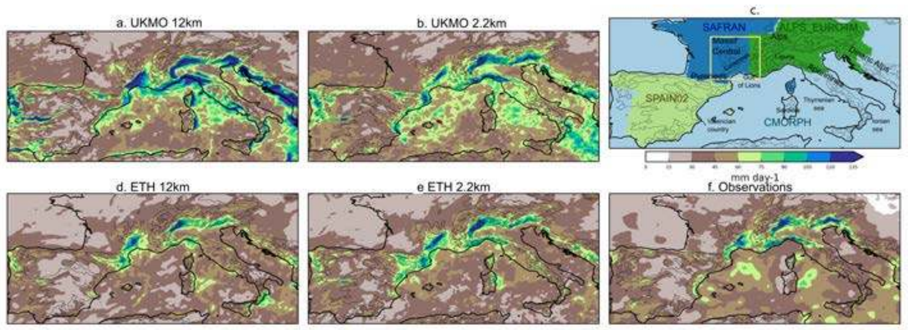


Fig. 12 Average of values above the 99th percentile of all days in autumn (SON). a. UKMO 12 km, b. UKMO 2.2 km, d. ETH 12 km, e. ETH 2.2 km; f. available observations (composite of CMORPH and gridded regional products, as shown in panel c). Yellow area in panel c shows the domain of the case study in Fig. 13 and Fig. 14 (mm/day).

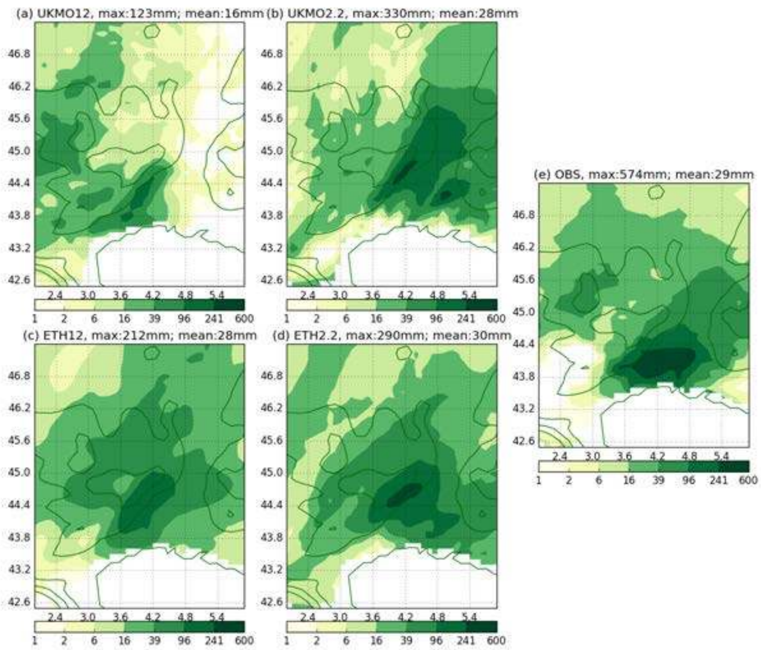


Fig. 13 2-day total precipitation between 08/09/2002 and 09/09/2002. The 12 km models, 2.2 km models and SAFRAN observations are respectively on the left, centre and right. Upper and lower row are for UKMO and ETH simulations. Green lines outline surface height above 500 and 1000 m for the UKMO 12-km simulation on which all models and observations are regridded. Maximum and spatial mean are also given. The domain corresponds to the box in Fig. 12.

