Final Draft

of the original manuscript

Ciarlo`, J.; Coppola, E.; Fantini, A.; Giorgi, F.; Gao, X.; Tong, Y.; Glazer, R.; Torres Alavez, J.; Sines, T.; Pichelli, E.; Raffaele, F.; Das, S.; Bukovsky, M.; Ashfaq, M.; Im, E.; Nguyen-Xuan, T.; Teichmann, C.; Remedio, A.; Remke, T.; Bülow, K.; Weber, T.; Buntemeyer, L.; Sieck, K.; Rechid, D.; Jacob, D.:

A new spatially distributed added value index for regional climate models: the EURO-CORDEX and the CORDEX-CORE highest resolution ensembles.

In: Climate Dynamics. Vol. 57 (2021) 1403 - 1424.

First published online by Springer: 13.08.2020

https://dx.doi.org/10.1007/s00382-020-05400-5

A new spatially distributed Added Value Index for Regional Climate Models: the EURO-CORDEX and the CORDEX-CORE highest resolution ensembles

James M. Ciarlo` (1,2), Erika Coppola (1), Adriano Fantini (1), Filippo Giorgi (1), XueJie Gao (3,4), Yao Tong (5), Russell H. Glazer (1), Jose Abraham Torres Alavez (1), Taleena Sines (1), Emanuela Pichelli (1), Francesca Raffaele (1), Sushant Das (1), Melissa Bukovsky (6), Moetasim Ashfaq (7), Eun-Soon Im (8), Thanh Nguyen-Xuan (8), Claas Teichmann (9), Armelle Remedio (9), Thomas Remke (9), Katharina Bülow (9), Torsten Weber (9), Lars Buntemeyer (9), Kevin Sieck (9), Diana Rechid (9), Daniela Jacob (9)

Affiliations:

(1) The Abdus Salam International Center for Theoretical Physics (ICTP), strada costiera 11, 34135 Trieste, Italy
(2) National Institute of Oceanography and Applied Geophysics - OGS
(3) Climate Change Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
(4) University of Chinese Academy of Sciences, Beijing, China
(5) Yingkou Meteorological Bureau, Yingkou, China
(6) National Center for Atmospheric Research, Boulder, CO, USA
(7) Oak Ridge National Laboratory, Oak Ridge, TN, USA
(8) The Hong Kong University of Science and Technology, Hong Kong, China
(9) Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht, Hamburg, Germany

Keywords Regional Climate Model; Added Value; CORDEX; EURO-CORDEX; CORDEX-CORE; Downscaling Signal

Outline

- New spatial metric to assess added value of RCM
- Precipitation added value in EURO-CORDEX and CORDEX-CORE ensemble
- Added value is greater in regions of complex topography and coastal areas
- A greater added value for extreme precipitation
- Added value is positive throughout most regions

Abstract

The added value of using Regional Climate Models (RCMs) to downscale data from General Circulation Models (GCMs) has often been questioned and researched. Although several studies have used different methods to identify (and in some cases quantify) the added value, there is still a need to find a general metric that quantifies the added value of any variable. This paper builds on past studies to propose a new metric of added value in the simulation of present-day climate which measures the difference in the probability density functions (PDFs) at each grid-cell between a model and an observation source, and then compares the results of the RCM and GCM in order to spatially compute the added value index. The same method is also adapted to quantify the climate change downscaling signal in a way that is consistent with the present-day metric. These new metrics are tested on the daily precipitation output from the EURO-CORDEX and CORDEX-CORE projection ensembles and reveal an overall positive added value of RCMs, especially at the tail-end of the distribution. Higher added value is obtained in areas of complex topography and coast-lines, as well as in tropical regions. Areas with large added value in present-day climate are consistent with areas of significant climate change downscaling signal in the RCP 8.5 far future simulations, and when the analysis is repeated at a low-resolution. The use of different resolution observations shows that the added value tends to decrease when models are compared to low-resolution observation datasets.

1. Introduction

Many institutions (using several RCMs) have completed numerous high-resolution (0.11° and 0.22°) climate projections over regions worldwide as part of the COordinated Regional climate Downscaling EXperiment (CORDEX; Giorgi et al., 2009; Jones et al., 2011; Gutowski et al. 2016). In particular, 55 simulations were completed within the EURO-CORDEX initiative (Jacob et al., 2013; 2020). Downscaling low-resolution data to a high-resolution using a Regional Climate Model (RCM) is a computationally costly process, and despite the statistical analysis and validation of these simulations against various observation sources (Sanchez-Gomez et al., 2009; Kjellström et al., 2010; Lenderink, 2010; Jacob et al., 2011, 2014; Kotlarski et al., 2014; Aalbers et al., 2018), a comprehensive assessment of the added value provided by such a downscaling process has not been carried out, so that the added value issue is still a point of debate (Di Luca et al., 2013; Hong et al., 2014; Laprise, 2014; Xue et al., 2014; Di Luca et al., 2015; Torma et al., 2015; Di Luca et al., 2016; Giorgi et al., 2016; Prein et al., 2016; Soares et al., 2018; Qiu et al., 2019).

Although one might argue that higher resolution should in principle improve all aspects of a simulation, the added value of downscaling depends on the variable and regional context of interest. For example, a higher resolution is always better at resolving complex topography and coastlines, and consequently the intensity and spatial distribution of precipitation over such regions should be improved when downscaled. Similarly, extreme precipitation events are most often very localized in space and time, and thus increasing resolution should lead to better simulations. The simulation of fine-scale circulations and their effects on regional climates, such as due to sea breezes or mesoscale convective systems, would also in general benefit from increased resolution (e.g. Rummukainen 2016; Giorgi 2019).

While the effects of improved horizontal resolution in such cases is easily observed, it may not translate into more accurate or credible climate change information (Barsugli et al. 2013). This raises the issue of how to assess improvements in downscaled simulations over those provided by the forcing reanalyses or GCMs, and thus, how to assess their added value. Towards this goal, there have been many attempts to identify the added value of a RCM compared to the driving GCM (e.g. Giorgi et al., 1994; Kanamitsu & Kanamaru, 2007; Coppola et al., 2010; Kanamitsu & DeHaan, 2011; Di Luca et al., 2013; Torma et al., 2015; Di Luca et al. 2016; Giorgi et al., 2016; Lucas-Picher et al., 2017; Fantini et al., 2018; Soares et al., 2018). In particular, one of the metrics often used to quantify added value is the probability distribution function (PDF) of a given variable (e.g. Torma et al., 2015; Fantini et al., 2018), as it describes the complete characteristics of the variable.

One of the quantitative metrics used to measure how a model reproduces observed PDFs is the Kolmogorov-Smirnov distance (Chakravarti et al., 1967), which Torma et al. (2015) used to compare the maximum difference between the Cumulative Distribution Functions (CDFs) of a model and an observation CDF. Fantini et al. (2018) employed a similar metric, the Kullback-Leibler divergence (Kullback and Leibler, 1951), which compares the mean difference of two PDFs. Both metrics, applied to daily precipitation PDFs, indicated that high-resolution RCMs performed better than the coarse-resolution driving models. Fantini et al. (2018) also showed that the greatest added value was found in regions of complex topography, such as the Alps, Italy, and Norway. Instead of focusing on differences, Soares et al. (2018) used the Perkins Skill Score (Perkins et al., 2007) to measure the common area between the simulated and observed distribution, which was then used to compare the gain (or loss) as a result of high-resolution downscaling. They showed that added value was present throughout the European region (especially for extreme precipitation) with some of the highest values obtained in the Alpine region.

The temporal correlation skill has proven effective to assess the spatial distribution of an added value metric in a point-by-point analysis (Kanamitsu and Kanamaru, 2007; Kanamitsu & DeHaan, 2011; Prein et al. 2016). In these studies, a substantial geographical variability of the added value metric was shown, even with areas of negative added value, thus highlighting the importance of showing the geographical distribution of relevant metrics. However, this correlation-based added value index cannot be used within the context of simulations driven by GCMs since no substantial temporal correlation can be expected with an observation time-series due to the lack of real-world data assimilation.

A good alternative is to use spatial correlation (Di Luca et al., 2016; Prein et al., 2016). Prein et al. (2016) used the fraction skill score (Roberts and Lean, 2008) and spatial correlation of each model with observations to compare the added value of low and high-resolution runs of RCMs. The study analysed European observation data-sets separately in order to visualise the spatial variation of added value. Di Luca et al. (2016) also used spatial correlation and the mean square error to quantify added value (Di Luca et al., 2013). These studies showed substantial improvements in the RCM simulations in most regions analysed, with some exceptions during different seasons.

A point-by-point analysis of PDFs can thus be an optimal solution to spatially assess the added value of a RCM, since it includes both a comprehensive representation of the characteristics of a variable and its geographical variation. Therefore this paper presents a new metric to

quantify the added value of a RCM with respect to its driving GCM based on a point-by-point PDF analysis of daily precipitation. We apply our approach to the European region via the large ensemble of RCM projections produced as part of the EURO-CORDEX program (Jacob et al., 2013; 2020) and on different continents via the ensemble of projections recently completed as part of the CORDEX-CORE program (Gutowski et al., 2016). The choice of precipitation is due to the availability of high-resolution observation data in Europe and the rest of the world, and to be able to compare with past studies (Torma et al., 2015; Giorgi et al., 2016; Prein et al., 2016; Fantini et al., 2018). Moreover, precipitation is strongly affected by topography and by fine-scale spatial and temporal processes, and thus downscaling can be especially useful in improving its simulation.

Quantification of the added value for a present-day simulation can be a relatively straightforward task if appropriate observations are available, but it is difficult to quantify the existence of added value in a future climate simulation. A novel way we propose to assess the potential for added value in climate change signals, is through the use of the same metric as for the present-day simulations but applied to the RCM and GCM change signals. This allows us to identify when and where the change signals diverge and how different they are (Giorgi et al., 2016). If these differences are shown to be large over the same locations where an added value was proven in the present climate validation exercise, then one could assume that the RCM projection could potentially be more accurate compared to the GCM's. The proposed methods are described in the next section.

2. Methods

We introduce here a new method for quantifying the added value of a variable and representing it spatially. This method stems from the spatial downscaling signal described by Giorgi et al. (2016) and the spatial correlation skill mentioned in Rummukainen (2016). Other studies (Kanamitsu & DeHaan, 2011; Torma et al., 2015; Fantini et al., 2016) use different metrics to describe the difference between simulated and observed PDFs, however, these are based only on parts of the distribution. Instead our method quantifies the added value by computing the absolute values of the differences across the entire PDF distributions, so that these differences do not cancel each other out. We then apply this method at each grid-point of the model domain so that we provide information on the spatial distribution of the added value.

For a variable of interest (in this case daily precipitation, including dry days), the method requires data from a RCM, the driving GCM, and an observation source (OBS; ideally of high-resolution) for the same time-period and frequency. Once the three datasets are interpolated onto a common grid, the PDFs can be calculated in a consistent way so that each grid point (for the 3 data-sets) has its own distribution, resulting in a grid of PDFs (hereafter referred to as PDF-grid). In order to ensure a fair comparison, the bin size should be identical for each grid point, however the number of bins must be independent to properly represent the different PDFs. In this paper, a bin-size of 1 mm/day is used in order to resolve high precipitation events in the tail-end of the PDFs, since the analysis is focused on wet extremes. The calculation of the added value index (see below) obviously depends on the bin size, and in the Appendix we present a sensitivity analysis of our results to a range of bin sizes. Furthermore, the grid-point maximum necessary for the computation of each PDF is taken as the maximum of all datasets at that grid point.

The resulting PDF-grid for a model is compared to the PDF-grid of the OBS by using the sum of the absolute differences between the model (*M*) and the observation (*O*) across all bin values (ν_t), divided by the sum of *O*. Here, we refer to this as the Relative Probability Difference, *D* (described in equation 1 and Figure 1), where *N* is the number of events in the dataset for a given bin ν, and $\Delta\nu$ is the bin size of the variable. This calculation is done for both the RCM and GCM and the resulting plots describe the spatial distribution of D_M with respect to the observations. In this manner, the difference value D_M is a unitless quantity which represents the compounded discrepancies between the distributions. A smaller value of D_M indicates a better performance by the model.

$$D_M = \frac{\Sigma_{\nu=1}^{\nu_t} |(N_M - N_O)\Delta\nu|}{\Sigma_{\nu=1}^{\nu_t} (N_O \Delta\nu)} \qquad (1)$$

The Added Value index (A_i) is thus quantified by comparing D_{GCM} to D_{RCM} (equation 2), where a positive (negative) index represents an improvement (degradation) of the RCM results compared to the GCM ones, as suggested by Di Luca et al. (2015). The quantity A_i is also unitless, and is given by

$$A_i = D_{GCM} - D_{RCM} \qquad (2)$$

A problem can arise when the PDF of the GCM is missing some bin-data, for which the corresponding RCM and OBS bin-data exist. This is common, for example, at the tail-end of the distribution which GCMs tend to fail to capture (Fantini et al., 2018, Torma et al., 2015). Such cases represent an important contribution to the added value calculation, but they cannot be quantified properly by this method because in such situations D_{GCM} is always equal to 1, while D_{RCM} can exceed this value and thus produce a misleading negative value to A_i. Therefore, a conditional assumption is introduced by which if N_{GCM} of a specific bin is zero, but the corresponding N_{RCM} and N_O are non-zero, that bin contributes 0 relative probability difference to the final D_{RCM}, thereby ensuring a positive contribution to the index A_i. In other words, we assume that the RCM adds value to the GCM if it simulates events in bins for which the observations have events and the GCM does not simulate any, regardless of how many events the RCM simulates. The inverse situation can obviously occur (although in fact it rarely does), in which the RCM misses data in a bin where both the OBS and GCM simulate events. Also in this case the same procedure is applied, so that neither the RCM or the GCM are favored. We acknowledge that this is an assumption based on a subjective assessment that it is more important to capture the existence of events in a bin than to exactly simulate the number of such events, an assumption especially important for the tail end of the distribution which is characterized by small numbers of rare events.

Some studies (Torma et al., 2015; Prein et al., 2016; Fantini et al., 2018) have shown how GCMs do not resolve precipitation extremes as well as RCMs. For this reason, the method above can also be modified to focus on a particular segment of the distribution, for example the 95-100 percentile interval. In such a case, the percentile values of the observation dataset are used as thresholds for the PDFs, and the part of the complete PDF (as in Figure 1) that contributes to this percentile interval would be the only data included in equations 1 and 2.

Since the 95th percentile varies from one grid-point to another, the threshold applied must be specific to that grid-point and cannot be the field-mean over the analysis domain. The 95-100 percentile interval is not an arbitrary choice, as studies have shown that substantial added value in a RCM can be found at the tail-end of the precipitation distribution (Torma et al., 2015; Fantini et al., 2018).

In an analogous way, a climate change downscaling signal, (A_{DS} in equation 4) can be defined from the change between a PDF in a future climate period and a corresponding PDF in a historical period of the simulation. In this case, instead of comparing the model data to an observation dataset, we compare the future data (*f*) to the historical data (*h*) of the same simulation (as shown in equation 3). This is similar to the method described by Giorgi et al. (2016). In this case, the conditional assumption applied to equation 2 (where a model does not resolve a particular bin) cannot be applied, as this data is not compared to any observations. The larger the value of this downscaling signal, the more different the projected and reference PDFs are, and the magnitude of A_{DS} is proportional to this difference. The climate change downscaling signal, A_{DS} is described in the same manner as A_i, i.e. a unitless quantity expressed as.

$$D_{Mf} = \frac{\Sigma_{\nu=1}^{\nu_t} \left| (N_{Mf} - N_{Mh}) \Delta\nu \right|}{\Sigma_{\nu=1}^{\nu_t} (N_{Mh} \Delta\nu)} \qquad (3)$$

$$A_{DS} = D_{GCMf} - D_{RCMf} \qquad (4)$$

The quantity D_{Mf} in equation 3 describes a relative climate change signal within a given model *M*; N_{Mf} is the value of the future period PDF at bin ν for model *M*; and N_{Mh} is the corresponding bin value in the historical period PDF of the same model *M*. The A_{DS} is the difference of the D_{Mf} signals of the RCM and GCM, i.e. it is based on the climate change signals in the driving and downscaling models (hence climate change downscaling signal). Here, large positive or negative values of A_{DS} indicate a larger climate change downscaling signal, hence a greater difference between the RCM and GCM resulting in the potential for added value. A_{DS} values close to 0 describe a weak downscaling signal. The sign of A_{DS} does not quantify which model is 'better', but rather how different the two PDFs are. A positive (negative) value of A_{DS} indicates a situation where the climate change signal of the GCM (RCM) in a given segment of the PDF is greater than that of the RCM (GCM). When the analysis is restricted to a specific percentile interval (such as 95-100, as mentioned above), since no observation data is included in this comparison, the percentile threshold is obtained from the historical data-set.

2.1 Simulated data

For our analysis we use two GCM-RCM projection ensembles. The first is the EURO-CORDEX ensemble (Jacob et al., 2013) of 55 RCM simulations at 0.11° (Table 1). This consists of 130-year climate projections (from 1970 to 2100) for the Representative Concentration Pathway, RCP 8.5 (Moss et al., 2008), with an incomplete matrix of 12 RCMs driven by 8 different GCMs (one should note that simulations run by MOHC-HadGEM2-ES do not include the year 2100). The analysis is carried out on daily precipitation, with a special focus on the higher percentiles of the distributions. The second data-set is the CORDEX-

CORE ensemble (described in Table 2; Mearns et al. 2017; Remedio et al. 2019; Coppola et al. *submitted*; Teichmann et al. *submitted*), which includes 0.22° resolution simulations run by two RCMs, the RegCM4 (Giorgi et al., 2012) and REMO2015 (Jacob et al., 2012; Remedio et al., 2019), each driven by three GCMs, for 8 non-European CORDEX domains: Africa; North, Central, and South America; East, South-East, and South Asia; and Australasia.

