From Ségolène Berthou:

I would like to make a few comments on this article, which is a big piece of effort, is very interesting and complements a similar analysis by Demory et al. (2020). It's always reassuring to have similar results with different pieces of code and types of analysis. I would like to point at a few differences between your article and Demory et al. (2020):

We thank Ségolène Berthou for making the effort of reviewing, and for all valuable comments. Responses follow in red below. In the markup version of the revised manuscript substantial changes are also marked in red.

- Demory et al. analyse precipitation on a 50km scale (except for CMIP5), whereas you mix all model resolutions. Klingaman et al. (2017) emphasize that regridding models changes the precipitation distribution as you point out at lines 128. But they argue that models should be compared on similar grids at different scales: a 12km model is meant to be good at 12km, at 50km and at 200km. A 200km model is not meant to be good at 12km. If you use observations only on a 25km scale (as I believe E-OBS is), you cannot expect CMIP5/6 to be good. Similarly, you show that 12km overestimates intense precipitation but this is compared with E-OBS which has a coarser scale than 12km model. In Demory et al., we showed that 12km models overestimated intense precipitation even when regridded at a 50km scale against observation regridded at 50km. Maybe you should include more discussion on this or deserve a few figures to a comparison of everything on a 200km scale, one on a 50km scale.

We see what you mean; it is, of course, in a sense unfair to compare models of different resolutions. We assume that models of higher resolution will perform better than models of lower resolution; and a model on 12 km will be extra good if the observations are also on 12 km. On the other hand, when you are about to use data from climate models the choice is for example between GCM and RCM, or between RCM of low resolution and RCM of high resolution. Or perhaps you are thinking about if it's worth the effort of making atmosphere only GCM runs to increase the resolution instead of just using standard GCM results. Then you will use the data of choice and perhaps compare it to observations, other models etc. Therefore we made the active choice of using this method because it allows us to preserve the model output on its native grid.
Nevertheless, we see the need of also comparing at common grids. We have now included analyses when all data are regridded to a 0.5°×0.5° grid and a 2°×2°grid.

- You use averaged distributions across grid-points whereas we first pool the data across the region and then plot the distribution. Both methods are equivalent in a flat homogeneous region but not in region with varied topography. You may be smoothing out more the tail of the distribution than we do. Both methods are valid, I'm just highlighting a difference.
- We use a new set of bins compared to Klingaman (2017) and Berthou (2018), defined in Berthou et al. (2019) for two reasons: – we wanted pure exponential increase in the bin size

so that all the bins have the same size in a log scale and area below the curve is the mean. It's not quite the case in Klingaman and Berthou but it does not make a huge difference. – The other reason was that the Klingaman method had too many bins at the start of the distribution for E-OBS, which does not have a continuous precipitation distribution. I wonder how you managed to have such a smooth distribution for E-OBS, maybe the newer version is improved. Or the spatial averaging of distributions does the job. The equation and the difference between the two sets of bins is shown in Fig. S5 here: https://agupubs.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1029%2F2019GL083544&file=grl59801-sup-0001-agusuppinfo_revised.pdf

Unfortunately there was an error in the method section describing the ASoP analysis. We actually pooled all grid points across the region prior to ASoP calculations. We have made changes accordingly in the text. An updated version of the section describing ASoP analysis is provided below.
Regarding the bins; we find the arguments for using exponential bin sizes (as used in Berthou et al. 2019) interesting and especially in the case of E-OBS that does not have continuous intensity distribution. In order to increase the readability of the figures, we applied a filter to the resulting distributions to reduce the noise. We've made sure that the smoothed data did not affect the interpretation of the results. However, we failed to include this procedure in the description of ASoP analysis. This has now been corrected for (see text below).

Other comments:
- From your explanation in the method section and the y-axis on the ASoP figures, it seems like you are computing the fractional contribution. This would mean that you care about the shape of the distribution only. However, the figures do show some curves almost always above E-OBS and the integral of the differences is not 0 but >0 (e.g. Fig. 2 SC and ME): this cannot happen if you normalise each curve by mean precipitation, unless you are normalising all curves by mean precipitation in E-OBS? In Demory et al. 2020, we chose to use actual contributions as we wanted information of both mean and distribution at the same time, to show which bins contribute to mean biases. From your discussion, it seems like you are also discussing actual contributions. Please clarify what you did.

The labels on the Y-axis were not correct unfortunately. All ASoP figures (except Fig. 4) show actual contributions and not fractional contributions. We have updated the figures and clarified in figure texts what is shown (please see attached figures).

Updated text in Method section, describing ASoP analysis:
"To investigate the effect of model grid resolution on the full distributions of daily precipitation intensities, we use the ASoP (Analysing Scales of Precipitation) method (Klingaman et al., 2017; Berthou et al., 2018). ASoP involves splitting precipitation distributions into bins of different intensities and then provides information of the contribution from each precipitation intensity separately to the total mean precipitation rate (i.e. given by all intensities taken

together). In the first step, precipitation intensities are binned in such a way that each bin contains a similar number of events, with the exception of most intense events, which are rare. The actual contribution (in mm) of each bin to the total mean precipitation rate is obtained by multiplying the frequency of events by the mean precipitation rate. The sum of the actual contributions from all bins gives the total mean precipitation rate. The fractional contribution (in %) of each bin is further obtained by dividing the actual contributions by the mean precipitation rate. In this case, the sum of all fractional contributions is equal to one, thus the information provided by fractional contributions is predominantly about the shape of the distribution. Taking the absolute differences between two fractional distributions and sum over all bins gives a measure of the difference in the shapes of the precipitation distributions. This is here called the "Index of fractional contributions". Since E-OBS precipitation intensities, in contrast to model data, are not continuous the resulting ASoP factors for E-OBS tend to be noisy, especially for lower intensities. In order to facilitate the interpretation of the results, the regionally averaged ASoP factors for E-OBS were smoothed to some extent by using a simple filter.
The ASoP method is here applied to grid points pooled over target regions (Fig. 1) separately and the result is a distribution for each model showing the probability of different precipitation intensities based on daily precipitation. Most results presented here concern the actual contributions, both to limit the number of figures and because these factors conveniently provide information on both shape of distributions as well as the mean values. The ASoP distributions of all analysed models are used to compare model behaviour and performance. In particular to see how changing the grid resolution affects different parts of the distribution, for example if contributions from low and high precipitation intensities are different."

- I agree with the sentence lines 19-21 but I think it applies to models of ~50km: PRIMAVERA-HR, CORDEX-44, CORDEX-11 since you show that CMIP5/6 have very different precipitation distributions and clearly overestimate small intensities. Orographic and coastal regions (AL, FR, IP, MD,) exhibit strong differences (as shown in your Fig. 4). So I would add:
"Once reaching ~50km resolution, the difference between different models is often larger than between the low- and high-resolution versions of the same model, which makes it difficult to quantify the improvement. In this sense the quality of an ensemble is depending more on the models it consists of rather than the average resolution of the ensemble."
We change the sentence accordingly.

-You could also include CMCC in the PRIMAVERA ensemble

We tried to get daily pr data of CMCC from the CEDA archive, but didn't manage to get it.

- In the accepted version of Demory et al., we consider 45 CORDEX HR and 26 CORDEX LR, so I think sentence line 24-25 is not valid. However, you have other strengths in your study, e.g. comparing the spread between resolution and between models. I think a strong common conclusion of our studies that you highlighted well is that it is best to carefully design an ensemble (across all high-resolution models available (>=50km)) rather than to take an ensemble of opportunity to have a good representation of precipitation distribution.

You're right. That was perhaps a bit exaggerated. We change the sentence to:

The results presented here are in line with previous similar studies. To these studies we add details about the spread between resolutions and between models.

- Many of the CMIP6 models have almost not wet days in the IP. Is this a bug or real? In which case it is quite worrying: these models are then very dry in this region.

This is a bug. Wrong versions of figures 6-9 were accidently inserted in the manuscript. This is now corrected.

- You could make use of the E-OBS ensemble rather than just mean in your ASoP figures (although it's already a crowded figure)

Individual E-OBS members are available upon request, but as we understand it these are useful if you want to sample uncertainty when you use E-OBS as forcing. E-OBS writes: "The individual ensemble members are mainly intended for users who require the uncertainty in the gridded fields to propagate through to various other applications. ..."

If we were looking at specific events this could perhaps be interesting, but since we look at climatologies we don't see the use of crowding this figure even more.

References:

Berthou, S., Kendon, E., Rowell, D. P., Roberts, M. J., Tucker, S. O., & Stratton, R. A. (2019). Larger future intensification of rainfall in the West African Sahel in a convectionâ˘AR˘ permitting model. Geophysical Research Letters, 46, 13299– 13307. https://doi.org/10.1029/2019GL083544

Interactive comment on Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2020-31, 2020.

Interactive comment on "The importance of model resolution on simulated precipitation in Europe –from global to regional model" by Gustav Strandberg and Petter Lind

Anonymous Referee #1

This study analyses precipitation characteristics over Europe from a wide range of model ensembles, including Global Climate Models (CMIP5, CMIP6, PRIMAVERA)and Regional Climate Models (CORDEX). The precipitation characteristics include daily precipitation distributions based on the ASoP diagnostics developed by Klingaman et al

(2017), as well as statistical metrics such as number of wet days, number of heavy precipitation days, intensity of wet days, intensity of heaviest precipitation day. The aim of this study is three-fold: 1) investigate differences between model ensembles, and between models within each ensemble, by using a wide range of ensembles from CMIP5, CMIP6, PRIMAVERA and CORDEX; 2) evaluate model performance against observations, using E-OBS data; 3) investigate the role of resolution in precipitation characteristics over Europe, by selecting only models available at both low and high resolution versions. I have several comments regarding this study, as described below. Some of them would require more analyses and restructuring of the paper, but I think it would also greatly improve it.

We thank you for making the effort of reviewing the paper and for all constructive suggestions. Responses to comments follow below in red. In the markup version of the revised manuscript substantial changes are also marked in red.

1) The authors have made an impressive work by analysing such a huge amount of simulations. This is very complementary to the work by Demory et al (2020), which have analysed daily precipitation over Europe in CMIP5, PRIMAVERA (high-resolution) and CORDEX (low and high resolutions) compared to high-quality observational datasets over Europe. This work has now been revised by focusing more on EUR-11 (which is a newer ensemble than EUR-44), and by including also spatial distribution of precipitation and Taylor diagrams, which confirm the results shown by the precipitation distribution. The paper is now accepted and should appear soon. I suggest to refer to this study already in the introduction. Iles et al (2020) could also be referred to in the introduction as another study evaluating a range of GCMs and RCMs at various resolutions, considering the atmosphere-only UPSCALE simulations. The fact that this study and Demory et al find similar results, despite using slightly different methods, give strength to these two studies and should be discussed further.

The Introduction has been expanded with a paragraph discussing Demory et al., and Iles et al., as well as other similar studies using CORDEX data:

A few studies have been made investigating how model resolution affects the simulated precipitation in the CORDEX ensembles, comparing 50 km and 12.5 km grid spacing. A clear result is that precipitation generally increases with higher resolution, which sometimes means that the bias increases when precipitation is added to already wet models (Kotlarski et al., 2014; Casanueva et al., 2016); something that is also seen in simulations with global models (e.g. Thackeray et al., 2018). An overall improvement of mean precipitation is not seen the high resolution CORDEX simulations, except for regions with complex topography (Korlarski et al., 2014; Casanueva et al., 2016; Prein et al., 2016). Prein et al. (2016) looked at local precipitation on short time scales. They find that 12.5 km simulations better represents extreme and mean precipitation, also when simulations are aggregated to 50 km. They note, however, that the results are highly dependent on which observations the simulations are compared with. They also note that improvements are on the ensemble as a whole, and not necessarily for each individual model. In similar studies as the present Iles et al. (2019) and Demory et al. (2020) compare CORDEX simulations with simulations from CMIP5 and Primavera. They see that precipitation increases with resolution so that CMIP5 underestimates precipitation amounts and CORDEX overestimates it, when compared to E-OBS, and that the effect of resolution is largest in complex topography. They also find that Primavera performs similarly to CORDEX when run on the same resolution, which is interesting regarding that the Primavera models are developed for low resolution. Iles et al. (2019) also find considerable inter-model differences meaning that improvements are seen on the ensemble level rather that for individual models.

2) The authors have managed to combine their results into well-designed figures. However, I feel the 3 goals should not be addressed with the same method. The authors have indeed decided to perform the analyses on the model native grids. This is a good choice for showing what each ensemble is able to simulate at its own resolution, and could be used for addressing aim 1) written above, as long as the models are not com-pared to each other. A clean comparison could only be done on a common coarser grid, as emphasised by Klingaman et al, 2017. Evaluating results on native grid not only shows the potential of the model physics but also includes the technical aspect of doing analyses on a finer grid. This

technical aspect can be evaluated by regridding the data on a coarser grid and see how the results are affected by such a regridding. Evaluating results on common grids would show the impact of the model physics, its internal resolution solely (Na et al, 2020), and allows a direct assessment and inter-comparison of the results across resolutions (Demory et al, 2020; Iles et al, 2020; see also Torma et al, 2015 (their Fig 3-6)). I would therefore suggest to redo analyses on a common coarser grid to verify the results shown on native grid. I believe this would strengthen the results. One way to answer all 3 aims of the study could be to split it into two parts: the first part would address 1) and 2) on native grids, considering observations available at various resolutions (such as low-resolution satellite data on grids similar to CMIP); the second part would evaluate the impact of resolution by regridding all data on a common coarser grid.

Thanks for pushing us in this direction. We have now included analyses where all data are regridded to two common grids 0.5°×0.5° grid and a 2°×2°grid.

3) The models are evaluated against E-OBS. E-OBS is a good product that tries to gather the highest number of stations currently available. This is particularly the case over Scandinavian regions, or Germany. However, there are still many regions where the station density is low (e.g. France, Italy, Spain, Switzerland, Austria). Over these regions, it would be better to use national gridded datasets, available at much higher resolution (see Demory et al, 2020 for details). I understand the authors may not want to go in that direction, as it adds a lot of processing time and the definitions of the regions would be slightly different than in the current study. I would therefore suggest to include a discussion on this (and eventually an intercomparison with observational results of Demory et al if feasible). Moreover, for aim 1) of the study, I would suggest the authors to use another lower resolution dataset, such as satellite observations, using a resolution closer to CMIP models. This would give an additional range of observational uncertainty.

Thank you, this is indeed an important and interesting issue. We agree that it would be valuable if regional and/or national observational datasets (with assumed higher quality than E-OBS) could be included for each of the investigated sub-regions, as for example in Demory et al 2020. As suggested we have now included a separate section with a discussion of observations and their associated uncertainties, including E-OBS. To emphasize the importance of high-quality observations and to partly put our results into perspective, we have also included an ASoP analysis comparing to another high-resolution (1x1 km) dataset covering Scandinavia - called NGCD (Nordic Gridded Climate Dataset). There we can clearly see the impact of including such observations, increasing the confidence in the high-resolution RCM model ensemble. We have not included any satellite data as these often (at least the ones we are aware of) has limited coverage or lower quality over high latitudes.

4) Please verify the use of model resolution when you actually refer to model horizontal grid spacing. The model effective resolution is typically 4 to 8 times the model horizontal grid spacing (Skamarock, 2004; Klaver et al, 2019).

Thanks for reminding us about this. We tried to straighten up the terminology so that we use "grid spacing" when talking about distances in km and "resolution" in more general statements, like comparing high and low resolution models.

5) Most analyses have been performed annually. It would be good to show them seasonally as well (at least DJF and JJA), as the processes driving precipitation are different and RCMs depend more on GCMs in DJF than JJA (e.g. Hall, 2014; Prein et al,2016; Fernandez et al, 2019).

We have now also included analyses of DJF and JJA. However we could not present results for all regions, seasons and resolutions as this would mean at least a 12 fold increase of the number of figures.

6) The abstract needs to be revised. It writes very general conclusions as it stands. See detailed suggestions below. This is true as well for the entire text. Some sentences area bit hard to read, and in many places, it reads like general statements or approximative sentences. I provided some suggestions for some of them below, but a careful review of the language would clarify the text and be beneficial to the final paper.

The abstract is rewritten to be more precise, and so is the rest of the text. We hope in a satisfactory way. Thanks for the detailed comments.

7) For reproducibility of the results, it appears important to list the models that were considered for the study.

We have inserted a new Table 1 listing the GCMs and a new Table 2 listing the RCMs.

Detailed comments:

Title: the importance of model ’horizontal’ resolution... from global to regional ’models’

Changed as suggested.

L. 10: model ’horizontal’ resolution

Changed as suggested.

L. 17-18: I find this conclusion too general. This depends on seasons, and most of the analyses have been performed annually.

The abstract is rewritten to be more precise.

L. 20: I don’t agree with this. The authors have shown here that the improvement is systematic across models but that there is a large inter-model variability.

This is rephrased to: “Even though higher resolution improves the simulated precipitation in a systematic way, the inter-model variability is still large. This means that the quality of an ensemble depends also on the models it consists of and not only the average resolution of the ensemble.”

L. 21: I agree with this, but I think it cannot be generalised for all resolutions. The authors have shown here that the averaged resolution of CMIP5 and CMIP6 anyway is too low to capture the characteristics of precipitation, at least against E-OBS and other higher resolution ensembles.

We imply that this is valid for the resolutions used in RCMs. To make this clearer we start the section with “Once reaching ~50 km …”.

L. 22: again, this depends on the season and the authors have mostly worked with annual means.

The abstract is rewritten to be more precise.

L. 23: different RCMs driven by the same GCM give different results, but the same RCM driven by different GCMs also give different results (e.g. Vautard et al, 2020).

That's true, and we know this of course. We show it in Fig. 10 and mention it at a few different times. We change the sentence to : "The result of a RCM simulation depends on the driving GCM, but the difference in simulated precipitation between an RCM and the driving GCM depends more on the choice of RCM and less on the down-scaling itself; as different RCMs driven by the same GCM may give different results."

If Vautard et al., 2020 is published before this goes to print we will add a reference to that.

L. 24-25: Given the complementarity to Demory et al (2020), this sentence needs to be rewritten.

This is changed to: "The results presented here are in line with previous similar studies. To these studies we add details about the spread between resolutions and between models."

L. 28: delete 'precipitation extremes' in 'precipitation extremes (heavy precipitation events)' -> heavy precipitation events

Changed as suggested

L. 34: see also Ban et al, 2015

A reference to Ban et al., 2015 is added.

L. 38-39: could the authors add references to support this sentence?

We added references to Champion et al., 2011; Zappa et al., 2013.

L. 40: 'statistically': remove

Changed as suggested

L. 40: 'decreasing' -> 'refining'

Changed as suggested

L. 45: these papers are among many others (e.g. Delworth et al, 2012; Kinter et al,2013; Roberts et al, 2018 and references therein)

We added these references.

L. 47: please also refer to more recent studies

We added references to Dai 2006; Stratton and Stirling, 2012; Gao et al., 2017

L. 53-54: Please be careful not to suggest that climate change response in RCM versus GCM may be solely due to resolution. They also depend on the forcings. For example, Boe et al, 2020 and Gutierrez et al, 2020 show the impact of different aerosol treatments between GCM and RCM that may explain part of the different climate change response.

Thanks for pointing this out. We added the sentence: “Differences in the treatment of aeorosols are also identified as a reason for differences is climate response between RCMs and GCMs (Boé et al., 2020; Gutiérrez et al., 2020).”

L. 61: please add a reference

We added a reference to Iorio et al., 2004

L. 62: check the study by Vergara-Temprado et al, 2019. They show that it is possible to turn off convection scheme at such resolution and get appropriate results.

These are interesting results, but don’t change the fact that most simulations on 10 km parameterize convection. We changed the sentence to: “Even at grid spacings of around 10 km convection is usually not resolved by the model dynamics but is instead parameterized (although it might be possible to turn off the parameterization already at this kind of resolution (Vergara-Temprado et al., 2019)).”

