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.