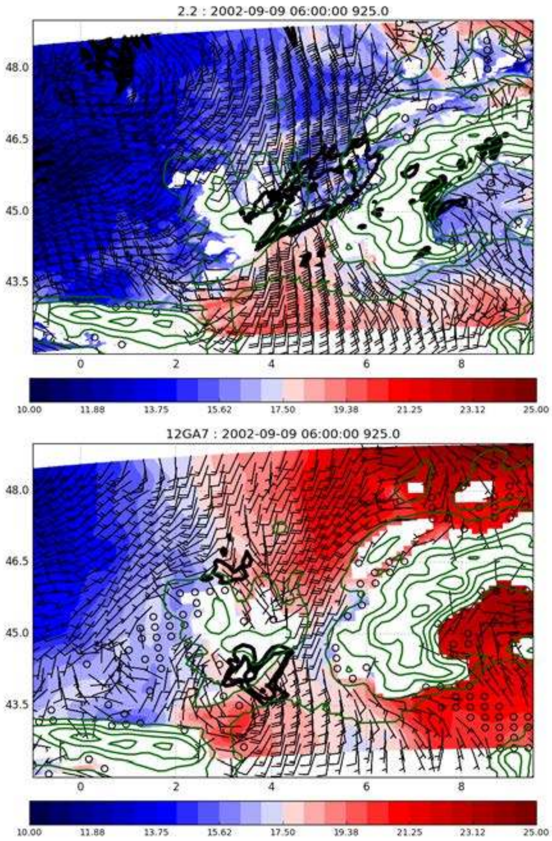


Fig. 14 UKMO 2.2 km (upper panel) and 12 km (lower panel) and model-simulated snapshots of 3h-accumulated precipitation (thick black lines; 10, 20, 50 mm/3h), 925 hPa wind (barbs; knots) and virtual temperature (colour shading). White space mask when 925 hPa isobar is below ground.

Model sensitivity to run-off and soil hydraulic physics – Supplementary Information

Ségolène Berthou Elizabeth Kendon Steven Chan
Nikolina Ban David Leutwyler Christoph Schär
Giorgia Fosser

December 20, 2017

Land surface physics sensitivity study

The Joint UK land environmental simulation [JULES; Best et al, 2011] is the the standard land surface model (LSM) for the Met Office UM modelling system. Within JULES, soil characteristics are parametrised with either the Van Genuchten [van Genuchten, 1980] or Brooks-Corey [Brooks and Corey, 1964] scheme. The 2.2 km and 12 km Met Office UM simulations here use Van Genuchten hydraulics, however, this leads to soil infiltration rates being too low as there is an incompatibility with the soil properties defined in the models. Recent internal Met Office work suggests that the Brooks-Corey hydraulics may give improved results, and hence this is tested here. In addition, the 12 km and 2.2 km Met Office UM models have a different set-up in the treatment of the saturated layers of soil: in the 2.2 km model, if a layer is saturated, the excess of water is transferred to the layer above. If it is the surface layer, then it disappears in the surface run-off term and is not available any more. This is the standard set-up of the operational forecast model. In the 12 km model, if a layer is saturated, the excess of water is transferred to the layer below. If it is the bottom layer, then the rainfall disappears in the sub-surface run-off term and is not available any more. This is the standard climate set-up. Here we test the effect changing the treatment of saturated layers in the 2.2 km model, from transfer up (giving excess water run-off at the surface) to down (giving excess water run-off at the sub-surface).

The configuration of the sensitivity simulations conducted here are given in Table 1. All simulations are driven by the ERA-Interim reanalysis [Dee et al, 2011]; hence, lateral boundary inter-annual and intra-seasonal variability of all simulations are the same.

The magnitude of the precipitation differences between the simulations for the first model summer (Fig. 1) increases from west to east (note the sharper colours). However, the differences are generally noisy without clear regional patterns. Opposite signed differences often occur in spatial scales less than

100 km; this suggests the difference are caused by uncertainties in movement of precipitation systems. Such uncertainties increase eastward downstream as LBC influence weakens.

The 2.2km simulations have warm summer surface temperature in Central Europe and the Balkans [Kendon et al, 2017]. Shown in Fig. 2a-c are summer surface air temperatures for the first completed model summer. The use of sub-surface runoff cools surface temperature across continental Europe by less than 0.5 °C, but larger cooling ($1 - 1.5\,^\circ$C) are seen in the Balkans. The use of Brooks-Corey soil characteristics [Brooks and Corey, 1964] has the opposite effect; northern continental Europe is warmed by about 0.5 °C.

The differences with E-OBS [Haylock et al, 2008] temperature for the same summer is illustrated in Fig. 2d. Differences with E-OBS are much larger than the differences between model simulations. In central and south east Europe, there are very large warm biases up to $5 - 6\,^\circ$C. However, the different JULES physics can make notable differences to those temperature biases. Cooling due to switching to downward movement of excess water and sub-surface run-off reduces the warm biases in Central Europe and the Balkans by about $20-25\%$. This may be explained by less moisture being lost as surface run-off, leading to more moisture being available for evapotranspiration, which in turn cools the surface. Switching to Brooks-Corey soil hydraulics exacerbates existing warm temperature biases in central Europe, but has little impact on the warmest biases in the Balkans. In this case, infiltration rates are increased, but it appears not enough to prevent the loss of moisture as surface run-off, especially given the intense and intermittent nature of rainfall in the 2.2 km convection-permitting model.