L. 63: ’certain’: which ones?

Mainly the diurnal cycle. We changed the sentence to: “However, models with parameterized convection often exhibit common biases in the diurnal precipitation cycle”

L. 66: ’giving’ -> ’simulating’

Changed as suggested

L. 68: that is true for models with parametrised convection, please also refer to Vergara et al, 2019 (also in L. 71).

We change the sentence to: “A deficiency of parameterized convection is that it starts too early (e.g. Dai and Trenberth, 2004; Dai, 2006; Brockhaus et al., 2008; Vergara-Temprado et al., 2019).” And also add a reference to Vergada-Temprado et al., 2019 on L.71.

L. 77-78: 12km is not high resolution for RCMs, it is its new standard resolution within CORDEX

What we refer to here are simulations with “convective permitting resolution” which is <5 km (e.g. Coppola et al., 2018)

L. 79: spell out HighResMIP

Changed as suggested

L. 95: high-resolution PRIMAVERA models are available at higher resolution than 40km at mid-latitude (with is the common referenced latitude), or please specify at which latitude this refers to. I would suggest to use the mid-latitude grid spacing (at 50 degreeN), as it is the mid latitude of the European domain (so comparable to EURO-CORDEX grid spacings). It would be clearer to use the term horizontal grid spacing here.

This was not so much a matter of latitudes, but a writing mistake. Never the less, it's a good suggestion to spell out mid-latitude grid spacing. We change as suggested "The models used in this study are a selection of CMIP5 global models (~100-300 km mid-latitude horizontal grid spacing); the high (~25-50 km mid-latitude) and low (~80-160 km mid latitude) resolution versions of the PRIMAVERA global models and the first models from CMIP6 (~100-300 km); and a selection of CORDEX regional models (at 12.5 and 50 km mid-latitude grid spacing)."

L. 95-97 & Table 1: why not considering the full ensembles? How were the models selected? Why are there 5 PRIMAVERA LR and 4 HR?

We selected the models for which we at the time could get daily precipitation. Since we thought that we got ensembles of reasonable sizes we decided not to track down individual models that were not available in common storages. The Primavera LR and HR ensembles are of different resolutions because HadGEM3-GC31 was run at three resolutions. Only one (25 km) was considered as HR, the other two (60 & 130 km) were considered as LR.

Figures 2-3-5: I refer to the revised figures. What does 'act' mean? Please clarify the x-axis 'precipitation bins' and y-axis 'precipitation contribution' labels.

We have updated these figures, and hopefully the titles and axis annotations are more clear now. The figure labels have also been updated to more clearly describe the figure contents.

Figures 2-3: specify in the caption that the thick lines are for ensemble means, and that the bottom panels are differences with E-OBS.

Thanks. The figure labels have been updated accordingly.

Figures 2-3-4: E-OBS is written in Table 1 to be available at 2 resolutions. Which is shown on these figures?

In these figures we use E-OBS with the highest resolution (0.1 deg). It is now specified in the figure labels.

Figures 6-7-8-9-10: I guess E-OBS is shown here at its 2 available resolutions, which one is which?

Correct. We added: “E-OBS at 0.25° (grey) and 0.1° km (black).”

L. 150: bottom left panel for the Alps. Also, CMIP6 upper end seems to be around 50mm/day and CORDEX HR over 100mm/day.

Correct, bottom left and bottom right was mixed up. The sentence is changed

Figure 3: The spread is much larger in CORDEX than CMIP6 in JJA. It shows that CORDEX is not so sensitive to the GCM boundary conditions but to different parametrisation schemes in JJA. The spread is determined by the min and max values for bothEUR-11 and EUR-44. So are these min and max values only represented by 1 RCM, or1 RCM-GCM simulation? If the spread is represented by min and max values, wouldn’t it be better to plot the median instead of the mean?

Indeed, the spread defined by max/min values is very sensitive to possible “outliers” that might not be a good representation of the ensemble spread. It is not entirely clear what the best way would be to indicate the spread of such relatively small ensembles (without the use of more sophisticated statistical techniques like bootstrapping). We have changed from max/min to instead show the 5-95 percentile range. We further agree that median values would be more appropriate than mean values and thus have changed accordingly.

Figure 4: It seems biased to consider EUR-11 as the reference and compare observations to that reference, possibly because, although EUR-11 has a higher resolution, their mean climate seems too wet against high density observations as shown by Demory et al (2020), although I agree observations have undercatch errors. If E-OBS are considered too low resolution and not trustable, considering datasets with higher density stations as the reference would be necessary here. Moreover, the ensembles are clearly compared to each other in this figure,

with respect to EUR-11. It would be good to see this analysis performed on a common grid to evaluate how it affects the conclusions. It could be done both at 50km for EUR-11, EUR-44 and PRIMAVERA, and then redone for all datasets at 150 (or even 300km), as done in Torma et al, 2015.Why writing the E-OBS total annual mean in the box if EUR-11 is used as a reference?

We have now included analysis on common grids (at two different resolutions, 0.5^{o}x0.5^{o} and 2^{o}x2^{o}), although not presented in the format as shown in Fig. 4 (see Figures S1 and S2 in Supplementary). The interpolation to common grids of course have an effect but the overall conclusions are not seriously impacted. Further on, as mentioned above, we included another, regional high-quality, data set in an ASoP analysis to emphasize the importance of such data sets and possible impact on the results. Still, we are limited for most regions to the E-OBS data as reference while acknowledging its inherent uncertainties. Regarding Fig.4 your concerns about having EUR-11 as reference is understandable and we have changed to E-OBS as reference instead.

L. 181: more strongly biased lower -> more negatively biased: I suggest not to use the word 'bias' when compared to an ensemble, which is itself biased.

We changed to: "Region total seasonal precipitation (averaged within each ensemble), are either mostly in the range of +/- 20 % from CORDEX HR (e.g. eastern Europe, EA) or with larger negative values…".

L. 193-194: Observations have uncertainties but EUR-11 could also rain too much along coastlines and over topography.

True, we changed to "…both factors contributing to uncertainties in quality and representativeness of observational and simulated data."

L. 211-212: Fig. 5 shows results for the annual mean, so this conclusion may be different at seasonal means (at least between DJF and JJA), so I would suggest to show these seasonally as well. Moreover, the delta in grid spacing between CORDEX LR(50km) and HR (12.5km) is similar for all models (delta=4), so the impact of resolution is potentially more similar (although it depends on models). This is more complex for the PRIMAVERA models that have various deltas between the LR and HR versions. I counted that deltas vary between 2 for most models, 3 for a couple and 5.4 for theHadGEM3 model (https://www.primavera-h2020.eu/modelling/our-models/).

Moreover, note that PRIMAVERA HR uses exactly the same tuning parameters as their LR version, so the effect of resolution solely is seen here (this is not the case for the CORDEX ensembles that may use different model versions). Something that could be interesting to show here is whether, depending on their delta in grid spacing, some PRIMAVERA models show larger differences than some others. But I would not generalise, based on ensemble means, that resolution in CORDEX has more effect than resolution in PRIMAVERA. It would be good to see the spread of the ensembles on figure 5.

A good point. We added the sentences: "Some differences between the CORDEX and PRIMAVERA ensembles should be noted. The PRIMAVERA models use the same tuning parameters for both the LR and HR version, but on the other hand the differences in resolution between LR and HR varies between models. The CORDEX ensembles have the same difference in resolution for all models, but the LR and HR simulations may be run with different models versions. Hence, all differences between PRIMAVERA and CORDEX ensembles can't be generalised to be attributed by resolution alone."

We also plotted the absolute difference in the precipitation indices between LR and HR against the ratio LR/HR. It turns out the the correlation is weak, e.g. the spread within CORDEX ensemble is large although all models have the same ratio.

In Figure 5 the absolute values for each model have now been included as well (in addition to the ensemble means) showing the ensemble spread.

L. 216-218: I agree with this hypothesis, and yet you found greater differences in CORDEX (driven by same low-resolution GCMs) than in PRIMAVERA (L. 211-212).I think this highlights the need for analyses on a common grid, based on seasonal means, and taking into account the fact that CORDEX and PRIMAVERA have different deltas in grid spacing.

A description of winter and summer is included in the text. Our analysis on common grids and of resolution delta doesn't suggest that this explains the differences. Rather, the conclusion is that for high intensities model resolution and performance is more important than the driving GCM. We don’t know the full answer. This section was also meant to show that there are unresolved issues and to point to possible future studies.

We added the following:”Still, the largest differences are seen in the CORDEX ensemble where the LR and HR models are run with the same coarse resolution GCM. This suggests that (regional) model resolution and performance is what determines high precipitation rates, rather than the driving GCM. ”

L. 226-227: this is not a sentence/question: please rephrase.

We changed to: “When do intense precipitation events occur in the high-resolution models? the kind of events that are rarely seen or absent in the low resolutions simulations.”

Figures 6-9: I considered the revised figures. I still do not understand why some values are not shown. For example: Fig. 6 top left: For one of the CMIP5, only the 10th and90th percentiles are shown, nothing else it seems. For some other CMIP5 and CMIP6models, the boxes are drawn but not the whiskers.

In small regions like the Alps and in models of coarse resolution the number of data points are actually too few to make good statistics. This means that calculation of percentiles can be difficult. Since this only happens in some regions for a small number of models we consider this a major problem.

Figures 6-7: it seems that CMIP5, CORDEX LR and HR have a larger variability, so is the variability of CORDEX driven by the variability of CMIP5? This could be answered by looking at the seasonal means (DJF and JJA).

We don’t agree that the variability is large in CMIP5, rather the variability increases with resolution. The signal is the same for the individual seasons, but less pronounced since the potential number of days is smaller when divided over four seasons instead of counted over the whole year.

L. 227-246: Again for these analyses, the metrics can be analysed for each ensemble on their native resolution, but if the ensembles are compared to each other, as written in the text, then the analyses need to be redone on a common grid.

Yes, descriptions of summer and winter are now included and figures of this when relevant. The analysis now also include data on common grids.

L. 232-233: Would it be possible to show this with seasonal means?

Yes, descriptions of summer and winter are now included and figures of this when relevant.

L. 245: isn't it 20 mm/day instead of 10?

Yes, we changed to 20.

L. 271: rephrase 'negative for some models and positive for some' as it reads too vague

We changed to: "The differences are small, mainly within ±10 days year^{-1}."

Figure 10: This intercomparison needs to be performed on a common grid

This is now done.

L. 281: left -> right

Changed as suggested

L. 283: right -> left

Changed as suggested

L. 283-284: Note that ECMWF HR is 25km grid spacing output at 50km, and LR is 50km output at 100km grid spacing. The delta in grid spacing is therefore 2, and the output are regridded to coarser resolution. This may impact the results.

We added a new Fig 12 showing the correlation between difference and delta.

L. 300-301: Demory et al have revised the manuscript with a focus on EUR-11.

Thanks for pointing that out we change to: "In a similar study Demory et al. (2020) compares PRIMAVERA models with CORDEX LR and CORDEX HR."

L. 306: give extremes that are heavier and more frequent -> simulate more intense and more frequent heavy precipitation. I would avoid the term 'extremes' with such low-resolution models, and refer instead to 'heavy' or 'intense'.

We change to: "They conclude that high resolution models systematically give intense precipitation that is heavier and more frequent."

L. 308: overestimation compared to E-OBS

Changed as suggested

L. 315: CMIP6 and CMIP5

Changed as suggested

L. 318-319: this is probably particularly the case for JJA (as shown in fig 3), but for this conclusion it would be good to see DJF and JJA for fig 6-9.

Information about DJF and JJA for Figs 6-9 are now included in text or in the supplementary.

L. 320: not only. E-OBS is not based on the full network of rain gauges over some other countries, such as France.

We added: “E-OBS is not based on the full network of rain gauges in all countries, which could also lead to undercatch.”

L. 332: scale -> grid

Changed as suggested

L. 336: will have -> has

Changed as suggested

L. 340-341: yes but PRIMAVERA tends to be drier than CORDEX in all seasons.

We changed to: “…furthermore GCMs and RCMs of comparable resolution simulate comparable precipitation climates, even though PRIMAVERA is often drier than CORDEX.”

L. 343: to -> too

Changed as suggested

L. 344: agree -> agrees

Changed as suggested

L. 345: the quantification can be done if performed on common grids

For the individual models it is possible on common grids, which we now do. On the ensemble level it’s more difficult. By difficult we mean that it’s not so obvious how resolution influences the ensemble mean because the actual model members used impact the ensemble mean more than the resolution of the members.

L. 348-349: this can depend on seasons

Yes, we addedd the following to the end of the sentence: ”especially for heavy precipitation and particularly in summer.”

L. 350-351: this needs to be rephrased, as Demory et al have evaluated CMIP5,CORDEX LR/HR and PRIMAVERA HR

We changed to: “The results presented here are in line with previous similar studies using different methods (Demory et al., 2020; Iles et al., 2020) To these studies details are added about the spread between resolutions and between models.”

Proper acknowledgement needs to be given to the PRIMAVERA, CORDEX, and CMIP modelling groups

Changed as suggested

There are several typos in the text, please check carefully (e.g. L.8: effects -> affects; L. 20: in depending -> depends; L. 180: region -> regional (and remove comma afterwards); L. 218: were -> where; many others)

These and others are corrected. We apologise for the lack of proof reading, as a reviewer it’s annoying to have to correct typos.

L. 139 and 141: below/above c: are these typos?

We removed the “c” for circa as it only confuses.

References:

*Ban N., Schmidli, J., and Schär, C.: Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster?, Geophys. Res. Lett., 42, 1165–1172, https://doi.org/10.1002/2014GL062588, 2015.

*Boé, J., Somot, S., Corre, L., and Nabat, P.: Large differences in Summer climate change over Europe as projected by global and regional climate models: causes and consequences, Clim. Dynam., 54, 2981–3002, https://doi.org/10.1007/s00382-020-05153-1, 2020.

*Delworth TL, Rosati A, Anderson W, Adcroft AJ, Balaji V, Benson R, Dixon K, GriffiesSM, Lee HC, Pacanowski RC, Vecchi GA, Wittenberg AT, Zeng F, Zhang R (2012)Simulated climate and climate change in the GFDL CM2.5 high-resolution coupledclimate model. J Clim 25:2755–2781. doi:10.1175/JCLI-D-11-00316.1

Fernández, J., Frías, M. D., Cabos, W. D., Cofiño, A. S., Domínguez, M., Fita, L.,Gaertner, M. A., García-Díez, M., Gutiérrez, J. M., Jiménez-Guerrero, P., Liguori, G.,Montávez, J. P., Romera, R., and Sánchez, E.: Consistency of climate change projec-tions from multiple global and regional model intercomparison projects, Clim. Dynam.,52, 1139-1156, https://doi.org/10.1007/s00382-018-4181-8, 2019.

*Gutiérrez, C., Somot, S., Nabat, P., Mallet, M., Corre, L., van Meijgaard, E., Perpiñán,O., and Gaertner, M. A.: Future evolution of surface solar radiation and photovoltaic potential in Europe: investigating the role of aerosols, Environ. Res. Lett.,15, 034035, https://doi.org/10.1088/1748-9326/ab6666, 2020.

Hall,A.:Projectingregionalchange,Science,346,1460–1462,https://doi.org/10.1126/science.aaa0629, 2014.

*Kinter III JL, Cash B, Achuthavarier D, Adams J, Altshuler E, Dirmeyer P, Doty B,Huang B, Jin EK, Marx L, Manganello J, Stan C, Wakefield T, Palmer T, HamrudM, Jung T, Miller M, Towers P, Wedi N, Satoh M, Tomita H, Kodama C, Nasuno T,Oouchi K, Yamada Y, Taniguchi H, Andrews P, Baer T, Ezell M, Halloy C, John D,Loftis B, Mohr R, Wong K (2013) Revolutionizing climate modeling with project Athena:a multi- institutional, international collaboration. Bull Am Meteorol Soc 94:231–245.doi:10.1175/BAMS-D-11-00043.1

Klaver R, Haarsma R,Vidale PL, Hazeleger W. Effective resolution inhigh resolutionglobal atmospheric models forclimate studies. Atmos Sci Lett. 2020;21:e952.Na, Y., Fu, Q., & Kodama, C. (2020).Precipitation probability and itsfuture changes from a global cloudâ˘AˇRresolving model and CMIP6 simula-tions.Journal of Geophysical Research:Atmospheres,125, e2019JD031926.https://doi.org/10.1029/2019JD031926

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., FoxMaule, C., van Meijgaard, E., Déqué, M., Nikulin, G., Vautard, R., Colette, A.,Kjellström, E., and Jacob, D.: Precipitation in the EUROâ˘AˇRCORDEX 0.11◦and0.44◦simulations: high resolution, high benefits?, Clim.Dynam., 46, 383-412,https://doi.org/10.1007/s00382-015-2589-y, 2016.

Skamarock, W.C., 2004. Evaluat-ing mesoscale NWP models using kinetic energy spectra. Monthly weather review,132(12), pp.3019-3032.

*Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., Chang,P., Christensen, H. M., Danilov, S., Demory, M.-E., Griffies, S. M., Haarsma, R., Jung,T., Martin, G., Minobe, S., Ringler, T., Satoh, M., Schiemann, R., Scoccimarro, E., Stephens, G., and Wehner, M. F.: The Benefits of Global High Resolution for ClimateSimulation: Process Understanding and the Enabling of Stakeholder Decisions at theRegional Scale, B. Am. Meteorol. Soc., 99, 2341–2359, https://doi.org/10.1175/BAMS-D-15-00320.1, 2018.

Torma, C., Giorgi, F., and Coppola, E.: Added value of regional climate modelingover areas characterized by complex terrainâ ˘A ˇTPrecipitation over the Alps, J. Geophys.Res. - Atmos., 120, 3957–3972, https://doi.org/10.1002/2014JD022781, 2015.

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Coppola, E., Bülow, K.,Corre, L., van Meijgaard, E., Nogherotto, R., et al.: Assessment of the large EURO-CORDEX regional climate simulation ensemble, J. Geophys. Res. – Atmos., in press.

*Vergara-Temprado, J., Ban, N., Panosetti, D., Schlemmer, L., and Schär, C. (2019).Climate models permit convection at much coarser resolutions than previously consid-ered. J. Clim., JCLI-D-19-0286.1. doi:10.1175/JCLI-D-19- 0286.1.

Interactive comment on “The importance of model resolution on simulated precipitation in Europe –from global to regional model” by Gustav Strandberg and Petter Lind

Anonymous Referee #2

The paper “The importance of model resolution on simulated precipitation in Europe –from global to regional model” by Strandberg and Lind assesses the ability of a large set of climate models in simulating precipitation (particularly extremes) in European subregions. The authors find that models with coarse grid spacings underestimate the amount and frequency of extreme precipitation but that the variability between models can be larger than the sensitivity to grid spacing. The novel contribution of this study is the inclusion of global climate model data in their analysis since very similar and more detailed analyses have been done with regional models over Europe. I have two major concerns with this manuscript.

We thank the reviewer for making the effort for reviewing our paper and for all comments. Responses follow below in red. In the markup version of the revised manuscript substantial changes are also marked in red.

First, it does not account for the spatial dependence of extreme precipitation. I argue that the authors can obtain the same results by first aggregating E-Obs observations to a coarser grid and then comparing the aggregated extreme precipitation with the original E-Obs data. They would also see that the coarser version of E-Obs “underestimates” extreme frequency and magnitude. Coarse-resolution models should not reproduce the magnitude of extreme events on local scales since they model aggregated rainfall over large areas (e.g., 100x100 km).

The main objective of the paper is not really to focus in precipitation extremes but rather the full distributions (which includes aspects of extremes). We are aware that extremes may not always be well represented in observations, depending on multiple factors including the spatial and temporal character of such events, and we try to acknowledge these weaknesses in the observations in the discussions of the results. As you say the model grid resolution sets limits to what the model can actually resolve but we argue that it is still important to show to what extent different models, from GCMs to RCMs, exhibit similarities and differences in the full precipitation distributions for different regions and seasons.

My second concern is the use of E-Obs for this analysis. E-Obs has very low station density over large parts of Europe and heavily underestimates extreme precipitation. There are other observational datasets available that are far more appropriate for the presented analysis. More details on these comments including relevant literature is provided below.