The method requires that all data, i.e. RCM, GCM, and observations are defined on the same horizontal grid. This raises two issues; interpolating the GCM to a higher resolution grid may create unrealistic values, while interpolating the RCM to a lower resolution grid degrades the spatial signal and the PDF (Prein et al., 2016). The latter is especially true at the tail-end of the distribution (Torma et al., 2015), where the largest added value is expected. To account for both issues, the analysis is conducted on two grids (using distance-weighted average interpolation), the RCM grid (0.11°) which allows us to have a more accurate representation of the spatial distribution of the index, and a 1.00° grid to ensure that the results are inter-comparable.

2.2 Observation Sources

The added value calculations are dependent on the observation data used as reference, thus multiple observation datasets are used to test the method. These are reported in Table 3, and additional information on station density can be found in Prein et al. (2016) and Fantini et al. (2018). The time period available for the different datasets is not uniform so a different time period is used for each dataset. The analysis of the EURO-CORDEX data is compared to two observation sources: the EOBS v20e and a composite of 9 sub-regional observations, hereafter referred to as 'European Composite Observations' (ECO).

The CORDEX-CORE analysis is also carried out using multiple observation sources (reported in Table 4). The CPC data-set is used to assess the added value compared to low-resolution observations, while the TRMM dataset provides a comparison between satellite and station-based observations. The regional observation datasets GCOSGHCN, IMD, and APHRODITE were combined into a single data-source, hereafter referred to as Global Composite Observations (GCO). Similarly to the European observation sources, the time periods used here are different for each dataset. However, since the indices are calculated using the entire dataset and not on a year by year basis, this should not affect the basic conclusions of the analysis.

3. EURO-CORDEX Analysis

3.1 Added value for the present-day validation

Figure 2 shows the relative probability differences for the GCM and RCM ensembles, and the resulting added value index for the EURO-CORDEX ensemble. The relative probability difference of the GCM ensemble is shown to be significantly larger than that of the RCM ensemble virtually everywhere, resulting in a positive added value throughout the EURO-CORDEX domain. This is particularly the case in areas of complex topography, where the added value is therefore maximum.

Figure 2 also provides a comparison of the added value calculated on the 0.11° and 1.00° grids. Clearly, the results are consistent for both the high and low resolution grids, and the geographical distribution of the added value is also maintained (although with less spatial detail) for the lower resolution grid. The added values calculated using the ECO and EOBS observation datasets are very similar, although slightly smaller in EOBS over some regions (e.g. Scandinavia, the British Isles, France, and the Carpatians).

Figures S1-S6 show the added value plots for individual simulations, and show a greater dependency on the GCM field than the RCM field (as also reported by Di Luca et al., 2016). The ensemble members show a large predominance of positive added value, although some members exhibit some areas of negative added value. This latter result mostly occurs due to a low relative probability difference (i.e. good performance) obtained by the driving GCM, such as HadGEM2-ES and MPI-ESM-LR (Figures S1 and S4). Conversely, the simulations providing the highest added value (Figures S3 and S6) are the ones driven by NorESM1-M and CNRM-CM5, where both GCMs display the highest relative probability difference (low performance). The only exception is the ALADIN53 driven by CNRM-CM5 which displays a very high relative probability difference (low performance) compared to the other RCMs.

Although the added value at 0.11° resolution (Figure 2) is larger over areas of complex topography (for both ECO and EOBS), the signal appears to be smaller around the highest peaks. For example, over the Alpine region this may be attributed to localized areas with a low density of stations in the observation source (Isotta et al. 2014) which produce an apparent reduction of added value. Another reason might be the lack of an undercatch gauge correction, which is especially relevant during windy and snowy conditions and can account for up to 30% underestimation of real precipitation by gauge data (e.g. Adam & Lettenmaier, 2003).

The added value shown in Figure 2 is calculated using the entire PDF. If different percentiles of the distribution are considered, the resulting added value may be quite different. Figure 3 shows the added value as a function of the percentile interval. There are two possible choices of intervals, the first keeps one end of the interval fixed to zero and moves the other end from zero to 100 (0-*x*), and the second keeps the far end fixed to 100 and moves the near end from 0 to 100 (*x*-100).

When the 0^{th} percentile is included (case 0-x), the added value of the RCM ensemble-mean gradually increases with the upper-bound threshold. This suggests that a higher added value is found at the tail-end of the distribution. The intervals that do not include the 0^{th} percentile (case *x*-100), show a substantially higher added value, even for the lower percentile intervals. This implies that the RCMs perform less adequately at the 0^{th} percentile.

When omitting the 0^{th} percentile (case 5-100), the added value is relatively constant until about the 50^{th} percentile and then decreases gradually until it reaches a minimum around the 90^{th} percentile, after which the added value increases sharply when compared to ECO. To understand this behaviour, in Figure 4 we show the observed and simulated PDFs over different sub-regions covered by the ECO data-set. It can be seen that while the RCM PDF reproduces quite well the observed one, the GCM overpredicts the frequency of low intensity events and underpredicts that of high intensity ones. In other words, there is a point in which the GCM PDF intersects the observed PDF, and this point is located around the 90-95^{th} percentile of the observed distribution. For this reason, as the percentile interval approaches

this intersection the relative probability difference of the GCM ensemble will be closer to that of the RCM, thus resulting in the dip seen in Figure 3.

Figure 5 shows the geographical distribution of the added value calculated using ECO (at 0.11°) at different percentile intervals. Here, the '0-*x*' intervals are positive throughout the PDF spectrum and increase in magnitude at higher intervals. This is consistent with the results presented in Figure 3. When the 0^{th} percentile is omitted, the '*x*-100' intervals show a variability that is also consistent with that of Figure 3. The 50-100 percentile interval has a larger magnitude than the 0-50 interval, and at higher intervals the added value decreases slightly in many regions due to the GCM PDF crossing the observed PDF, as explained earlier and shown in Figure 4. The added value increases again (and peaks) at the 99-100 percentile interval, after which a second slight decrease in the added value is observed (as explained below).

The added value compared to the EOBS data shows similar results when looking at the same regions covered by ECO (Figure 3). However, the results in the other areas are very different (as seen in Figure 6), and in many regions, e.g. Russia, we even see a decrease in the added value. This is likely caused by the low density of station observations over these areas (Haylock et al., 2008). The pronounced spatial diversity in these results also illustrates the importance of using an observation data-set of equal or higher resolution than the model's throughout the entire analysis domain when assessing the added value.

Another interesting feature of the added value of the ensemble mean is the slight decrease at the 99.9-100 fraction when compared to ECO (Figure 5). The number of events occurring above the 99.9 percentile threshold tends to be very small, with numerous bins having zero events. Despite the improvement in the RCM representation of the tail, the magnitude of these extreme events is often different from the observations (and would thus correspond to a different bin value). Since this added value metric is comparing the frequency of the events in each bin, some of these cases would not be comparable. This means that the non-zero events for the GCM, RCM, and observations above this threshold may not always coincide in the same bin, which results in a more negative apparent added value. Since the frequency of events at this extreme percentile interval is very small compared to the rest of the distribution, this problem does not influence the calculations for the entire distribution.

A similar effect is also seen in the 99.9-100 added value compared to EOBS (Figure 6), which shows a large positive added value for the 99.9-100 percentile interval fraction in many areas, but a large negative added value in others. This is likely a combined effect from the small number of events occurring at this percentile interval, and the low station density of some areas.

Our results indicate that the best observation source to use in order to assess the 0.11° EURO-CORDEX simulations is the ECO, since all the observation data-sets are of the same horizontal resolution as the model or finer. Figure 7 shows the added value of each ensemble member at the 99-100 percentile interval compared to ECO. This portion of the PDF is where the ensemble mean shows the highest added value. The positive added value is consistent in all RCM members, with the NorESM-1 driven simulations displaying the greatest values.

3.2 Added value for the climate change projections

An observation-based analysis cannot be used to quantify the added value of future simulations. To address this issue, the downscaling signal described by Giorgi et al. (2016) is combined with our method, as described in Section 2, to provide a downscaling signal based on the PDFs. As an illustrative example, the far future time slice (2080-2099) is compared to the 1995-2014 reference period. The 90, 95, and 99 '*x*-100' percentile intervals are shown in Figure 8, together with the added value compared with ECO for the same intervals.

Here, the downscaling signal near complex topographic regions and coastal areas becomes increasingly visible at the higher percentile intervals, which is consistent with the added value in the same regions. The strongest downscaling signal is found in the 99-100 percentile interval, and is visible in areas such as Scandinavia, the British Isles, and mainland Europe, where the latter shows a pronounced RCM signal (negative) that does not always appear to be linked to topography or coastal areas. This implies that the RCM projects a larger climate change signal than those obtained by the GCM, as also shown by Coppola et al. (*submitted*). The high added value obtained over these regions when comparing to observations might suggest that the RCM climate change signal is more reliable than that of the GCM, similar to the Realised Added Value described by Di Virgilio et al. (2020).

This downscaling signal is also most pronounced at the highest percentile intervals, as the change in daily precipitation is greatest for extreme events (where also here it is dominated by a strong RCM signal). Furthermore, the spatial structure of the P99 change signal appears similar to the one seen in Coppola et al. (*submitted*), and also conforms with the downscaling signal reported by Giorgi et al. (2016). Once again, the higher percentile intervals show a stronger signal, not only because the precipitation change is larger at the extremes but also because GCMs tend to underpredict the tail of the distribution.

4. CORDEX-CORE Analysis

4.1 Added value for the present-day validation

We now move to the analysis of the CORDEX-CORE ensemble described in Section 2.1. Consistent with the EURO-CORDEX results, the added value of the complete daily precipitation distribution (Figure S9) is mostly positive in all regions, with the most positive values occurring in complex topographical areas. However, a few notable exceptions show a negative added value, such as areas of western North America, Sahara, South Asia, and Australia. This negative added value is attributed to the lower percentile intervals (as explained in Section 3.1, and shown in Figures 3 and S7), since a lower added value in these intervals would carry a greater weight on the overall distribution.

The added value of higher percentile intervals (Figures S10-S13) is consistent with this assessment. Figure 9 focuses on the added value of the 99-100 percentile interval of daily precipitation compared to the four observational datasets (described in Section 2.2). This percentile interval was shown to have the most positive added value in the EURO-CORDEX analysis (see Section 3.1) and this is confirmed in the CORDEX-CORE ensemble (Figures S10-S13). The added value is strongly positive in the tropics, characterized by the occurrence of more intense precipitation events than in mid-latitudes, which are evidently not captured by the GCMs.

The added value with respect to CHIRPS (Figure 9) shows some areas of high negative values over African countries such as the Democratic Republic of Congo and South Africa. This may be at least partly due to the data sparsity in these areas (Funk et al., 2015), which would especially dampen the tail-end of the distributions, and thus favour the GCMs. It is important to note that the rest of the dataset was shown to be reliable (Funk et al, 2015). This negative added value is also visible when using the GCO dataset (especially for APHRODITE), and these regions also correspond to areas of low station density (Yatagai et al., 2009). Similarly, lower station densities (Menne et al., 2012) likely contribute to the area of negative added value in western North America. Furthermore, these areas of negative added value correspond to areas with very low moisture around the world. The aridity of these regions (and hence the larger number of dry-days; Daly et al., 1994) likely contributes to this low added value. This is somewhat similar to the added value associated with the EOBS (as explained in Section 3.1, and shown in Figure 6), and to a smaller degree with ECO (Figure 6).

All regional observations show a significant increase in the added value for the higher percentile intervals. The CPC shows a stronger signal where positive added value is found, and more areas with negative added value than the other observation sources. This wide variability in added value (similar to the case of low station density areas in EOBS in Figure 6), is attributed to the low resolution of the data-set. Out of all the data-sets, the added value compared to TRMM shows the most consistently positive signals geographically.

4.2 Added value for the climate change projections

The climate change downscaling signal for the CORDEX-CORE analysis exhibits results similar to the EURO-CORDEX analysis (Section 3.2). Figure 10 compares the 99-100 percentile interval climate change downscaling signal with the added value compared to TRMM (which was found to produce the strongest added value). The 99-100 percentile fraction not only shows the highest added value, but also the strongest climate change downscaling signal.

While both signals are spatially very similar, they are not identical. A distinct topographical influence is also visible in the climate change downscaling signal, while a very strong RCM signal dominates over the equatorial regions. Once again, this implies that the RCMs are projecting a larger change in events than the GCMs in locations of strong added value with respect to observations.

5. Conclusions

In this paper, a new method for quantifying the added value of RCMs is described and tested using the EURO-CORDEX and CORDEX-CORE ensembles of GCM-driven RCM projections. The method is based on the intercomparison of PDFs for a given variable, in this paper daily precipitation, at the grid point level. It requires the comparison of GCM and RCM PDFs with corresponding observed data-sets at the same horizontal resolution and can be applied not only to estimate the added value in present-day climate but also the potential added value in the future projections. In our study we also tested the robustness of the results to different observation data-sets. An important caveat of our method is that, if at a given bin the observations have events and the RCM (GCM) simulates events while the GCM (RCM) does not, then the RCM (GCM) adds value regardless of how many events it simulates. Thus, we assume that it is more important for a model to capture events in a given bin where there are observations than to reproduce the exact number of observed events. This situation occurs in particular towards the tail end of the distributions which are often not captured by the GCMs.

The RCM added value was found to be predominantly positive for the EURO-CORDEX ensemble mean, and became larger when assessing only the higher percentile intervals of the daily precipitation distribution (despite a higher uncertainty due to the decrease in frequency). This was also generally true for the CORDEX-CORE regions where the most positive added value was produced for the 99-100 percentile interval. The contribution of the lowest percentiles of the PDF substantially reduced the added value of the overall distribution due to the higher frequency of these events.

The observation sources used for comparison had a significant influence on the added value obtained. Higher resolution observations were more adequate in the identification of added value at fine scales, since these were more comparable to the model resolution and also had a better record of extreme events. Low-station density in the station-based gridded observations, which smooth out especially the tails of the distributions, could potentially produce a ‘false low or negative added value’. Overall, this method supports previous studies (Fantini et al., 2018; Torma et al., 2015) in showing that RCMs provide added value by better representing extreme events.

The method was also used to produce a PDF-based climate change downscaling signal for future simulations, which was found also to increase at higher percentile intervals and in areas characterized by complex topography. The CORDEX-CORE ensemble showed this signal to be strongest in the equatorial regions.

The method described in this study explicitly demonstrates that RCMs provide an added value for precipitation in complex topographical regions, coastal areas and islands, as well as in tropical regions, especially for the tail-end of the distribution (extremes), as a result of the higher resolution of the downscaling models. Although the method was only used to assess precipitation at this stage, it can be used to quantify the added value of any variable provided reliable high-resolution observation data-sets are available.

Acknowledgements

The RegCM simulations for the ICTP institute have been completed thanks to the support of the CINECA supercomputing center, Bologna, Italy and the ISCRA projects HP10BDU7TR and HP10BQCFJ2. The authors would like to thank Graziano Giuliani and Ivan Girotto for their constant support in the preparation of the simulations used in this paper.

The authors would also like to thank the CMIP5 and EURO-CORDEX community, as well as the ESGF for providing access to their database where most of the data is available. The majority of the EURO-CORDEX contributions have funded by the “Producing RegIoNal ClImate Projections Leading to European Services" (PRINCIPLES, C3S_34b Lot2) as part of the Copernicus Climate Change Service (C3S). The work was also partially supported by the EUCP H2020 project (GA number: 776613).

The study was also supported by the Oak Ridge Leadership Computing Facility and the National Climate-Computing Research Center at the Oak Ridge National Laboratory, the Climate Change Research Center, Institute of Atmospheric Physics; the Chinese Academy of Sciences, Beijing, China and University of Chinese Academy of Sciences, Beijing, China; the Yingkou Meteorological Bureau, Yingkou, China; the National Center for Atmospheric Research, Boulder, CO, USA; and the Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht, Hamburg, Germany; all of whom provided access to their simulation data.

The observations were provided by MeteoSwiss, Santander meteorology group, Meteo-France, Met office UK, METNO Norway, SMHI, Hungarian Meteorological Service, DWD Germany, and CETEMPS University of L’Aquila. The CPC Global Unified Precipitation data was provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at https://www.esrl.noaa.gov/psd/.

References

- Aalbers EE, Lenderink G, van Meijgaard E, van den Hurk BJJM (2018) Local-scale changes in mean and heavy precipitation in Western Europe, climate change or internal variability? Clim Dyn 50:4745. https://doi.org/10.1007/s00382-017-3901-9
- Adam JC, & Lettenmaier DP (2003) Adjustment of global gridded precipitation for systematic bias, J. Geophys. Res., 108(D9), 4257, doi:10.1029/2002JD002499.
- Barsugli JJ, et al. (2013) The Practitioner's Dilemma: How to Assess the Credibility of Downscaled Climate Projections, Eos Trans. AGU, 94(46), 424.
- Chakravarti IM, Laha RG, and Roy J (1967) Handbook of Methods of Applied Statistics, Volume I, pp. 392–394, John Wiley, Hoboken, N. J.
- Chen M, Xie P, et al. (2008) CPC Unified Gauge-based Analysis of Global Daily Precipitation, Western Pacific Geophysics Meeting, Cairns, Australia, 29 July - 1 August, 2008.
- Coppola E, Giorgi F, Rauscher S, & Piani C (2010) Development of regional climate model weights based on the model’s “mesoscale signal”. Climate Res, 44, 121–134.