To illustrate the above further, surface air temperatures within the Central Europe Pannonian Basin for the first full model year are shown in Fig. 3. This region is chosen due to its positive surface air temperature bias, low-lying continental location. Inter-model differences are much smaller than the differences with E-OBS. The seasonal cycle and intra-seasonal variability are generally well-captured by all model simulations. Around late spring, warm biases relative to E-OBS begin to develop for all simulations. The biases are somewhat less severe in SubSfcRunOff simulation, but they are slightly more severe in the BrooksCoreySoil simulation. Despite of the warm biases, model summer intra-seasonal temperature variabilities remain well captured. In pre-summer period, temperature biases are less clear, but individual model days can have large temperature biases (exceeding 3 °C).

Similar temperature changes are seen in regions where summer temperature biases are smaller. Daily temperature changes in France are shown in Fig. 4. Similar to the Pannonian Basin, the seasonal and intra-seasonal variability are well captured by all model simulations for the full model year. However, there is a clear shift from a pre-spring cold temperature bias to a summer warm bias.

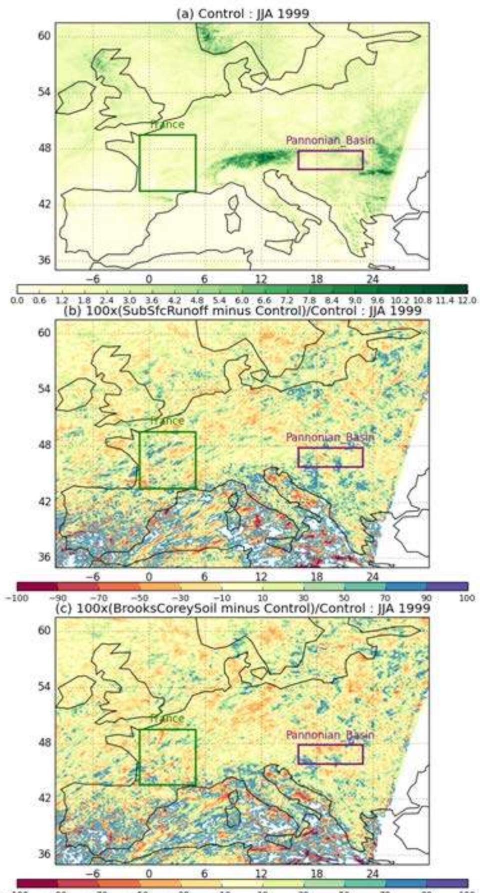
The similar trends in both regions suggest that the origin of the temperature biases are unrelated to models developing its internal variability, and the model summer is well constrained by the lateral boundary conditions. The onset of warm bias in both regions are abrupt between May and June 1999.

Overall, these sensitivity experiments with the 2.2 km model suggest that removing the inconsistency between the soil hydraulics and soil properties in the Met Office UM model (by using Brooks-Corey equations) has little impact on the warm temperature bias seen over central Europe and the Balkans in summer. Switching the treatment of saturated layers from upward to downward transfer (with sub-surface excess water run-off) reduces the bias, but does not remove it completely. Thus it appears, it is the intense/intermittent nature of rainfall in the 2.2 km model which is mostly responsible for the dry soil conditions and associated warm temperature bias in this region. Other regions such as the UK are less sensitive, as soil moisture is not close to the critical value for limiting evaporation.