It is true that E-OBS is inherently associated with uncertainties and the quality is highly dependent on the underlying station density as you say. We intended here to keep the model-observation comparison consistent for all sub-regions by using the same observational data set and hence constrained the comparison to E-OBS solely. We have included a separate section (Sec 2.2) with a discussion of observations and related uncertainties. Furthermore, to highlight the importance of high-quality data sets, we have included in one of the ASoP analyses a regional high-resolution data set (Nordic Gridded Data set, NGCD) that covers the Scandinavian region (see Fig. 3 in Supplementary material). It is seen that NGCD has higher contributions for both low and high precipitation intensities, providing more confidence in especially the RCMs (at least over this region).

General Comments:

1.I have major concerns with your approach to compare extreme precipitation. Extreme precipitation is strongly scale dependent and largest on point scales (e.g. measured by precipitation gauges) and decreases on larger spatial-scales. E-OBS for example has way weaker extreme precipitation than other regional datasets in Europe that feature higher resolution and a higher station density (e.g. Prein and Gobiet 2017). If you compare extreme precipitation on the model native grid, you mix the model ability in simulating extreme precipitation with the spatial scale on which the model simulates extremes. E.g., extreme precipitation in a 100 km grid spacing model should not match observed extreme precipitation on a 25 km grid. In this case the only way to do a fair comparison is to aggregate the 25 km grid observations to the 100km model grid. This aggregation does not introduce large biases such as you state for interpolation (in Line 127-128).

We have now also included analyses where all data are regridded to a 0.5°×0.5° and a 2°×2° grid. This makes it possible for us to separate the effect of model physics from the effect of just having more data points.

2. E-Obs should be used with care for extreme precipitation (Haylock et al. 2008). There are other/regional datasets in Europe that are much better suited for the assessment of extreme precipitation (see Prein and Gobiet et al. 2017).

As mentioned in the response above we have included one other regional data set for the region of Scandinavia (the NGCD data set, see Fig. S3 in Supplementary). However, we would like to emphasize again that extreme precipitation is not the main focus of the study, rather a more holistic approach in the investigation of the model's representation of precipitation over Europe.

3. You are missing to discuss and to refer relevant literature on the ERUO-CORDEX simulations that performed very similar analysis as you present. Kotlarski et al. (2014), Casanueva et al. (2016), and Prein et al. (2016) address similar questions and come to fairly similar conclusions. The novelty of your analysis is that you also include GCM data, which is a valuable contribution but does not change the major conclusions. You should also take a look at Thackeray et al. (2018) who show a highly relevant analysis of model grid spacing and extreme precipitation on a global-scale.

The Introduction has been expanded with a paragraph discussing Demory et al., and Iles et al., as well as other similar studies using CORDEX data:

A few studies have been made investigating how model resolution affects the simulated precipitation in the CORDEX ensembles, comparing 50 km and 12.5 km grid spacing. A clear result is that precipitation generally increases with higher resolution, which sometimes means that the bias increases when precipitation is added to already wet models (Kotlarski et al., 2014; Casanueva et al., 2016); something that is also seen in simulations with global models (e.g. Thackeray et al., 2018). An overall improvement of mean precipitation is not seen the high resolution CORDEX simulations, except for regions with complex topography (Korlarski et al., 2014; Casanueva et al., 2016; Prein et al., 2016). Prein et al. (2016) looked at local precipitation on short time scales. They find that 12.5 km simulations better represent extreme and mean precipitation, also when simulations are aggregated to 50 km. They note, however, that the results are highly dependent on which observations the simulations are compared with. They also note that improvements are on the ensemble as a whole, and not necessarily for each individual model. In similar studies as the present Iles et al. (2019) and Demory et al. (2020) compare CORDEX simulations with simulations from CMIP5 and Primavera. They see that precipitation increases with resolution so that CMIP5 underestimates precipitation amounts and CORDEX overestimates it, when compared to E-OBS, and that the effect of resolution is largest in complex topography. They also find that Primavera performs similarly to CORDEX when run on the same resolution, which is interesting regarding that the Primavera models are developed for low resolution. Iles et al. (2019) also find considerable inter-model differences meaning that improvements are seen on the ensemble level rather that for individual models.

4. Please be careful with the use of model resolution. In most cases you refer to model grid spacing. Model resolution depends on the numeric diffusion in the model and models with the same grid spacing can have different resolutions. The effective resolution of a model is typically 4-8 times its grid spacing (e.g., Skamarock 2004).

Thanks for reminding us about this. We tried to straighten up the terminology so that we use “grid spacing” when talking about distances in km and “resolution” in more general statements, like “comparing high and low resolution models”.

5. There are many typos and grammar errors in the document. Please consider using a proofreader before resubmitting the document.

Typos are corrected. We apologise for the lack of proof reading, as a reviewer it’s annoying to have to correct typos.

Literature:

Casanueva, A., Kotlarski, S., Herrera, S., Fernández, J., Gutiérrez, J.M.,Boberg, F., Colette, A., Christensen, O.B., Goergen, K., Jacob, D. and Keuler, K., 2016.Daily precipitation statistics in a EURO-CORDEX RCM ensemble: added value of rawand bias-corrected high-resolution simulations. Climate dynamics, 47(3-4), pp.719-737.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. and New, M.,2008. A European daily highâ ˘A ˇRresolution gridded data set of surface temperatureand precipitation for 1950–2006. Journal of Geophysical Research: Atmospheres,113(D20).

Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goer-gen, K., Jacob, D., Lüthi, D., Van Meijgaard, E. and Nikulin, G., 2014. Regional climatemodeling on European scales: a joint standard evaluation of the EURO-CORDEX RCMensemble. Geoscientific Model Development, 7, pp.1297-1333.

Prein, A.F. and Gobiet, A., 2017. Impacts of uncertainties in European gridded precip-itation observations on regional climate analysis. International Journal of Climatology,37(1), pp.305-327.

Prein, A.F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Maule,C.F., Van Meijgaard, E., Déqué, M., Nikulin, G. and Vautard, R., 2016. Precipitationin the EURO-CORDEX 0.11 and 0.44â´LŸ simulations: high resolution, high benefits?.Climate dynamics, 46(1-2), pp.383-412.

Skamarock, W.C., 2004. Evaluating mesoscale NWP models using kinetic energyspectra. Monthly weather review, 132(12), pp.3019-3032.

Thackeray, C.W., DeAngelis, A.M., Hall, A., Swain, D.L. and Qu, X., 2018. On the con-nection between global hydrologic sensitivity and regional wet extremes. GeophysicalResearch Letters, 45(20), pp.11-343.

The importance of model horizontal resolution on simulated precipitation in Europe – from global to regional models

Gustav Strandberg[1,2], Petter Lind[1,2,3]

[1]Rossby Centre, Swedish Meteorological and Hydrological Institute, SMHI, Norrköping, SE-602 19, Sweden
[2]Bolin Centre for climate research, Stockholm University, Stockholm, SE-106 91, Sweden
[3]Department of meteorology, Stockholm University, Stockholm, SE-106 19, Sweden

Correspondence to: Gustav Strandberg (gustav.strandberg@smhi.se)

Abstract. Precipitation is a key climate variable that affects large parts of society, especially in situations with excess amounts. Climate change projections show an intensified hydrological cycle through changes in intensity, frequency, and duration of precipitation events. Still, due to the complexity of precipitation process and its large variability in time and space, weather and climate models struggle to represent it accurately. This study investigates the simulated precipitation in Europe in a range of climate model ensembles that cover a range of model horizontal resolution. The ensembles

used are: Global climate models (GCMs) from CMIP5 and CMIP6 (~100-300 km horizontal resolution), GCMs from the PRIMAVERA project at low (~80-160 km) and high (~25-50 km) resolution and CORDEX regional climate models (RCMs) at low (~50 km) and high (~12.5 km) resolution. The aim is to investigate the differences between models and model ensembles in the representation of the precipitation distribution in its entirety and through analysis of selected standard precipitation indices, for different seasons and different regions of Europe. In addition, the model ensemble performances are compared to gridded observations from E-OBS.

The impact of model resolution on simulated precipitation is evident. Overall, in all seasons and regions the largest differences are seen for moderate and high precipitation rates, where the largest contribution is seen in the RCMs with highest resolution (i.e. CORDEX 12.5 km) and lowest in the CMIP GCMs. However, when compared to E-OBS the high-resolution models most often overestimate high-intensity precipitation amounts, especially the CORDEX 12.5 km resolution models. An additional comparison to a regional data set of high-quality lends, on the other hand, more confidence to the high-resolution model results. The effect of resolution is larger for precipitation indices describing heavy precipitation (e.g. maximum one-day precipitation) than for indices describing the large-scale atmospheric circulation (e.g. the number of precipitation days), especially in regions with complex topography and in summer when precipitation is predominantly caused by convective processes. Importantly, the systematic differences between low resolution and high resolution remain also when all data are regridded to common grids of 0.5°×0.5° and 2°×2° prior to analysis. This shows that the differences are effects of model physics and better resolved surface properties and not due to the different grids on which the analysis is performed. PRIMAVERA high resolution and CORDEX low resolution give similar results as they are of similar resolution.

Within the PRIMAVERA and CORDEX ensembles there are clear differences between the low- and high-resolution simulations. Once reaching ~50 km the difference between different models is often larger than between the low- and high-resolution versions of the same model. Even though higher resolution most often improves the simulated precipitation in comparison to observations, the inter-model variability is still large, particularly in summer when smaller scale processes and inter-actions are more prevalent and model formulations (such as convective parameterizations) become more important.

The result of an RCM simulation depends on the driving GCM, but the difference in simulated precipitation between an RCM and the driving GCM depends more on the choice of RCM, and the model physics of that model, and less on the down-scaling itself; as different CORDEX RCMs driven by the same GCM may give different results. The results presented here are in line with previous similar studies. To these studies we add details about the spread between resolutions and between models.

1 Introduction

Precipitation is a key climate variable affecting the environment and human society in different ways and on different temporal and spatial scales. In particular, heavy precipitation events may lead to large damages caused by floods or landslides, while the absence of precipitation may cause droughts and has impact on water- and hydropower supply. In recent decades there has therefore been extensive study, and considerable advancement in our understanding, of the response of extreme precipitation to climate change (O'Gorman, 2012; Kharin et al. 2013; Donat et al., 2016; Pfahl et al. 2017). For example, it is widely held through theoretical considerations and model experiments that extremes will respond differently than changes in mean precipitation (e.g. Allen and Ingram 2002; Pall et al 2007; Ban et al., 2015).

Still, the simulation of precipitation in weather and climate models is challenging because of the wide range of processes involved that acts and interacts on widely different temporal and spatial scales. An accurate representation of precipitation in models requires skill in simulating (1) the large-scale circulation, (2) interaction of the flow with the surface, and, (3) convection and cloud processes. With the typical horizontal grid resolution of O (100 km) of global climate models (GCMs) point (1) can to a large extent be properly represented but less so for (2) and (3) (e.g. van Haren et al., 2015; Champion et al., 2011; Zappa et al., 2013). In particular, atmospheric convective processes are not resolved and needs to be treated with convection parameterizations. As the range of scales resolved is broadened through refining the horizontal grid spacing the simulation of precipitation generally improves. This is achieved through more realistic representation of surface characteristics (such as topography, coastlines and inland lakes and water bodies) and through more accurately solving the motion equations resulting in more accurate horizontal moisture transport and moisture convergence (Giorgi and Marinucci 1996; Gao et al. 2006; Prein et al. 2013a). Indeed, GCMs with ~25-50 km grid spacing show promise to improve simulation of precipitation (van Haren et al., 2015; Delworth et al., 2012; Kinter et al., 2013; Haarsma et al., 2016; Roberts et al., 2018a; Baker et al., 2019).

Dynamical down-scaling of GCMs with regional climate models (RCMs) allows for even finer grids which leads to more detailed information of and further improvements in regional and local climate features, for example spatial patterns and distributions of precipitation in areas of complex terrain (Rauscher et al., 2010; Di Luca et al., 2011; Prein et al., 2013b). This can also have important implications for climate change signals. Giorgi et al. (2016) found that an ensemble of RCMs at ~12 km resolution showed consistently an increase in summer precipitation over the Alps region which contrasted to the forcing GCMs that instead showed a decrease. The different responses were attributed to increased convective rainfall in the RCMs due to enhanced potential instability by surface heating and moistening at high altitudes not captured by the GCMs. Differences in the treatment of aeorosols are also identified as a reason for differences in climate response between RCMs and GCMs (Boé et al., 2020; Gutiérrez et al., 2020). RCMs are constrained by the lateral boundary conditions provided by the forcing GCM and studies of RCM ensembles have shown that the choice of forcing GCM have introduced the major part of the overall uncertainty in regional climate (e.g. Déqué et al., 2007; Kjellström et al., 2011). This effect is relatively more important for large-scale precipitation systems, for example frontal systems associated with extra-tropical cyclones. In seasons and regions when smaller scale processes like convection dominate, for example in summer over mid-latitudes, simulated precipitation is to a larger degree dependent of the RCM itself, in terms of grid resolution and sub-grid scale parameterizations (e.g. Iorio et al., 2004). A recent study investigated the effects of model resolution on local precipitation on short time scales and found that the 12.5 km simulations better represent daily and sub-daily extreme and mean precipitation, also when simulations are aggregated to 50 km (Prein et al., 2016). They note, however, that the results are highly dependent on which observations the simulations are compared with, and that improvements are seen for the ensemble mean, and not necessarily for each individual model. In similar studies as the present one Iles et al. (2019) and Demory et al. (2020) compare simulations from the CORDEX, CMIP5 and PRIMAVERA ensembles. The results show that precipitation increases with resolution and that, when compared to E-OBS, CMIP5 underestimates precipitation amounts while CORDEX overestimates it, and the effect of grid resolution is largest in areas with complex topography. They also find that PRIMAVERA performs similarly to CORDEX when run on the same resolution, which is interesting regarding that the PRIMAVERA models are developed for low resolutions. Iles et al. (2019) concluded from the considerable inter-model differences that improvements are seen for the ensemble mean rather that for individual models.

Although increased grid resolution often leads to improved simulation of precipitation convection is usually not resolved by the model dynamics, even at grid spacings of around 10 km, but is instead parameterized (although it might be possible to turn off the parameterization already at this kind of resolution (Vergara-Temprado et al., 2019)). The choice of convection parameterization can have various effects on the occurrence and amount as well as on the onset timing and location (e.g. Dai et al., 1999; Dai 2006; Stratton and Stirling, 2012; Gao et al., 2017). Commonly, models with parameterized convection exhibit biases in the diurnal precipitation cycle (Liang, 2004; Brockhaus et al., 2008; Gao et al. 2017), sometimes regardless of increases in grid resolution (Dirmeyer et al., 2012). In addition, models of coarse resolution often suffer from simulating precipitation over too large area compared to observations, and usually also too many days with weak precipitation (the "drizzle" problem) (e.g. Dai, 2006, Stephens et al., 2010). At sufficiently high resolution (< 4 km) models start to largely resolve deep convection enabling the parameterization to be turned off, so called "convection-permitting" models (Prein et al., 2015; Vergada-Temprado et al., 2019). Convection-permitting regional climate models (CPRCMs) are widely shown to reduce, at least to some extent, these biases, most evidently by improving the match of the diurnal cycle to observations (e.g. Prein et al., 2013a; Ban et al., 2014; Brisson et al., 2016; Gao et al., 2017; Leutwyler et al., 2017; Belušić et al. 2020) and better representation of sub-daily high-intensity precipitation events (e.g. Ban et al., 2014; Kendon et al., 2014; Fosser et al., 2015; Lind et al., 2020) than models with parameterized convection. A major drawback using these high-resolution climate models is the very high computational cost, making their use in ensembles to only recently emerge (Coppola et al., 2018).

The aim of this study is to:

i. Investigate to what extent a large number of global and regional climate models can reproduce observed daily precipitation climatologies and characteristics over Europe.

ii. Investigate how model horizontal grid resolution in either global or regional models affect the simulated precipitation in Europe; are there systematic differences and if so, are these persistent for different parts of Europe and for different seasons.

To this end, GCMs of standard resolution from the CMIP5 (Climate Model Intercomparison Project phase 5, Taylor et al., 2012) are compared with GCMs which participated in the HighResMIP (High Resolution Model Intercomparison Project, Haarsma et al., 2016) experiment within the H2020-EU-project PRIMAVERA. These models are: ECMWF-IFS (Roberts et al., 2018b), HadGEM3-GC31 (Roberts et al., 2019), MPI-ESM1.2 (Gutjahr et al., 2019), CNRM-CM6.1 (Voldoire et al., 2019) and EC-Earth3P (Haarsma et al., 2020). Furthermore, the first results from the CMIP6 (Climate Model Intercomparison Project phase 6, Eyring et al., 2016) GCMs are included in the analysis. The GCMs are compared with RCMs from CORDEX (COordinated Regional Downscaling EXperiment, Gutowski et al., 2016). This allows for comparisons of different generations of models, global versus regional models and the impact of model horizontal grid resolutions. For a few cases, the same model version has been applied at two different grid resolutions which allows for investigating the impact of resolution alone. The simulated daily precipitation is analysed both in terms of precipitation intensity distributions and through a collection of standard precipitation-based indices.

2 Models and Methods

2.1 Global and regional models

The models used in this study are a selection of CMIP5 global models (corresponding to ~100-300 km horizontal grid spacing at mid-latitudes); the high (~25-50 km) and low (~80-160 km) resolution versions of the PRIMAVERA global models and the first available runs from CMIP6 (~100-300 km); and finally, a selection of CORDEX RCMs (at 12.5 and 50 km). The low-resolution versions in each model ensemble is called LR, and the high-resolution HR. Note that not the full CMIP5, CMIP6 and CORDEX ensembles are used, but rather "ensembles of opportunity" for which daily precipitation were easily available. Table 1 lists the GCM ensembles used. Table 2 lists the GCM RCM combinations used in the CORDEX ensembles. The simulated precipitation for all models is analysed over the PRUDENCE regions in Europe (Fig. 1; Christensen & Christensen, 2007). Prior to analysis all grid points over sea are filtered out, and then for each region and model we calculate precipitation characteristics for all remaining land grid points. The simulations are analysed on their native grids, because this is the kind of data that users of climate simulations will face, and since all interpolation may alter precipitation characteristics (Klingaman et al., 2017). Nevertheless, to investigate all aspects of changed resolution it is sometime necessary to compare simulations on a common grid. In these cases, the results are also aggregated to two common grids with 2°×2° and 0.5°×0.5° grid spacing respectively.

2.2 Observations

Climate model evaluation exercises often rely, when possible, on gridded reference data sets. In this study daily precipitation sums in models are compared with data from E-OBS version 19.0e at 0.1° and 0.25° grid spacing (Cornes et al., 2018). E-OBS comprise daily station values interpolated onto a grid that spans the entire European continent. The main advantage of using E-OBS is the large geographical coverage at a relatively high resolution available over an extended (climatological) time period. It enables a consistent model-observation comparison over the whole continental part of Europe, with its varying climatological and environmental characteristics.

Gridded products, such as E-OBS, involves spatial analysis and interpolation of point measurements onto a regular grid, and are inherently associated with uncertainties originating from both non-climatic influences (e.g. inaccuracies in measurement devices or relocation of measurement sites) and from sampling issues associated with weather and environmental conditions, for example in situations with snowfall in windy conditions (Kotlarski et al. 2019; Rasmussen et al., 2012). The quality of such data sets largely depends on the availability of stations to base the interpolation on, implying that in regions where station density is low the quality of the gridded product is also lower (Herrera et al. 2019). For precipitation this is of even greater importance due to its highly heterogeneous character in both time and space, in particular for high-intensity precipitation events (extremes). These are often local in character (temporally and spatially), even in cases when embedded in larger (synoptic) scale precipitation systems, and can thus be heavily undersampled (Herrera et al. 2019; Prein and Gobiet 2017). Furthermore, mountainous areas act as strong forcing of precipitation giving rise to large spatial variability over the terrain. Combined with the lack of dense networks of stations in these regions, and usually also a higher occurrence of snowfall, makes it very difficult to achieve highly reliable data over mountains (e.g. Hughes et al. 2017; Lundquist et al. 2019).