- Coppola E, Nogherotto R, Ciarlo` J, et al. (*submitted*). Assessment of the European climate projections as simulated by the large EURO-CORDEX regional climate model ensemble. Submitted to JGR
- Coppola E, Raffaele F, Giorgi F, et al. (*submitted*) Climate hazard indices projections based on CORDEX-CORE, CMIP5 and CMIP6 ensemble. Clim Dyn. Submitted to this issue.
- Daly, C., R.P. Neilson, and D.L. Phillips, 1994: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountanious Terrain. J. Appl. Meteor., 33, 140-158
- Di Luca A, de Elia R & Laprise R (2013) Potential for small scale added value of RCM's downscaled climate change signals. Clim Dyn 40, 14151433.
- Di Luca A, de Elía R & Laprise R (2015) Challenges in the Quest for Added Value of Regional Climate Dynamical Downscaling. Curr Clim Change Rep 1: 10. https://doi.org/10.1007/s40641-015-0003-9
- Di Luca, A., Argüeso, D., Evans, J. P., de Elía, R., and Laprise, R. (2016), Quantifying the overall added value of dynamical downscaling and the contribution from different spatial scales, *J. Geophys. Res. Atmos.*, 121, 1575– 1590, doi:10.1002/2015JD024009.
- Di Virgilio, G., Evans, J. P., Di Luca, A., Grose, M. R., Round, V., & Thatcher, M. (2020). Realised added value in dynamical downscaling of Australian climate change. Climate Dynamics, 54(11), 4675-4692. https://doi.org/10.1007/s00382-020-05250-1
- Fantini A, et al. (2018). Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-CORDEX and EURO-CORDEX experiments against high resolution observations. Clim Dyn 51: 877. https://doi.org/10.1007/s00382-016-3453-4
- Fantini A (2019) Climate change impact on flood hazard over Italy. PhD Thesis. University of Trieste. http://hdl.handle.net/11368/2940009
- Funk C, & Hoell A (2015) The leading mode of observed and CMIP5 ENSO-residual sea surface temperatures and associated changes in Indo-Pa-cific climate. J Climate, 28, 4309–4329, doi:10.1175 /JCLI-D-14-00334.1.
- Funk C, Peterson P, Landsfeld M et al. (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2, 150066 doi:10.1038/sdata.2015.66
- Giorgi F, Shields-Brodeur C, and Bates GT (1994) Regional climate change scenarios over the United States produced with a nested regional climate model. J Clim, 7, 375–399.
- Giorgi F, Jones C & Asrar GR (2009) Addressing climate information needs at the regional level: the CORDEX framework. WMO Bulletin, 58, 3, 175 – 183.
- Giorgi F, et al. (2012) RegCM4: model description and preliminary tests over multiple CORDEX domains. Climate Research. 52, 7-29.
- Giorgi F, Torma C, Coppola E, Ban N, Schär C, Somot S (2016) Enhanced summer convective rainfall at Alpine high elevations in response to climate warming. Nat Geosc 9, 584-589.
- Giorgi F (2019) Thirty years of regional climate modeling. Where are we and where are we going next? J Geophys Res 124, 5606-5723.
- Gutowski WJ, 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, & Tangang F (2016) WCRP COordinated Regional Downscaling EXperiment (CORDEX): a

diagnostic MIP for CMIP6, Geoscientific Model Development, Vol. 9, Issue 11, pp. 4087-4095, doi:10.5194/gmd-9-4087-2016

- Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, & New M (2008) A European daily high resolution gridded data set of surface temperature and precipitation for 1950–2006. J Geophys Res, 113, D20119, doi:10.1029/2008JD010201
- Herrera S et al. (2015) Update of the Spain02 gridded observational dataset for Euro-CORDEX evaluation: assessing the effect of the interpolation methodology. International Journal of Climatology, doi:19.1002/joc.4391
- Hong SY, Kanamitsu M (2014) Dynamical downscaling: fundamental issues from an NWP point of view and recommendations. Asia-Pacific J Atmos Sci 50:83–104.
- Isotta et al. (2014) The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. Int J Climatol, 34 (5), 1657-1675.
- Jacob D, Elizalde A, Haensler A, Hagemann S, Kumar P, Podzun R, Rechid D, Remedio AR, Saeed F, Sieck K et al. (2012) Assessing the Transferability of the Regional Climate Model REMO to Different COordinated Regional Climate Downscaling EXperiment (CORDEX) Regions. Atmosphere, 3, 181–199, doi:10.3390/atmos3010181.
- Jacob D, Petersen J, Eggert B, Alias A, Christensen JH et al. (2013) EURO-CORDEX: new high resolution climate change projections for European impact research. Reg Environ Change 14:563–578
- Jacob D, Petersen J, Eggert B et al. (2014) EURO-CORDEX: New high-resolution climate change projections for European impact research. Regional Environmental Change, 14, 563-578.
- Jacob D, Teichmann C, Sobolowski S, et al. (2020) Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community. Regional Environmental Change, 20, 51.
- Johansson B (2002) Estimation of areal precipitation for hydrological modelling. PhD Thesis. Earth Sciences Centre, Goteborg University, Report nr. A76.
- Jones C, Giorgi F & Asrar G (2011) The Coordinated Regional Downscaling Experiment: CORDEX An international downscaling link to CMIP5. CLIVAR Exchanges No.56, Vol. 16, 34-40
- Kanamitsu M & DeHaan L (2011) The Added Value Index: A new metric to quantify the added value of regional models. J Geophys Res 116 (D11106): 1-10.
- Kanamitsu M & Kanamaru H (2007) Fifty- seven- year California Reanalysis Downscaling at 10 km (CaRD10). Part I: System detail and validation with observations. J Clim, 20, 5553–5571, doi:10.1175/2007JCLI1482.1.
- Kjellström E, Boberg F, Castro M, Christensen JH, Nikulin G, & Sanchez E (2010) On the use of daily and monthly temperature and precipitation statistics as a performance indicator for regional climate models. Climate Research, 44(2-3), 135-150. Doi: 10.3354/cr00932.
- Kotlarski S, Keuler K, Christensen OB, Colette A, Déqué M, et al. (2014) Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. Geoscientific Model Development, 7, 1297–1333,https://doi.org/10.5194/gmd-7-1297-2014, 2
- Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22:79–86

- Kummerow C, Barnes W, Kozu T, Shiue J, Simpson J (1998) The tropical rainfall measuring mission (TRMM) sensor package. J Atmos Ocean Technol 15:809–817. https://doi.org/10.1175/1520-0426(1998)015%3c0809:TTRMMT%3e2.0.CO;2
- Laprise R (2014) "Comment on 'The added value to global model projections of climate change by dynamical downscaling: A case study over the continental US using the GISS- ModelE2 and WRF models' by Racherla et al." Journal of Geophysical Research: Atmospheres 119(7):3877-3881. doi:10.1002/ 2013JD019945.2012.
- Lenderink G (2010) Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. CR 44:151-166 doi: 10.3354/cr00946
- Lucas-Picher P, Laprise R & Winger K (2017) Evidence of added value in North American regional climate model hindcast simulations using ever-increasing horizontal resolutions. Climate Dynamics, 48(7-8), pp.2611-2633.
- Mearns, L.O., et al., 2017: The NA-CORDEX dataset, version 1.0. NCAR Climate Data Gateway, Boulder CO. https://doi.org/10.5065/D6SJ1JCH
- Menne MJ, Durre I, Vose RS, Gleason BE, & Houston TG (2012) An overview of the Global Historical Climatology Network-Daily Database. Journal of Atmospheric and Oceanic Technology, 29, 897-910, doi.10.1175/JTECH-D-11-00103.1
- Mohr M (2009) Comparisons of versions 1.1 and 1.0 of gridded temperature and precipitation data for Norway. Technical report Met. No note 19.
- Pai DS, Sridhar L, Rajeevan M, et al (2014) Development of a new high spatial resolution (0.25°×0.25°) Long Period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. Mausam 65:1–18
- Perkins SE, Pitman AJ, Holbrook NJ, & McAneney J (2007) Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. J Climate, 20, 4356–4376, https://doi.org/10.1175/JCLI4253.1
- Perry M et al. (2009) The generation of daily gridded datasets of temperature and rainfall for the UK. Met Office Climate Memorandum No. 24.
- Prein, A.F., Gobiet, A., Truhetz, H. et al. Precipitation in the EURO-CORDEX 0.11◦ and 0.44◦ simulations: high resolution, high benefits?. Clim Dyn 46, 383–412 (2016). https://doi.org/10.1007/s00382-015-2589-y
- Qiu L, Im ES, Hur J, Shim KM (2019) Added value of very high resolution climate simulations over South Korea using WRF modeling system. Climate Dynamics, In press, https://doi.org/10.1007/s00382-019-04992-x.
- Rauthe M et al. (2013) A Central European Precipitation Climatology – part I: generation and validation of a high resolution gridded daily data set (HYRAS). Meteorologische Zeitschrift, 22: 235-256.
- Remedio AR, Teichmann C, Buntemeyer L, et al. (2019) Evaluation of New CORDEX Simulations Using an Updated Köppen–Trewartha Climate Classification. Atmosphere, 10, 726, https://doi.org/10.3390/atmos10110726.
- Richard Moss, Mustafa Babiker, Sander Brinkman, et al. (2008) Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies. Intergovernmental Panel on Climate Change, Geneva, 132 pp.

- Roberts NM, Lean HW (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Mon Weather Rev 136(1):78–97.
- Rummukainen M (2016) Added value in regional climate modelling. Climatic Change. 7: 145-159.
- Sanchez-Gomez E, Somot S, Déqué M (2009) Ability of an ensemble of regional climate models to reproduce the weather regimes during the period 1961-2000. Clim Dyn, 33(5):723-736, doi:10.1007/s00382-008-0502-7
- Soares PMM, Cardoso RM (2018) A simple method to assess the added value using high-resolution climate distributions: application to the EURO-CORDEX daily precipitation. Int J Climatol 38: 1484–1498
- Szalai S et al. (2013) Climate of the greater Carpathian region. Final Technical Report. www.carpatclim-eu.org
- Teichmann C, Jacob D, Teichmann C, Remedio AR, Buelow K, Remke T, et al. (*submitted*) Assessing mean climate change signals in the global CORDEX-CORE ensemble. Clim Dyn. Submitted to this issue.
- Torma C, et al. (2015) Added value of regional climate modelling over areas characterized by complex terrain – Precipitation over the Alps. J Geophys Res Atmos 3957-3972.
- Vidal JP, et al. (2010) A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30: 1627-1644.
- Xue Y, Janjic Z, Dudhia J, Vasic R, De Sales F (2014) A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. Atmos Res 147–148:68–85. doi:10.1016/j.atmosres. 2014.05.001, http://linkinghub.elsevier.com/retrieve/pii/S0169809 514002.
- Yatagai AO, Arakawa K, Kamiguchi H, et al. (2009) A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges. SOLA, 5, 137-140, DOI:10.2151/sola.2009-035.

Appendix

The method described in Section 2 is based on the difference between two PDFs, therefore the selection of the PDF bin-size is a very important process. Since a smaller bin-size is analogous to a higher horizontal resolution, it should allow a better representation of the details of the PDF. However, the effect of varying the bin-size on this new added value method requires some testing.

One example simulation for all RCMs used in this study (each driven by a different GCM and compared to the EURO4M data-set) was used to assess the dependence of the added value on the bin-size (Figures A1 and A2). The results show a decrease in magnitude of the added value as the bin-size increases. This happens as a result of aggregating a larger number of events and thus smearing out the details of the distributions. The sign of the added value changes in some cases, but in these cases the magnitude of the added value is very low. As a result of this test, in order to obtain the best possible resolution of the PDFs and the most informative outcome from this new method, the bin-size of 1 mm/day is used.

Supplementary Information

The EURO-CORDEX added value analysis presented in this study was performed at resolutions of 0.11° and 1.00°. The ensemble means are based on results conducted on each RCM member (and their corresponding driving GCM); these results are presented in Figures S1 to S6. The variability of added value with percentile fraction, and the PDFs of ECO regions at low resolutions can be seen in Figures S7 and S8 respectively.

The CORDEX-CORE added value ensemble mean of the complete precipitation PDFs is shown in Figure S9, while the percentile fraction analysis of the added value for each observation source is displayed in Figures S10 to S13.

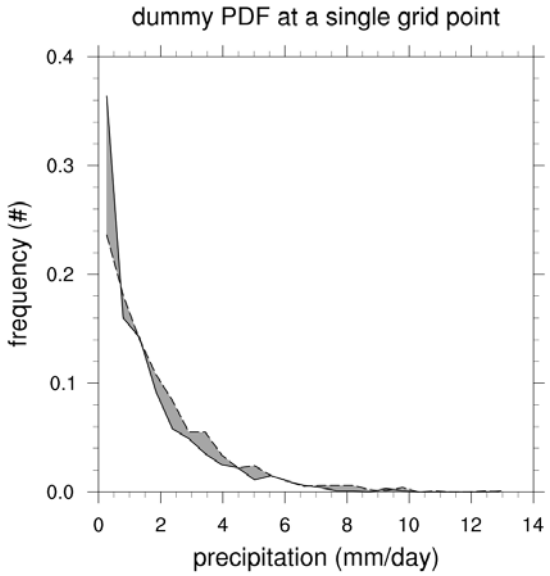

Fig. 1. An illustrative plot of the precipitation distribution of a single grid point. The lines describe the distribution of a hypothetical model and an observation data-set. The shaded area represents the Sum of the Relative Probability Difference between the model and observations (D_M).

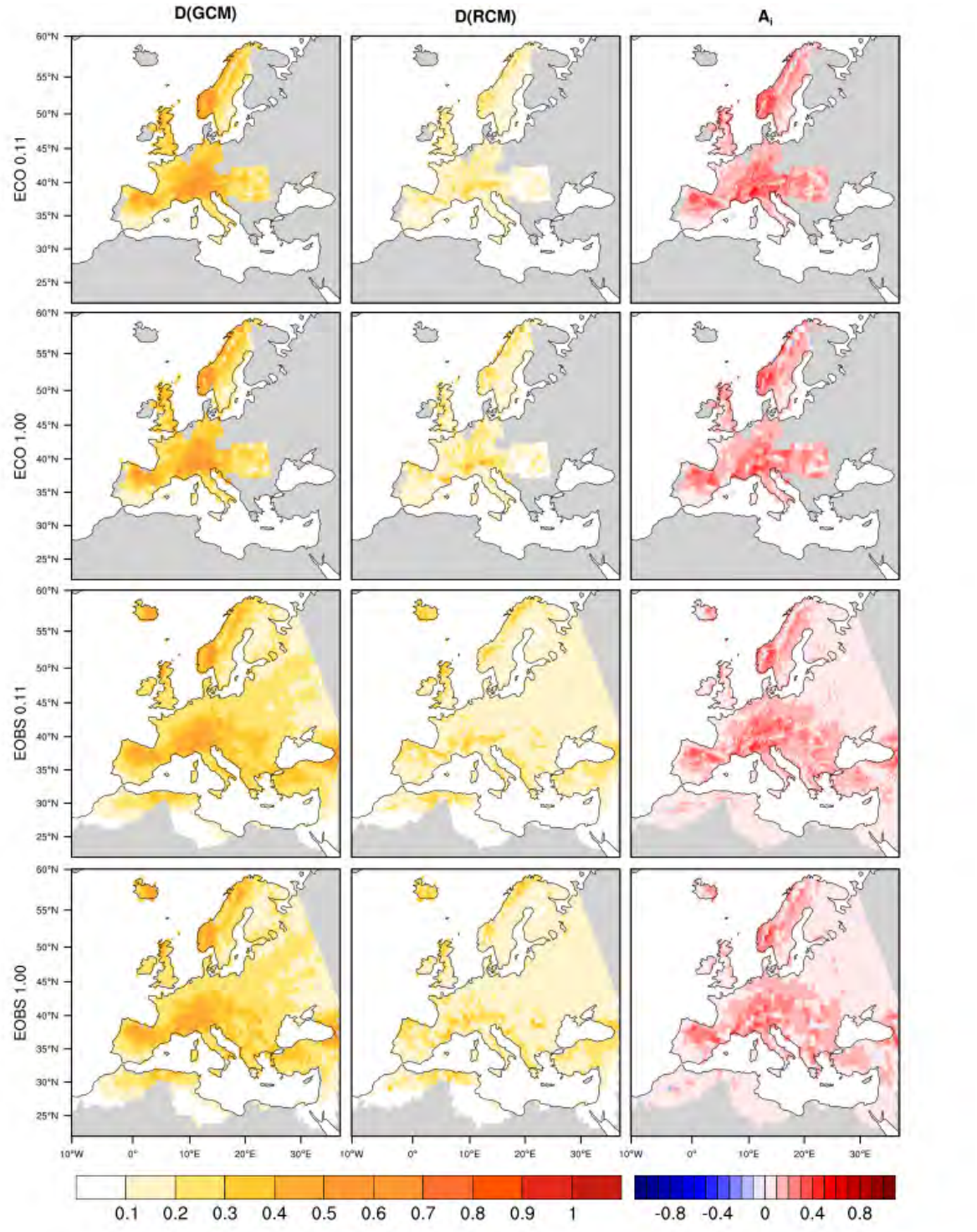

Fig. 2. Relative probability difference (D) for the GCM (left) and RCM (mid) ensemble, and added value (A_i) for the RCM ensemble (right) compared to ECO at 0.11° (top) and 1.00° (mid-top) and EOBS at 0.11° (mid-bottom) and 1.00° (bottom).