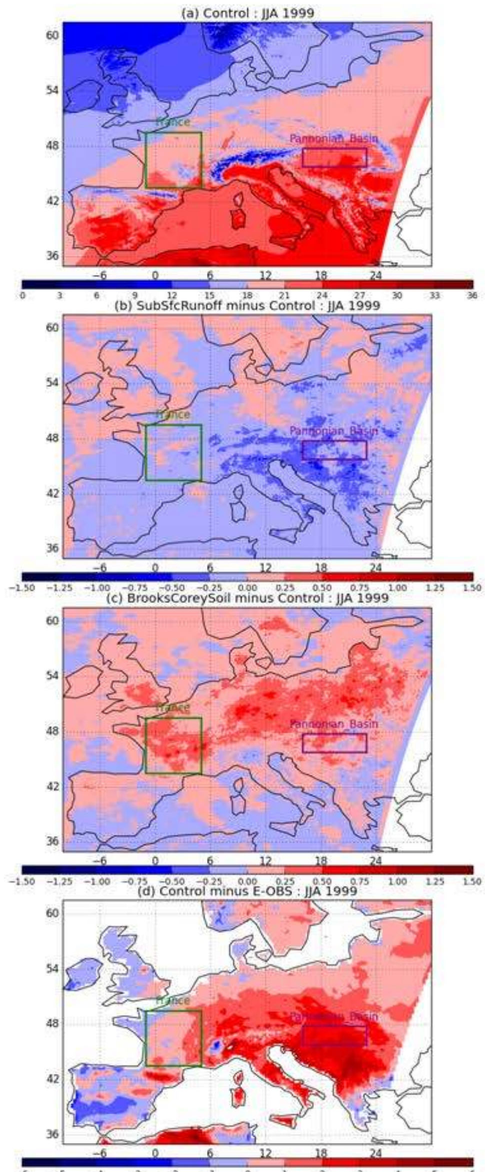
It should be noted that these results correspond to a single summer, and further years of simulation are required to draw robust conclusions. In particular, the deepest soil layers are still spinning up in the first few years of simulation, which may impact the results.

Table 1: The configurations of LSM sensitivity simulation

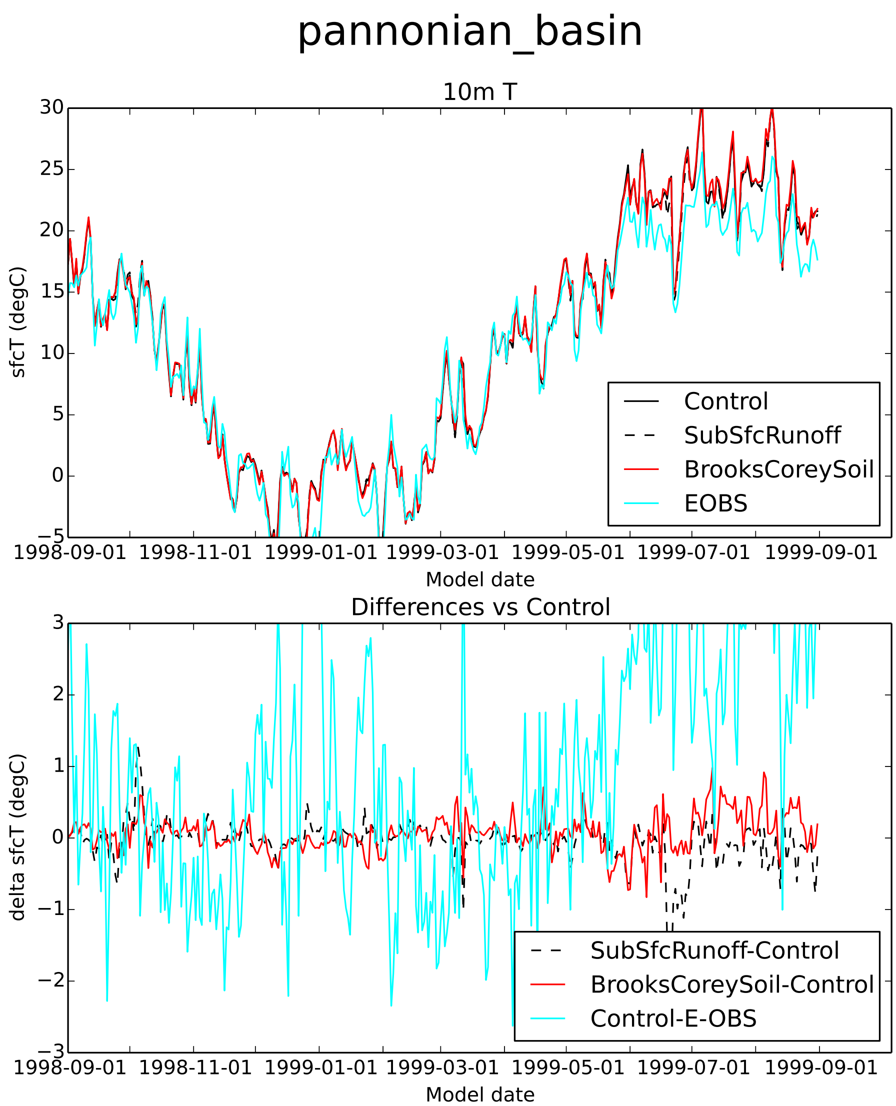
Simulation	Soil hydraulics	Excess water run-off
Control	Van Genuchten [van Genuchten, 1980]	Surface
SubSfcRunOff	Van Genuchten [van Genuchten, 1980]	Sub-surface
BrooksCoreySoil	Brooks-Corey [Brooks and Corey, 1964]	Surface