The quality of E-OBS varies over Europe (see Fig. 1 in Cornes et al. 2018); the station density is for example very high over Scandinavia, Germany and Poland, while it is lower in Eastern Europe and in the Mediterranean region. Gridded regional or national data sets may offer higher quality as these are generally based on a denser station network and are often also provided with higher spatial and/or temporal resolution compared to E-OBS (Kotlarski et al. 2019, Prein and Gobiet 2017). Here, we limit the comparison to E-OBS only. However, to assess the impact of high-quality regional data, an additional analysis of the precipitation distributions was performed, using ASoP analysis (see Sec. 2.3), comparing models and E-OBS against the NGCD (Nordic Gridded Climate Dataset, Lussana et al. 2018) data set. NGCD is based on daily station data for precipitation and temperature, interpolated onto a 1x1 km grid covering Scandinavia.

2.3 ASoP and precipitation indices

To investigate the effect of model grid resolution on the full distributions of daily precipitation intensities, we use the ASoP (Analysing Scales of Precipitation) method (Klingaman et al., 2017; Berthou et al., 2018). ASoP involves splitting precipitation distributions into bins of different intensities and then provides information of the contributions from each precipitation intensity separately to the total mean precipitation rate (i.e. given by all intensities taken together). In the first step, precipitation intensities are binned in such a way that each bin contains a similar number of events, with the exception of the most intense events, which are rare. The actual contribution (in mm) of each bin to the total mean precipitation rate is obtained by multiplying the frequency of events by the mean precipitation rate. The sum of the actual contributions from all bins gives the total mean precipitation rate. The fractional contribution (in %) of each bin is further obtained by dividing the actual contributions by the mean precipitation rate. In this case, the sum of all fractional contributions is equal to one, thus the information provided by fractional contributions is predominantly about the shape of the distribution. Taking the absolute differences between two fractional distributions and sum over all bins gives a measure of the difference in the shapes of the precipitation distributions. This is here called the "Index of fractional contributions". Since E-OBS precipitation intensities, in contrast to model data, are not continuous, the resulting ASoP factors for E-OBS tend to be noisy, especially for lower intensities. In order to facilitate the interpretation of the results, the regionally averaged ASoP factors for E-OBS were smoothed to some extent by using a simple filter.

The ASoP method is here applied to grid points pooled over target regions (Fig. 1) separately and the result is a distribution for each model showing the probability of different precipitation intensities based on daily precipitation. Most results presented here concern the actual contributions, both to limit the number of figures and because these factors conveniently provide information on both shape of distributions as well as the mean values. The ASoP distributions of all analysed models are used to compare model behaviour and performance. In particular to see how changing the grid resolution affects different parts of the distribution, for example if contributions from low and high precipitation intensities are different.

In addition to ASoP, a number of indices based on daily precipitation (listed in Table 3) are calculated for the same regions. For each model, the indices are calculated separately for each grid point within a region (land points only), and the values are then pooled to calculate percentiles representing the region.

This also means that the calculated model spread reflects geographical and not temporal variability. The index percentiles are represented by box plots (Sect. 3).

3 Results

3.1 ASoP analysis

3.1.1 Annual precipitation

Since the ASoP results are very similar between CMIP5 and CMIP6 GCMs (not shown), the results presented here include only one of these ensembles, CMIP6. Figure 2 presents the actual contributions (normalized bin frequency × mean bin rate) for annual daily precipitation over four of the PRUDENCE regions: Scandinavia, mid-Europe, the Alps and the Mediterranean. In general, the model ensembles have higher amounts of precipitation compared to E-OBS, signified by larger contributions at low (< 2-3 mm day^{-1}) and moderate-to-high (> 5-10 mm day^{-1}) intensities. An exception is the CMIP6 ensemble that instead shows lower contributions for moderate-to-high precipitation intensities, i.e. above 10-20 mm day^{-1} (Scandinavia, mid-Europe and the Alps) or between 5-20 mm day^{-1} (Mediterranean). CMIP6 also tends to have the largest overestimates of contributions from the lower intensities (below 5 mm day^{-1}). Another consistent feature is that the probabilities for the higher intensities (above 15 mm day^{-1}) increase with increasing grid resolutions of respective model ensemble, and consequently the contributions become increasingly larger than E-OBS (Fig. 2). This is most evident for the Alps region where the CMIP6 models (100-300 km grid spacing) clearly give smaller contributions than E-OBS and the PRIMAVERA models (25-160 km), the latter having smaller contributions than the CORDEX LR models (50 km) and the CORDEX HR models (12.5 km). The higher resolution models peak at higher intensities and have wider distributions with larger contributions from high-intensity daily rates. The sensitivity of model grid resolution to precipitation amounts and variability in association with areas with complex and steep topography (e.g. Prein et al., 2015) is most likely the main reason for the large differences between model ensembles in the Alps region. For example, the upper end of the CMIP6 distributions is around 50 mm day^{-1} while corresponding part in CORDEX HR models is around 100 mm day^{-1} (bottom right panel in Fig. 2). To further verify the results, the same analysis was performed after all data had been interpolated (conservatively) to two common grids; one at $2^{\circ}\times2^{\circ}$ resolution and one at $0.5^{\circ}\times0.5^{\circ}$ degree resolution (Figs. S1 and S2 in Supplementary). The interpolation to either grid has an overall small impact on the results. With the coarser grid ($2^{\circ}\times2^{\circ}$) the ASoP actual contributions have relatively larger contributions from the bulk part and a smaller contribution from the highest intensities, as expected from the smoothing effect of interpolation. These results provide increased confidence in the conclusions drawn from analysis on native grids.

3.1.2 Seasonal precipitation

Further insight can be gained by investigating seasonal differences (Fig. 3). In winter (DJF) the model ensemble means generally overestimate total mean precipitation compared to E-OBS (i.e. total areas under the curves showing differences are positive). The bulk of the distributions are slightly shifted to higher precipitation rates and also to higher contributions (except for the Mediterranean region). The largest inter-ensemble differences are seen for the Mediterranean where CORDEX HR shows the largest shift from E-OBS towards contributions from higher precipitation rates, and PRIMAVERA is similar to CORDEX LR. In summer (JJA), the ensemble means show larger contributions from intensities above 10-15 mm/day than E-OBS, especially in CORDEX HR. However, as this is in many cases compensated by lower contributions from rates between 2-10, the total mean precipitation biases are smaller than in winter. While the CORDEX ensemble means indicate larger total mean precipitation in France and Mediterranean, CMIP6 produces in all regions higher contributions from low-to-moderate (< ~5 mm/day) compared to E-OBS and lower contributions from higher intensities. Furthermore, there is a tendency in all regions of a larger spread within each model ensemble in JJA than in DJF (see coloured shadings in Fig. 3). Even though it is a very crude estimate of the spreads (the 5-95 percentile range in respective model ensemble), it can be argued that the differences in part is related to the seasonally prevailing weather conditions. In winter the North Atlantic storm track is in its active phase with frequent passings of synoptic weather systems over Europe. These features are generally well represented in climate models – hence larger consistency with associated precipitation across models. In summer, on the other hand, synoptic activity is reduced and convective processes (either as isolated or organized systems or embedded in larger scale features like fronts) become more prominent in precipitation events. Sensitivity to model grid resolution and physics parameterizations (e.g. convection parameterization) is larger during this season. The larger summertime spread in ensembles seen in Fig. 3 might then reflect larger uncertainties associated with model resolution and formulation. It is further noted that the ensemble spread is not increased as much (from winter to summer) over northern/north-western Europe which is relatively more affected by synoptic scale events during summer compared to southern parts of Europe (not shown).

Model ensemble differences for all regions and seasons are summarized in Figure 4, with E-OBS as reference. In spring (MAM) and winter (DJF) all ensembles have higher total mean precipitation in all regions. In summer (JJA) and autumn (SON) biases are also mostly on the positive side but smaller (primarily for GCM ensembles), and in some regions close to zero or slightly negative (e.g. the Alps, East Europe, Iberian Peninsula). Often there is an indication of a positive correlation between differences in mean (x-axis in Fig. 4) and differences in fractional contributions (y-axis, which indicates overall differences in the shape of the distributions), as seen for example in France or Mid-Europe regions. However, there are also cases with large differences in the shape but small total mean precipitation biases, for example the CMIP ensembles in JJA and SON over the Alps, suggesting compensating effects from different parts of the precipitation distribution. The overall spread is also highly variable between the regions; Scandinavia, Mid- and East-Europe and the British Isles are characterized by relatively smaller inter-ensemble differences, while in the Alps and Mediterranean the spread is large. The spread is in some regions dominated by inter-seasonal differences, e.g. in Mid-Europe and France, where typically the largest differences (in terms of both total means and distribution shapes) occur in DJF and MAM and smaller spreads in JJA and SON. In the Alps, Iberian Peninsula and the Mediterranean regions, however, the relatively larger inter-ensemble differences lead to an increased overall spread. Here, CORDEX HR further exhibits the largest differences to the GCM ensembles and also often larger deviations from E-OBS. These latter regions are either characterized by complex and steep topography (e.g. the Alps and the Pyrenees), large fraction of coastal areas and/or by relatively dry environments dominated by precipitation of convective nature (particularly for the warmer months). These factors most likely play important roles for the larger differences seen between the low resolution CMIP GCMs and the higher resolution PRIMAVERA GCMs and CORDEX RCMs, as well as contributing to larger uncertainties in, and lower quality and representativeness of, observational data. In contrast, in almost all seasons over the British Isles, the CORDEX HR biases in total precipitation compared to E-OBS are among the smallest with respect to the other ensembles (the difference in the shape is similar). Finally, it is noted that for all regions PRIMAVERA HR and CORDEX LR give comparable distributions as they are of similar resolution.

To summarize, we can conclude that, in comparison to E-OBS, most model ensembles exhibit larger contributions for most precipitation intensities, but most consistent for low (< ca 3 mm day^{-1}) and moderate-to-high (> ca 10 mm day^{-1}). The larger contributions occur predominantly in DJF while in summer there are often lower contributions than in E-OBS for moderate intensities (leading to smaller biases in total means). In general, the CORDEX ensembles, and most often PRIMAVERA, show a shift towards larger contributions from higher intensities compared to CMIP ensembles, especially in areas with complex orography as in the Alps. The higher model grid resolution does not always lead to improvements, i.e. closer agreements to E-OBS. However, it is worth re-emphasizing that the quality of E-OBS observations can be significantly lower in certain regions (e.g. mountainous areas or areas with low density of gauges) and seasons (especially in wintertime when the fraction of snowfall is largest which is more sensitive to wind induced undercatch) (Prein and Gobiet, 2017; Herrera et al., 2019), thus complicating the assessment of model behaviour in comparison to observations. To further highlight this issue, we have included an ASoP analysis for the Scandinavia region (Fig. S3) including a regional high-quality high-resolution gridded observational data set; NGCD (Lussana et al., 2018). In both DJF and JJA, the model ensembles still overestimate contributions from the bulk of the intensity distribution, however, NGCD has higher contributions from low intensities compared to E-OBS, reducing the model ensemble bias. More interestingly, NGCD shifts towards larger contributions for high intensities, > 10 mm day^{-1}, in effect lending more credibility to the CORDEX HR ensemble and less to the others.

3.1.3 Effect of grid resolutions – a one-to-one comparison

For multi-model ensembles, the sensitivity to model grid resolutions can generally only be assessed qualitatively since other aspects, such as differences in model formulation, also contribute to differences in model performance. In other words, it cannot be definitely stated to what extent differences in performance comes from higher resolution or from other differences in the model code. For the PRIMAVERA models, however, it is possible to directly compare low- and high-resolution model versions. In CORDEX ensembles this is also possible to some extent for a few models where low- and high-resolution versions of RCMs have been forced by the same parent GCMs. This is the case for nine RCM-GCM combinations (6 different RCMs driven by 4 different GCMs). Note that, in contrast to PRIMAVERA, CORDEX LR-HR "pairs" may not use the same version of the common model, which could also influence the results in addition to change in grid resolution. Further, the magnitude of the grid resolution change (the *delta* value) is the same for CORDEX models (*delta*=4), while for PRIMAVERA models it varies between approximately 2 and 5. Figure 5 shows the one-to.one comparison for DJF and JJA for selected regions. For CORDEX models the high-resolution model versions generally generate, in both seasons, larger contributions from precipitation intensities above ca 10 mm day^{-1}. This is sometimes accompanied by lower contributions from lower rates as seen in for example in Scandinavia and in the Alps in DJF. Similar results are seen for PRIMAVERA although not as consistently; e.g. over the British Isles and the Alps in JJA about half the models show increased contributions in the HR models over the bulk part, the other half showing instead lower contributions (although for higher rates most HR models show larger contributions). In fact, for many regions there is a larger spread in JJA within each model ensemble and also between the individual LR vs HR responses compared to DJF. It could be argued that this effect is related to precipitation events being of more convective nature in summer and thus larger sensitivity to model grid resolution as well as model physics. In winter, CORDEX RCMs are to a larger extent being influenced by the forcing GCMs and therefore, as there is only four different GCMs used in the nine RCM-GCM combinations shown here, tends to exhibit more similar responses in this season.

3.2 Selected precipitation-based indices

3.2.1 Model ensemble comparison

Figure 6 shows the number of precipitation days (RR1, Table 3) as simulated by all models for each PRUDENCE region. The number of precipitation days does not differ much between the model ensembles. There are clear differences between individual models, but it is difficult to establish any significant differences between the model ensembles. This is the case both for regions with a higher occurrence of precipitation days (e.g. SC) and regions with fewer precipitation days (e.g. IP). All models show about the same number of precipitation events over the whole year, which may suggest that the large-scale weather patterns are not influenced that much by higher resolution; also, when looking at individual seasons the differences between ensembles are small (Fig. S4). Note, however, that the large-scale circulation in the RCMs to a large extent is governed by the driving GCM which have typical resolutions of around 200 km. Interpolating the data to a common grid prior to analysis does not have a large impact on RR1 (Fig. S5). Most models overestimate the number of precipitation days compared to observations. It is a well-known feature of climate models, particularly those that use parameterized convection, that they tend to have too many wet days (e.g. Dai, 2006; Stephens et al., 2010).

The number of days with large precipitation amounts, above 10 mm day^{-1} and 20 mm day^{-1}, become more frequent with higher model resolution. For example, the number of days with precipitation over 20 mm (R20mm, Table 3) increases from just a few in CMIP5 to 5-10, or even more, in CORDEX HR (Fig. 7). The 10th to 90th inter-percentile range increases, due to a larger increase in the 90th percentile. Generally, the spread is larger for models with high resolution. This could partly be explained by higher number of data points in the high-resolution models (i.e. larger number of grid points); a high-resolution model is more likely to better represent the spatial variations of precipitation within a region while in coarser scale models precipitation fields are smoother due to fewer grid points. The differences between resolutions remain, however, also when all data are interpolated to two common grids of 0.5°×0.5° and 2°×2° resolutions. The median and spread is similar in all ensembles also when interpolated to another grid. In small regions such as AL the coarsest grid gives to few points, which means that it's difficult to calculate the 10th and 90th percentiles. The spread in CORDEX HR increases when interpolated to 2°×2° because the points with high values are not balanced by as many points close to the median (a 0.5°×0.5° grid contains 16 times more points than a 2°×2° grid). Compared to E-OBS the average number of days with more than 20 mm day^{-1} is more accurately simulated in the high-resolution ensembles, but the spread is highly exaggerated. The PRIMAVERA models have an average similar to E-OBS and also a more similar spread. The signal is the same for the individual seasons, but less pronounced since the potential number of days is smaller when divided over four seasons instead of counted over the whole year (Fig S6). The effect of resolution is therefore clearest in the season where most days occur, which means winter in western Europe and summer in central Europe.

The fact that the number of wet days is similar between LR and HR models (Fig. 6) but with increased frequency of (heavy) precipitation in HR models (Fig. 7) suggests that, for the latter, the precipitation intensity on the wet days is higher. This is shown in the simple precipitation intensity index (SDII, Table 3, Fig. 8). SDII is indeed affected by resolution, at least between CMIP5/6 and CORDEX; the wet day average precipitation is larger in the HR simulations compared to LR models, and also the intra-model spread (spread between models within the ensemble) is larger. For all regions, SDII is higher in the HR models. Perhaps, the relative increase in SDII is higher in regions with large spatial variations (for example because of complex orography or coastlines) such as IP and AL. The median SDII values in high-resolution models are in all regions closer to E-OBS than the low-resolution models, even though the model spread is generally larger in the climate models than in E-OBS. The differences between ensembles remain both for the median and the spread when the data are regridded to common grids. Also, for individual seasons it is clear that SDII increases with higher resolution, but the SDII values do not vary much with season (Fig. S7).

The higher intensities for extreme precipitation in high-resolution models compared to low-resolution models are also seen in the maximum one-day (Rx1day, Table 3, Fig. 9) and maximum five-day precipitation (not shown). There is a clear increase in both intensities and intra model spread in the high-resolution models. It can be discussed if this increase is an improvement since the CORDEX HR models give a maximum one-day precipitation that is significantly larger than E-OBS. On the other hand, it can be discussed if E-OBS is able to reliably represent these extremes (Hofstra et al., 2009; Prein and Gobiet, 2017). The medians and the spreads remain more or less the same also when regridded to common grids. In small regions such as AL the spread is reduced because the number of data points is small when regridded to a coarse grid. In regions with large spatial variations (e.g. between coast and mountain) such as IP the spread increases because high values are not balanced by as many points with values close to the median. In winter the effect of higher resolution is mainly seen in regions with complex topography, while in summer there is a clear signal in all regions (Fig 10). This reflects that higher resolution makes the largest difference in complex topography and for convective precipitation events.

3.2.2 One-to-one comparison

We let the mid-Europe region (ME) represent the whole domain, as the same conclusions can be made for all regions, only with small differences in the number of models that give significant differences. A one-to-one comparison is made of the selected indices for the models where there is both a low and a high grid resolution version (Fig. 11). The LR and HR versions are compared with a Welsh's t-test (Welsh, 1947) at the 0.05 significance level to see if the simulated indices are significantly different. This corroborates the analysis above, and adds some further detail by quantifying the differences.

Although the difference in the number of precipitation days (RR1, Fig. 11, top row) is significant for most models it is not clear how it is affected by resolution. The differences are small, mainly within ±10 days year^{-1}, and the difference between LR and HR is in some cases negative and in some positive. The differences between different models are larger than the differences between resolutions. It is clear, however, that all models overestimate the number of precipitation days compared to E-OBS. This is true also when the data is regridded to common grids, but three models and E-OBS get insignificant differences when regridded to 2°×2° instead of only one model at the native grids.

The number of days with precipitation more than 20 mm (R20mm, Fig. 11, second row) is significantly different between HR and LR for all models and E-OBS. For the CORDEX models R20mm is higher in most HR versions, while the difference is less clear in the PRIMAVERA models. All simulations with the RCA4 RCM, regardless of the driving GCM, clearly show higher R20mm in the HR version compared to the LR versions, which indicates that the difference in the index mainly is a result of the changed grid resolution in the RCM. The differences between LR and HR remain also when regridded to common grids which means that this is an effect of differences in model physics. CORDEX LR is close to E-OBS, while CORDEX HR generally overestimates R20mm.

The simple precipitation intensity index (SDII, Fig. 11, third row) is significantly different in one out of four PRIMAVERA models and four out of nine CORDEX models. Differences are small, tenths of mm day^{-1}, for most models. Most significant differences disappear when regridded to 0.5°×0.5° and all disappear when regridded to 2°×2° suggesting that the resolution does not affect SDII much in these model pairs. We still see a difference between CMIP GCMs and CORDEX RCMs (cf. Fig 8).

The maximum one-day precipitation (Rx1day, Fig. 11, bottom row) is significantly different in the HR version in all but one model (a PRIMAVERA model). The HR versions have higher precipitation values and larger spread in all but two PRIMAVERA models and one CORDEX model. Especially the CORDEX HR models have a higher maximum one-day precipitation. This seems to be driven by the RCM rather than the driving GCM. As an example, three RCMs are forced with the MPI-ESM-LR GCM. When forced by this GCM the Rx1day in the CCLM4-8-17 RCM is lower in the HR version, while in REMO2009 and RCA4 HR RCMs Rx1day is higher. In RCA4 the difference is particularly large, regardless of the driving GCM. That the differences are results of differences in model physics is supported by the fact that the difference remain also when the data is regridded to common grids.