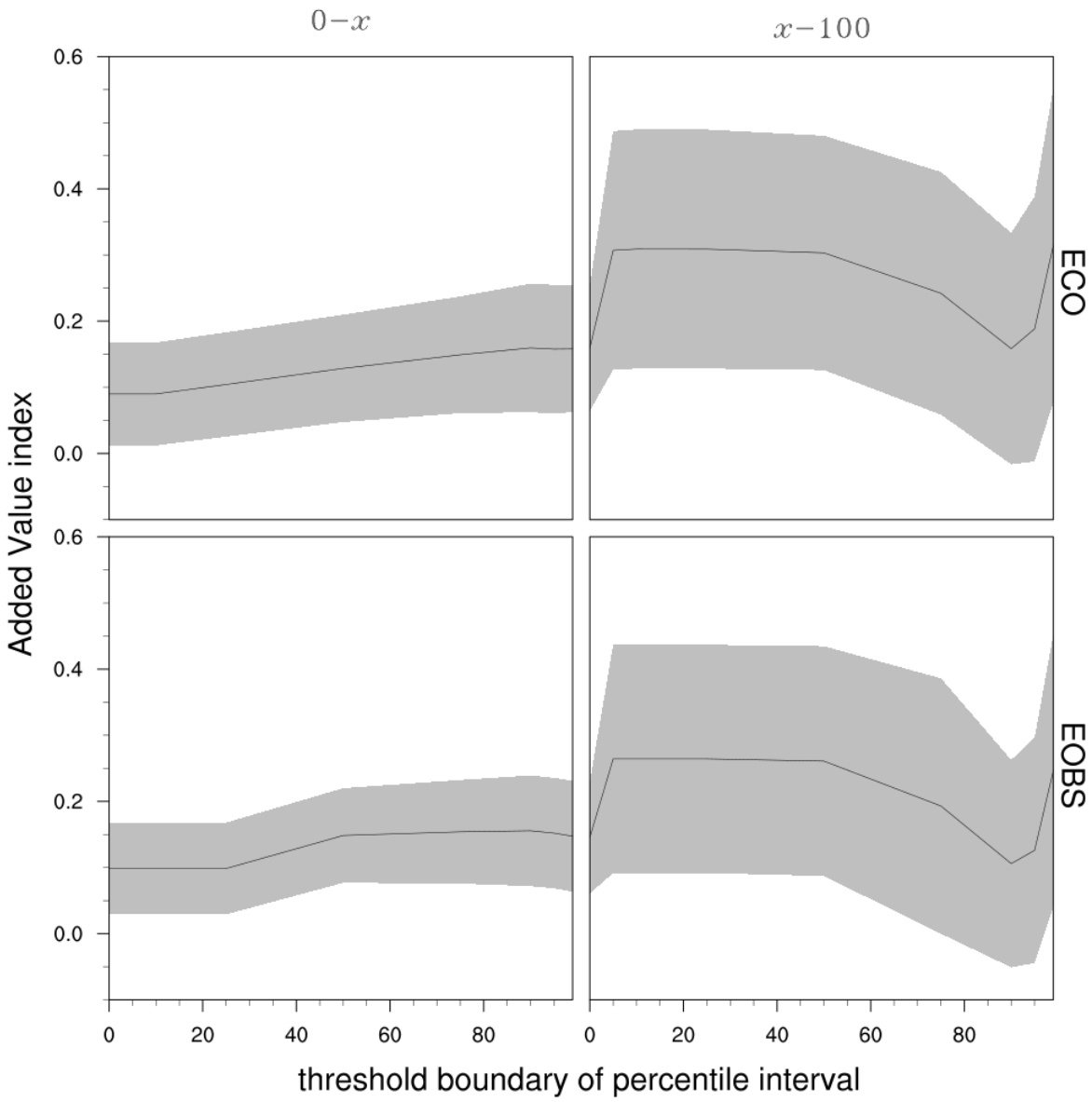

Fig. 3. The variability of spatial mean added value index at different percentile intervals compared to ECO (top) and EOBS (bottom) at 0.11°. The EOBS data in this figure has been masked to match the locations of ECO. Each point *x* describes the added value of the percentile fraction '0-*x*' (left), and '*x*-100' (right). The shaded area shows the standard deviation of the data.

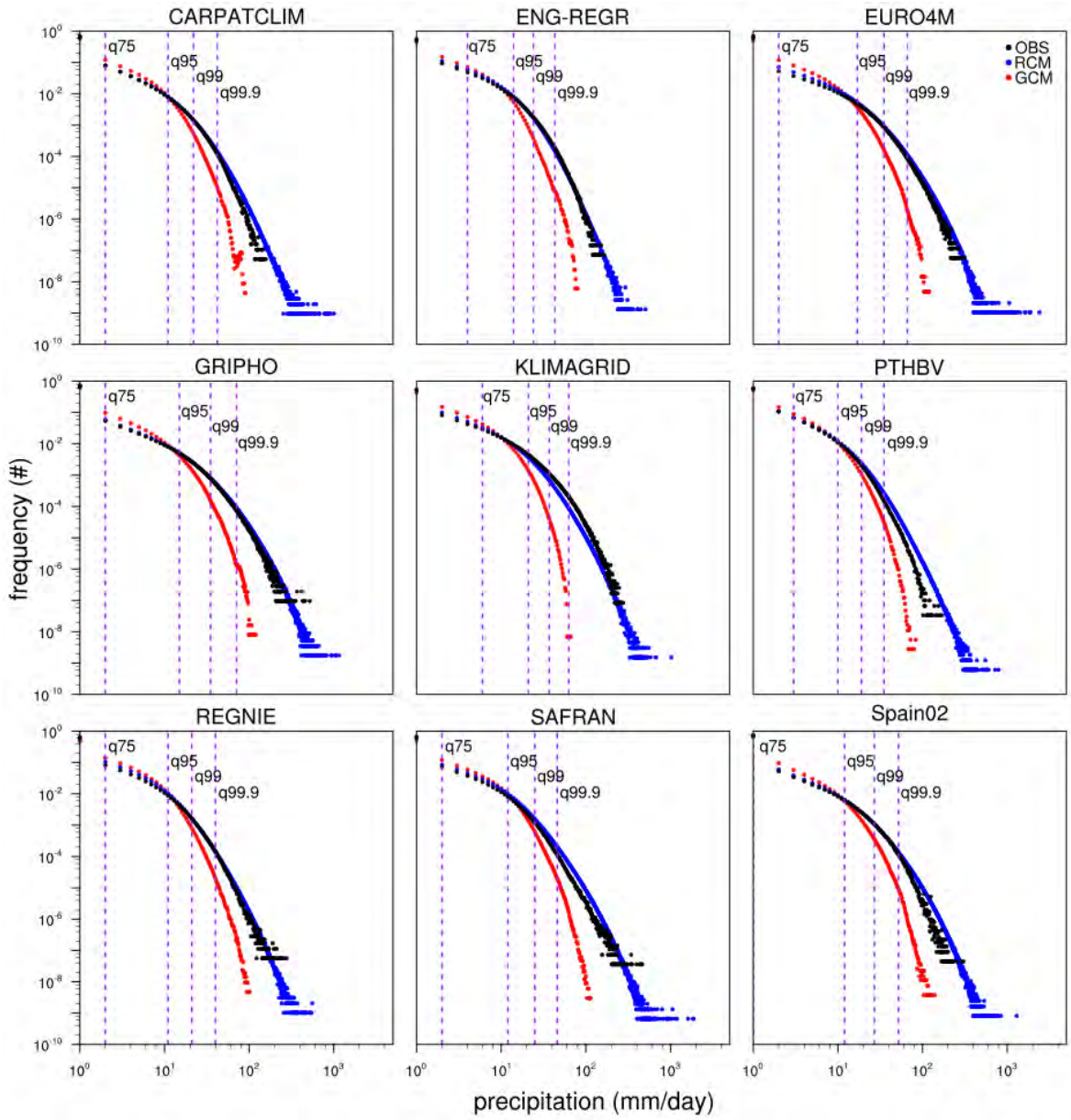

Fig. 4. PDFs of the RCM and GCM ensemble member data compared to all 9 regional observations at 0.11°. Each PDF includes a marker for the 75th, 95th, 99th and 99.9th percentiles.

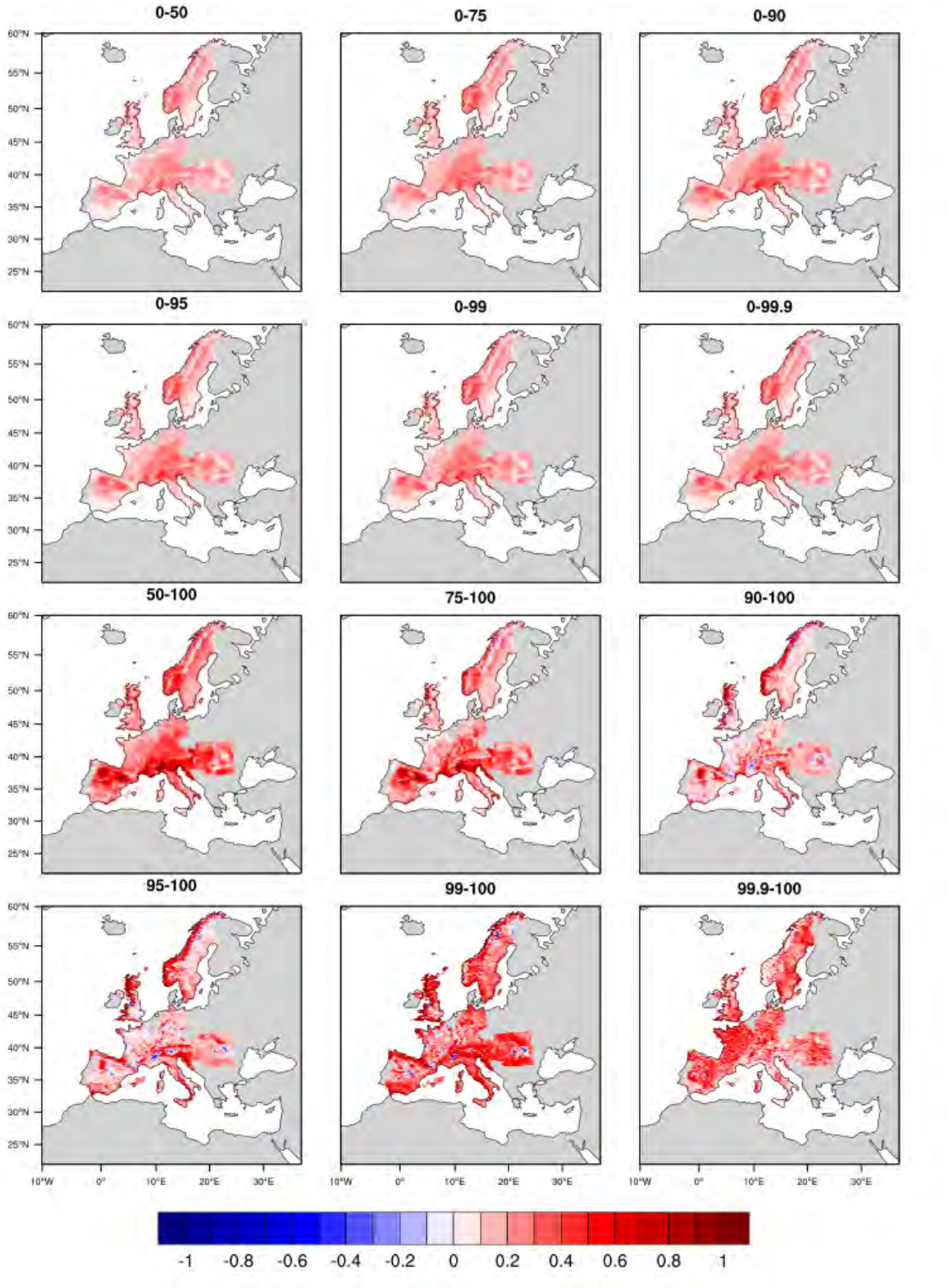

Fig. 5. Added value for RCM ensemble-mean at different percentile intervals compared to ECO at 0.11°.

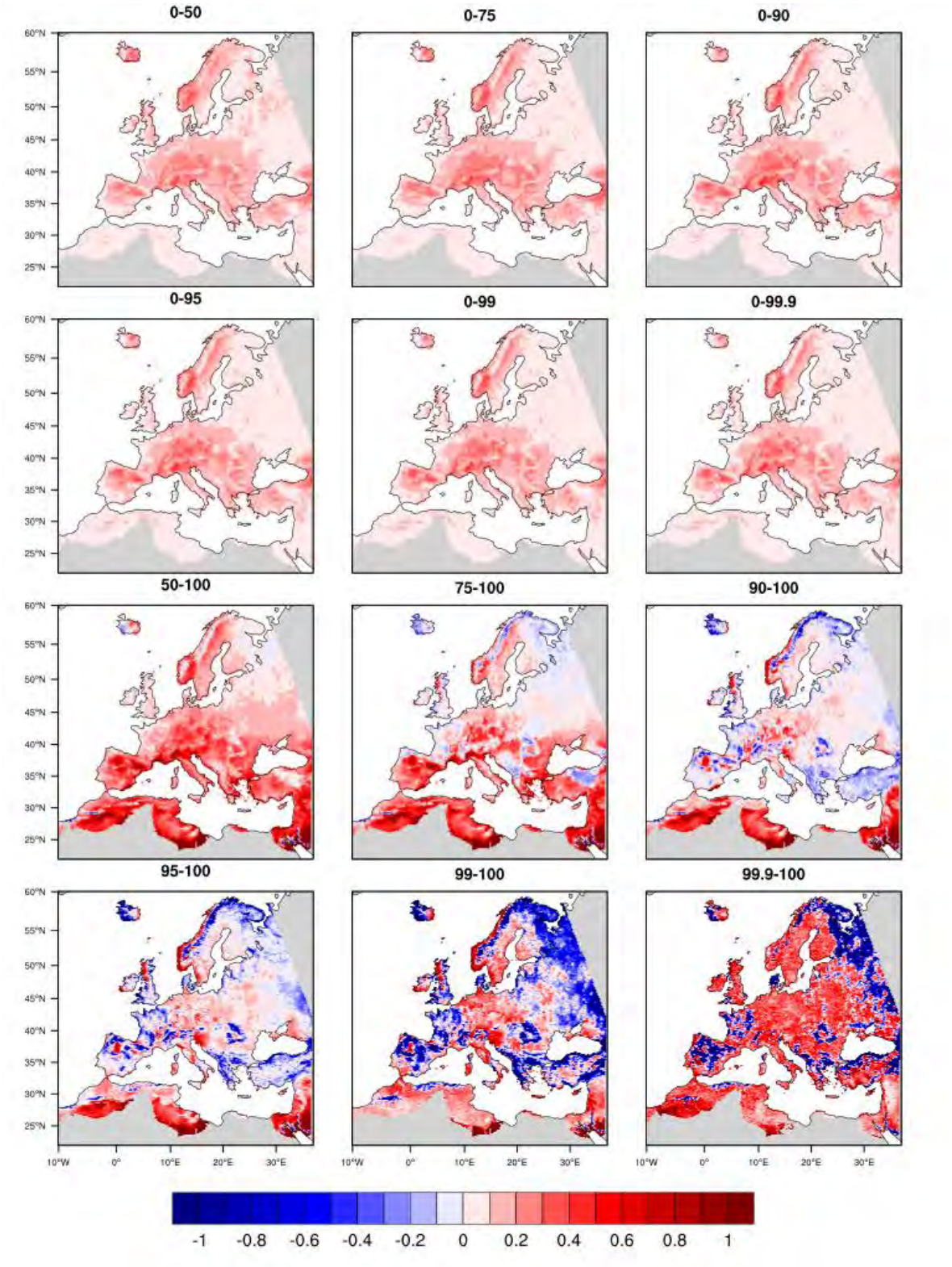

Fig. 6. Added value for RCM ensemble mean at different percentile intervals compared to EOBS at 0.11°.

Fig. 7. Added value for 99-100 percentile fraction of the EUR-11 ensemble members compared to the ECO at a resolution of 0.11°.