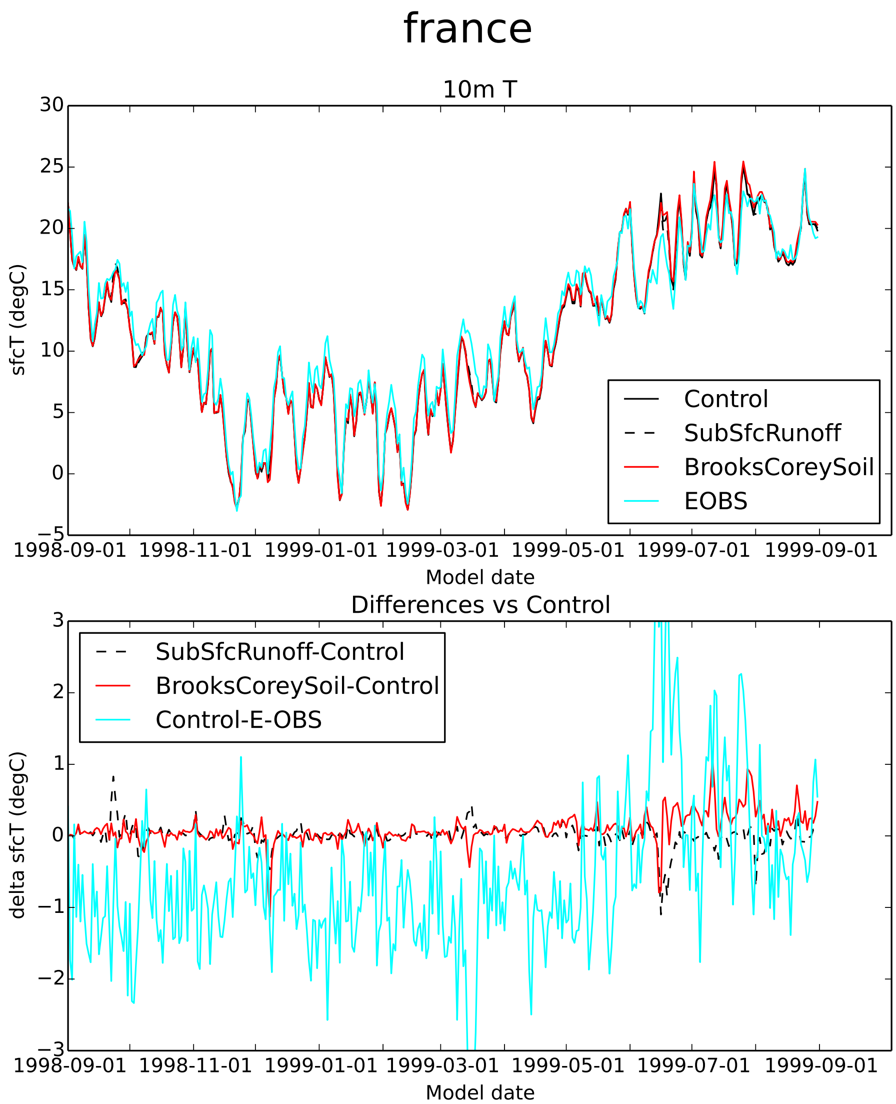
Supplementary Figure 1: Precipitation differences between the control and sensitivity simulations for the first model summer (June-July-August, 1999). Panel a shows the daily-averaged precipitation for the control simulation, and panels b and c show the percentage differences between SubSfcRunOff/Brooks-Corey simulations and the control simulation.