The one-to-one comparison of selected indices shows that there are significant differences between the LR and HR models and that these are results of differences in model performance and not only difference in the number of data points. It also shows that for some indices the largest difference occurs between CMIP5/6 and PRIMAVERA HR, rather than between PRIMAVERA and CORDEX. This means that some of the differences seen in Figures 6-10 are not as clear in figure 11. The comparison also shows that even though there are significant differences between LR and HR it is for some cases difficult to establish significant differences between two ensembles since the difference between two different models are often larger than the differences between the LR and HR version of the same model.

It should be noted that the CORDEX RCMs are not always run with the same model version in the LR and HR simulations. Model differences could thus explain some of the differences between LR and HR. Since we don't have LR and HR simulations with all model versions we can't quantify this effect, only acknowledge it. It should also be noted that the difference in horizontal grid spacing varies between models. For CORDEX RCMs the resolution *delta* (LR/HR) is always 4 (50 km/12.5 km), but for PRIMAVERA it varies between 2 and 5. The *delta* value is larger in CORDEX than in most PRIMAVERA models, which could potentially mean that the effect of resolution is overestimated for the CORDEX RCMs. Figure 12 shows how the absolute differences in RR1, R20mm, SDII and Rx1day between the LR and HR version of the PRIMAVERA and CORDEX models described above correlates to the *delta* value in the ME region. There is no clear relation between the *delta* value and the size of the difference. CORDEX models that all have the same *delta* value span from small to large differences. The spread between PRIMAVERA models is also quite large. This again suggests that the response of a model to increased resolution depends on the model itself and not only on the magnitude of the resolution change.

4 Discussion and conclusions

This study investigates the importance of model resolution on the simulated precipitation in Europe. The aim is to investigate the differences between models and model ensembles, but also to evaluate their performance compared to gridded observations. In a similar study Demory et al. (2020) compare PRIMAVERA models with CORDEX LR and CORDEX HR. They come to the conclusion that

CORDEX indisputably improves the data from the driving CMIP5 models, but that the differences between CORDEX LR and PRIMAVERA are generally small. Both ensembles perform well, but tend to overestimate precipitation in winter and spring. The largest differences between the ensembles are for high precipitation intensities, in especially summer, where PRIMAVERA gives less heavy precipitation which makes it agree more with observations than CORDEX. Iles et al. (2020) compare the effect of resolution on extreme precipitation in Europe in CMIP5 GCMs and CORDEX RCMs. They conclude that high resolution models systematically produce higher frequencies of high-intensity precipitation events. Our interpretation of this, given the results in our study, is that in some cases also the overestimation of precipitation compared to E-OBS increases with higher resolution. The findings in this study support the conclusions from the above-mentioned studies, and add details based on a wider range of model ensembles and precipitation metrics. The fact that we come to the same conclusions as Iles et al. (2019) and Demory et al (2020) with slightly different methods give strength to these conclusions.

The ASoP analysis in this study shows that all model ensembles have larger contributions from heavy precipitation in winter compared to E-OBS, and that the higher values become most prominent for the ensemble with the highest grid resolution, CORDEX HR. The biases compared to E-OBS are generally smaller in summer. The PRIMAVERA ensemble is in good agreement with observations and has smaller bias than CORDEX for many regions. CMIP5 and CMIP6 mostly underestimate contributions from moderate-to-high precipitation intensities in summer while overestimating low-intensity events. Overall, in the summer season, the spread is large between ensembles and between models within the ensembles. This is indicative of large uncertainties which are most likely related to uncertainties in how models are able to treat smaller scale precipitation events involving convection. With respect to E-OBS, the ASoP results partly show that high resolution does not necessarily mean better. However, in coastal regions and regions with steep or complex topography there are uncertainties in both models and observations. Particularly in winter observations suffer from undercatch when precipitation falls as snow during windy conditions and in summer, smaller scale convective precipitation may be smoothed considerably or missed completely by ground rain gauges (which E-OBS is based on). E-OBS is not based on the full network of rain gauges in all countries, which could also lead to undercatch.

Therefore, it is not always obvious which model or ensemble of models is closest to reality. When compared to NGDC, a regional data set of high-quality, the difference between CORDEX HR and observations is reduced, which gives more confidence to the high-resolution model results.

It is clear that the horizontal resolution of a model has a large effect on precipitation, mostly on the heavier precipitation and in areas with complex and steep orography. The number of precipitation days does not depend much on resolution as this is mostly depending on large scale weather patterns and not so much on local topography and convection. For heavy precipitation events, which often are more local and short-lived in character, model resolution is more important. The high-resolution models better resolve such events and distinguish better between different parts of a region. Thus, extreme precipitation is more intense and more frequent in the HR models compared to the LR models in this study. With the same amount of wet days this means that precipitation intensifies so that the wet days get wetter. The largest impact of increased model scale resolution on precipitation is most evident for the coarser scale models; increasing the resolution from CMIP5/6 to PRIMAVERA HR has a greater effect than increasing from CORDEX LR/PRIMAVERA HR to CORDEX HR. This does not, however, mean that increased resolution gets less and less worthwhile; further refining the grid until convection-permitting resolutions are reached (less than ~5 km grid spacing), in which case convection parameterizations may be turned off, has a large positive effect (e.g. Prein et al. 2015). This is not shown here as the smallest grid spacing in models in this study is 12.5 km. The effect of higher resolution is seen in regions with small amounts of precipitation as well as regions with high amounts of precipitation, and in regions with small and large geographical differences. The higher percentiles change more than the low percentiles for all studied indices. Increasing resolution has about the same effect on both GCMs and RCMs, furthermore GCMs and RCMs of comparable resolution simulate comparable precipitation climates, even though PRIMAVERA is often drier than CORDEX.

It is worth to note that the differences between different RCM simulations, and how they respond to differences in resolution, may very well be explained by the driving GCM and the state of the atmospheric general circulation in them (Kjellström et al., 2018, Sørland et al., 2018). Higher resolution is expected to give a better described and more detailed climate, with for example deeper cyclones and more intense local showers; in a sense with more pronounced weather events. If two models are in different states, for example when it comes to where storm tracks cross Europe, and if these states are pronounced, that may lead to even larger model differences. Instead of a weak storm track in the south and a weak storm track in the north in the low-resolution model, we may now instead have strong storm tracks, which mean that the difference between the models increases. Still, the largest differences are seen in the CORDEX ensemble where the LR and HR models are run with the same coarse resolution GCM. This suggests that (regional) model resolution and performance is what determines high precipitation rates, rather than the driving GCM. To fully answer that would require an analysis of the circulation patterns in the different models. This is not done here, but should be a topic for further studies.

The differences between LR and HR largely remain also when the results are regridded to common grids of 0.5°×0.5° and 2°×2° which means that the HR version performs differently than the LR version of the same model, mainly because of better representations of topography and convection. The largest seasonal differences are seen for the heavy precipitation (R20mm, Rx1day). Heavy precipitation events usually occur locally in summer which makes it more sensitive to model resolution. Difference in resolution has a larger impact on heavy precipitation in summer than in winter.

Higher resolution does not necessarily mean better results. If a model is already too wet the increase in heavy precipitation that is induced by the higher resolution means that the HR version agrees less with observations that the LR version. For the individual model it is possible to quantify the difference and improvement between LR and HR. On the ensemble level this is more difficult. The difference between different models is often larger than between LR and HR versions of the same model. In this sense the quality of an ensemble is depending more on the models it consists of rather than the average resolution of the ensemble. Furthermore, when downscaling with an RCM, the simulated extreme precipitation, and the differences between GCM and RCM, depends more on the used RCM and less on the down-scaling itself, especially for heavy precipitation and particularly in summer.

Acknowledgements

The authors would like to thank Ségolène Berthou and two anonymous reviewers for giving valuable comments on the manuscript. This work has been funded by the PRIMAVERA project, which is funded by the European Union's Horizon 2020 programme, Grant Agreement no. 641727PRIMAVERA. This work used JASMIN, the UK collaborative data analysis facility. Some analyses were performed on the Swedish climate computing resource Bi provided by the Swedish National Infrastructure for Computing (SNIC) at the Swedish National Supercomputing Centre (NSC) at Linköping University. We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (http://www.uerra.eu) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (https://www.ecad.eu). We thank the modelling groups that run models and provide data within CMIP5, CMIP6, PRIMAVERA and CORDEX.

Data: The data are stored on the Jasmin infrastructure, http://www.ceda.ac.uk/projects/jasmin/. The simulations are part of the High Resolution Model Intercomparison project (HiResMIP) and will be uploaded to the ESGF: https://esgf-node.llnl.gov. Scripts for analysing the data will be available from the corresponding authors upon reasonable request.

References

Allen, M., and Ingram, W.: Constraints on future changes in climate and the hydrologic cycle. Nature 419, 228–232 (2002). https://doi.org/10.1038/nature01092, 2002.

Baker, A. J., Schiemann, R., Hodges, K. I., Demory, M.-E., Mizielinski, M. S., Roberts, M. J., Schaffrey, L. C., Strachan, J. and Vidale P. L.: Enhanced Climate Change Response of Wintertime North Atlantic Circulation, Cyclonic Activity, and Precipitation in a 25-km-Resolution, Global Atmospheric Model. J. Climate, 32, 7763–7781, https://doi.org/10.1175/JCLI-D-19-0054.1, 2019.

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

Ban N., Schmidli, J., and Schär, C.: Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster?, Geophys. Res. Lett., 42, 1165–1172, https://doi.org/10.1002/2014GL062588, 2015.

Belušić, D., de Vries, H., Dobler, A., Landgren, O., Lind, P., Lindstedt, D., Pedersen, R. A., Sánchez-Perrino, J. C., Toivonen, E., van Ulft, B., Wang, F., Andrae, U., Batrak, Y., Kjellström, E., Lenderink, G., Nikulin, G., Pietikäinen, J.-P., Rodríguez-Camino, E., Samuelsson, P., van Meijgaard, E. and Wu., M.: HCLIM38: a flexible regional climate model applicable for different climate zones from coarse to convection-permitting scales, Geosci. Model Dev., 13, 1311–1333, doi: 10.5194/gmd-13-1311-2020, 2020.

Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C. and Fosser, G.: Pan-European climate at convection-permitting scale: a model intercomparison study, Clim. Dyn., doi:10.1007/s00382-018-4114-6, 2018.

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

Brockhaus. P, Lüthi, D. and Schär, C.: Aspects of the diurnal cycle in a regional climate model, Meteor. Z., 17: 433-443, doi:10.1127/0941-2948/2008/0316, 2008.

Boé, J., Somot, S., Corre, L., and Nabat, P.: Large differences in Summer climate change over Europe as projected by global and regional climate models: causes and consequences, Clim. Dynam., 54, 2981–3002, https://doi.org/10.1007/s00382-020-05153-1, 2020.

Champion, A. J., Hodges, K. I., Bengtsson, L. O., Keenlyside, N. S., and Esch, M.: Impact of increasing resolution and a warmer climate on extreme weather from Northern Hemisphere extratropical cyclones, Tellus A: Dynamic Meteorology and Oceanography, 63, 5, 893-906, DOI: 10.1111/j.1600-0870.2011.00538.x, 2011.

Christensen, J. H. and Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Climatic Change 81, 7–30, https://doi.org/10.1007/s10584-006-9210-7, 2007.

Coppola, E., Sobolowski, S., Pichelli, E., Raffaele, F., Ahrens, B., Anders, I., Ban, N., Bastin, S., Belda, M., Belusic, D., Caldas-Alvarez, A., Cardoso, R. M., Davolio, S., Dobler, A., Fernadez, J., Fita, L., Fumiere, Q., Giorgi, F., Görgen, K., Güttler, I., Halenka, T., Heinzeller, D., Hodnebrog, Ø., Jacob, D., Kartsios, S., Katragkou, E., Kendon, E., Khodayar, S., Kunstmann, H., Knist, S., Lavín-Gullón, A., Lind, P., Lorenz, T., Maraun, D., Marelle, L., van Meijgaard, E., Milovac, J., Myhre, G., Panitz, H.-J., Piazza, M., Raffa, M., Raub, T., Rockel, B., Schär, C., Sieck, K., Soares, M. M., Somot, S., Srnec, L., Stocchi, P., Tölle, M. H., Truhetz, H., Vautard, R., de Vries, H. and Warrch-Sagi, K.: A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean, Clim. Dyn. 55, 3–34, https://doi.org/10.1007/s00382-018-4521-8, 2018.

Cornes, R., van der Schrier, G., van den Besselaar, E. J. M., Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Datasets, J. Geophys. Res. Atmos., 123. doi:10.1029/2017JD028200, 2018.

Dai, A.: Precipitation characteristics in eighteen coupled climate models, J. Climate, 19: 4605-4630, doi:10.1175/JCLI3884.1., 2006.

Dai, A., Trenberth, K. E.: The diurnal cycle and its depiction in the community climate system model, J. Climate, 17: 930-951, doi:10.1175/1520-0442, 2004.

Dai, A., Giorgi, F., and Trenberth, K. E.: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States, J. Geophys. Res., 104(D6), 6377– 6402, doi:10.1029/98JD02720, 1999.

Delworth, T. L, Rosati, A,, Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., Dixon, K., Griffies, S.M., Lee, H. C., Pacanowski, R. C., Vecchi, G. A., Wittenberg, A. T., Zeng, F., and Zhang, R.: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupledclimate model, J. Clim. 25, 2755–2781, doi:10.1175/JCLI-D-11-00316.1, 2015.

Demory, M.-E., Berthou, S., Sørland, S. L., Roberts, M. J., Beyerle, U., Seddon, J., Haarsma, R., Schär, C., Christensen, O. B., Fealy, R., Fernandez, J., Nikulin, G., Peano, D., Putrasahan, D., Roberts, C. D., Steger, C., Teichmann, C., and Vautard, R.: Can high-resolution GCMs reach the level of information provided by 12–50 km CORDEX RCMs in terms of daily precipitation distribution?, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-370, in review, 2020.

Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M. and van den Hurk, B.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, Climatic Change 81, 53–70, https://doi.org/10.1007/s10584-006-9228-x, 2007.

Dirmeyer, P. A., Cash, B. A., Kinter, J. L., Jung, T., Marx, L., Satoh, M., Stan, C., Tomita, H., Towers, P., Wedi, N. and Achuthavarier, D.: Simulating the diurnal cycle of rainfall in global climate models: Resolution versus parameterization. Clim. Dyn. 39(1–2):399–418, 2012.

Di Luca, A., de Elía, R. and Laprise, R.: Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations, Clim. Dyn. 38, 1229–1247, https://doi.org/10.1007/s00382-011-1068-3, 2011.

Donat, M., Lowry, A., Alexander, L., O'Gorman, P. A. and Maher, N.: More extreme precipitation in the world's dry and wet regions, Nature Clim. Change 6, 508–513, https://doi.org/10.1038/nclimate2941, 2016.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016.

Fosser, G., Khodayar, S., Berg, P.: Benefit of convection permitting climate model simulations in the representation of convective precipitation, Clim. Dyn., 44: 45-60, doi:10.1007/s00382-014-2242-1, 2015.

Gao, X., Xu, Y., Zhao, Z., Pal, J. S. and Giorgi, F.: On the role of resolution and topography in the simulation of East Asia precipitation. Theor. Appl. Climatol. 86, 173–185: https://doi.org/10.1007/s00704-005-0214-4, 2006.

Gao, Y., Leung, L. R., Zhao, C., and Hagos, S.: Sensitivity of U.S. summer precipitation to model resolution and convective parameterizations across gray zone resolutions, J. Geophys. Res. Atmos., 122: 2714-2733, doi:10.1002/2016JD025896, 2017.

Giorgi, F., and Marinucci, M. R.: A Investigation of the Sensitivity of Simulated Precipitation to Model Resolution and Its Implications for Climate Studies, Mon. Wea. Rev., 124, 148–166, https://doi.org/10.1175/1520-0493(1996), 1996

Giorgi, F., Torma, C., Coppola, E., Ban, N., Schär, C. and Somot, S.: Enhanced summer convective rainfall at Alpine high elevations in response to climate warming, Nature Geosci. 9, 584–589, https://doi.org/10.1038/ngeo2761, 2016.

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., and Stössel, A.: Max Planck Institute Earth System Model(MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), Geosci. Model Dev., 12, 3241–3281, https://doi.org/10.5194/gmd-12-3241-2019, 2019.

Gutiérrez, C., Somot, S., Nabat, P., Mallet, M., Corre, L., van Meijgaard, E., Perpiñán,O., and Gaertner, M. A.: Future evolution of surface solar radiation and photovoltaic potential in Europe: investigating the role of aerosols, Environ. Res. Lett.,15, 034035, https://doi.org/10.1088/1748-9326/ab6666, 2020.

Gutowski Jr., W. J., Giorgi, F., Timbal, B., Frigon, A., Jacob, D., Kang, H.-S., Raghavan, K., Lee, B., Lennard, C., Nikulin, G., O'Rourke, E., Rixen, M., Solman, S., Stephenson, T., and Tangang, F.: WCRP COordinated Regional Downscaling EXperiment (CORDEX): a diagnostic MIP for CMIP6, Geosci. Model Dev., 9, 4087-4095, doi:10.5194/gmd-9-4087-2016, 2016.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Haarsma, R., Acosta, M., Bakhshi, R., Bretonnière, P.-A. B., Caron, L.-P., Castrillo, M., Corti, S., Davini, P., Exarchou, E., Fabiano, F., Fladrich, U., Fuentes Franco, R., García-Serrano, J., von Hardenberg, J., Koenigk, T., Levine, X., Meccia, V., van Noije, T., van den Oord, G., Palmeiro, F. M., Rodrigo, M., Ruprich-Robert, Y., Le Sager, P., Tourigny, É., Wang, S., van Weele, M., and Wyser, K.: HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR. Description, model performance, data handling and validation, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-350, in review, 2020.

Herrera, S, Kotlarski, S, Soares, PMM, et al. Uncertainty in gridded precipitation products: Influence of station density, interpolation method and grid resolution. Int J Climatol. 2019; 39: 3717–3729. https://doi.org/10.1002/joc.5878

Hofstra, N., Haylock, M., New, M. and Jones, P. D.: Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature, J. Geophys. Res.,114, D21101, doi:10.1029/2009JD011799, 2009.

Hughes, M., Lundquist, J.D. & Henn, B. Dynamical downscaling improves upon gridded precipitation products in the Sierra Nevada, California. Clim Dyn 55, 111–129 (2020). https://doi.org/10.1007/s00382-017-3631-z

Iles, C. E., Vautard, R., Strachan, J., Joussaume, S., Eggen, B. R., and Hewitt, C. D.: The benefits of increasing resolution in global and regional climate simulations for European climate extremes, Geoscientific Model Development Discussion, https://doi.org/10.5194/gmd-2019-253, 2019.

Iorio, J.P., Duffy, P.B., Govindasamy, B.,Khairoutdinov, M.,and Randall, D.: Thomson S. L., et al.: Effects of model resolution and subgrid-scale physics on the simulation of precipitation in the continental United States, Climate Dynamics 23, 243–258, https://doi.org/10.1007/s00382-004-0440-y, 2004.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., Senior, C. A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, Nat. Clim. Change 4: 570–576, doi:10.1038/nclimate2258, 2014.

Kharin, V. V., Zwiers, F. W., Zhang, X. and Wehner, M.: Changes in temperature and precipitation extremes in the CMIP5 ensemble, Climatic Change 119, 345–357 (2013). https://doi.org/10.1007/s10584-013-0705-8, 2013.

Kinter III, J. L., Cash, B., Achuthavarier, D., Adams, J., Altshuler, E., Dirmeyer, P., Doty, B., Huang, B., Jin, E. K., Marx, L., Manganello, J., Stan, C., Wakefield, T., Palmer, T., Hamrud, M., Jung, T., Miller, M., Towers, P., Wedi, N., Satoh, M., Tomita, H., Kodama, C., Nasuno, T., Oouchi, K., Yamada, Y., Taniguchi, H., Andrews, P., Baer, T., Ezel,l M., Halloy, C., John, D., Loftis, B., Mohr, R., and Wong, K.: Revolutionizing climate modeling with project Athena: a multi-institutional, international collaboration, Bull. Am. Meteorol. Soc., 94, 231–245, doi:10.1175/BAMS-D-11-00043.1, 2013.