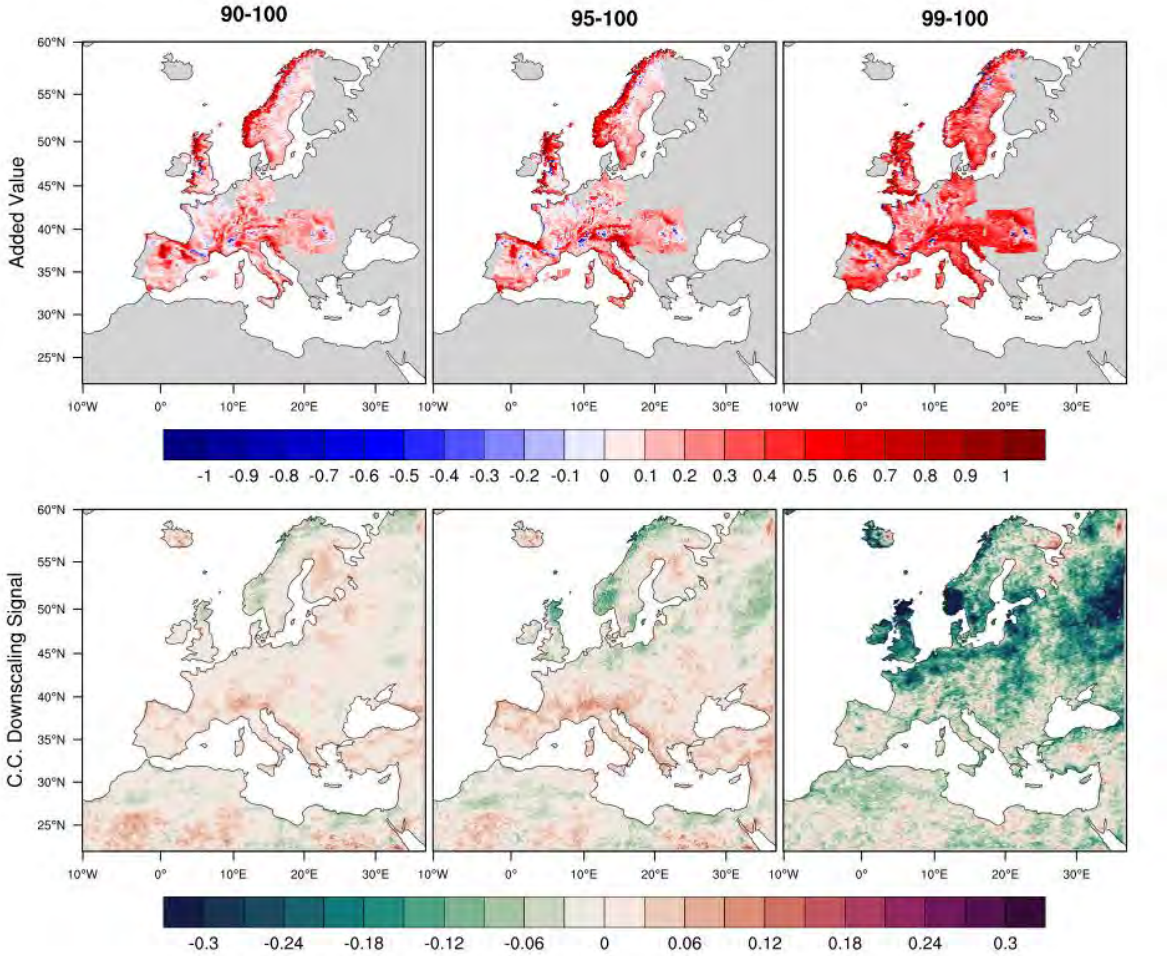

Fig. 8. Added value (top) and Climate Change Downscaling Signal (bottom) of the EURO-CORDEX ensemble for the RCP 8.5 far future at different percentile intervals and at a resolution of 0.11°.

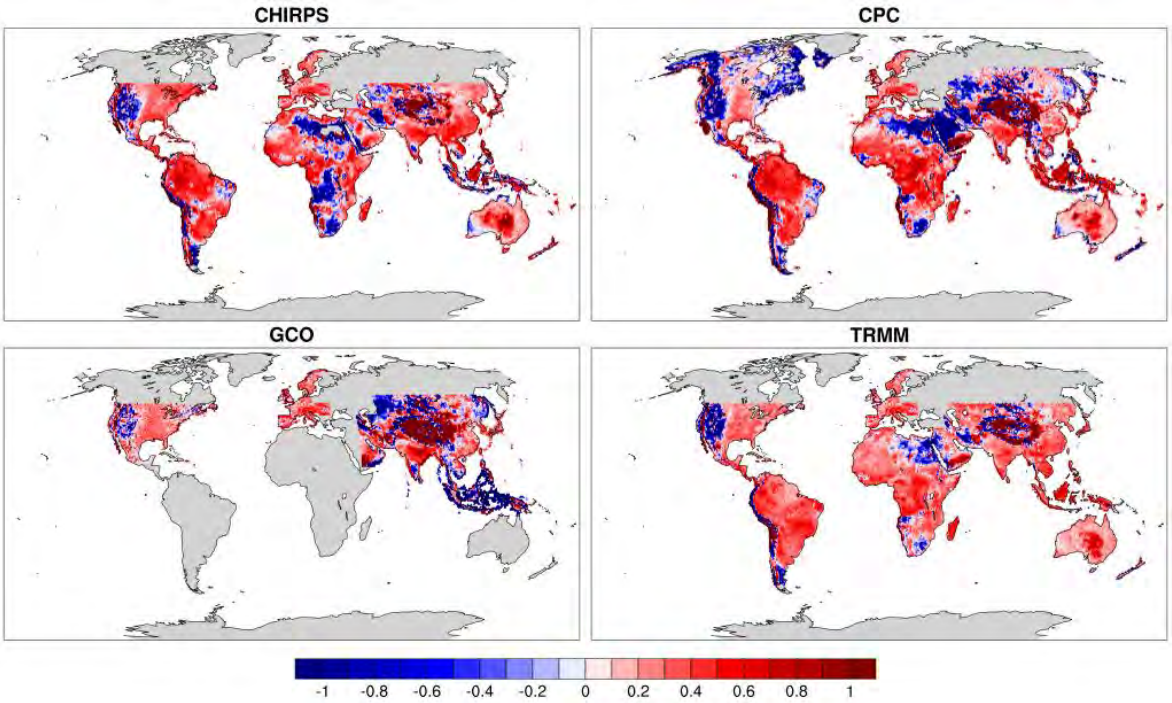

Fig. 9. Added value for the 99-100 percentile fraction of precipitation for the CORDEX-CORE ensemble members compared to CHIRPS, CPC, GCO (Regionals), and TRMM, at 0.22°. The Europe data used is the added value compared to the ECO as in Figure 2.

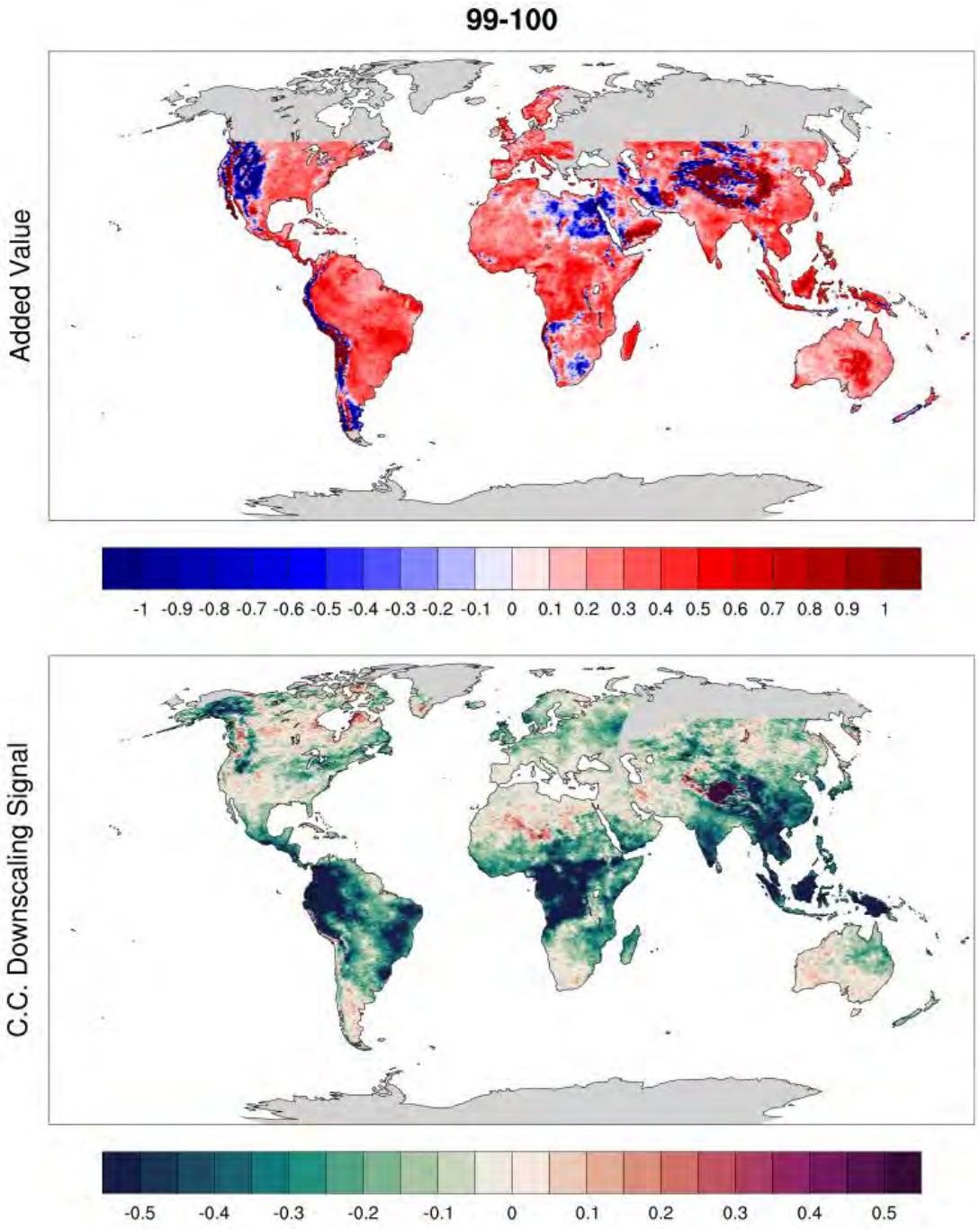

Fig. 10. RCM ensemble means of the CORDEX-CORE at 99-100 percentile intervals at 0.22°. (top) Added value compared to TRMM, and (bottom) climate change downscaling signal for the RCP 8.5 far future.

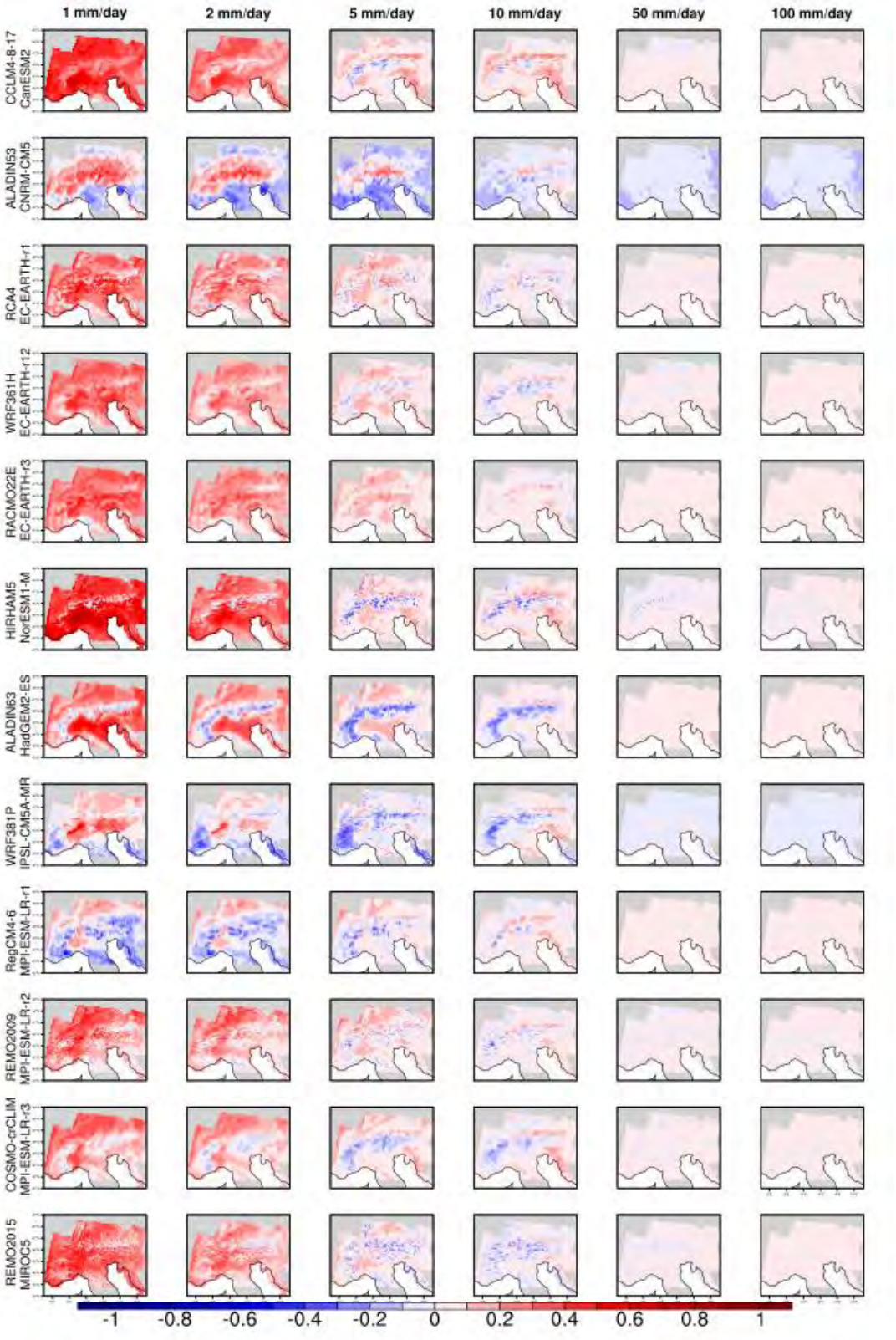

Figure A1. Added value for some model examples at different PDF (with full distribution) bin sizes compared to EURO4M at 0.11°

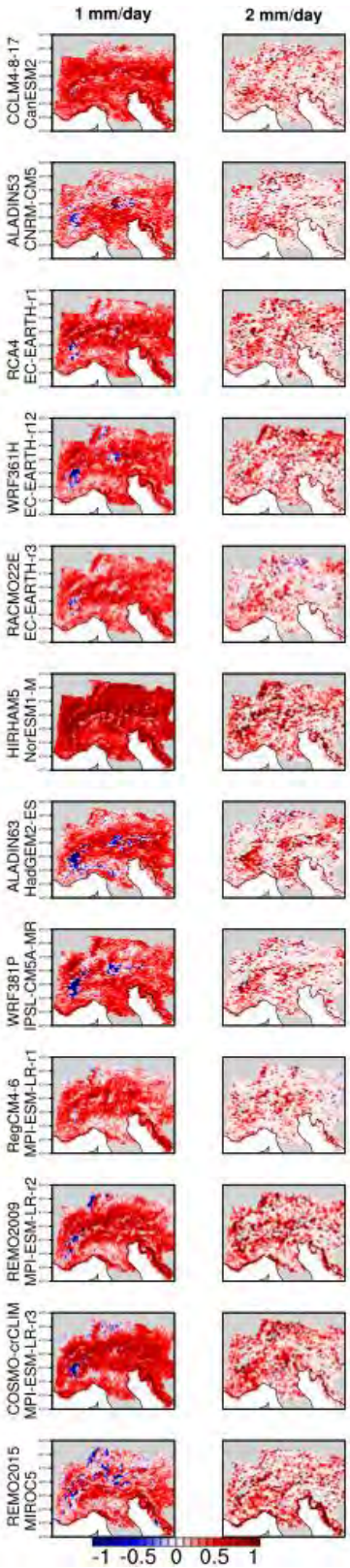

Figure A2. Added value of 99-100 percentile interval for some model examples at different PDF bin sizes compared to EURO4M at 0.11°.