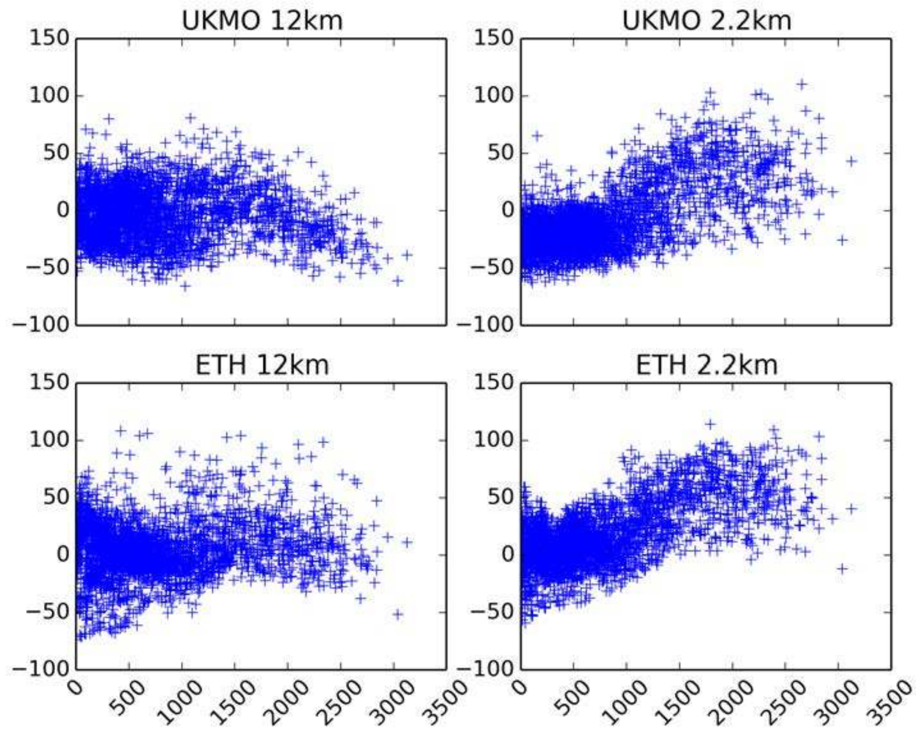
Supplementary Figure 2: Same as in Fig. 1, but for 1.5m temperature. Unlike Fig. 1, °C differences are shown instead. In addition, we also show the biases of the control simulation relative to gridded E-OBS air temperature observations [Haylock et al, 2008] in the bottom panel. The purple and green boxes indicate regions where the spatial averages of 1.5 m temperatures are taken – Fig. 3 for Pannonian Basin and Fig. 4 for France.



Supplementary Figure 3: Model-simulated and E-OBS area-averaged 1.5-m air temperatures in the Pannonian Basin region (see Fig. 2). The actual temperatures are shown in the upper panel, and the differences relative to the control simulation are shown in the lower panel. Note the differences between E-OBS and control simulation are control minus E-OBS (i.e. biases relative E-OBS).



Supplementary Figure 4: Same as in Fig. 3, but for France.



Supplementary Figure 5: Mean bias (%) against height (m) in summer for the ALPS_EURO4M domain (regridded on the UKMO 12km grid) for the UKMO 12km, UKMO2.2km, ETH 12km and ETH2.2km models.

References

Best MJ, Pryor M, Clark DB, Rooney GG, Essery RLH, Mnard CB, Edwards JM, Hendry MA, Porson A, Gedney N, Mercado LM, Sitch S, Blyth E, Boucher O, Cox PM, Grimmond CSB, Harding RJ (2011) The joint uk land environment simulator (jules), model description - part 1: Energy and water fluxes. Geosci Model Devel 4():677–699, DOI 10.5194/gmd-4-677-2011

Brooks RJ, Corey AT (1964) Hydraulic properties of porous media. Colorado State University, Fort Collins, Colorado, United States, hydrological Papers 3

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hölm EV, Isaksen L, Kallberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park PK, Peubey C, de Rosnay P, Tavolato C, Thêpaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597, DOI 10.1002/qj.828

van Genuchten MT (1980) A closed-formequation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am 44:892–898, DOI 10.2136/sssaj1980.03615995004400050002x

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A european daily high-resolution gridded dataset of surface temperature and precipitation. J Geophys Res 113:D20,119, DOI 10.1029/2008JD010201

Kendon EJ, Chan SC, Berthou S, Fosser G, Roberts MJ, Fowler HJ (2017) Introducing the met office 2.2-km europe-wide convection-permitting regional climate simulations. In: European Geophysical Union General Assembly 2017, Vienna, Austria