Kjellström, E., Nikulin, G., Hansson, U., Strandberg, G. and Ullerstig, A.: 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations, Tellus A: Dynamic Meteorology and Oceanography, 63:1, 24-40, DOI: 10.1111/j.1600-0870.2010.00475.x, 2011.

Kjellström, E., Nikulin, G., Strandberg, G., Christensen, O. B., Jacob, D., Keuler, K., Lenderink, G., van Meijgaard, E., Schär, C., Somot, S., Sørland, S. L., Teichmann, C., and Vautard, R.: European climate change at global mean temperature increases of 1.5 and 2 °C above pre-industrial conditions as simulated by the EURO-CORDEX regional climate models, Earth Syst. Dynam., 9, 459–478, https://doi.org/10.5194/esd-9-459-2018, 2018.

Klingaman, N. P., Martin, G. M., and Moise, A.: ASoP (v1.0): a set of methods for analyzing scales of precipitation in general circulation models, Geoscientific Model Development, 10(1), 57–83. https://doi.org/10.5194/gmd-10-57-2017, 2017.

Kotlarski, S., Szabó, P., Herrera, S., et al. Observational uncertainty and regional climate model evaluation: A pan-European perspective. Int J Climatol. 2019; 39: 3730–3749. https://doi.org/10.1002/joc.5249

Leutwyler, D., Lüthi, D., Ban, N., Fuhrer, O. and Schär, C.: Evaluation of the convection-resolving climate modeling approach on continental scales, J. Geophys. Res. Atmos., 122, 5237– 5258, doi:10.1002/2016JD026013, 2017.

Liang, X.-Z., Li, L., Dai, A., and Kunkel, K. E.: Regional climate model simulation of summer precipitation diurnal cycle over the United States, Geophys. Res. Lett., 31, L24208, doi:10.1029/2004GL021054, 2004.

Lind, P., Belušić, D., Christensen, O. B., Dobler, A., Kjellström, E., Landgren, O., Lindstedt, D., Matte, D., Pedersen, R. A., Toivonen, E., and Wang, F.: Benefits and added value of convection-permitting climate modeling over Fenno-Scandinavia, Climate Dynamics, accepted, 2020.

Lundquist, J., M. Hughes, E. Gutmann, and S. Kapnick, 2019: Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. Bull. Amer. Meteor. Soc., 100, 2473–2490, https://doi.org/10.1175/BAMS-D-19-0001.1.

Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Tveito, O. E., and Andersen, J.: seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day, Earth Syst. Sci. Data, 10, 235–249, https://doi.org/10.5194/essd-10-235-2018, 2018

O'Gorman, P.: Sensitivity of tropical precipitation extremes to climate change, Nature Geosci. 5, 697–700, https://doi.org/10.1038/ngeo1568, 2012.

Pall, P., Allen, M. R. and Stone, D. A.: Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming. Clim. Dyn. 28, 351–363, https://doi.org/10.1007/s00382-006-0180-2, 2007.

Pfahl, S., O'Gorman, P. and Fischer, E.: Understanding the regional pattern of projected future changes in extreme precipitation. Nature Clim. Change 7, 423–427, https://doi.org/10.1038/nclimate3287, 2017.

Prein, A. F. and Gobiet, A.: Impacts of uncertainties in European gridded precipitation observations on regional climate analysis, Int. J. Climatol., 37, 305-327, doi:10.1002/joc.4706, 2017

Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K. and Georgievski, G.: Added value of convection permitting seasonal simulations. Clim. Dyn., 41 (9-10): 2655-2677. doi:10.1007/s00382-013-1744-6, 2013a.

Prein, A. F., Holland, G. J., Rasmussen, R. M., Done, J., Ikeda, K., Clark, M. P. and Liu, C. H.: Importance of Regional Climate Model Grid Spacing for the Simulation of Heavy Precipitation in the Colorado Headwaters. J. Climate, 26: 4848–4857, doi: 10.1175/JCLI-D-12-00727.1, 2013b.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schidli, J., van Lipzig, N. P. M. and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53: 323– 361. doi:10.1002/2014RG000475, 2015.

Prein, A.F., Gobiet, A., Truhetz, H. et al. Precipitation in the EURO-CORDEX 0.11∘0.11∘ and 0.44∘0.44∘ simulations: high resolution, high benefits?. Clim Dyn 46, 383–412 (2016). https://doi.org/10.1007/s00382-015-2589-y

Rasmussen, R., Baker, B., Kochendorfer, J., Myers, T., Landolt, S., Fischer, A., Black, J., Thériault, J., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Cristanelli, S. and Gutmann, A.: How well are we measuring snow: the NOAA/FAA/NCAR winter precipitation test bed. Bull. Am. Met. Soc., 93. doi:10.1175/BAMS-D-11-00052.1, 2012.

Rauscher, S.A., Coppola, E., Piani and Giorgi F.: Resolution effects on regional climate model simulations of seasonal precipitation over Europe. Clim. Dyn. 35, 685–711, https://doi.org/10.1007/s00382-009-0607-7, 2010.

Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., Chang,P., Christensen, H. M., Danilov, S., Demory, M.-E., Griffies, S. M., Haarsma, R., Jung,T., Martin, G., Minobe, S., Ringler, T., Satoh, M., Schiemann, R., Scoccimarro, E., Stephens, G., and Wehner, M. F.: The Benefits of Global High Resolution for ClimateSimulation: Process Understanding and the Enabling of Stakeholder Decisions at theRegional Scale, B. Am. Meteorol. Soc., 99, 2341–2359, https://doi.org/10.1175/BAMS-D-15-00320.1, 2018a.

Roberts,C.D., Senan R. , Molteni F. , Boussetta S., Mayer M. , and Keeley, S. P. E.: Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP, Geosci. Model Dev., 11, 3681–3712, https://doi.org/10.5194/gmd-11-3681-2018, 2018b.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., Jackson, L. C., Kuhlbrodt, T., Mathiot, P., Roberts, C. D., Schiemann, R., Seddon, J., Vannière, B., and Vidale, P. L.: Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments, Geosci. Model Dev., https://doi.org/10.5194/gmd-12-4999-2019, 2019.

Sørland, S. L., Schär, C., Lüthi, D. And Kjellström E.: Bias patterns and climate change signals in GCM-RCM model chains, Environ. Res. Lett., 13, 074017, https://doi.org/10.1088/1748-9326/aacc7, 2018.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P. and Haynes, J.: Dreary state of precipitation in global models. J Geophys Res, 115, D24211. doi:10.1029/2010JD014532, 2010.

Stratton, R.A. and Stirling, A.J.: Improving the diurnal cycle of convection in GCMs, Q.J.R. Meteorol. Soc., 138, 1121-1134, doi:10.1002/qj.991, 2012.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485-498, DOI:10.1175/BAMS-D-11-00094.1, 2012.

van Haren, R., R. J. Haarsma, G. J. Van Oldenborgh, and W. Hazeleger, 2015: Resolution Dependence of European Precipitation in a State-of-the-Art Atmospheric General Circulation Model. J. Climate, 28, 5134–5149, https://doi.org/10.1175/JCLI-D-14-00279.1.

Vergara-Temprado, J., Ban, N., Panosetti, D., Schlemmer, L., and Schär, C.: Climate models permit convection at much coarser resolutions than previously considered, J. Clim., JCLI-D-19-0286.1. doi:10.1175/JCLI-D-19- 0286.1, 2019.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, P., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J., Douville, H., Franchisteguy, L., Ethé, C., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y.,Msadek,R., Ribes, A., Sanchez-Gomez, E., and Terray,L.: Evaluation of CMIP6 DECK Experiments with CNRM-CM6-1, J. Adv. Model. Earth Syst., 11, 2177–2213, https://doi.org/10.1029/2019MS001683, 2019.

Welch, B. L.: The generalization of 'students' problem when several different population variances are involved, Biometrika, Volume 34, Issue 1-2, January 1947, Pages 28–35, https://doi.org/10.1093/biomet/34.1-2.28, 1947.

Zappa, G., Shaffrey, L. C., and Hodges, K. I.: The Ability of CMIP5 Models to Simulate North Atlantic Extratropical Cyclones, J. Climate, 26, 5379–5396, https://doi.org/10.1175/JCLI-D-12-00501.1, 2013.

Tables

Ensemble	**Model**	**Contact institute**	**Atmo-spheric grid spacing**
CMIP5	ACCESS1-0	Commonwealth Scientific and Industrial Research Organisation, Australia, and Bureau of Meteorology	N96
CMIP5	ACESS1-3	Commonwealth Scientific and Industrial Research Organisation, Australia, and Bureau of Meteorology	N96
CMIP5	CanESM2	Canadian Centre for Climate Modelling and Analysis	T63
CMIP5	CMCC-CESM	Centro Euro-Mediterraneo per i Cambiamenti Climatici	96x48
CMIP5	CMCC-CM	Centro Euro-Mediterraneo per i Cambiamenti Climatici	480x240
CMIP5	CMCC-CMS	Centro Euro-Mediterraneo per i Cambiamenti Climatici	192x96
CMIP5	CSIRO-Mk3-6-0	Australian Commonwealth Scientific and Industrial Research Organization (CSIRO) Marine and Atmospheric Research in collaboration with the Queensland Climate Change Centre of Excellence (QCCCE)	T63
CMIP5	FGOALS-g2	Institute of Atmospheric Physics, Chinese Academy of Sciences and Tsinghua University	128x60
CMIP5	GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory	144x90
CMIP5	GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	144x90
CMIP5	HadCM3	Met Office Hadley Centre	96x73
CMIP5	HadGEM2-CC	Met Office Hadley Centre	N96
CMIP5	HadGEM2-ES	Met Office Hadley Centre	N96
CMIP5	IPSL-CM5A-LR	Institut Pierre Simon Laplace	96x96
CMIP5	IPSL-CM5A-MR	Institut Pierre Simon Laplace	144x143

CMIP5	MPI-ESM-LR	Max Planck Institute for Meteorology	T63
CMIP5	MPI-ESM-MR	Max Planck Institute for Meteorology	T63
CMIP5	NorESM1-M	Norwegian Climate Centre	144x96
CMIP6	ACCESS-CM2	Commonwealth Scientific and Industrial Research Organisation, Australia, and Bureau of Meteorology	192x145
CMIP6	ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation, Australia, and Bureau of Meteorology	192x145
CMIP6	CESM2-FV2	The National Center for Atmospheric Research	144x96
CMIP6	CESM2	The National Center for Atmospheric Research	288x192
CMIP6	CESM2-WACCM-FV2	The National Center for Atmospheric Research	144x96
CMIP6	CESM2-WACCM	The National Center for Atmospheric Research	288x192
CMIP6	EC-Earth3	EC-Earth-Consortium	512x256
CMIP6	EC-Earth3-Veg	EC-Earth-Consortium	512x256
CMIP6	GFDL-CM4	NOAA Geophysical Fluid Dynamics Laboratory	360x180
CMIP6	INM-CM4-8	Institute for Numerical Mathematics, Russian Academy of Science	180x120
CMIP6	INM-CM5-0	Institute for Numerical Mathematics, Russian Academy of Science	180x120
CMIP6	MIROC6	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, The University of Tokyo, National Institute for Environmental Studies, RIKEN Center for Computational Science	T85
CMIP6	MPI-ESM-1-2-HAM	Max Planck Institute for Meteorology	192x96
CMIP6	MPI-ESM1-2-LR	Max Planck Institute for Meteorology	192x96
CMIP6	MRI-ESM2-0	Meteorological Research Institute, Tsukuba	320x160
CMIP6	NorCPM1	Norwegian Climate Centre	320x384
CMIP6	NorESM2-LM	Norwegian Climate Centre	144x96
CMIP6	NorESM2-MM	Norwegian Climate Centre	288x192
CMIP6	SAM0-UNICON	Seoul National University	288x192
PRIMAVERA	CNMR-CM6-1	CNRM-CERFACS	256x128
PRIMAVERA	CNRM-CM6-1-HR	CNRM-CERFACS	720x360
PRIMAVERA	EC-Earth3	EC-Earth-Consortium	512x256

PRIMAVERA	EC-Earth3-HR	EC-Earth-Consortium	1024x512
PRIMAVERA	IFS-HR	European Centre for Medium-Range Weather Forecasts	720x360
PRIMAVERA	IFS-LR	European Centre for Medium-Range Weather Forecasts	360x180
PRIMAVERA	HadGEM3-GC31-HM	Met Office Hadley Centre	1024x720
PRIMAVERA	HadGEM3-GC31-LM	Met Office Hadley Centre	192x144
PRIMAVERA	HadGEM3-GC31-MM	Met Office Hadley Centre	432x324
PRIMAVERA	MPIESM-1-2-HR	Max Planck Institute for Meteorology	384x192
PRIMAVERA	MPIESM-1-2-XR	Max Planck Institute for Meteorology	768x384

Table 1. The GCM ensembles used in this study and the GCMs they consist of. Grid spacing is given in the same format is in the meta data for each model.

Institue	**RCM**	**Driving GCM**									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CLMcom	CCLM4-8-17	x	x		x		x		x	xo	
CNRM	ALADIN53		x								
CNRM	ALADIN63		x								
DMI	HIRHAM5				xo		x				x
GERICS	REMO2015	x	x		x		x		x		x
IPSL	WRF331F							xo			
KNMI	RACMO22E				xo		o				x
MPI-CSC	REMO2009									xo	
SMHI	RCA4	o	o	o	xo	o	xo	xo	o	xo	o
UHOH	WRF361H						x			x	
HMS	ALADIN52		o								

Table 2. RCM GCM combinations used in this study. Euro-CORDEX simulations at 0.11° (~12.5 km) are marked with "x" and at 0.44° (~50 km) are marked with "o". The driving GCMs are: 1) CanESM2, 2) CNRM-CM5, 3) CSIRO-Mk3-6-0, 4) EC-Earth, 5) GFDL-ESM2M, 6) HadGEM2-ES, 7) IPSL-CM5A-MR, 8) MIROC5, 9) MPI-ESM-LR, 10) NorESM1-M

Short	**Long name**	**Definition**	**Unit**

name			
RR1	Wet days index	Number of days with precipitation sum equal to or more than 1 mm	Days year^{-1}
R20mm	Very heavy precipitation days index	Number of days with precipitation sum more than 20 mm	Days year^{-1}
SDII	Simple daily intensity index	Average precipitation sum on days with precipitation sum equal to or above 1 mm	mm day^{-1}
Rx1day	Highest one day precipitation amount	Precipitation amount on the day with highest amount	mm day^{-1}

Table 3. Definitions of indices

Figures

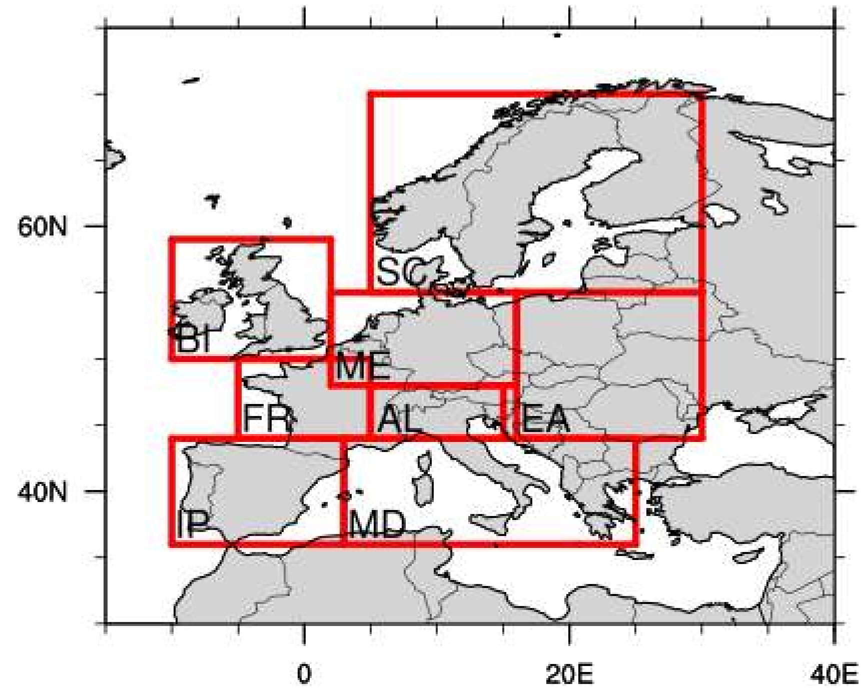


Figure 1: The regions for which precipitation data is analysed: Scandinavia (SC), British Isles (BI), Mid-Europe (ME), France (FR), The Alps (AL), Eastern Europe (EA), Iberian Peninsula (IP) and the Mediterranean (MD).

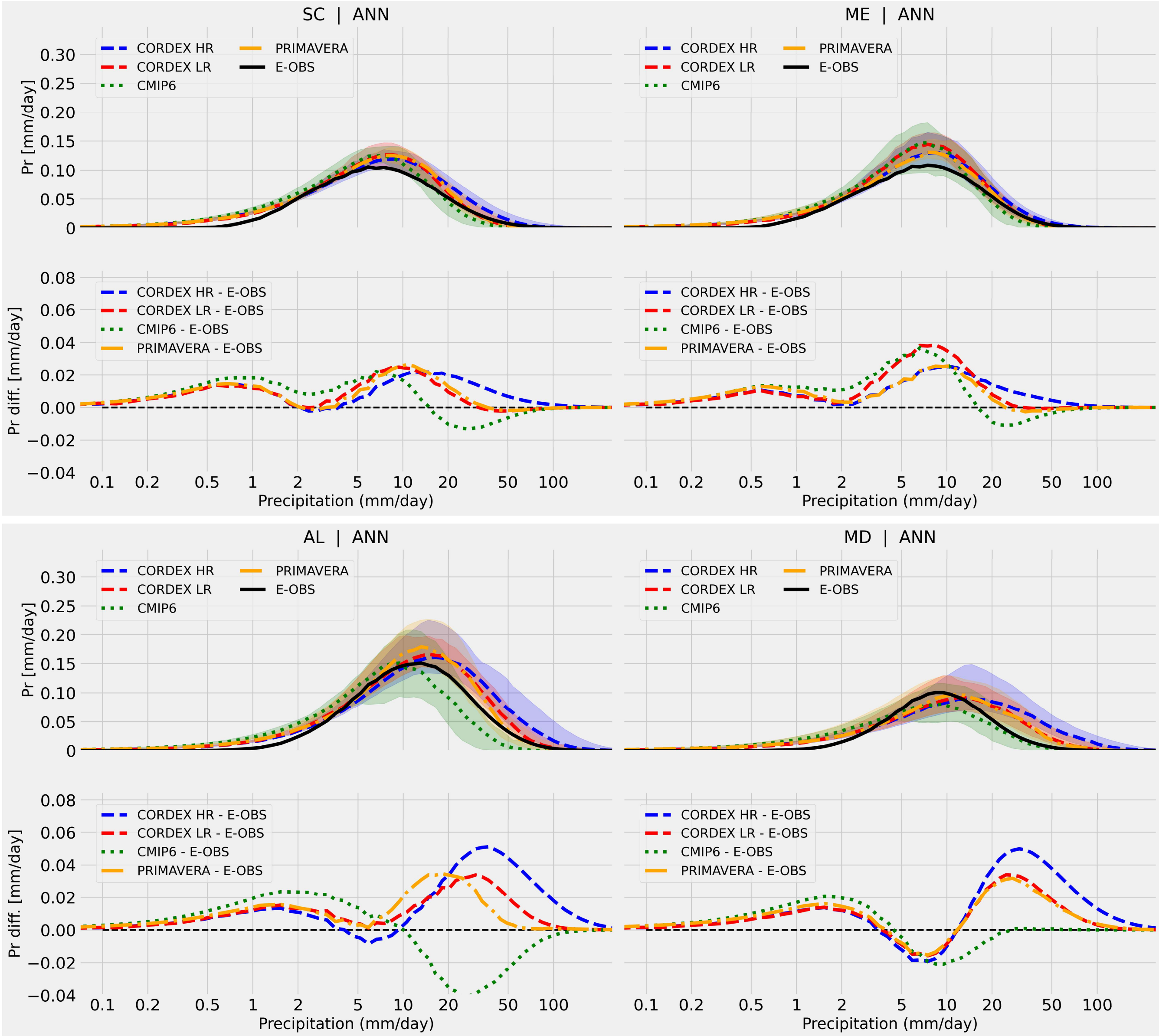


Figure 2: The panels show the actual contribution (to the total mean precipitation, y-axis) per precipitation intensity bin (x-axis), based on annual (ANN) daily precipitation values in the CMIP6 (green dotted lines and shading), PRIMAVERA (orange dashed-dotted lines and shading), CORDEX low resolution (red dashed lines and shading) and CORDEX high resolution (blue dashed lines and shading) ensembles. The displayed regions are Scandinavia (SC, top left), mid-Europe (ME, top right), the Alps (AL, bottom left) and the Mediterranean (MD, bottom right). Coloured shadings represent the 5-95 percentile range in respective ensemble. Black solid lines are E-OBS (0.1° resolution) observations.