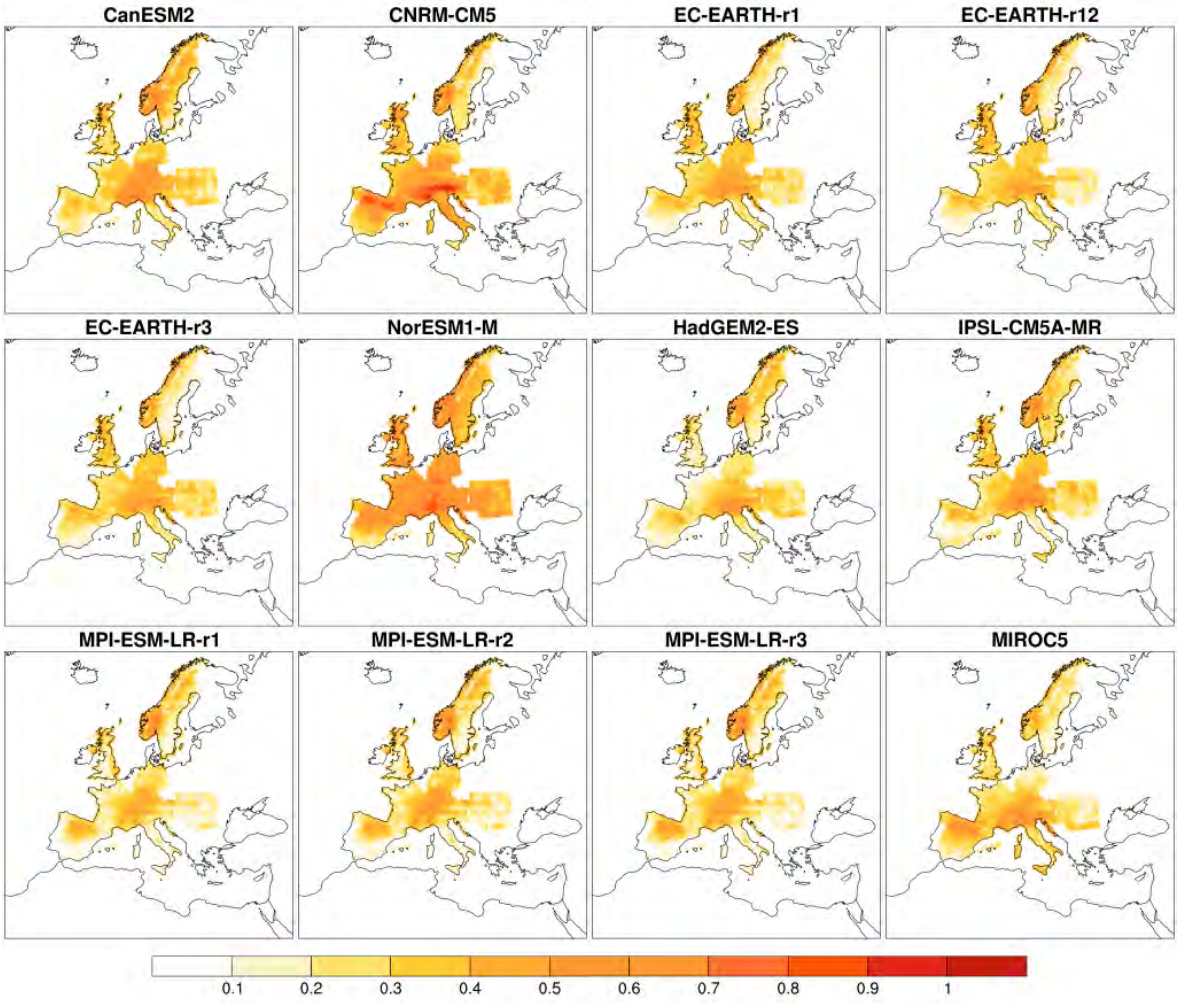

Fig. S1. Relative probability difference for each driving GCM ensemble member compared to ECO at a resolution of 0.11°.

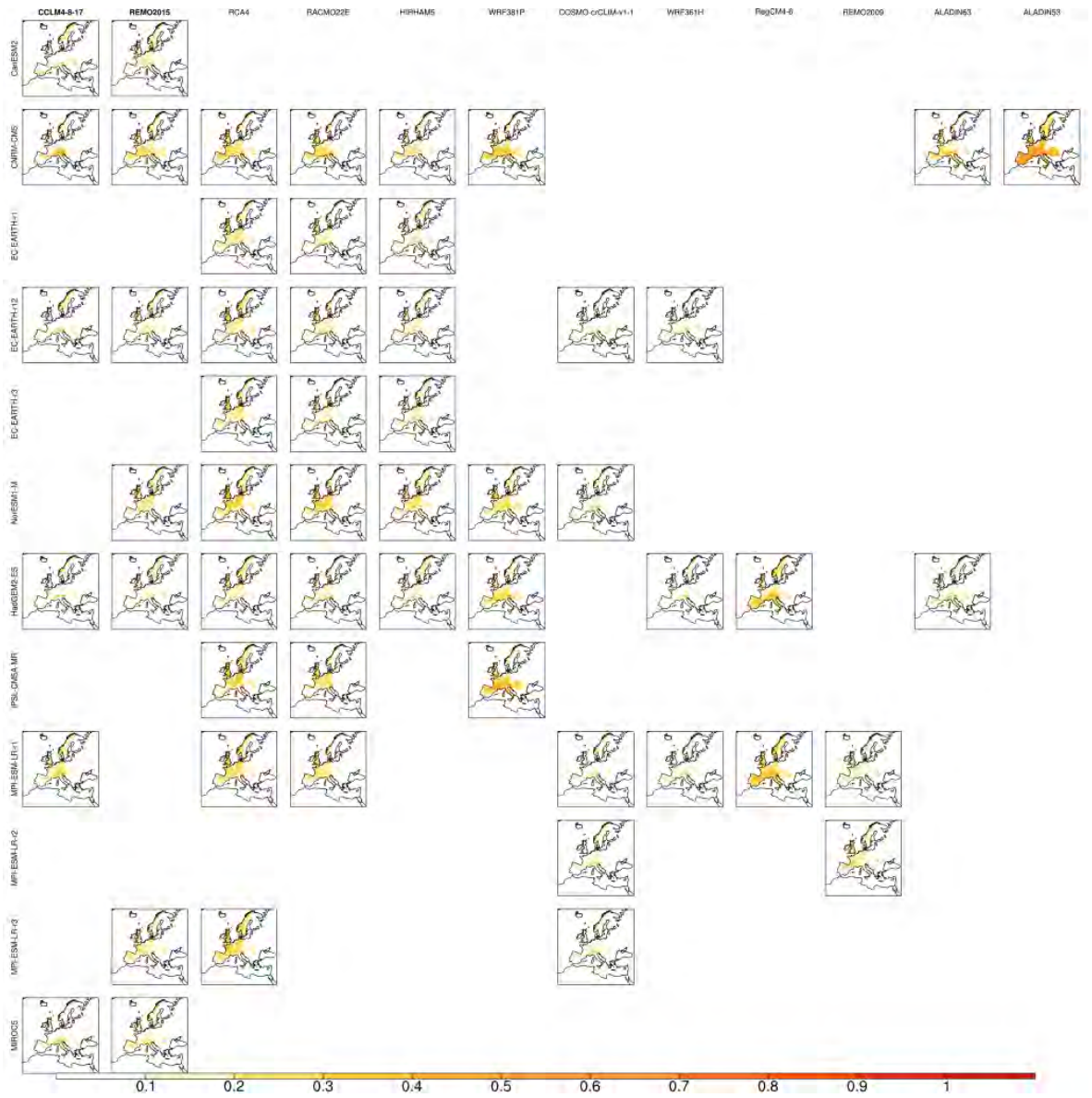

Fig. S2. Relative probability difference for each RCM ensemble member compared to ECO at a resolution of 0.11°.

Fig. S3: Added value for each RCM ensemble member compared to ECO at a resolution of 0.11°.

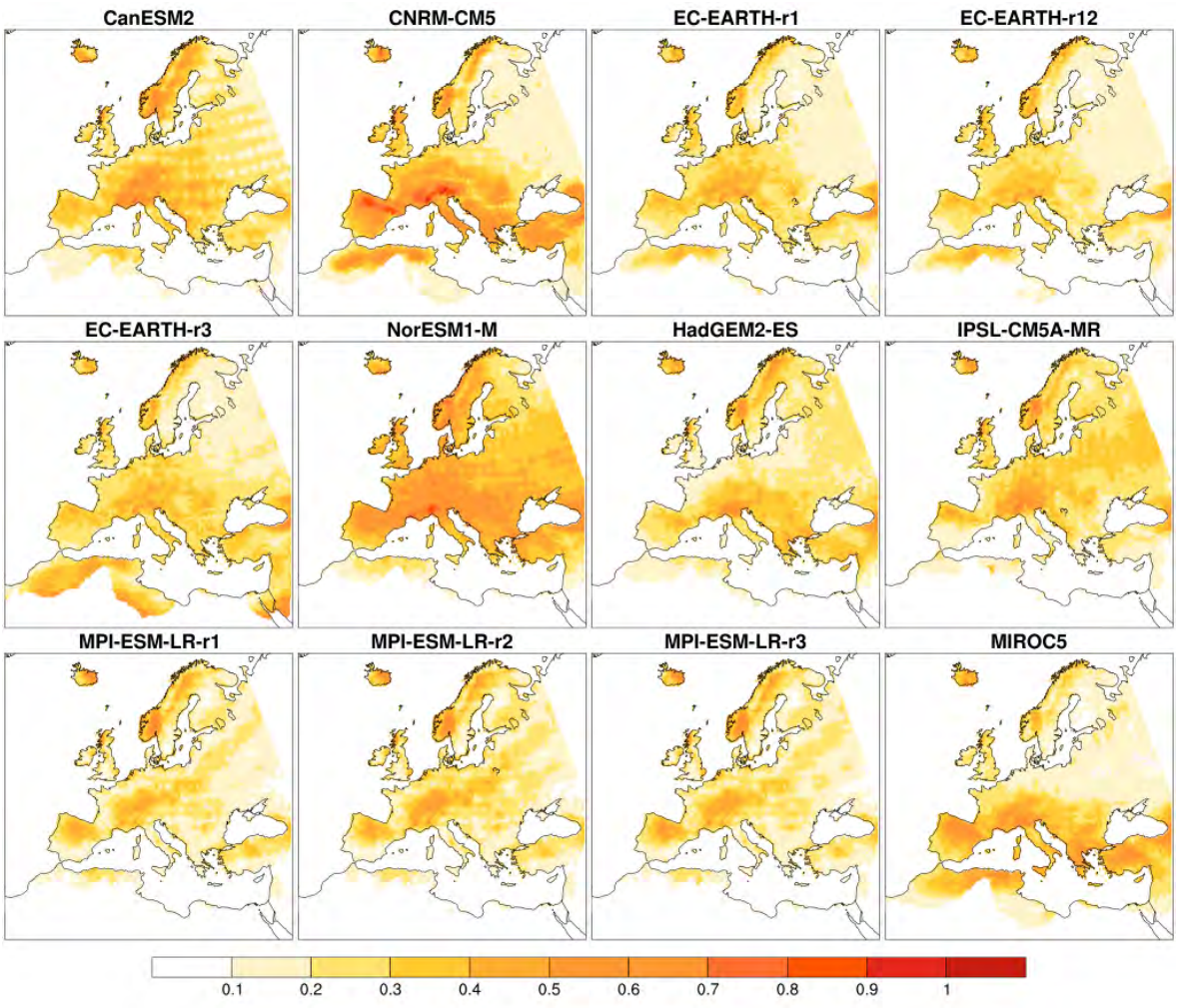

Fig. S4. Relative probability difference for each driving GCM ensemble member compared to EOBS at a resolution of 0.11°.

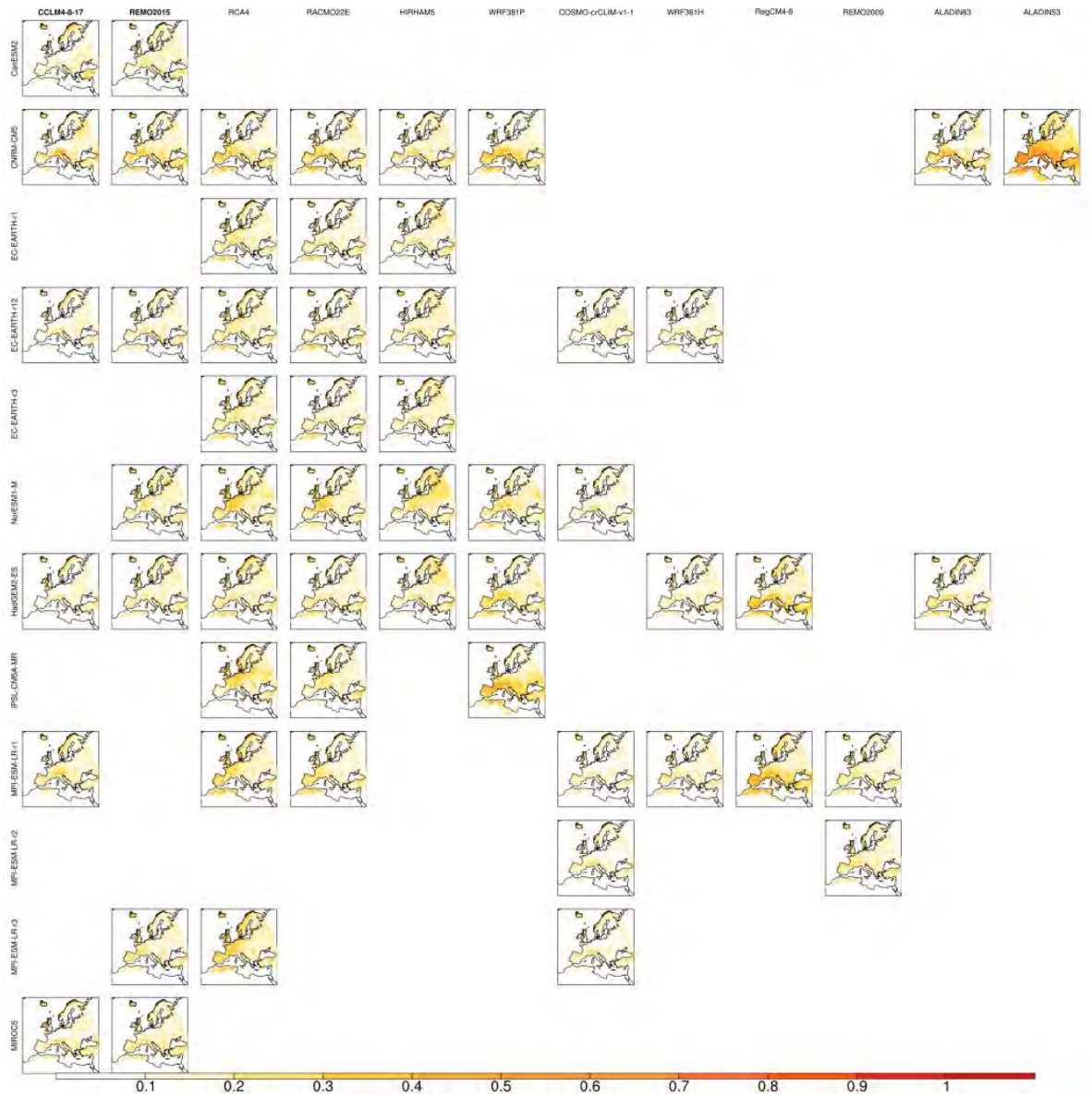

Fig. S5. Relative probability difference for each RCM ensemble member compared to EOBS at a resolution of 0.11°.

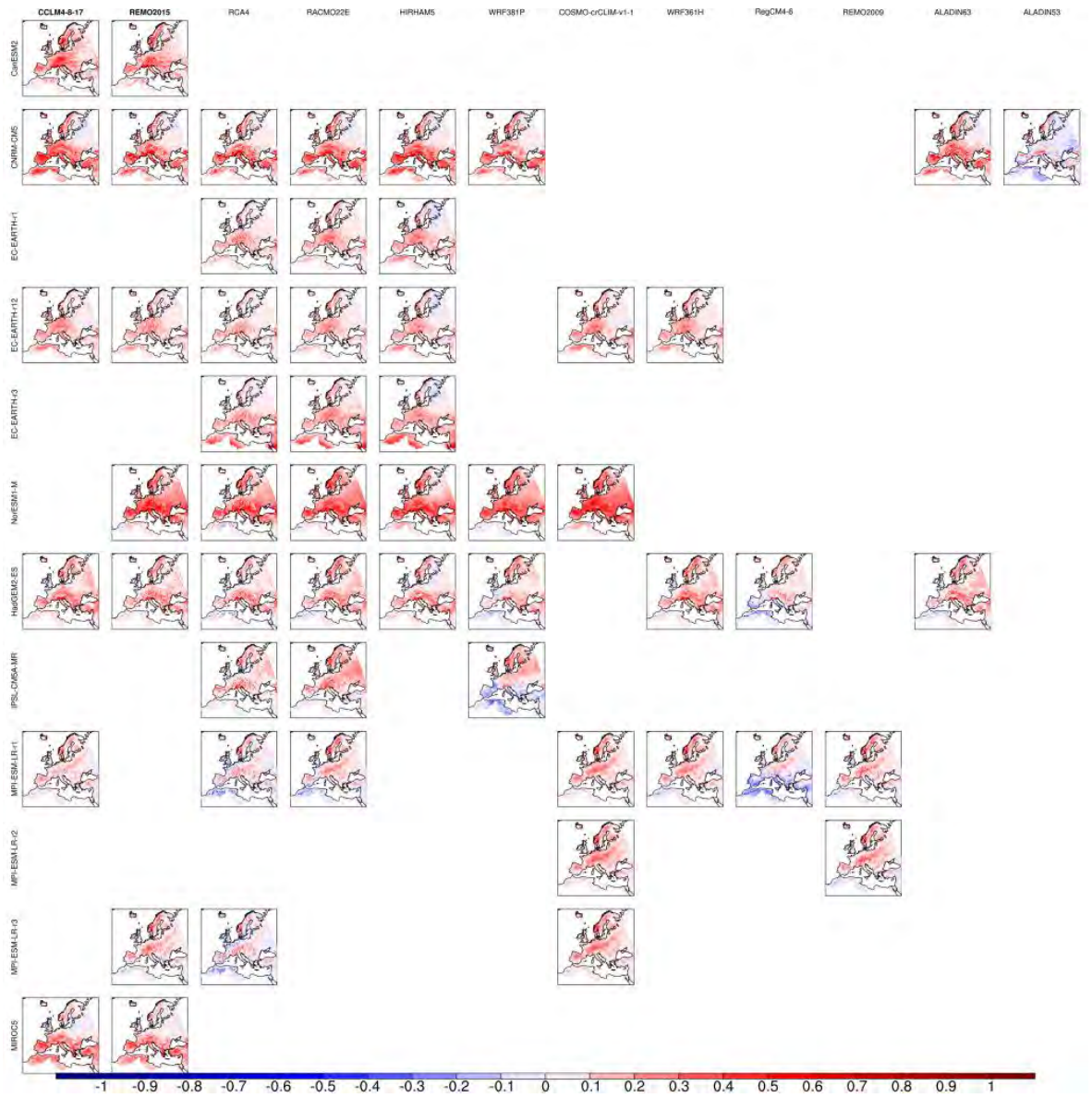

Fig. S6: Added value for each RCM ensemble member compared to EOBS at a resolution of 0.11°.