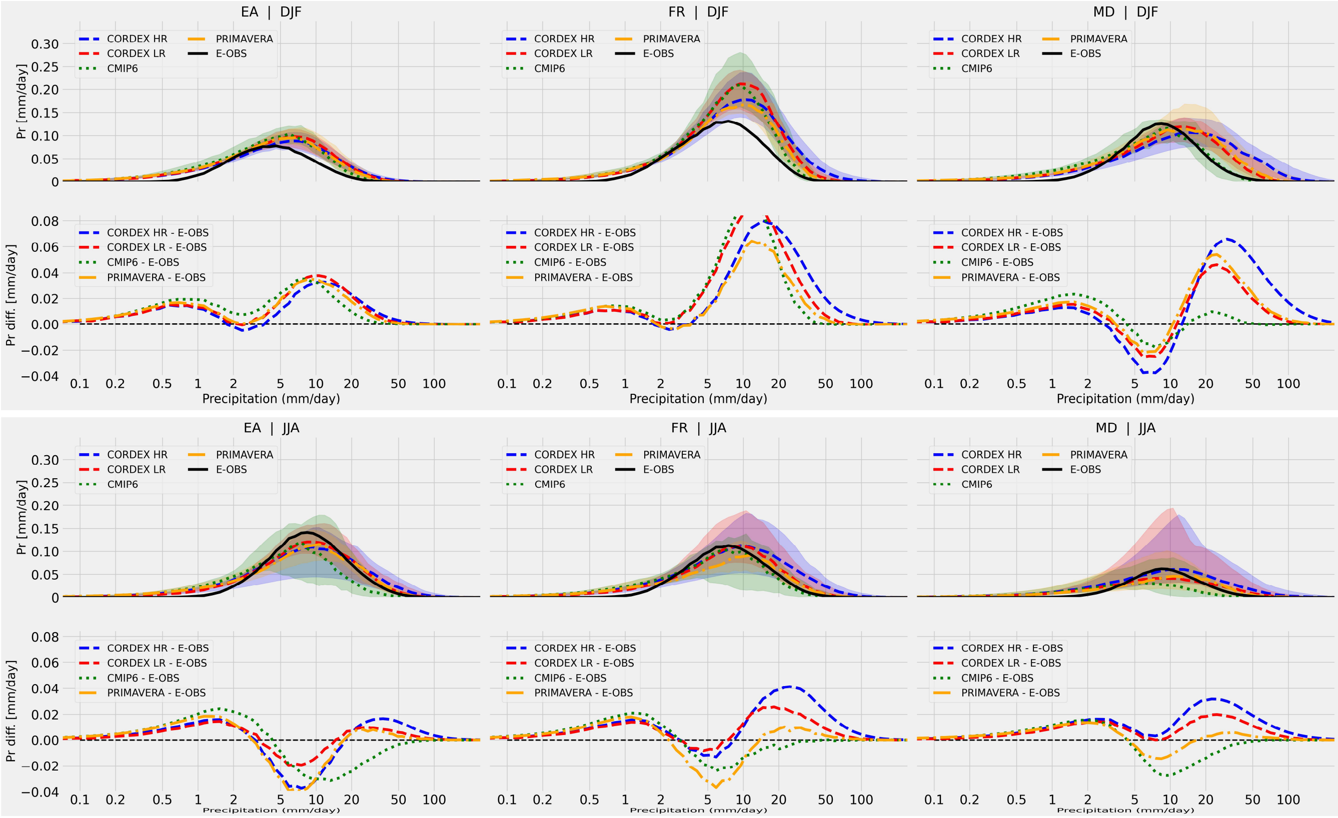


Figure 3: Same as in Fig. 2 but for DJF (top row) and JJA (bottom row) daily precipitation values and for the eastern Europe (EA, left), France (FR, middle) and the Mediterranean (MD, right) regions. Coloured shadings represent the 5-95 percentile range in respective ensemble. Black solid lines are E-OBS (0.1º resolution) observations.

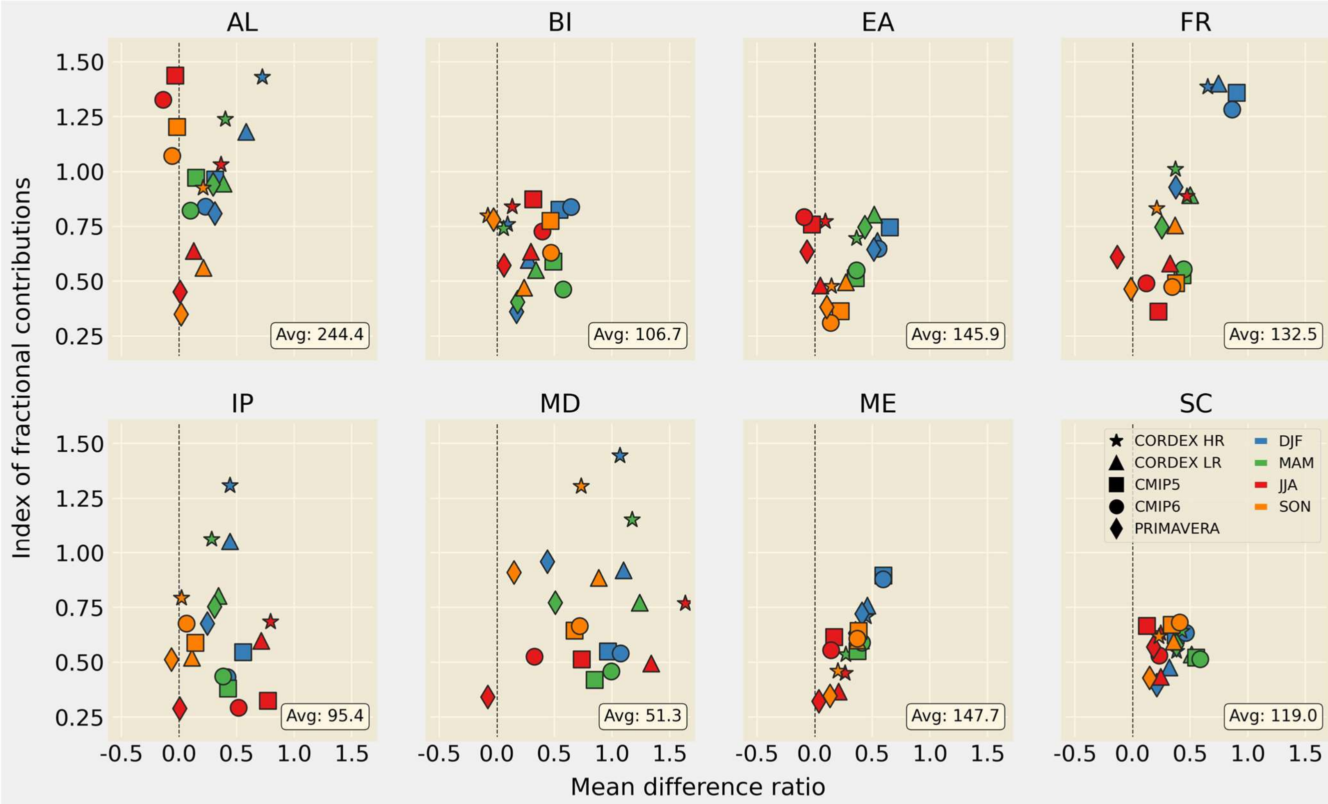


Figure 4: The index of fractional contributions (y-axis) plotted as a function of the fractional difference in seasonal total precipitation (x-axis). E-OBS (0.1º resolution) is the reference data set and E-OBS average annual total precipitation (in mm year^{-1}) is shown in lower right in each panel.

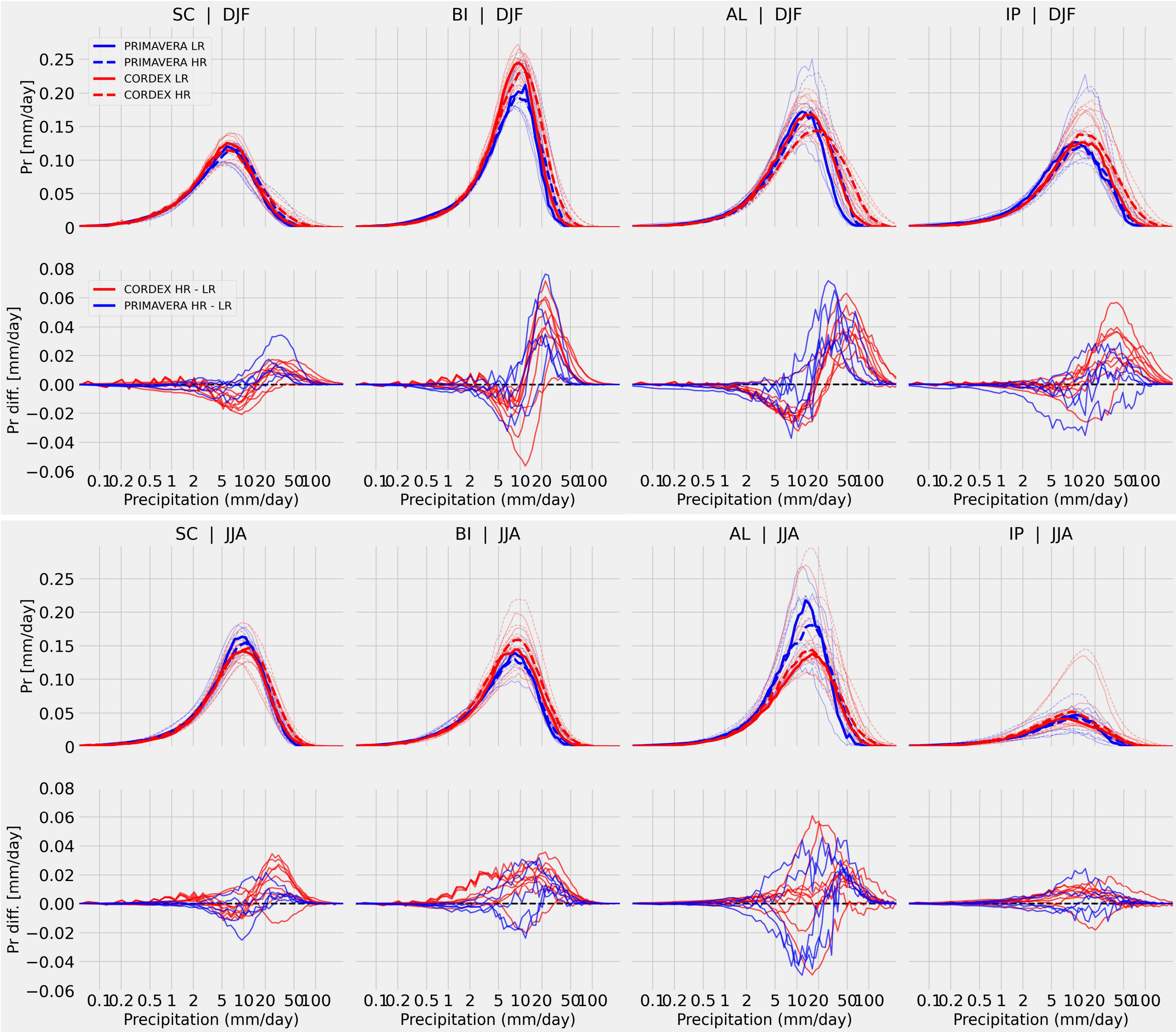


Figure 5: The panels show the actual contribution (to the total mean precipitation, y-axis) per precipitation intensity bin (x-axis), based on DJF (top row) and JJA (bottom row) daily mean precipitation values in CORDEX and PRIMAVERA models for the Scandinavia (SC), British Isles (BI), the Alps (AL) and Iberian Peninsula (IP) regions. Thin lines in upper part of each panel represent each individual model while the thick lines represent the ensemble means. In the lower part of each panel each line represents differences between respective high- and low-resolution model pair.

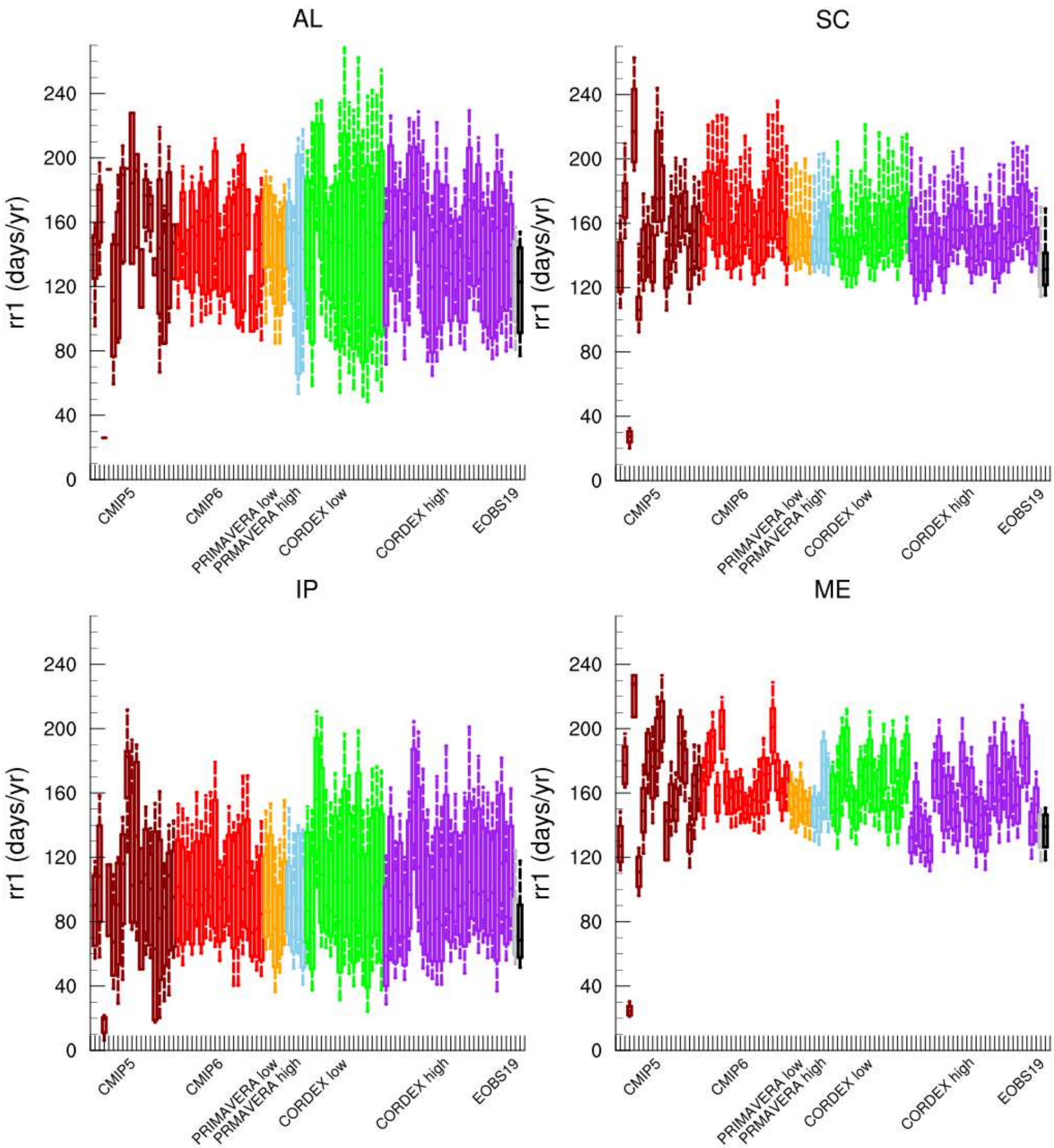
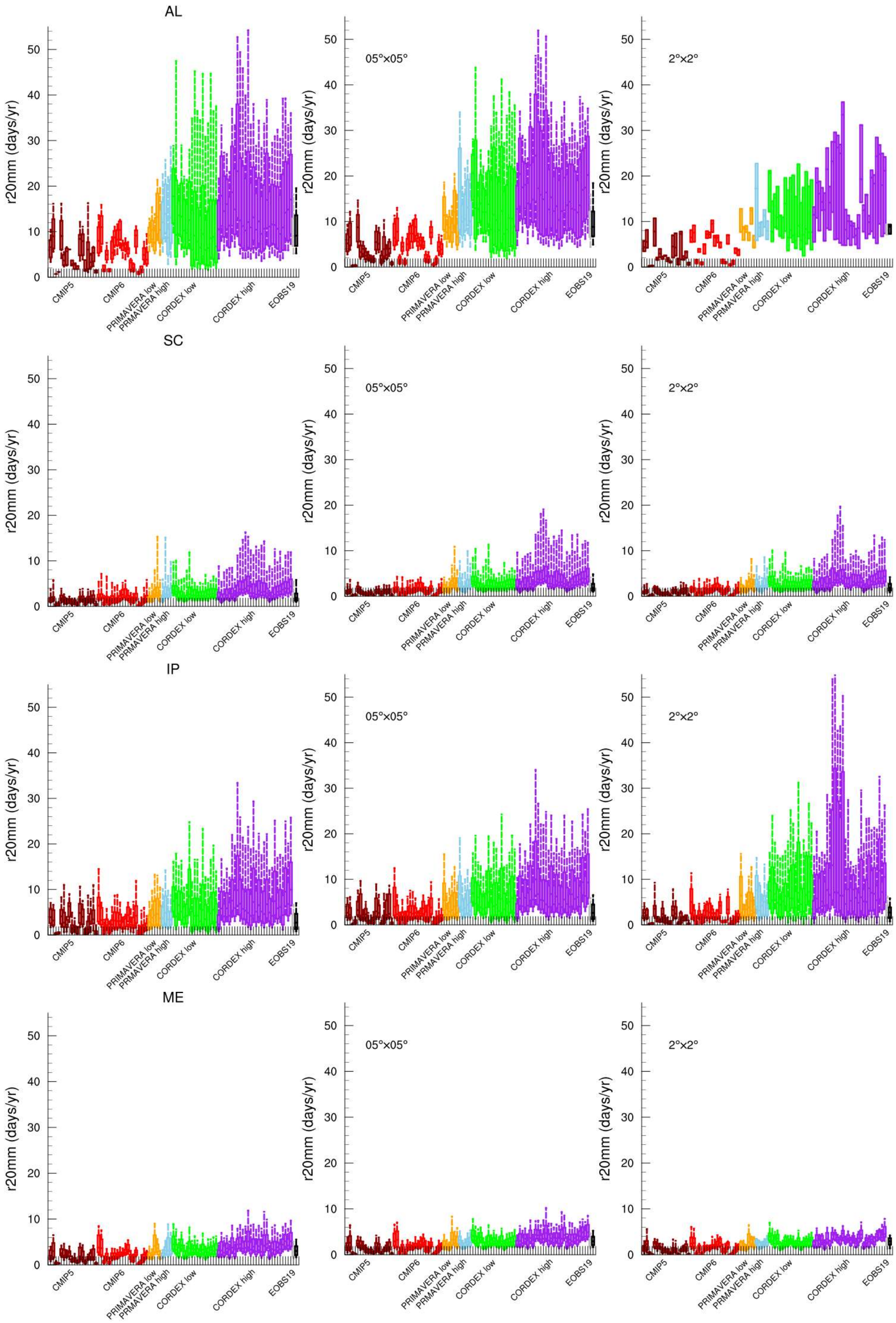
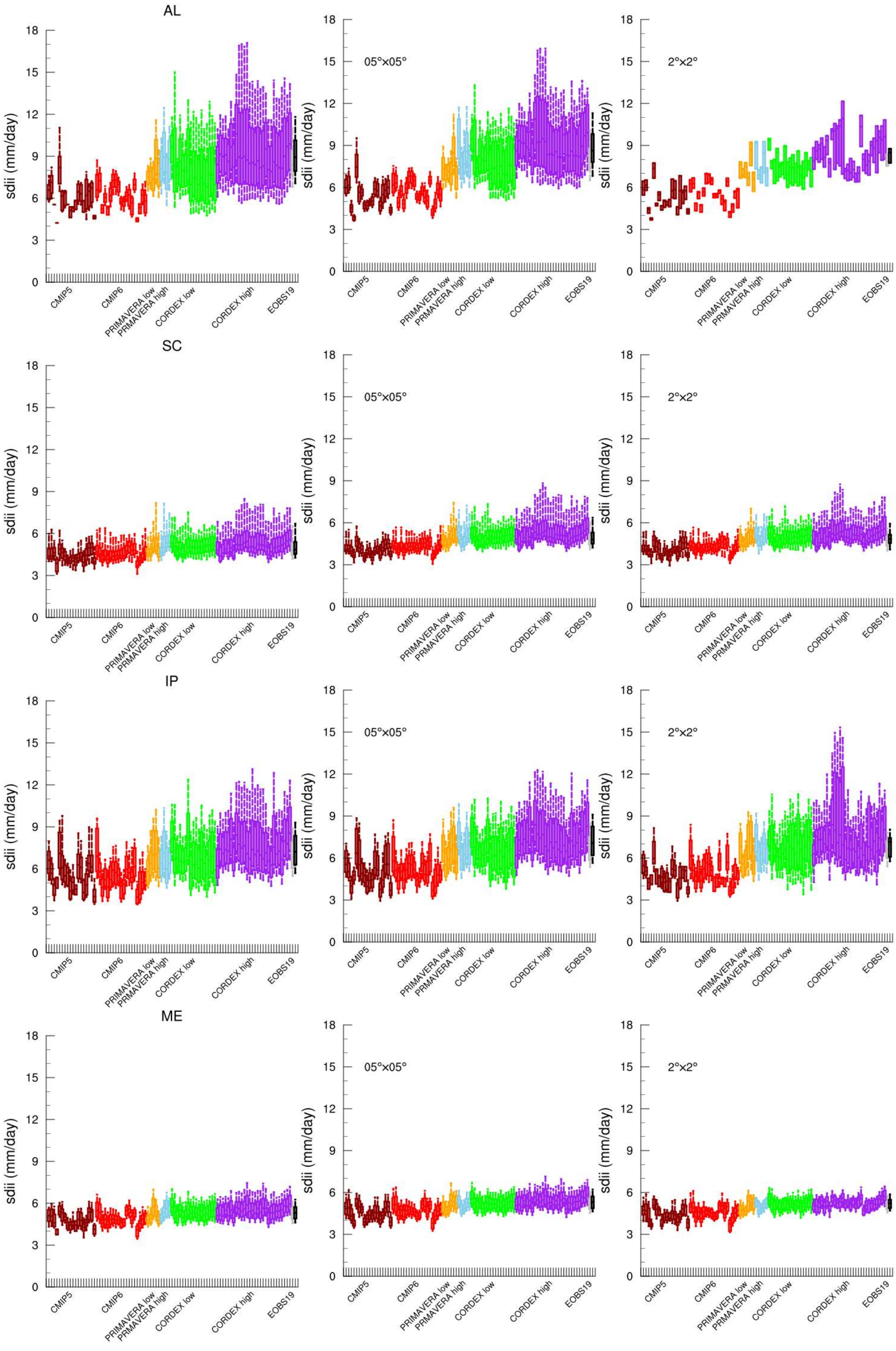