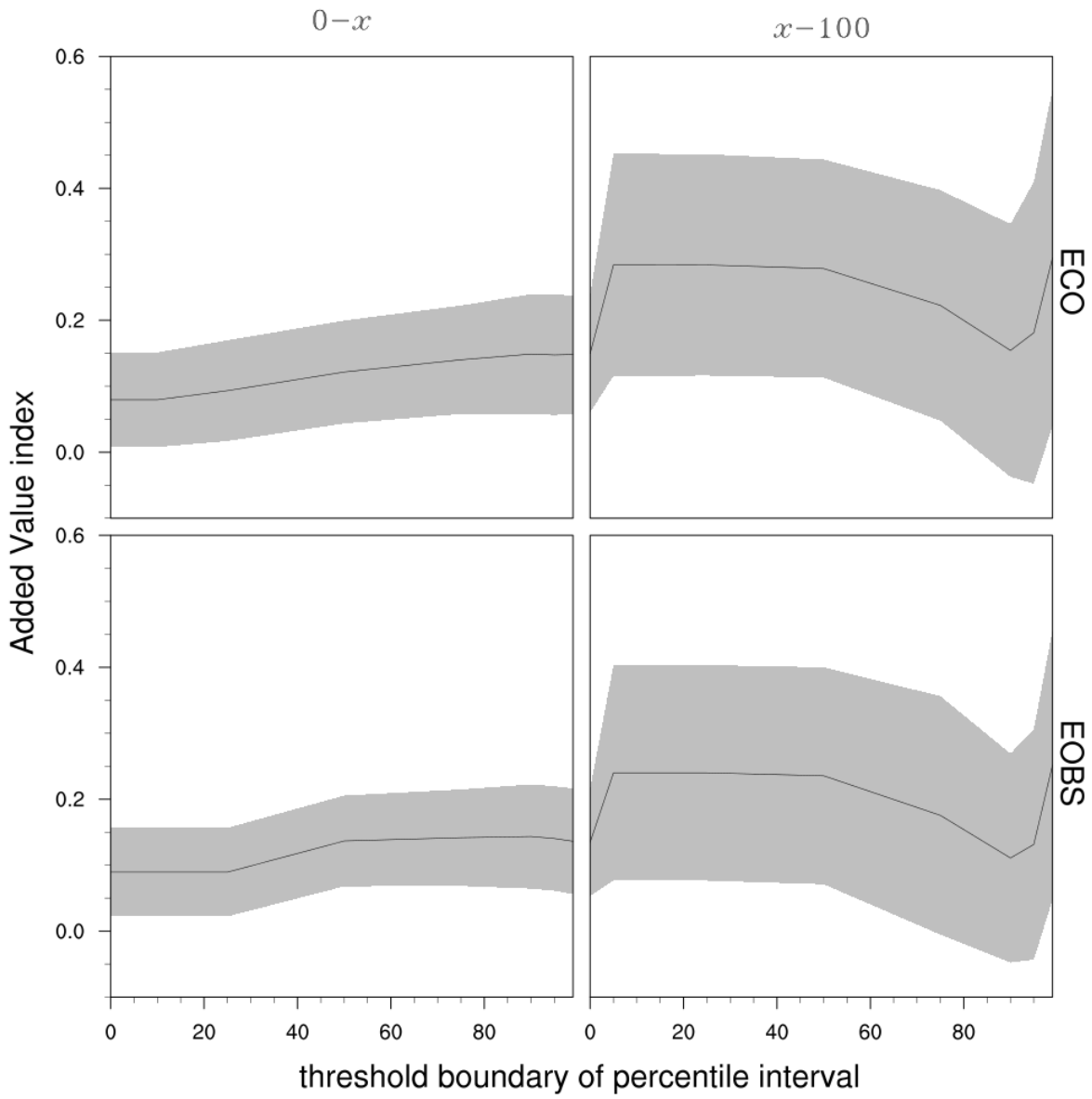

Fig. S7. The variability of spatial mean added value index at different percentile intervals compared to ECO (top) and EOBS (bottom) at 1.00°. The EOBS data in this figure has been masked to match the locations of ECO. Each point *x* describes the added value of the percentile fraction '0-*x*' (left) and '*x*-100' (right). The shaded area shows the standard deviation of the data.

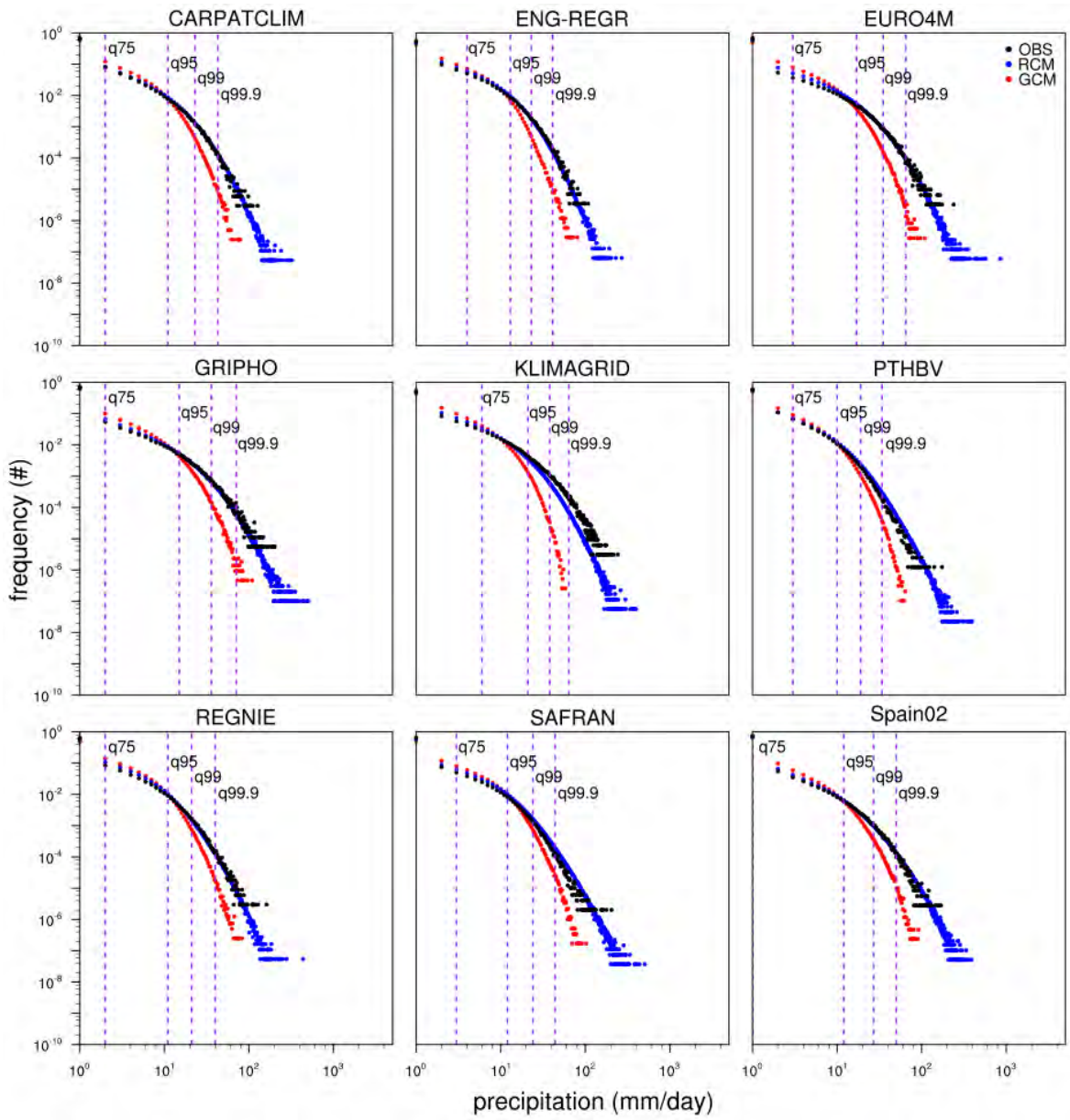

Fig. S8. PDFs of the RCM and GCM ensemble member data compared to all 9 regional observations at 1.00°. Each PDF includes a marker for the 75th, 95th, 99th and 99.9th percentiles.

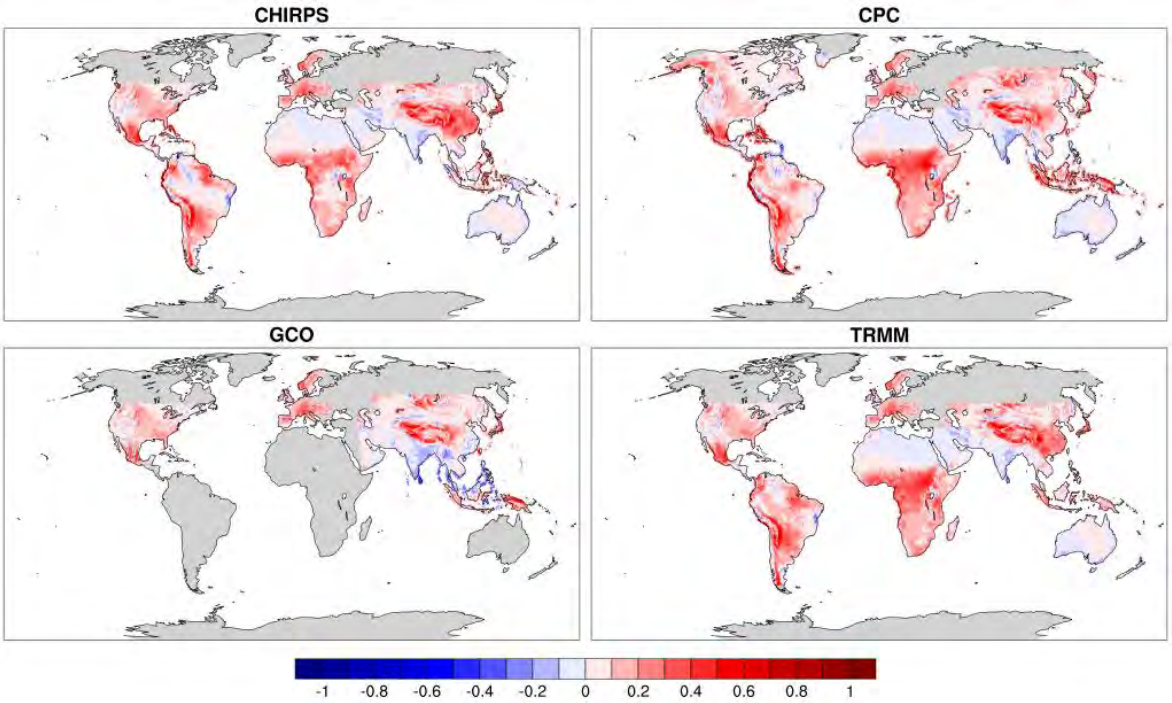

Fig. S9. Added value for precipitation of the CORDEX-CORE ensemble members compared to CHIRPS, CPC, GCO, and TRMM at 0.22°. The Europe data used is the added value compared to ECO as in Figure 2.

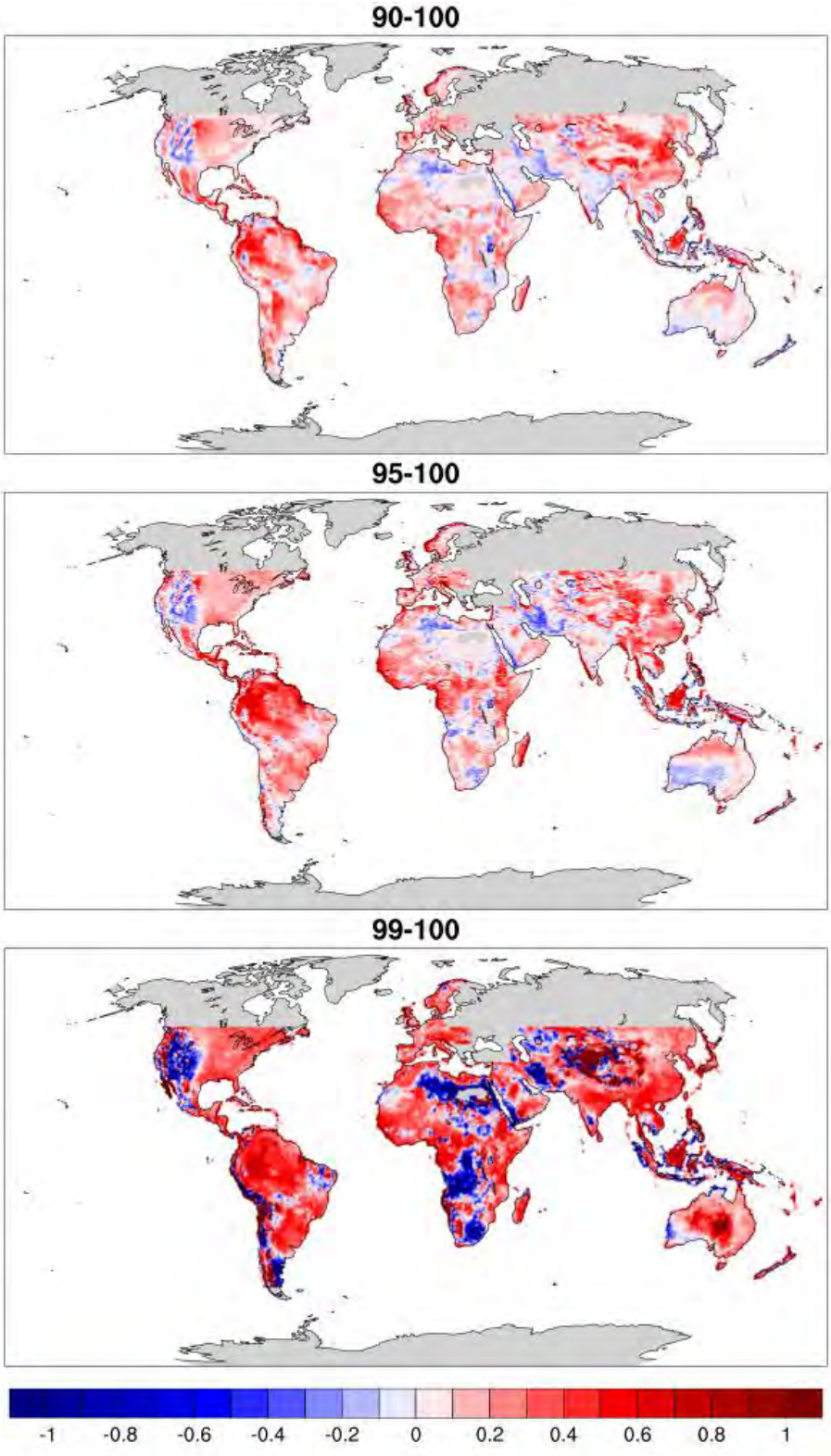

Fig. S10. Added value for RCM ensemble means of the CORDEX-CORE at different percentile intervals compared to CHIRPS at 0.22°. The Europe data used is as Figure 5.

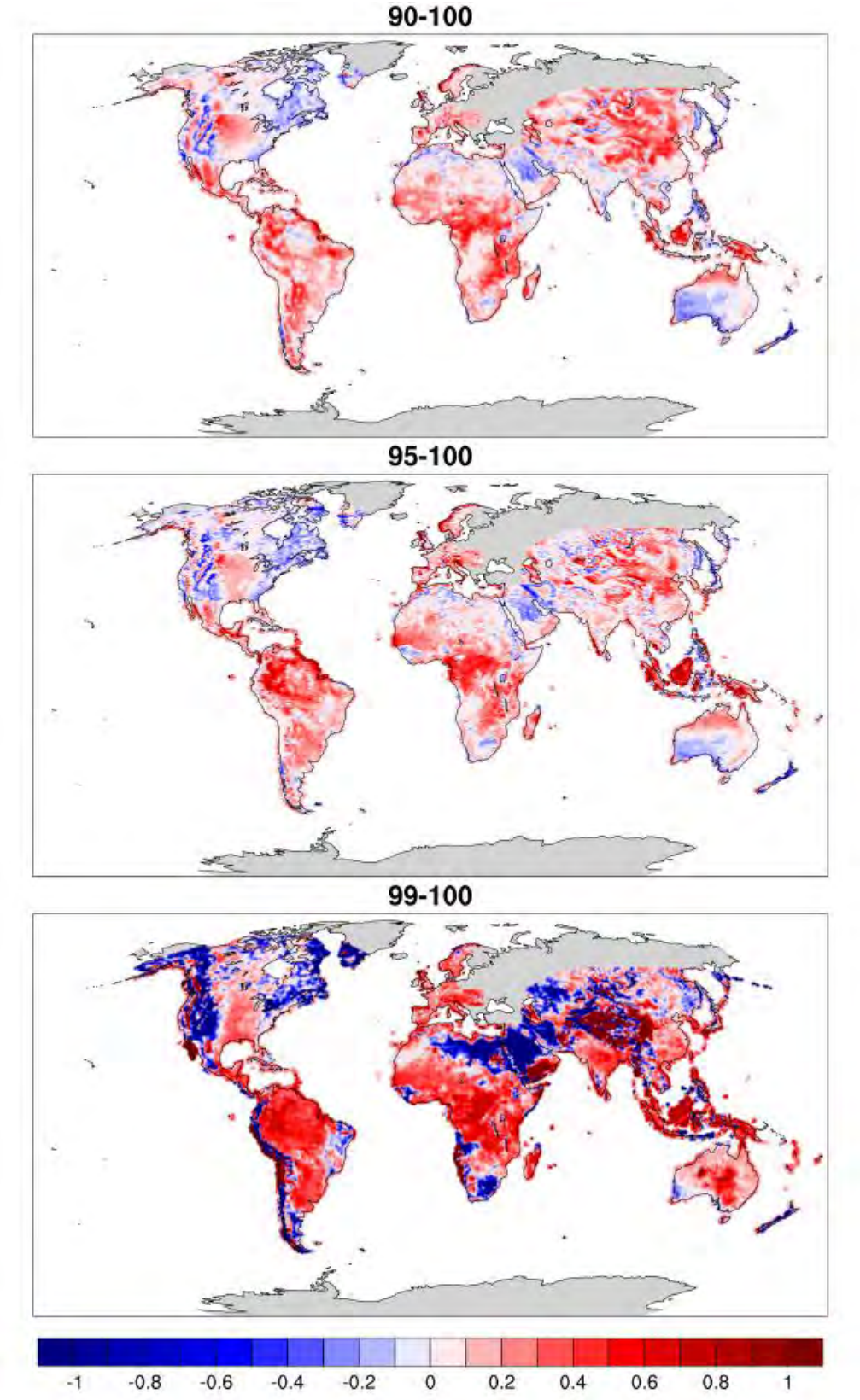

Fig. S11. Added value for RCM ensemble means of the CORDEX-CORE at different percentile intervals compared to CPC at 0.22°. The Europe data used is as Figure 5.