Figure 6. Number of precipitation days (RR1 (days year^{-1}]) in the Alps (AL, top left), Scandinavia (SC, top right), the Iberian Peninsula (IP, bottom left) and mid-Europe (ME, bottom right) for individual models in the CMIP5 (brown), CMIP6 (red), PRIMAVERA LR (orange), PRIMAVERA HR (light blue), CORDEX LR (green) and CORDEX HR (purple) ensembles as well as E-OBS at 28 (grey) and 11 km (black). Boxes mark the 25th and 75th percentile, with the median inside; whiskers go from the 10th to the 90th percentile.



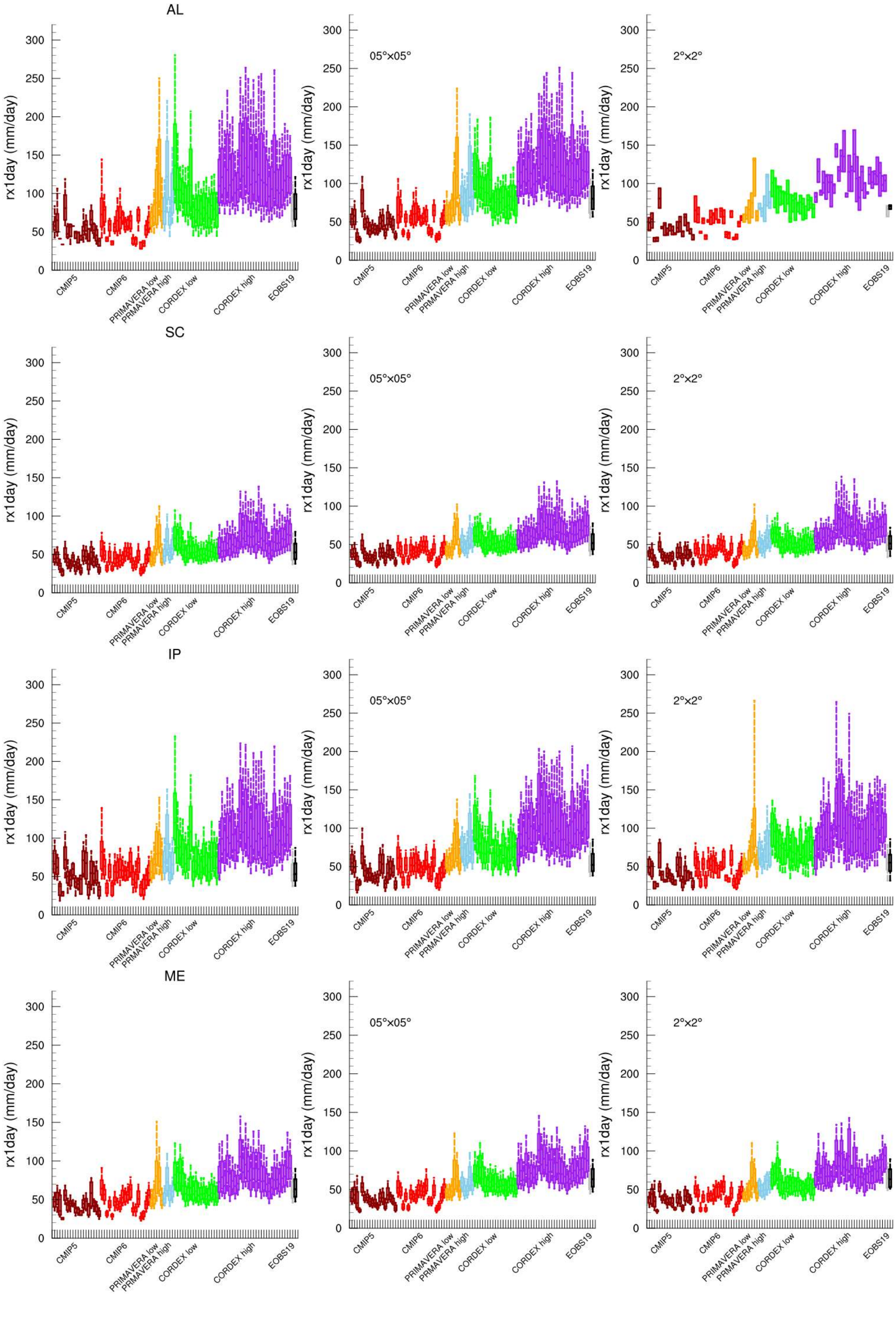
AL
SC
IP
ME
05°×05°
2°×2°
r20mm (days/yr)
CMIP5
CMIP6
PRIMAVERA low
PRIMAVERA high
CORDEX low
CORDEX high
EOBS19

Figure 7. Same as Figure 6 but for the number of days with precipitation amount over 20 mm (R20mm (days year^{-1})). Left column: model data on their original grids, centre column: all data regridded to 0.5°×0.5° grid, right column: all data regridded to 2°×2° grid.



AL
SC
IP
ME
05°×05°
2°×2°
sdii (mm/day)
0
3
6
9
12
15
18
CMIP5
CMIP6
PRIMAVERA low
PRMAVERA high
CORDEX low
CORDEX high
EOBS19

Figure 8. Same as Figure 7 but for the simple precipitation intensity index (SDII (mm day^{-1})).



AL
SC
IP
ME
05°×05°
2°×2°
rx1day (mm/day)
0
50
100
150
200
250
300
CMIP5
CMIP6
PRIMAVERA low
PRMAVERA high
CORDEX low
CORDEX high
EOBS19

Figure 9. Same as Figure 7 but for the maximum one day precipitation (Rx1day (mm day^{-1})).

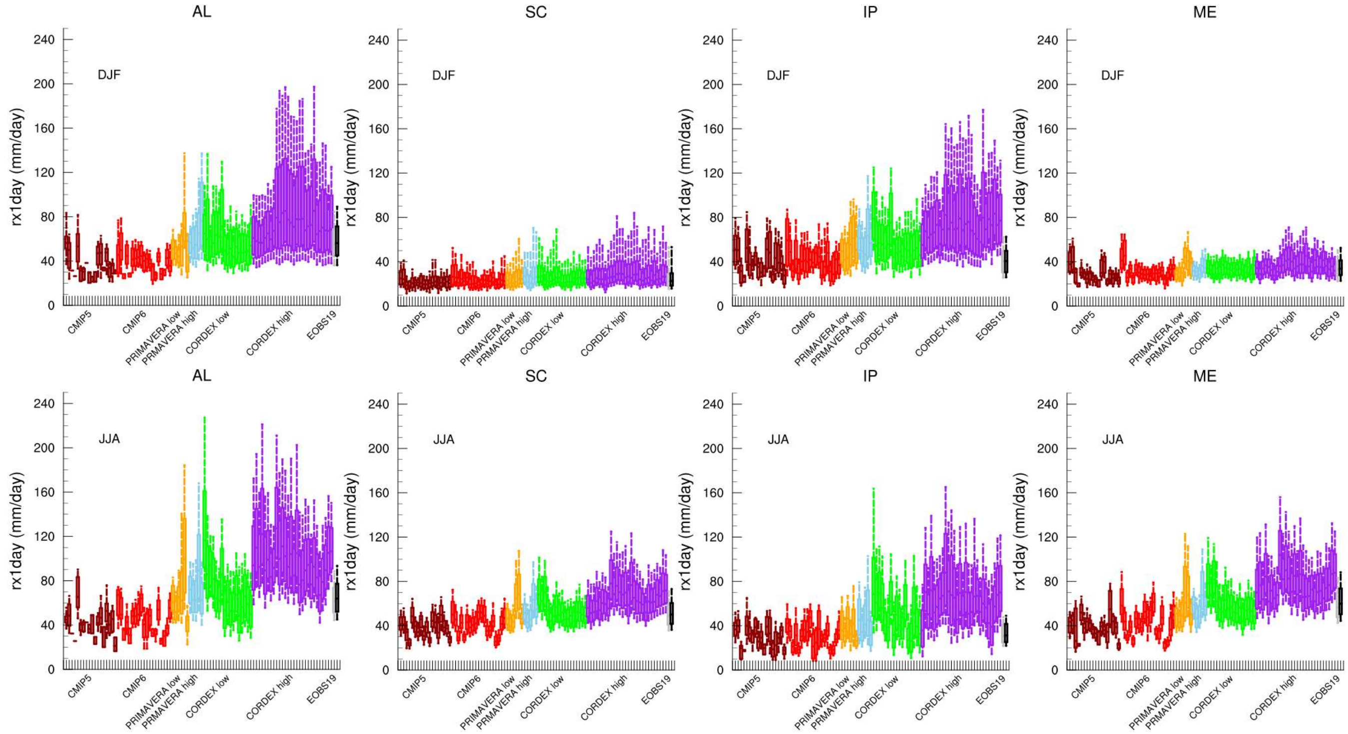


Figure 10. Same as Figure 6 but for the maximum one-day precipitation (Rx1day (mm day^{-1})), top row: winter (DJF), bottom row: summer (JJA).

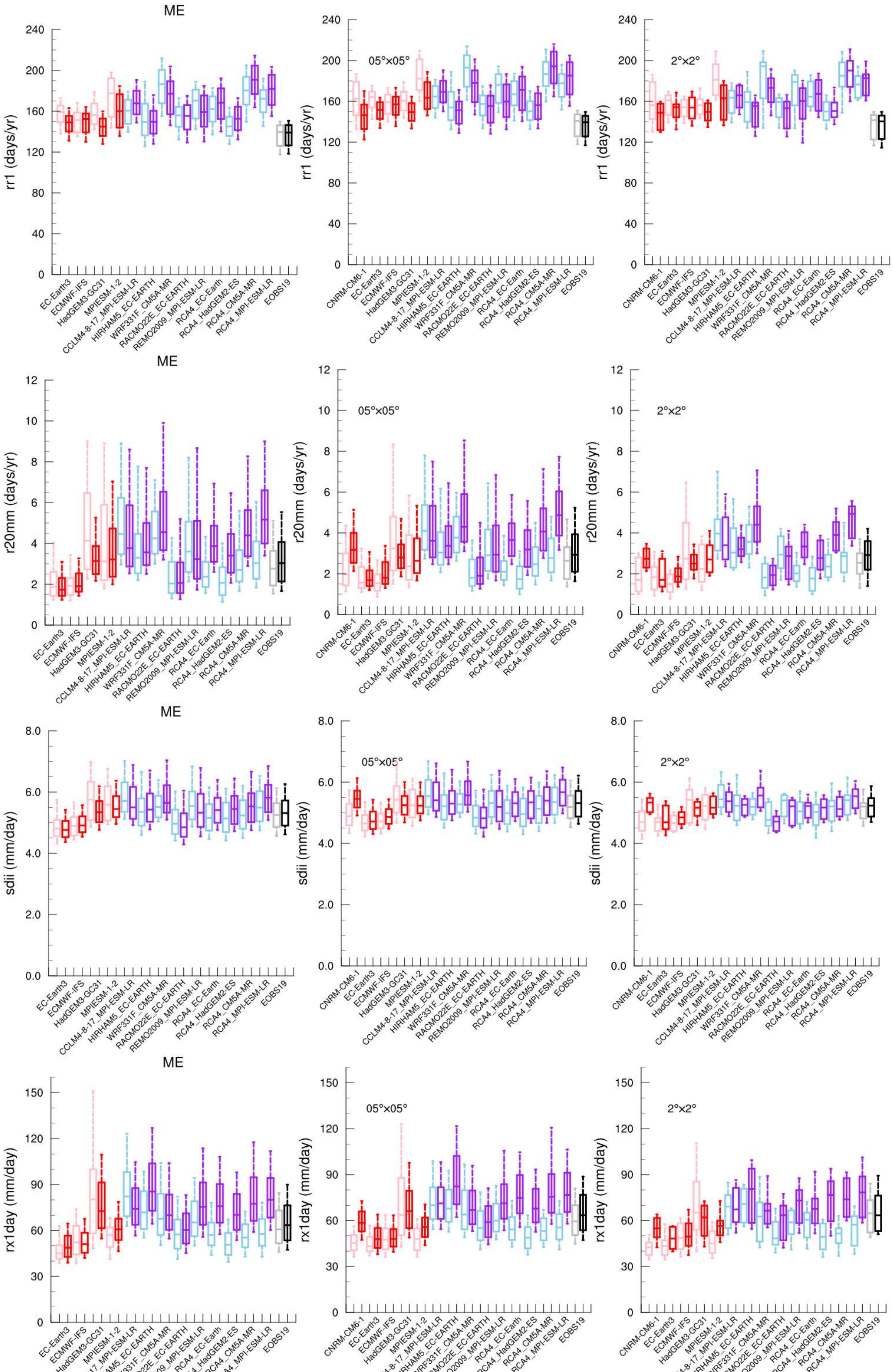

ME
05°×05°
2°×2°
rr1 (days/yr)
r20mm (days/yr)
sdii (mm/day)
rx1day (mm/day)
CNRM-CM6-1
EC-Earth3
ECMWF-IFS
HadGEM3-GC31
MPIESM-1-2
CCLM4-8-17_MPI-ESM-LR
HIRHAM5_EC-EARTH
WRF331F_CM5A-MR
RACMO22E_EC-EARTH
REMO2009_MPI-ESM-LR
RCA4_EC-Earth
RCA4_HadGEM2-ES
RCA4_CM5A-MR
RCA4_MPI-ESM-LR
EOBS19

Figure 11. Number of precipitation days (RR1 (days year^{-1}), first row), number of days with precipitation amount over 20 mm (R20mm (days year^{-1}), second row), simple precipitation intensity index (SDII (mm day^{-1}), third row), maximum one day precipitation (Rx1day (mm day^{-1}), fourth row) in the Mid-European region (ME) in the PRIMAVERA LR (pink) and HR (red) models, CORDEX LR (light blue) and HR (purple) models as well as E-OBS LR (grey) and HR (black). Left column: model data on their original grids, centre column: all data regridded to 0.5°×0.5° grid, right column: all data regridded to 2°×2° grid. Boxes mark the 25th and 75th percentile, with the median inside; whiskers go from the 10th to the 90th percentile. If the the high-resolution version of a model is significantly different from the low-resolution version this is marked with a vertical line in the high-resolution boxes.

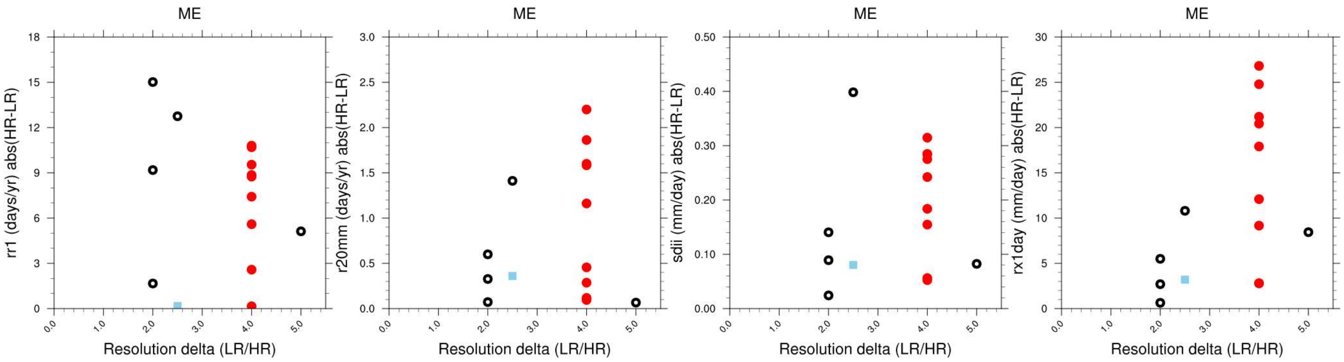


Figure 12. Absolute difference between HR and LR version of PRIMAVERA (black rings), CORDEX (red circles) and E-OBS (blue squares) in precipitation days (RR1 (days year^{-1}), first column, number of days with precipitation amount over 20 mm (R20mm (days year^{-1}), second column), simple precipitation intensity index (SDII (mm day^{-1}), third column), maximum one day precipitation (Rx1day (mm day^{-1}), fourth column) in the Mid-European region (ME). X-axes show the resolution delta (LR/HR) for each model (example: 50 km grid spacing divided by 12.5 km equals 4).