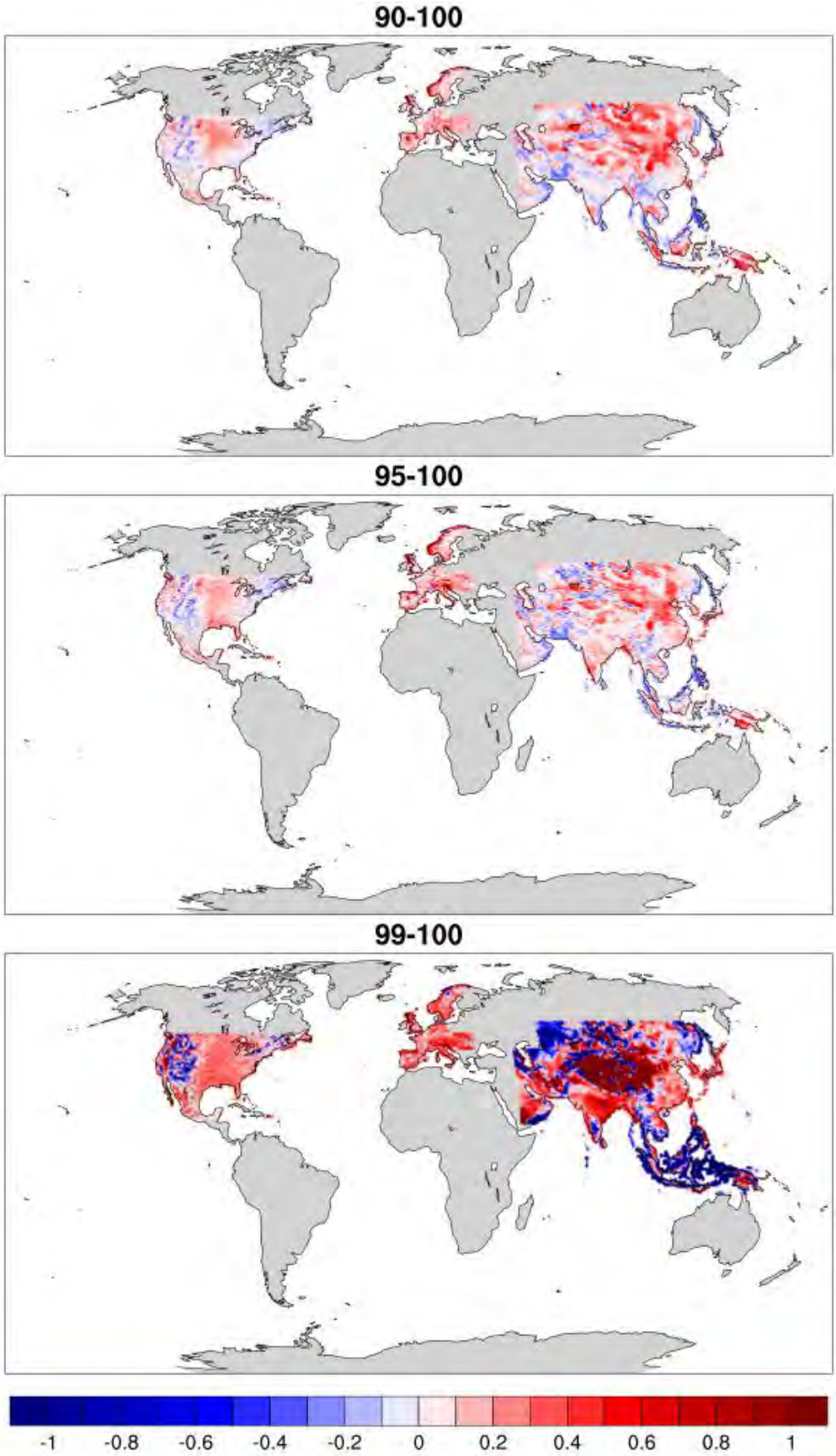

Fig. S12. Added value for RCM ensemble means of the CORDEX-CORE at different percentile intervals compared to GCO at 0.22°. The Europe data used is as Figure 5.

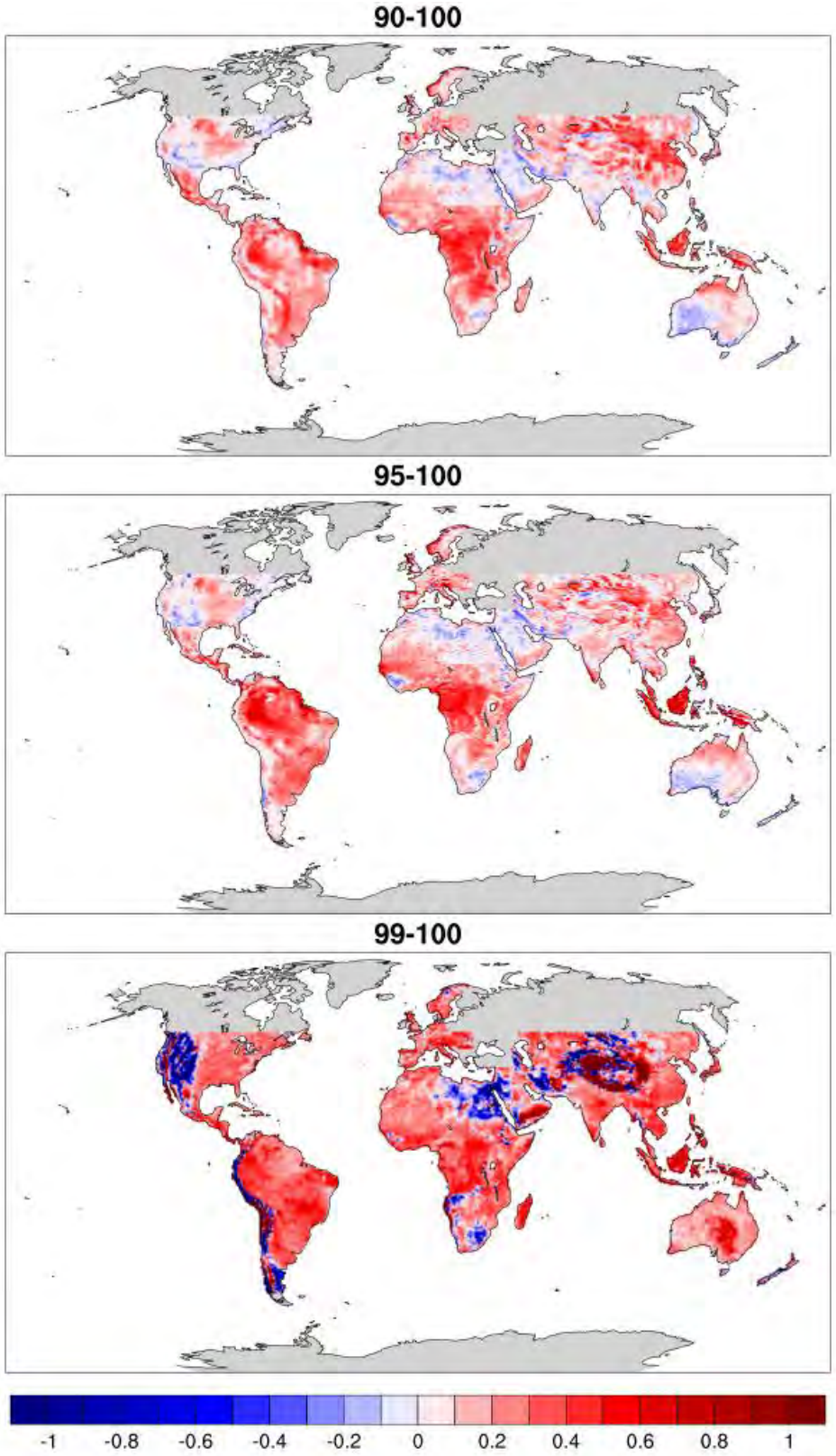

Fig. S13. Added value for RCM ensemble means of the CORDEX-CORE at different percentile intervals compared to TRMM at 0.22°. The Europe data used is as Figure 5.

Table 1. EURO-CORDEX RCM ensemble members and their corresponding driving GCMs (with variant label) used for this analysis.

Driving CMIP5 GCM	Variant	RCM
CCCma-CanESM2	r1i1p1	CLMcom-CCLM4-8-17
CCCma-CanESM2	r1i1p1	GERICS-REMO2015
CNRM-CERFACS-CNRM-CM5	r1i1p1	CLMcom-CCLM4-8-17
CNRM-CERFACS-CNRM-CM5	r1i1p1	SMHI-RCA4
CNRM-CERFACS-CNRM-CM5	r1i1p1	CNRM-ALADIN53
CNRM-CERFACS-CNRM-CM5	r1i1p1	CNRM-ALADIN63
CNRM-CERFACS-CNRM-CM5	r1i1p1	KNMI-RACMO22E
CNRM-CERFACS-CNRM-CM5	r1i1p1	GERICS-REMO2015
CNRM-CERFACS-CNRM-CM5	r1i1p1	DMI-HIRHAM5
CNRM-CERFACS-CNRM-CM5	r1i1p1	IPSL-WRF381P
ICHEC-EC-EARTH	r12i1p1	CLMcom-CCLM4-8-17
ICHEC-EC-EARTH	r12i1p1	CLMcom-ETH-COSMO-crCLIM-v1-1
ICHEC-EC-EARTH	r12i1p1	DMI-HIRHAM5
ICHEC-EC-EARTH	r12i1p1	KNMI-RACMO22E
ICHEC-EC-EARTH	r12i1p1	GERICS-REMO2015
ICHEC-EC-EARTH	r12i1p1	UHOH-WRF361H
ICHEC-EC-EARTH	r1i1p1	DMI-HIRHAM5
ICHEC-EC-EARTH	r1i1p1	KNMI-RACMO22E
ICHEC-EC-EARTH	r3i1p1	DMI-HIRHAM5
ICHEC-EC-EARTH	r3i1p1	KNMI-RACMO22E
ICHEC-EC-EARTH	r1i1p1	SMHI-RCA4
ICHEC-EC-EARTH	r12i1p1	SMHI-RCA4
ICHEC-EC-EARTH	r3i1p1	SMHI-RCA4
MOHC-HadGEM2-ES	r1i1p1	CLMcom-CCLM4-8-17
MOHC-HadGEM2-ES	r1i1p1	DMI-HIRHAM5
MOHC-HadGEM2-ES	r1i1p1	KNMI-RACMO22E
MOHC-HadGEM2-ES	r1i1p1	SMHI-RCA4
MOHC-HadGEM2-ES	r1i1p1	GERICS-REMO2015

MOHC-HadGEM2-ES	r1i1p1	ICTP-RegCM4-6
MOHC-HadGEM2-ES	r1i1p1	UHOH-WRF361H
MOHC-HadGEM2-ES	r1i1p1	IPSL-WRF381P
MOHC-HadGEM2-ES	r1i1p1	CNRM-ALADIN63
IPSL-IPSL-CM5A-MR	r1i1p1	SMHI-RCA4
IPSL-IPSL-CM5A-MR	r1i1p1	IPSL-WRF381P
IPSL-IPSL-CM5A-MR	r1i1p1	KNMI-RACMO22E
MIROC-MIROC5	r1i1p1	CLMcom-CCLM4-8-17
MIROC-MIROC5	r1i1p1	GERICS-REMO2015
MPI-M-MPI-ESM-LR	r1i1p1	CLMcom-CCLM4-8-17
MPI-M-MPI-ESM-LR	r1i1p1	KNMI-RACMO22E
MPI-M-MPI-ESM-LR	r1i1p1	SMHI-RCA4
MPI-M-MPI-ESM-LR	r1i1p1	MPI-CSC-REMO2009
MPI-M-MPI-ESM-LR	r1i1p1	ICTP-RegCM4-6
MPI-M-MPI-ESM-LR	r1i1p1	UHOH-WRF361H
MPI-M-MPI-ESM-LR	r1i1p1	CLMcom-ETH-COSMO-crCLIM-v1-1
MPI-M-MPI-ESM-LR	r2i1p1	MPI-CSC-REMO2009
MPI-M-MPI-ESM-LR	r2i1p1	CLMcom-ETH-COSMO-crCLIM-v1-1
MPI-M-MPI-ESM-LR	r3i1p1	SMHI-RCA4
MPI-M-MPI-ESM-LR	r3i1p1	GERICS-REMO2015
MPI-M-MPI-ESM-LR	r3i1p1	CLMcom-ETH-COSMO-crCLIM-v1-1
NCC-NorESM1-M	r1i1p1	DMI-HIRHAM5
NCC-NorESM1-M	r1i1p1	KNMI-RACMO22E
NCC-NorESM1-M	r1i1p1	GERICS-REMO2015
NCC-NorESM1-M	r1i1p1	SMHI-RCA4
NCC-NorESM1-M	r1i1p1	IPSL-WRF381P
NCC-NorESM1-M	r1i1p1	CLMcom-ETH-COSMO-crCLIM-v1-1

Table 2. CORDEX-CORE RCM ensemble members for each domain (excluding Europe) and their corresponding driving GCMs used for this analysis. The 55 EURO-CORDEX model ensemble described in Table 1 was included with these domains.

Domains	**RCMs**	**Driving GCMs**		
AFR-22	RegCM4	NorESM1-M	MPI-ESM-MR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
NAM-22	RegCM4	GFDL-ESM2M	MPI-ESM-LR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
CAM-22	RegCM4	GFDL-ESM2M	MPI-ESM-MR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
SAM-22	RegCM4	NorESM1-M	MPI-ESM-MR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
EAS-22	RegCM4	NorESM1-M	MPI-ESM-MR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
SEA-22	RegCM4	NorESM1-M	MPI-ESM-MR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
WAS-22	RegCM4	NorESM1-M	MPI-ESM-MR	MIROC5
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES
AUS-22	RegCM4	NorESM1-M	MPI-ESM-MR	HadGEM2-ES
	REMO2015	NorESM1-M	MPI-ESM-LR	HadGEM2-ES

Table 3. Observation datasets used to assess the added value of the EURO-CORDEX ensemble.

Region	Source	Resolution	Period used	Reference
Europe	EOBS v20e	0.10°	1995-2014	Haylock et al. (2008)
Alps	EURO4M	5 km	1995-2008	Isotta et al. (2014)
Spain	Spain02	0.11°	1995-2010	Herrera et al. (2015)
France	SAFRAN	8 km	1995-2013	Vidal et al. (2010)
UK	ENG REG	0.11°	1995-2010	Perry et al. (2009)
Norway	KLIMAGRID	1 km	1995-2008	Mohr (2009)
Sweden	PTHBV	4 km	1995-2011	Johansson (2002)
Carpathians	CARPATCLIM	0.10°	1995-2010	Szalai et al. (2013)
Germany	REGNIE	1 km	1995-2014	Rauthe et al. (2013)
Italy	GRIPHO	12 km	2001-2014	Fantini (2019)

Table 4. Observation datasets used to assess the added value of the CORDEX-CORE ensemble.

Domain	Source	Res.	Period used	Reference
Global	CPC	0.50°	1995-2014	Chen et al. (2008)
Global	CHIRPS	0.05°	1995-2014	Funk & Hoell (2015)
Global	TRMM	0.25°	1998-2009	Kummerow et al. (1998)
NAM, CAM	GCOSGHCN	2551 st	1995-2005	Menne et al. (2012)
WAS	IMD	0.25°	1995-2014	Pai et al. (2014)
WAS, EAS, SEA	APHRODITE	0.25°	1995-2007	Yatagai et al. (2009)